Oxford textbook of rheumatology 2nd ed. edition: By by David Isenberg (Editor), et al Oxford University Press;

Sorry, no front pic! By OkdoKeY

Oxford Textbook of Rheumatology

Contents

Volume 1

6 Intervention, rehabilitation, and sports medicine

Mark Harries

6.7 Sports injuries
J. R. Jenner and M. Shirley Emerson

Contributors list

D. ADU
Consultant Renal Physician, Queen Elizabeth Hospital, Birmingham, UK
5.11.3 Classical polyarteritis nodosa, microscopic polyarteritis, and Churg-Strauss syndrome

SANJIV N. AMIN
Consultant Rheumatologist, Bombay Hospital, India
5.3.7 Mycobacterial diseases

ANDREW A. AMIS
Reader in Orthopaedic Biomechanics, Imperial College London, UK
2.7 Biomechanics of articulations and derangements in disease

BERNARD P. AMOR
Professor of Rheumatology, Hôpital Cochin, University of Paris V, France
5.5.5 Reactive arthropathy, Reiter's syndrome, and enteric arthropathy in adults

JANICE R. ANDERSON
Consultant Neuropathologist, Addenbrooke's Hospital, Cambridge, UK
4.10 Histopathology

BARBARA M. ANSELL
Consultant Rheumatologist, Windsor, Berkshire, UK
5.4.3 Juvenile rheumatoid arthritis (rheumatoid factor positive polyarthritis)

BARBARA M. ANSELL
Consultant Rheumatologist, Windsor, Berkshire, UK
5.17.4 Diseases of bone and cartilage in children

PAUL A. BACON
Arthritis and Rheumatism Council Professor of Rheumatology and Head of Department of Rheumatology, University of Birmingham Medical School, UK
5.11.3 Classical polyarteritis nodosa, microscopic polyarteritis, and Churg-Strauss syndrome

ROSAMONDE E. BANKS
ICRF Cancer Medicine Research Unit, St James's University Hospital, Leeds, UK
4.1 Acute phase response

ROSAMONDE E. BANKS
ICRF Cancer Medicine Research Unit, St James's University Hospital, Leeds, UK
4.3 Biochemistry

JOHN I. BALL
Nuffield Professor of Clinical Medicine, Nuffield Department of Clinical Medicine, University of Oxford; Consultant Physician, John Radcliffe Hospital, Oxford, UK
3.3 Specific immune responses

THOMAS G. BENEDEK
Professor of Medicine, University of Pittsburgh School of Medicine, Pennsylvania, USA
5.19.1 Rheumatic diseases and neoplasia

ROBERT M. BERNSTEIN
Consultant Rheumatologist, Manchester Royal Infirmary, UK
5.19.3 Rheumatic complications of drugs and toxins

A. K. BHALLA
Consultant Rheumatologist, Royal National Hospital for Rheumatic Diseases, Bath, Avon, UK
5.17.1 Osteoporosis and osteomalacia

ALLAN I. BINDER
Consultant Rheumatologist, Lister Hospital, North Hertfordshire NHS Trust, Stevenage, Hertfordshire, UK
1.3.9 An anaesthetic perspective

ALLAN I. BINDER
Consultant Rheumatologist, Lister Hospital, North Hertfordshire NHS Trust, Stevenage, Hertfordshire, UK
5.18.2 Cervical pain syndromes

ALLAN I. BINDER
Consultant Rheumatologist, Lister Hospital, North Hertfordshire NHS Trust, Stevenage, Hertfordshire, UK
6.5 Corticosteroid injection therapy

HOWARD A. BIRD
Reader in Rheumatology, University of Leeds, UK
4.1 Acute phase response

HOWARD A. BIRD
Reader in Rheumatology, University of Leeds, UK
4.3 Biochemistry

CAROL M. BLACK
Professor of Rheumatology, Academic Unit of Rheumatology and Connective Tissue Diseases, Royal Free Hospital Medical School, London, UK
5.8 Scleroderma and related disorders in adults and children

DAVID R. BLAKE
Arthritis and Rheumatism Council Professor of Rheumatology and Consultant Physician, St Bartholomew's and Royal London Hospital School of Medicine and Dentistry, Queen Mary and Westfield College, University of London, UK
2.5 The physiology of the joint and its disturbance in inflammation

TOMÁS S. BOCANEGRA
Senior Director, Arthritis Clinical Group, Clinical Research, G.D. Searle & Co., Skokie, Illinois, USA
5.3.9 Parasitic involvement

EMILIO BOUZA

Professor and Chief of the Division of Clinical Microbiology and Infectious Diseases, Hospital General Universitario 'Gregorio Maranon', Madrid, Spain
5.3.3 Osteomyelitis and associated conditions

S. J. BOWMAN
Senior Lecturer in Rheumatology, University of Birmingham, UK
3.3 Specific immune responses

BARRY BRESNIHAN
Professor of Rheumatology, University College Dublin, and Consultant Rheumatologist, St Vincent's Hospital, Dublin, Ireland
5.13.5 Sarcoidosis

PETER BROOKS
Professor of Medicine, University of New South Wales, St Vincent's Hospital, Sydney, Australia
3.5.1 Non-steroidal anti-inflammatory drugs

CAROLYN P. CACHO
Assistant Professor of Medicine, Case Western Reserve School of Medicine, Cleveland, Ohio, USA
1.3.11 The kidney

ANDREI CALIN
Consultant Rheumatologist, Royal National Hospital for Rheumatic Diseases, Bath, Avon, UK
5.5.1 Spondylarthropathy, undifferentiatated spondylarthritis, and overlap

ANDREI CALIN
Consultant Rheumatologist, Royal National Hospital for Rheumatic Diseases, Bath, Avon, UK
5.5.3 Ankylosing spondylitis

EVAN CALKINS
Emeritus Professor of Medicine, School of Medicine and Biomedical Sciences, State University of New York at Buffalo; Senior Physician and Coordinator, Geriatric Programs, Health Care Plan, Buffalo,New York, USA
1.1.3 The geriatric age group

JEFFREY P. CALLEN
Professor of Medicine (Dermatology); Chief, Division of Dermatology, University of Louisville School of Medicine, Kentucky, USA
5.13.3 Panniculitis

JEFFREY P. CALLEN
Professor of Medicine (Dermatology); Chief, Division of Dermatology, University of Louisville School of Medicine, Kentucky, USA
5.13.4 Neutrophilic dermatoses

T. E. CAWSTON
Professor of Rheumatology, The Medical School, University of Newcastle upon Tyne, UK
5.15 Osteoarthritis

JUSTIN COBB
Consultant Orthopaedic Surgeon, University College Hospital, London, UK
6.1 Surgery in adults

ADRIAN J. CRISP
Consultant Rheumatologist, Addenbrooke's Hospital, Cambridge, UK
5.17.2 Paget's disease of bone

CHAD DEAL
Associate Professor of Medicine, Case Western Reserve University School of Medicine, Cleveland, Ohio, USA
1.3.7 The endocrine system

A. M. DENMAN
Consultant Rheumatologist and Clinical Immunologist, Northwick Park Hospital, Harrow, Middlesex, UK
3.5.2 Antirheumatic drugs

A. M. DENMAN
Consultant Rheumatologist and Clinical Immunologist, Northwick Park Hospital, Harrow, Middlesex, UK
3.5.3 Therapeutic immunomodulation

CHRISTOPHER P. DENTON
Clinical Research Fellow, Academic Unit of Rheumatology, Royal Free Hospital, London, UK
5.8 Scleroderma and related disorders in adults and children

STANLEY C. DERESINSKI
Clinical Professor of Medicine, Stanford University; Director, AIDS Health Services; Associate Chief of Infectious Diseases, Santa Clara Valley Medical Center, San Jose, California, USA
5.3.10 Fungal arthritis

KOMAL B. DESAI
Children's Mercy Hospital, Kansas City, Missouri, USA
5.3.2 Pyogenic arthritis in children

MICHAEL J. DILLON
Consultant Physician and Senior Clinical Nephrologist, The Hospital for Sick Children, London, UK
5.11.8 Primary vasculitis in children

MICHAEL DOHERTY
Professor of Rheumatology, Nottingham University Medical School, UK
5.15 Osteoarthritis

FRANCES DORMON
Consultant in Anaesthetics and Intensive Care, Lister Hospital, Stevenage, Hertfordshire, UK
1.3.9 An anaesthetic perspective

GORDON W. DUFF
Lord Florey Professor of Molecular Medicine, Royal Hallamshire Hospital, Sheffield, UK
3.1 Cells and mediators

ALISON A. EDDY
Associate Professor of Pediatrics, The Hospital for Sick Children, University of Toronto, Ontario, USA

5.7.2 Systemic lupus erythematosus in children

ERKKI EEROLA
Head of Laboratory, Department of Medical Microbiology, Turku University, Finland
1.3.6 The gastrointestinal system

MICHAEL EHRENFELD
Clinical Senior Lecturer, Sackler School of Medicine, Tel Aviv University; Consultant Rheumatologist, Rheumatic Disease Unit, Sheba Medical Center, Tel Hashomer, Israel
1.3.8 Oncology and haematology

M. SHIRLEY EMERSON
Clinical Assistant, Peter Wilson Sports Injury Clinic, Addenbrooke's Hospital, Cambridge, UK
6.7 Sports injuries

PAUL EMERY
ARC Professor of Rheumatology and Rheumatologist, Leeds General Infirmary and St James's University Hospital, Leeds, UK
4.2 Haematology

ADEL G. FAM
Professor of Medicine and Head, Division of Rheumatology, Sunnybrook Health Science Centre, University of Toronto, Ontario, Canada
1.2.2.1 The upper limbs in adults

MARC FELDMANN
Head of Cytokine and Immunology Unit, Kennedy Institute of Rheumatology and Professor of Cellular Immunology, Charing Cross and Westminster Medical School, London, UK
5.4.1 Immunopathogenesis of rheumatoid arthritis

I. T. FERGUSON
Consultant Neurologist, Southmead Hospital, Bristol, UK
1.3.3 Neurological complications

ØYSTEIN FØRRE
Professor and Head, Centre for Rheumatic and Autoimmune Connective Tissue Disease, The National Hospital, Oslo University, Norway
4.7 Gene marker analysis

ANDREW O. FRANK
Consultant Physician in Rheumatology and Rehabilitation, Northwick Park Hospital, Harrow, Middlesex, UK
1.2.1.1 Spinal problems in adults

PATRICIA A. FRASER
Assistant Professor of Medicine, Harvard Medical School; Director of Pediatric Rheumatology, Brigham and Women's Hospital, Boston, Massachusetts, USA
5.1.2 Epidemiology of rheumatic diseases in selected non-European populations

CHARLES S.B. GALASKO
Professor of Orthopaedic Surgery, University of Manchester; Consultant Orthopaedic Surgeon, Manchester Children's Hospitals NHS Trust; Consultant Orthopaedic Surgeon, Salford Royal Hospitals NHS Trust, UK
1.2.1.2 Spinal problems in children

PATRICK J. GALLAGHER
Reader in Pathology, Southampton University Hospitals, UK
4.10 Histopathology

ALLAN GIBOFSKY
Professor, Departments of Medicine and Public Health, Cornell University Medical College, New York, USA
5.3.12 Rheumatic fever

DAFNA D. GLADMAN
Professor of Medicine, University of Toronto, Ontario, Canada
5.5.4 Psoriatic arthritis

DAVID N. GLASS
Professor of Pediatrics and Director, Division of Rheumatology, Children's Hospital Medical Center, Cincinnati, Ohio, USA
5.2 Nosology of the chronic inflammatory rheumatic diseases

ELIZABETH M. GRAHAM
Consultant Medical Ophthalmologist, St Thomas's Hospital and The Hospital for Sick Children, London, UK
1.3.10 The eye

CLIVE E.H. GRATTAN
Consultant Dermatologist, West Norwich Hospital, Norfolk, UK
5.11.4 Small vessel vasculitis

IAN D. GRIFFITHS
Consultant Rheumatologist, Freeman Hospital and Senior Clinical Lecturer, University of Newcastle upon Tyne, UK
1.2.3 Extra-articular features of rheumatic diseases

BRIDGET GRIFFITHS
Lecturer, Department of Rheumatology, University of Leeds, UK
4.2 Haematology

WOLFGANG L. GROSS
Professor of Internal Medicine and Rheumatology; Director, Department of Rheumatology, University of Lubeck and Clinic for Rheumatology in Bad Bramstedt, Germany
5.11.2 Wegener's granulomatosis

RENATE HÄFNER
Paediatric Rheumatologist, Rheuma-Kinderklinik, Garmisch-Partenkirchen, Germany
6.4 Rehabilitation of children

VEDAT HAMURYUDAN
Associate Professor of Medicine, Division of Rheumatology, Cerrahpa Ş a Medical Faculty, University of Istanbul, Turkey
5.11.7 Behçet's syndrome

TIM HARDINGHAM

Professor of Biochemistry, School of Biological Sciences, University of Manchester, UK
2.3 Articular cartilage

MARK HARRIES
Consultant Physician, Northwick Park and St Mark's NHS Trust; Honorary Clinical Director, British Olympic Medical Centre, Northwick Park Hospital, Harrow, Middlesex, UK
6.6 Sports medicine

DORIAN O. HASKARD
Sir John McMichael Professor of Cardiovascular Medicine, Royal Postgraduate Medical School and Honorary Consultant Physician, Hammersmith Hospital, London, UK
3.2 Lymphocyte traffic in inflammation

BRIAN HAZLEMAN
Consultant Rheumatologist, Addenbrooke's Hospital, Cambridge, UK
5.14 Soft-tissue rheumatism

SIMON HELFGOTT
Assistant Professor of Medicine, Harvard Medical School and Brigham and Women's Hospital, Boston, Massachusetts, USA
1.3.12 Psychiatric issues in rheumatology

PETER HOLLINGWORTH
Consultant Rheumatologist, Southmead Hospital, Bristol, UK
1.3.3 Neurological complications

ANGELA C. HORSFALL
Senior Research Fellow, Charing Cross Medical School and Honorary Senior Scientist, Kennedy Institute, London, UK
5.7.1 Systemic lupus erythematosus in adults

J. LAWRENCE HOUK
Professor of Clinical Medicine, University of Cincinnati Medical Center, Ohio, USA
4.6 Joint fluid

G. R.V. HUGHES
Head of the Lupus Arthritis Research Unit, Rayne Institute, St Thomas's Hospital, London, UK
5.7.3 The antiphospholipid antibody syndrome

GUNNAR HUSBY
Professor, Oslo Sanitetsforening Rheumatism Hospital, The National Hospital, University of Oslo, Norway
5.13.1 Amyloidosis

DAVID A. ISENBERG
Professor of Rheumatology, University College London, UK
1.3.5 The respiratory system

DAVID A. ISENBERG
Professor of Rheumatology, University College London, UK
5.7.1 Systemic lupus erythematosus in adults

SILVIU ITESCU
Assistant Professor of Surgical Science; Director, Transplantation Immunology, Department of Surgery, Columbia University College of Physicians and Surgeons, New York, USA
5.3.6 HIV infection

MARY ANNE JACKSON
Professor of Pediatrics, University of Missouri School of Medicine and Pediatrician in Infectious Diseases, Children's Mercy Hospital, Kansas City, Missouri, USA
5.3.2 Pyogenic arthritis in children

MALCOLM I.V. JAYSON
Professor of Rheumatology, Rheumatic Diseases Centre, University of Manchester, UK
5.18.1 Intervertebral disc disease and other mechanical disorders of the back

J. R. JENNER
Consultant in Rheumatology, Addenbrooke's NHS Trust, Cambridge, UK
6.7 Sports injuries

VICTORIA A. JOLLIFFE
Senior Registrar, West Norwich Hospital, Norfolk, UK
5.11.4 Small vessel vasculitis

ADRIAN JONES
Consultant Rheumatologist, Nottingham City Hospital, UK
5.15 Osteoarthritis

D. A. JONES
Professor of Sport and Exercise Sciences, University of Birmingham, UK
2.6 Skeletal muscle damage

GARY M. KAMMER
Professor of Internal Medicine/Rheumatology, The Bowman Gray School of Medicine, Wake Forest University, Winston-Salem, North Carolina, USA
1.3.11 The kidney

WALTER F. KEAN
Clinical Professor of Medicine (Rheumatology) and Head, Service of Rheumatology, McMaster University Medical Centre, Hamilton, Ontario, Canada
1.3.1.1 Pregnancy

WALTER F. KEAN
Clinical Professor of Medicine (Rheumatology) and Head, Service of Rheumatology, McMaster University Medical Centre, Hamilton, Ontario, Canada
1.3.1.2 Antirheumatic drugs in pregnancy and lactation

CAROL A. KEMPER
Clinical Assistant Professor of Medicine, Stanford University; Co-director, AIDS Health Services, Santa Clara Valley Medical Center, San Jose, California, USA
5.3.10 Fungal arthritis

MUNTHER A. KHAMASHTA

Deputy Director, Lupus Arthritis Research Unit, The Rayne Institute; Honorary Senior Lecturer, St Thomas's Hospital, UMDS, London, UK
5.7.3 The antiphospholipid antibody syndrome

PNINA LANGEVITZ
Senior Lecturer in Medicine, Tel Aviv University Medical School; Head of Rheumatic Disease Unit, Sheba Medical Center, Tel Hashomer, Israel
5.13.2 Familial Mediterranean fever

C. B. D. LAVY
Associate Professor, University of Malawi School of Medicine and Honorary Consultant, University College London Hospitals, UK
6.1 Surgery in adults

RONALD M. LAXER
Associate Professor of Pediatrics and Medicine, University of Toronto; Head, Division of Rheumatology, The Hospital for Sick Children, Toronto, Ontario, Canada
5.6.2 Systemic-onset juvenile chronic arthritis

ALISON M. LEAK
Consultant Rheumatologist, Thanet General Hospital, Margate, Kent, UK
1.3.10 The eye

KENG HONG LEONG
Consultant Rheumatologist, Department of Rheumatology and Immunology, Tan Tock Seng Hospital, Singapore
5.13.8 Hyperlipidaemias

MATTHEW H. LIANG
Professor of Medicine, Harvard Medical School, and Professor of Health Policy and Mangement, Harvard School of Public Health, Boston, Massachusetts, USA
1.1.5 Outcomes assessment in rheumatology

GEOFFREY O. LITTLEJOHN
Associate Professor of Medicine and Director of Rheumatology, Monash Medical Centre, Melbourne, Australia
5.19.2 Algodystrophy (reflex sympthetic dystrophy syndrome)

AVI LIVNEH
Senior Physician, Sheba Medical Center and Senior Lecturer, Sackler Faculty of Medicine, Tel Aviv University, Israel
5.13.2 Familial Mediterranean fever

DANIEL J. LOVELL
Associate Professor of Pediatrics, Children's Hospital Medical Center, University of Cincinnati Medical Center, Ohio, USA
1.1.7 Growth and skeletal maturation

C. R. LOVELL
Consultant Dermatologist, Royal United Hospital, Bath, Avon, UK
1.3.2 The skin

MICHAEL E. LUGGEN
Associate Professor of Clinical Medicine, University of Cincinnati Medical Center, Ohio, USA
1.3.4 The cardiovascular system

P. J. MADDISON
Consultant Rheumatologist, Gwynedd Hospitals NHS Trusts, Bangor, UK
1.3.2 The skin

P. J. MADDISON
Consultant Rheumatologist, Gwynedd hospitals NHS Trusts, Bangor, UK
4.5 Autoantibody profile

P. J. MADDISON
Consultant Rheumatologist, Gwynedd hospitals NHS Trusts, Bangor, UK
5.13.7 Multicentric reticulohistiocytosis

P. J. MADDISON
Consultant Rheumatologist, Gwynedd hospitals NHS Trusts, Bangor, UK
5.17.3 Diseases of bone, cartilage, and synovium

RAVINDER N. MAINI
Director, Kennedy Institute of Rheumatology; Head and Professor of Rheumatology, Charing Cross and Westminster Medical School; Consultant Physician, Charing Cross Hospital, London, UK
5.4.1 Immunopathogenesis of rheumatoid arthritis

PAUL MAPP
Research Fellow, Bone and Joint Research Unit, The London Hospital Medical College, UK
2.5 The physiology of the joint and its disturbance in inflammation

NEIL JOHN McHUGH
Consultant Senior Lecturer in Rheumatology, Royal National Hospital for Rheumatic Diseases, Bath, Avon, UK
5.12 Overlap syndromes in adults and children

NEIL JOHN MCHUGH
Consultant Senior Lecturer in Rheumatology, Royal National Hospital for Rheumatic Diseases, Bath, Avon, UK
5.13.8 Hyperlipidaemias

ELIZABETH D. MELLINS
Associate Professor of Pediatrics, Stanford University Medical Center, California, USA
5.6.1 Pauciarticular-onset juvenile chronic arthritis

MARC L. MILLER
Clinical Associate Professor of Medicine, Unversity of Vermont School of Medicine, USA
5.3.1 Pyogenic arthritis in adults

H. M. MOUTSOPOULOS
Professor and Chairman, Department of Pathophysiology, Medical School, University of Athens, Greece
5.10 Sjögren's syndrome

PATRICIA MUÑOZ
Assistant Professor in Clinical Microbiology and Infectious Diseases, Hospital General Maranón, Universidad Complutense, Madrid, Spain

5.3.3 Osteomyelitis and associated conditions

STANLEY J. NAIDES
Associate Professor, Division of Rheumatology, Department of Internal Medicine, University of Iowa College of Medicine, and Clinical Investigator, Veterans Affairs Medical Center, Iowa City, USA
5.3.5 Viral arthritis

BARBARA S. NEPOM
Research Associate Member, Virginia Mason Research Center, and Research Associate Professor, Department of Pediatrics, University of Washington School of Medicine, Washington, USA
5.6.1 Pauciarticular-onset juvenile chronic arthritis

S. P. NEWMAN
Professor of Health Psychology, University College London Medical School, London, UK
1.1.6 Psychological aspects of rheumatic disease

LAUREN M. PACHMAN
Professor of Pediatrics and Head, Division of Pediatric Immunology/Rheumatology, Northwestern University Medical School, Chicago, Illinois, USA
5.9.2 Polymyositis and dermatomyositis in children

ELISABETH PAICE
Consultant Rheumatologist, Whittington Hospital, London, UK
1.1.4 Principles of examination

G. S. PANAYI
Arthritis and Rheumatism Council Professor of Rheumatology, United Medical and Dental Schools of Guy's and St Thomas' Hospitals, London, UK
5.11.5 Polymyalgia rheumatica

E. PASCUAL
Professor of Rheumatology, University of Alicante School of Medicine, Spain
5.3.8 Brucellar arthritis

MURRAY H. PASSO
Associate Professor of Clinical Pediatrics, University of Cincinnati College of Medicine, Children's Hospital Medical Center, Ohio, USA
5.5.2 Spondylarthropathies in childhood

REIJO PELTONEN
Specialist in Infectious Diseases, Department of Medicine, Turku University Central Hospital, Finland
1.3.6 The gastrointestinal system

JOHN S. PERCY
Professor of Medicine, University of Alberta, Edmonton, Canada
1.1.1 The adult patient

ROSS E. PETTY
Professor of Pediatrics, University of British Columbia, Vancouver, Canada
1.1.2 Children and adolescents

RAFAL PLOSKI
Research Fellow, Centre for Rheumatic and Autoimmune Connective Tissue Disease, The National Hospital, Oslo University, Norway
4.7 Gene marker analysis

F. MICHAEL POPE
Head of Medical Research Council Connective Tissue Genetics Group, Addenbrooke's Hospital, Cambridge, UK
2.2 Molecular abnormalities of collagen and connective tissue

MORDECHAI PRAS
Professor of Medicine, Faculty of Medicine, Tel Aviv University, Israel
5.13.2 Familial Mediterranean fever

ANNE-MARIE PRIEUR
Consultant Paediatrician, Unite Fonctionnelle de Rhumatologie Pediatrique, Hôpital Necker-Enfants Malades, Paris, France
5.6.3 Rheumatoid factor- negative polyarthritis in children ('seronegative polyarthritis')

ANNE-MARIE PRIEUR
Consultant Paediatrician, Unite Fonctionnelle de Rhumatologie Pediatrique, Hôpital Necker-Enfants Malades, Paris, France
5.13.6 The chronic, infantile, neurological, cutaneous, and articular syndrome (CINCA)

J.-F QUIGNODON
Paediatric Radiologist, Hôpital des Enfants Malades, Paris, France
4.9.2 Imaging in children

DANIEL W. RAHN
Professor and Vice Chairman, Department of Medicine, Medical College of Georgia, Augusta, USA
5.3.4 Lyme disease

J. RECKLESS
Consultant Physician, Royal United Hospital, Bath, and Honorary Senior Lecturer in Biochemistry and in the Postgraduate Medical School, University of Bath, UK
5.13.8 Hyperlipidaemias

PETER RENTON
Consultant Radiologist, University College Hospital and The Royal National Orthopaedic Hospital, London, UK
4.9.1 Imaging in adults

CHARLES RICHARDSON
Faculty of Medicine and Dentistry, University of Birmingham, UK
4.2 Haematology

MALCOLM P. ROGERS
Associate Professor of Psychiatry, Harvard Medical School and Attending Psychiatrist, Division of Psychiatry, Brigham and Women's Hospital, Boston, Massachusetts, USA
1.3.12 Psychiatric issues in rheumatology

ANN K. ROSENTHAL
Assistant Professor of Medicine, Division of Rheumatology, Medical College of Wisconsin, Milwaukee, USA

5.16 Crystal arthropathies

JOAN M. ROUND
Honorary Research Fellow, School of Sport and Exercise Sciences, University of Birmingham, UK
2.6 Skeletal muscle damage

ANTHONY S. RUSSELL
Professor of Medicine, University of Alberta, Edmonton, Canada
1.1.1 The adult patient

OLIVER SANGHA
Harvard Medical School, Brigham and Women's Hospital, Boston, Massachusetts, USA
1.1.5 Outcomes assessment in rheumatology

HANS-GEORG SCHAIBLE
Professor of Physiology, University of Würzburg, Germany
2.8 The neurophysiology of pain

RAYFEL SCHNEIDER
Assistant Professor of Pediatrics, University of Toronto; Staff Rheumatologist, The Hospital for Sick Children, Toronto, Ontario, Canada
5.6.2 Systemic-onset juvenile chronic arthritis

DAVID G.I. SCOTT
Department of Rheumatology, Norfolk and Norwich Hospital, UK
5.11.1 Classification of vasculitis

DAVID G.I. SCOTT
Department of Rheumatology, Norfolk and Norwich Hospital, UK
5.11.6 Large vessel vasculitis

GEOFFREY SCOTT
Consultant Microbiologist, University College London Hospitals, London, UK
4.4 Microbiology and diagnostic serology

G. H. SEBAG
Paediatric Radiologist, Hôpital Robert Debré, Université Saint Louis, Paris, France
4.9.2 Imaging in children

DAVID D. SHERRY
Director, Pediatric Rheumatology, Children's Hospital and Medical Center, Seattle; Associate Professor of Pediatrics, University of Washington, USA
5.6.1 Pauciarticular-onset juvenile chronic arthritis

YEHUDA SHOENFELD
Professor and Head, Department of Medicine 'B', Chaim Sheba Medical Center, Tel-Hashomer, Israel
1.3.8 Oncology and Haematology

ALAN J. SILMAN
ARC Professor of Rheumatic Disease Epidemiology, University of Manchester, UK
5.1.1 Epidemiology and the rheumatic diseases

EARL D. SILVERMAN
Associate Professor of Pediatrics and Immunology, Division of Rheumatology, Hospital for Sick Children, University of Toronto, Canada
5.7.2 Systemic lupus erythematosus in children

ROBERT W. SIMMS
Director, Boston City Hospital Arthritis Clinic; Associate Professor of Medicine, Boston University School of Medicine, Massachusetts, USA
1.2.2.2 The lower limbs in adults

ROGER SMITH
Consultant Physician, Nuffield Orthopaedic Centre, Oxford, UK
2.4 Bone in health and disease

ENRIQUE ROBERTO SORIANO
Consultant Rheumatologist, Hospital Italiano de Buenos Aires, Argentina
5.12 Overlap syndromes in adults and children

TAUNTON R. SOUTHWOOD
Senior Lecturer and Consultant in Paediatric Rheumatology, Department of Rheumatology, University of Birmingham, UK
5.5.2 Spondylarthropathies in childhood

MARIANNE SPAMER
Physiotherapist, Rheuma-Kinderklinik, Garmisch-Partenkirchen, Germany
6.4 Rehabilitation of children

STEPHEN G. SPIRO
Consultant Physician, University College London Hospitals Trust, UK
1.3.5 The respiratory system

CLIFF. R. STEVENS
Head of Vascular Biology, Bone and Joint Research Unit, St Bartholomew's and The Royal London Hospital School of Medicine and Dentistry, London, UK
2.5 The physiology of the joint and its disturbance in inflammation

GEROLD STUCKI
Lecturer, Department of Rheumatology and Physical Medicine, University of Zurich, Switzerland
1.1.5 Outcomes assessment in rheumatology

MALCOLM SWANN
Consultant Orthopaedic Surgeon
6.2 Surgery in in children

IRA N. TARGOFF
Associate Professor of Medicine, University of Oklahoma Health Sciences Center; Staff Physician, Department of Veterans Affairs Medical Center and Assistant Member, Oklahoma Medical Research Foundation, Oklahoma City, USA
5.9.1 Polymyositis and dermatomyositis in adults

DOUGLAS THOMPSON
Principal Biochemist, Leeds General Infirmary, UK
4.1 Acute phase response

DOUGLAS THOMPSON
Principal Biochemist, Leeds General Infirmary, UK
4.3 Biochemistry

ANTOINE A. TOUBERT
Laboratoire d'Immunogénétique Humaine, INSERM U396, Hôpital Saint-Louis, Paris, France
5.5.5 Reactive arthropathy, Reiter's syndrome, and enteric arthropathy in adults

LYNNE TURNER-STOKES
Director, Regional Rehabilitation Unit, Northwick Park Hospital, London, UK
6.3 Rehabilitation of adults

ATHANASIOS G. TZIOUFAS
Consultant Rheumatologist, Department of Pathophysiology, Medical School, University of Athens, Greece
5.10 Sjögren's syndrome

ANTHONY UNSWORTH
Professor of Engineering and Director of the Centre for Biomedical Engineering, University of Durham, UK
2.7 Biomechanics of articulations and derangements in disease

W. B. VAN DEN BERG
Professor of Experimental Rheumatology and Head, Rheumatology Research Laboratory, University of Nijmegen, The Netherlands
3.4 Animal models of arthritis

MARK J. WALPORT
Professor of Medicine, Royal Postgraduate Medical School, Hammersmith Hospital, London, UK
3.1 Cells and mediators

W. WATSON BUCHANAN
Emeritus Professor of Medicine, McMaster University Faculty of Health Sciences, Sir William Osler Health Institute, Hamilton, Ontario, Canada
1.3.1.1 Pregnancy

W. WATSON BUCHANAN
Emeritus Professor of Medicine, McMaster University Faculty of Health Sciences, Sir William Osler Health Institute, Hamilton, Ontario, Canada
1.3.1.2 Antirheumatic drugs in pregnancy and lactation

RICHARD A. WATTS
Consultant Rheumatologist, Ipswich NHS Trust, UK
5.11.1 Classification of vasculitis

RICHARD A. WATTS
Consultant Rheumatologist, Ipswich NHS Trust, UK
5.11.6 Large vessel vasculitis

A. D. B. WEBSTER
Senior Lecturer, MRC Immunodeficiency Research Group, Royal Free Hospital Medical School, London, UK
5.3.11 Immunodeficiency

LAURA F. WEXLER
Professor of Medicine, University of Cincinnati Medical Center, Ohio, USA
1.3.4 The cardiovascular system

FRANCES B. WHEELER
The Bowman Gray School of Medicine, Wake Forest University, Winston-Salem, North Carolina, USA
1.3.11 The kidney

JOHN T. WHICHER
Professor, Research School of Medicine, University of Leeds, UK
4.1 Acute phase response

JOHN T. WHICHER
Professor, Research School of Medicine, University of Leeds, UK
4.3 Biochemistry

PATIENCE H. WHITE
Associate Professor of Pediatrics and Medicine, George Washington University; Chairman, Pediatric and Adult Rheumatology, Children's National Medical Center, Washington DC, USA
1.2.2.3 The arm and leg in children

BRYAN D. WILLIAMS
Professor of Rheumatology, University Hospital of Wales, Cardiff, UK
4.8 Immune complex detection

ROBERT WINCHESTER
Professor of Medicine and Pediatrics, College of Physicians and Surgeons, Columbia University, New York City, USA
5.3.6 HIV infection

JAY B. WISH
Professor of Medicine, Case Western Reserve University; Medical Director, Hemodialysis Services, University Hospitals of Cleveland, Ohio, USA
1.3.11 The kidney

FRANK A. WOLLHEIM
Professor of Rheumatology, Lund University, Sweden
5.4.2 Rheumatoid arthritis-the clinical picture

PATRICIA WOO
Professor of Paediatric Rheumatology, University College London, UK
1.1.7 Growth and skeletal maturation

PATRICIA WOO

Professor of Paediatric Rheumatology, University College London, UK
2.1 Molecular genetics and its relevance in rheumatology

PATRICIA WOO
Professor of Paediatric Rheumatology, University College London, UK
5.2 Nosology of the chronic inflammatory rheumatic diseases

B. P. WORDSWORTH
Clinical Reader in Rheumatology, Nuffield Department of Clinical Medicine, Oxford University, UK
3.3 Specific immune responses

HASAN YAZICI
Professor of Medicine and Chief, Division of heumatology, Cerrahpa§a Medical Faculty, University of Istanbul, Turkey
5.11.7 Behçet's syndrome

PIERRE YOUINOU
Professor of Immunology, Brest University Medical School, France
5.10 Sjögren's syndrome

ADAM YOUNG
Consultant Rheumatologist, City Hospital, St Albans, Hertfordshire, UK
4.11 Electrophysiology

SEBAHATTIN YURDAKUL
Professor of Medicine, Division of Rheumatology, Cerrahpa§a Medical Faculty, University of Istanbul, Turkey
5.11.7 Behçet's syndrome

JOHN B. ZABRISKIE
Associate Professor and Head, Laboratory of Clinical Microbiology/Immunology, Rockefeller University, New York, USA
5.3.12 Rheumatic fever

DEBORAH ZEMER
Sheba Medical Center, Tell Hashomer, Israel
5.13.2 Familial Mediterranean fever

Dose schedules are being continually revised and new side-effects recognized. Oxford University Press makes no representation, express or implied, that the drug dosages in this book are correct. For these reasons the reader is strongly urged to consult the pharmaceutical company's printed instructions before administering any of the drugs recommended in this book.

Preface to the Second Edition

Four years ago, in 1993, the first edition of this textbook was published to considerable acclaim. The many rapid developments in the basic science of rheumatology, imaging of joints, bones, and soft tissues, and exciting advances in treatment for some of the previously most intractable rheumatic diseases, have persuaded us that a second edition is timely and will contain sufficient new material both to stimulate and inform the reader.

The second edition has benefited, we believe, from the rearrangement of some chapters, the expansion of many others which have been brought up to date, and the addition of several completely new chapters. Our contributors have also been asked to provide expanded reference lists to facilitate access to the original sources. This approach ensures freshness of ideas and style, which is complemented by the improved quantity and quality of the colour figures. We also wished to make this a textbook to which rheumatologists could refer as a guide to their management of both the common and more unusual rheumatic conditions. To facilitate this aim, algorithms of optimal treatment are provided in the clinical chapters with additional practical management suggestions, including a section on joint and soft tissue injection.

The textbook has been designed as an attractive and informative manual for rheumatology trainees, full-time clinicians, and academic rheumatologists. We have encouraged our authors to express their opinions freely, to bring areas of dispute into the open, and to present a balanced view overall. As in the first edition; this volume places special emphasis upon the perspective of the age of the patient when dealing with the presentation of various rheumatic conditions, and a number of chapters focus on paediatric rheumatology. Another important emphasis in this edition is on the overlap of rheumatology with other subspecialties. This is provided by a series of chapters co-authored by rheumatologists and colleagues with expertise in a wide range of related conditions.

It has been both a challenge and a pleasure for us to edit the second edition which we hope will build on the clarity and standards of production of the first.

Peter Maddison
David Isenberg
Patricia Woo
David Glass

Preface to the First Edition

These are exciting times for the study of rheumatology. Current research, incorporated in this textbook and integral to the editors' enthusiasm for their project, continuously increases our knowledge of the molecular basis of many of the rheumatic diseases. These advances more than justify a new textbook written to complement the successful *Oxford Textbook of Medicine*. As with the other volumes in this series, it is designed to be comprehensive and sufficient for both the trainee and the general physician who require up-to-date information.

The book begins with the variations in presentation of rheumatic symptoms at different ages. There follow chapters dealing with those syndromes, which are not easy to classify and which may be best considered as regional. Although the aetiology of these syndromes is still incomplete they are such a significant part of practice that they demand comprehensive cover. Special emphasis has been placed on back symptomatology both in children and in adults as this problem represents an especially large part of rheumatological practice. Chapters in the third part of Section 1 then discuss rheumatic disease in relation not only to general (internal) medicine but also to other specialties, including psychiatry, anaesthesia, obstetrics, and ophthalmology. The text provides both a comprehensive account of these extensive and pervasive interactions and more focused discussions, either on an area of particular interest to the authors or of special clinical interest; hence the variety of approaches ranges from the general view to the in-depth analysis of a specialist topic.

The next two sections, dealing primarily with basic science, include conceptual advances of relevance to rheumatology and provide an understanding of the rationale for newer therapies being introduced into clinical practice. We have not attempted to replace basic science textbooks; rather we have indicated good reviews on specific topics, concentrating in our volume on how the different areas interact within the context of rheumatic diseases. For example, genetic abnormalities of collagen are described in relation to diseases of cartilage and bone. The joint is treated as a functional unit, and its physiological and biomechanical disturbances are described in the context of a variety of diseases. Parts of the immune system currently thought to be important in the pathogenesis of chronic inflammation are highlighted in some detail. The sections finish with a review of available and innovative ways of controlling inflammation.

Volume 1 concludes with discussion of clinical laboratory practices. Considerable advances have been made in developing laboratory and imaging techniques for clinical assessment. In this section, guidance is given on the selection of appropriate investigations as well as an indication of future developments.

Volume 2 contains the necessary systematic and comprehensive review of the rheumatic diseases. We have tried to encompass rheumatic diseases met throughout the world together with their epidemiological and environmental influences. To this end, authors have been selected who have appropriate clinical experience and established reputations in teaching and research. Colour has been used in this volume to enhance the clinical descriptions while the authors provide a personal as well as an informative approach to management. The final section deals with the important aspects of surgery and rehabilitation. A comprehensive review of surgical techniques has not been attempted. Rather, the major principles have been established to enable appropriate referral as well as the early recognition of complications of surgery.

Throughout the text, considerable emphasis is placed on the age of presentation, thus ensuring that a paediatrician faced with a rheumatological problem is well catered for. Referencing has been selective rather than exhaustive; with the knowledge that computer searches are widely available, we feel that the space may be better used for clinical description.

No project of this magnitude is complete without acknowledgements. We would like to thank the staff of Oxford University Press for their unfailing help. Thanks are also due to our staff, Sheena Stewart, Carolyn Keith Haun, Louise Kittredge, Geraldine Brown, Ann Maitland, and Kate Young for their invaluable assistance in preparing the text.

Peter Maddison
David Isenberg
Patricia Woo
David Glass

Authors

P. J. Maddison
Consultant Rheumatologist, Gwynedd Hospitals NHS Trusts, Bangor, UK

David A. Isenberg
Professor of Rheumatology, University College London, London, UK

Patricia Woo
Professor of Paediatric Rheumatology, University College London and Institute of Child Health, London, UK

David N. Glass
Professor of Pediatrics and Director, Division of Rheumatology, Children's Hospital Medical Center, Cincinnati, Ohio, USA

1.1.1 The adult patient

Anthony S. Russell and John S. Percy

Introduction

In rheumatology, as in other subspecialties, the very core of any medical practice is the consultation, during which the physician gives advice to the patient about diagnostic and therapeutic procedures. The physician must recognize that the patient may believe that he or she has come 'for a blood test' or 'for an X-ray' or 'for treatment'. What the patient has come for, in fact, is advice, which may incorporate some if not all of the patient's expectations. This chapter explores the special problems the rheumatologist meets in attempting to fulfil these expectations.

In locomotor disease, it is not uncommon to meet patients who bring what can best be described as a loosely tied bundle of complaints that they wish to transfer to the doctor for instant solution. Trying to help, they may often have diagnostic suggestions, which they believe will ease this task. These suggestions may be of diverse origins and doubtful practicality based on folklore or superstition, often compounded by information from the communications media. The physician, a pragmatist by training and experience, should use these suggestions judiciously to direct the patient along the correct diagnostic and therapeutic pathway.

The experienced rheumatologist often makes a diagnostic and therapeutic management plan by an intuitive process in the first moments of meeting the patient. Dissection of this intuitive exercise is difficult. But, nevertheless, central to this process is the differentiation of systemic from local disease and serious from minor illness. Systemic disease is suggested by the patient looking ill, weight loss, dyspnoea, fever, neuropathy, lymphadenopathy, splenomegaly, rash, anaemia, raised erythrocyte sedimentation rate or C-reactive protein, and abnormal urinalysis.

The interview and examination that follow the initial contact serve to confirm or deny the initial impression. Any surprises that may arise have to be integrated by a feedback mechanism into the continuing process of assessment. For example the discovery of significant weight loss in a patient who sounded as though he had a rotator cuff syndrome will cause reflection and reassessment. Indeed, this is the mechanism by which physicians develop their clinical expertise and continue to learn (Schon 1987). One important purpose of the interview and physical examination, apart from identification of specific problems, is to gain the patient's confidence and thereby encourage compliance with the suggested investigations and treatment. The patient needs to know that the physician is considering all possibilities and is making a thorough assessment, not only of the presenting complaints, but also of the patient's overall physical and mental state. The difference in emphasis of the approach used in children and in elderly people will be discussed in Chapter 1.1.2 and Chapter 1.1.3.

Before the history and physical examination

The process of evaluation begins even before the patients begin to give their history. After reading the physician's referral letter and any accompanying documentation, our own initial assessment begins in the waiting room (Table 1).

Clinical problem	Possible diagnosis
Painful foot, elderly lady on diuretics	Gout
Young, sexually active; hot swollen joint	Gonococcal arthritis
Headache with diffuse aches and pains in an elderly person	Polymyalgia rheumatica
Antinuclear antibody-positive, but no symptoms	Referring physician's dilemma, not the patient's
On allopurinol with a high uric acid but no arthritis	Not gout
Postpubertal male with low back pain	Ankylosing spondylitis
Woman, 6 weeks postpartum; small joint arthritis	Rubella vaccination; rheumatoid arthritis

NB: intuitive diagnoses must be constantly subject to reflection and reassessment.

Table 1 Preconsultation intuitive observations

Many important signs may be noticed when asking for the patient. An accompanying friend or relative may respond, leading a disabled or reluctant patient forward and apparently wishing to take charge of the proceedings. Observation of such interpersonal reactions forms a significant part of the assessment. We also see whether the patient rises from a chair with difficulty, owing to weakness or stiffness, and how he or she walks towards us. Is a walking aid being used? If so, does it seem to be needed? On introduction, with a normal handshake, is there a flinch? Does the gait change in the walk from the waiting to the examining room? What is the patient's general demeanour. Even before the history taking begins, areas of focus will have already been defined in an intuitive manner. These will subsequently be organized in a systematic manner.

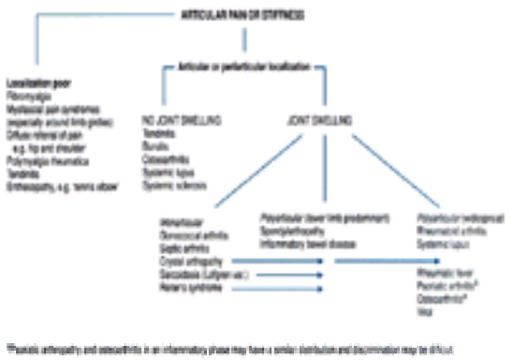
History

When taking the history, the physician should, in due course, attempt to focus the patient's complaints into a differential diagnostic scheme. It is obviously important to determine where the complaints are situated. Are they localized or generalized, episodic, constant, or progressive? Is there any diurnal variation and what, if any, are the aggravating or alleviating factors? In many conditions it is often possible to establish a fairly firm diagnosis from the history alone and, when symptoms are episodic, such a working diagnosis will have to suffice until the patient is seen during a symptomatic period. A 'wait and see' approach is not uncommon in rheumatology but the rationale must be carefully explained to the patient. Thus, the described pattern of joint involvement and characteristic symptoms will often allow the diagnosis of acute gout or palindromic rheumatism to be made with confidence, awaiting objective confirmation during the next episode.

Specific symptoms to be asked for include arthralgias, myalgias, joint swelling, morning stiffness, Raynaud's phenomenon, and skin rashes and paraesthesias.

If there is suspicion of a systemic disease, a full systematic medical inquiry must be embarked upon and any significant symptoms noted. Figure 1 shows what we believe to be the key points in our own intuitive diagnostic process. This becomes modified by the statistics of disease likelihood in our practice (Table 2). Even if the problem appears to be localized, one must be conscious of the fact that systemic diseases (rheumatic and non-rheumatic) often present with local symptoms, for

example carpal tunnel syndrome in rheumatoid arthritis and hypothyroidism (Table 3).



[a]Psoriatic arthropathy and osteoarthritis in an inflammatory phase may have a similar distribution and discrimination may be difficult.

Fig. 1 Aspects of differential diagnosis in rheumatology.

	University (%)	Private practice (%)
Rheumatoid arthritis	14.7	16.9
Degenerative arthritis	17.0	6.6
Ankylosing spondylitis	6.0	2.8
Crystalline arthropathy	3.8	4.1
Septic arthritis	0.7	0.5
Systemic lupus erythematosus	2.8	4.3
Other connective tissue diseases	3.1	—
Juvenile arthritis	0.5	—
Other[a]	51.7	64.7

[a]By courtesy of Dr A. Fitzgerald.
[b]Largely periarticular disorders but includes fibromyalgia.

Table 2 Distribution of rheumatic diseases as seen in a university clinic and in a busy private practice [a]

Symptom/sign	Illness
Weight loss/bone pain	Multiple myeloma
Carpal tunnel syndrome	Acromegaly Hypothyroidism Amyloid
Bone pain	Secondary tumour
Vasculitis (polyarteritis nodosa)	Hepatitis
Stiffness and difficulty in walking	Parkinson's disease
Chronic synovitis with bowel problems	Inflammatory bowel disease
Stiff fingers, shoulder pain	Diabetic cheiroarthropathy

Table 3 Primarily non-rheumatic illnesses presenting in the rheumatology clinic

Pain and stiffness

The majority of adult patients seeking a rheumatological consultation complain of pain and/or stiffness. The pain is usually localized descriptively to their joints, muscles, or bones and the clinician has to determine whether this localization is correct.

It is important to bear in mind that patients are not usually familiar with anatomy and physiology, and that it is often impossible, for example, to elucidate an anatomically correct description of pain and paraesthesias due to compression of the median nerve from a patient with pain in the hand and forearm. Similarly, patients may complain of pain in the hip 'right in the joint' and when asked to indicate the site of the pain, they will point to the greater trochanter or gluteal region, not the groin.

Nevertheless, if a patient presents with a somewhat localized problem (e.g. forearm pain with use), close attention to the anatomical peculiarities of the region and the relevant symptoms and signs will generally provide a clear diagnosis. Localized problems may be multifocal, suggesting a generalized disorder, but a careful examination (e.g. for bursitis and for tendinitis) after the history will help in this distinction. It is important to be wary of the diagnosis of diffuse or even focal osteoarthritis as the cause of symptoms, even when there may be supporting radiographic evidence. Although this condition may be present, it may not be the predominant cause of symptoms (Table 4).

Diagnosis	Clinical pointer
Periarticular shoulder pain	Referred to deltoid insertion
Tennis and golfer's elbow	Diffuse forearm pain on grip-ping
Carpal tunnel	Nocturnal paraesthesiae, often diffuse
Flexor tenosynovitis	Triggering and/or finger pain on gripping (pulp-pinch sign positive)
de Quervain's tenosynovitis	Positive Finkelstein test
Mechanical back pain	Tenderness over gluteals and sacroiliac ligaments frequent
Spondylitic back pain	Morning sacroiliac or gluteal tenderness; morning stiffness marked
Trochanteric bursitis	Nocturnal pain when lying on that side; focal, point tenderness
Hip synovitis	Usually groin and outer thigh pain, occasionally elsewhere
Anserine bursitis	Often nocturnal medial knee pain if knees are touching; localized tenderness

Other, less localized syndromes, e.g. calcified bursitis, pleural basitis, infrapatellar bursitis, should be immediately apparent on examination of the painful area.

Table 4 Clinical pointers in syndromes where pain is poorly localized

Pain is defined by the patient's subjective description. The clinician's task is to attempt to characterize this in terms that are in common medical usage. As indicated, the starting point for an assessment of pain is to describe localization. The quality, intensity, duration, and type of onset, as well as provoking, aggravating, or alleviating factors must be determined, together with the presence of a diurnal variation. However, such technical terms are usually best avoided with the patient. Important and revealing questions are 'What brings on the pain?', 'What can you do to relieve it?' Having become familiar with the customary responses to such questions the rheumatologist is more alert to the patient who seems unable or even unwilling to describe the problem clearly. It is often helpful to try and reproduce the symptoms by manipulation or pressure, particularly when the patient appears unable to pinpoint the location without help. Thus, forearm pain on gripping tightly

may help distinguish the tender origin of the common extensor in a patient with tennis elbow from a similarly tender trigger point found in fibromyalgia.

While the spread of pain outside the usual limits may reflect intensity, this rarely proves a problem with a superficial pain, which is commonly well localized; thus gout or septic arthritis may be associated with severe pain but the origin is obvious. A glomus tumour of the nail bed may cause considerable proximal radiation of pain but examination should localize the source. To locate the source of pain of deeper or visceral origin may be much more difficult because of the phenomenon of 'referral'. Some patterns of referred pain are precise and well known, for example shoulder pain from a diaphragmatic lesion or gallbladder; arm and/or neck pain from myocardial ischaemia. Others are more variable. Thus, the presence of sciatica may reflect nerve-root compression from a lumbar disc but much more commonly represents a referred distribution of pain from one or more regions around the lumbar spine. Experiments involving threads left temporarily after surgery around painful sensory structures or involving injection of hypertonic saline into ligaments and facetal joint structures have shown referral of pain from the lumbar spine into gluteal areas, to the posterior thigh, and even down the calf (Kellgren 1977). Involvement of the upper half of the lumbar spine tends to be referred to the anterior thigh. The only way to distinguish clearly true sciatica from such referred pain is on the basis of neurological signs, as associated symptoms (e.g. a cough or sneeze impulse) may be found with either. Sometimes the pain may be remarkably well localized but again at a point distant from the exciting lesion; thus interscapular pain may be seen with postural/mechanical problems in the cervical spine, and anterior chest pain in relation to inflammation involving the mid-dorsal spine. Sometimes relatively focal or diffuse tenderness may be associated with such referred pain syndromes. Pain referral may prove confusing to the physician but it is also often difficult to explain to a patient with, for example, pain well localized to the deltoid insertion that it is coming from a pericapsular lesion of the shoulder joint. Equally, deep visceral pain, such as a penetrating duodenal ulcer, may of course be referred to skeletal structures. Because of potentially similar patterns of pain referral a differentiation between skeletal and true visceral pain may have to be made on grounds other than the location of the discomfort.

Quality of pain

Skeletal pain may be caused by a variety of problems ranging from ischaemia, inflammation, and nerve entrapment to central factors that lead to a perception of pain where no evident external causes exist. Quite apart from aggravating or relieving factors, the quality of the pain may provide diagnostic pointers. Thus, a burning pain in the feet, especially at night, may suggest a neuropathy. The pain of rheumatoid arthritis is usually steady and aching not agonizing, excruciating, or terrifying. Such descriptions in a patient with rheumatoid arthritis would suggest an alternative explanation, for example, sepsis, fracture, nerve entrapment, or non-organic causes. Chronic pain syndromes are well recognized but their underlying nature remains controversial. Because of overlapping and insufficient criteria a single individual may be designated by one physician as having a myofascial pain, by another as referred pain with secondary depression, and by a third as a non-organic, that is central, pain syndrome. The descriptions used by patients with central pain tend to be extreme and we find that to have patients complete instruments such as the McGill pain questionnaire is helpful as a teaching aid in illustrating to students the range of terms used. Table 5 illustrates some of the terms from that questionnaire which distinguish between central and peripheral pain syndromes. Similar expressions are chosen by patients with fibromyalgia, which is one reason for its earlier name of 'pain magnification syndrome', but also by those with causalgia, reflex sympathetic dystrophy, and occasionally even periarticular shoulder pain. In addition, some individuals or groups of individuals tend to use these extreme terms out of a sense of frustration or anxiety, or because of ethnic convention (Zborowski 1952), and risk being labelled as neurotic because of them, when the presence of organic disease may be missed. The reality of such pain of central origin has often been dramatically demonstrated by the stage hypnotist with a variety of forms of posthypnotic suggestion. The severe quality of the rare thalamic pain syndrome is also well recognized. It is notable that Mark *et al*. observed with stereotactic thalamic surgery that in some, pain sensation could be abolished without loss of the unpleasant emotional effect and in others the emotional effect may disappear and yet the pain remain without its unpleasant qualities (Mark *et al*. 1960). Some of our current pharmacotherapeutic approaches reflect these findings. Therapy that is less conventional may have even more dramatic affects. Organic pain—for example hot metal or hand immersion at 4°C for 45 min—can be blocked both from consciousness and from detection by electroencephalography during some forms of meditation. The block appears to be below the level of the cortex (Davidson and Goleman 1977). While the various central pain syndromes may respond to narcotic analgesics, this is an inappropriate approach both because it decreases the production of endorphins and because it is likely to cement the abnormal pain-behaviour patterns. Some of these patients have post-traumatic pain syndromes with associated medicolegal implications. In any event, as a group they form a large proportion of patients in many rheumatological practices. It is clear that we urgently need better criteria both for diagnosis and classification of pain syndromes, as well as better approaches to treatment.

TABLE cellSpacing=0 cellPadding=0 align=left border=0 hspace="10" vspace="5">

Organic	Non-organic
Pounding	Flickering
Jumping	Shooting
Pricking	Lancinating
Sharp	Lacerating
Pinching	Crushing
Hot	Searing
Tender	Splitting
Nagging	Torturing
Spreading	Piercing
Annoying	Unbearable
Tiring	Exhausting
Fearful	Terrifying
Tight	Tearing

Table 5 Terms from the McGill pain scale that help distinguish between organic and non-organic pain syndromes

Stiffness

Stiffness is an important symptom in rheumatological disease. Typically, patients with rheumatoid arthritis complain of peripheral joint stiffness present for more than 30 min after awakening and recurring after periods of immobility. Patients with peripheral joint osteoarthritis may also have morning stiffness. This is usually of shorter duration but may also recur after rest. Peripheral stiffness may reflect non-articular inflammation and can be marked, as in plantar fasciitis, but is unusual with lesions around the shoulder. Stiffness for less than 5 min is generally of minor significance but could represent, for example, flat feet. Mechanical disorders of the lumbar spine may be associated with morning stiffness of short duration (less than 15 min) whereas patients with ankylosing spondylitis may say that their stiffness persists for several hours. Stiffness present for many hours, and even 'all day', in the absence of gross physical findings usually points to a functional condition such as fibromyalgia.

Is there a history of joint swelling or not?

The cardinal observation made by rheumatologists is that of joint swelling but patients can be amazingly unreliable, in being able to describe a joint swelling that does not exist and has not existed. This is most notorious in fibromyalgia and it is critical to be able to distinguish arthralgias, which may include a variety of lesions of bursas, tendons, and tendon sheaths, from true arthritis (i.e. a swollen joint observed by a competent physician). It is, therefore, a critical part of the subsequent examination to confirm the presence or absence of joint swelling (see below).

Comment

In addition to focusing on differential diagnosis, the history and presentation of symptoms should give the first indication of how patients should be advised to manage their problem(s). Some people will often make light of what is obviously severe rheumatoid arthritis whereas in another circumstance a patient with fibromyalgia will say that they are unable to carry out mundane tasks because of 'knives twisting in the muscles'. The physician's task is to ameliorate the diseases and to provide advice that will allow both types of patient to carry on with normal lives. The stoical patient may elicit a more sympathetic response from the physician but complaints from a medically trivial condition are no less valid or worthy of being taken seriously.

Physical examination

The physician recognizes disease entities from the physical findings in addition to symptoms and needs to integrate these two sets of data. Pain may be localized to joints, the spine, soft tissue, muscle, or bone. If it is localized in joints, can swelling be observed? If so, is the swelling severely inflammatory, moderately inflammatory, or non-inflammatory? This issue of joint swelling and its nature is one of the most critical aspects of the rheumatological examination and is reflected in the importance of teaching proper techniques for the 'wipe' or 'bulge' test in the knee, the proper palpation of the interphalangeal joints, metatarsophalangeal joints, etc. It is perhaps

due to inability to assess swelling in the hip joint that the assessment of the range of motion is most important in this joint. The pattern of joint involvement is very important: is it monoarticular, polyarticular, symmetrical, asymmetrical, widespread, or restricted to lower extremities? Is there a surrounding cellulitis? Severe monoarticular arthritis with or without surrounding cellulitis suggests a crystal-induced arthritis, septic arthritis, sometimes Reiter's syndrome, gonococcal arthritis, or the Löfgren variant of sarcoidosis. These conditions may affect more than one joint, in which case the distribution is usually asymmetrical and, in gout, may involve contiguous joints of the same limbs. Widespread inflammatory joint swelling of moderate or severe degree may be seen in rheumatoid arthritis, psoriatic arthritis, polyarticular crystal-deposit disease, systemic lupus erythematosus, and other diffuse diseases of connective tissue; systemic conditions such as Whipple's disease may also present in this fashion. Joint swelling, without much evidence of inflammation, usually indicates osteoarthritis, which may involve distal and proximal finger joints—a classical feature of this predominantly female familial condition. Involvement of the metacarpophalangeal joints, especially if seen with chondrocalcinosis, should suggest the possibility of previously unrecognized haemochromatosis in the family, and premature osteoarthritis in a variety of joints makes one also consider conditions such as alcaptonuria, Wilson's disease, or epiphyseal dysplasias.

During the course of the examination, fairly obvious findings may give substantial clues to the diagnosis. The facial rash of systemic lupus erythematosus, the scaling skin lesions accompanying psoriatic arthritis, and the specific lesion associated with Lyme arthritis are examples that readily come to mind and are discussed further in Chapter 5.3.4.

Tests

Think hard before you order them, inappropriate testing causes more problems than it solves.

Many of the tests that are available (see Section 4) are often abnormal in a range of rheumatic and non-rheumatic diseases and so, for example, the commonly used rheumatoid factor and antinuclear antibody tests are rarely sufficient in themselves to arrive at a precise diagnosis. Clinical experience teaches that the results of laboratory tests must be interpreted with caution and with continual reference to the clinical presentation. This is not because the tests used in the rheumatology subspecialty are less reliable than those of other disciplines but they are less specific and, more importantly, as many of these diseases are uncommon, the history and physical examination remain pre-eminent. Most of us would request a recent full blood picture and urinalysis on every patient, regardless of diagnosis. The erythrocyte sedimentation rate is still used by many despite the recognized pitfalls of false positive and negative results. It nevertheless remains difficult to support a diagnosis of polymyalgia rheumatics, for example, with a relatively normal sedimentation rate; on the other hand, it is comfortably reassuring to see a normal result in a patient in whom we have found no evidence of inflammatory disease. Selective lymphopenia of less than 1000/ml is the poor man's indication of systemic lupus or a recent virus infection. Thrombocytosis is common in active rheumatoid arthritis and the vasculitides.

As with other tests, it is important to know why the test is being ordered and what influence the result will have on diagnosis and management. A test for rheumatoid factor may be ordered in a patient with clinically obvious nodular rheumatoid arthritis not for diagnosis but for better definition of the condition. A serum uric acid level is of no value in the diagnosis of gout; there are simply too many non-gout individuals with raised levels and too many normal levels seen even during the acute attack to make this helpful. It is, nevertheless, an effective way to monitor compliance with hypouricaemic therapy. In a patient with mechanical back symptoms it is clear that a raised uric acid or a positive rheumatoid factor, for example, are not going to influence diagnostic considerations; they should, therefore, not have been ordered.

A consideration of pre- and post-test probabilities is critical. A positive fluorescent antinuclear antibody (FANA) test is about 97 per cent specific and 99 per cent sensitive for systemic lupus erythematosus. These figures sound, and are, excellent but it remains true that as systemic lupus has a prevalence of around 0.1 per cent of the population or even less, there are, in a given population, more normal people positive for antinuclear antibody than there are lupus patients. Thus, despite a 97 per cent specificity, the majority of individuals with a positive antinuclear antibody in a randomly screened population will not have systemic lupus. On the other hand, if a patient with photosensitivity, leucopenia, and an abnormal urinary sediment has a positive antinuclear antibody, one can be virtually certain that the diagnosis is systemic lupus. Using Bayes' theorem, it has often been calculated that if the pretest probability, that is the clinical likelihood, of a given patient having systemic lupus is over 50 per cent, a positive antinuclear antibody will increase this likelihood to a near certainty. If one orders the test 'as a screen' in a patient who has no clinical evidence of lupus, for example a tennis elbow, where the pretest likelihood of lupus is minimal, then a positive test will have no real import on this probability, perhaps increasing it to about 2 per cent, that is still 50:1 against the diagnosis of systemic lupus erythematosus. This would surely argue against ever doing the test as a screen. In the management of diffuse disease of connective tissue it is important to establish the broad category of the disease which is being treated but the precise subcategory may be therapeutically unimportant. For example a serological diagnosis of mixed connective tissue disease (MCTD) should not prevent routine monitoring of the urinalysis and aggressive treatment of any renal or any other end organ involvement which may occur. 'Anti-ENA' antibody tests (antibodies to extractable nuclear antigen) are unfortunately often performed routinely in patients with a positive FANA. Their role should surely be to try and help delineate clinical subtypes of disease where this information will be of help, either therapeutically or prognostically.

Sometimes laboratory tests may be helpful in reducing the need for more invasive procedures. Thus, a positive test for antineutrophil cytoplasmic antibodies in a patient with features of Wegener's granulomatosis may make a biopsy unnecessary. Similarly, anti-Jo-1 antibodies in a patient with muscle weakness and elevated enzymes may not only help avoid a biopsy but give potentially important prognostic information.

If a young male complains of nocturnal low back pain with stiffness lasting more than one hour in the morning with a restricted range of flexion in the lumbar spine, would testing for HLA-B27 be at all helpful? Under normal circumstances it would not. One would check the pelvic radiograph—it would be very uncommon, even in early ankylosing spondylitis, to see normal sacroiliac joints. Although, as routine interpretation is notoriously unreliable it is a good idea to develop the expertise to review these personally. In the late teens and early twenties and, of course, in children, the radiographs can be difficult to interpret and here HLA-B27 typing may be helpful. In a relative of a HLA-B27 ankylosing spondylitis proband, a negative HLA-B27 test indicates that any back symptoms are very unlikely to be ankylosing spondylitis related, just as a positive test in this context, even with normal radiographs, may suggest a forme fruste of the disease.

Problems of interpretation of the findings also abound in radiography. Mild asymptomatic radiographic changes are frequent, especially in the lumbar spine but also in the knee. It is all too easy to ascribe symptoms to these changes when examination would reveal an anserine bursitis or a similar localized, readily treatable problem. A more critical problem has arisen because of 'false positive' findings, for example of disc protrusion, seen with computed tomography scans of the lumbar spine and even more so with magnetic resonance imaging. Thus, a high proportion of asymptomatic subjects will have one or more abnormalities evident. This means that little reliance can be placed on these findings in symptomatic subjects unless the findings clearly fit with an anatomical lesion predictable with reasonable probability from the clinical examination. Even in a time of no financial constraints these investigations become useful in the management of patients with back pain only if it is important to clearly delineate either the nature of the pathological process or the extent and location of a lesion, for example prior to a surgical intervention or, more questionably, for medicolegal requirements.

It is tempting for an inexperienced physician to place more weight on the printed test results than on clinical judgement; all rheumatologists see too many patients with minor problems where a consultation has been stimulated by an incorrectly ordered test having proved positive. Thus, from a diagnostic perspective, tests are ordered where they can help support or deny the clinically derived diagnostic impression. A further indication would be to follow a patient's progress. Thus, in the treatment of patients with rheumatoid arthritis progression of erosive changes may, by themselves, influence therapeutic decisions. Bone mineral density measurements may be helpful in the diagnosis of demineralization. They will be indicated primarily when the results will be influential in helping the patient or physician reach a therapeutic decision. Thus, if the patient, for various reasons, is to be started on oestrogen therapy anyway there is little or no point in ordering the test. A 'baseline' measurement may be intellectually satisfying but as a failure to respond to oestrogen must be extraordinarily rare, repeated studies make little economic sense.

A management plan and the individual patient

It is clear that any management plan depends on the underlying disease process and the extent of its progression in the individual patient. There is now a clear consensus that the earlier remittive therapy is instituted in rheumatoid arthritis, the more likely it is to be successful. Thus, we would argue that at least all seropositive patients should be started on a remittive agent as soon as the diagnosis is made without subjecting the patient to a demonstration of the inefficacy of non-steroidal anti-inflammatory drugs alone. Similarly, the longer lupus nephritis has gone untreated, the more refractory it seems to be to treatment, which hastens our decision to add cytotoxic therapy to the management of this condition (Table 6).

How compliant	Decisions about parenteral versus oral therapies
Geographic	Are there problems of access to clinic; can the patient attend for a weekly series of gold shots or must visits be more intermittent?
Age of the patient	Is this a child, adult, or elderly person? Are there drug dose problems, drug interactions, or accompanying illnesses? A stroke might make postoperative rehabilitation difficult, deferring, for example, an otherwise indicated arthroplasty
Psychological status	Could the patient tolerate the complications of therapy, e.g. nausea with methotrexate or rash with gold?
Upper gastrointestinal problems, peptic ulceration	Difficulty taking non-steroidal anti-inflammatory drugs

Table 6 Fitting the management plan to the patient

In elderly patients, or those with impaired renal function from other causes, we check the serum creatinine before and either 2 or 7 days after the institution of therapy with non-steroidal anti-inflammatory drugs, depending on the drug half-life. Similarly, in elderly patients or in those who have had significant gastrointestinal haemorrhage or perforation, we would often combine use of such non-steroids with gastric protection. In rheumatoid arthritis this is a further argument in favour of the early use of remittive agents. In localized but painful conditions such as tendinitis and bursitis it is more economical for the patient to be treated with local steroid injections where appropriate rather than by oral non-steroidal drugs, although it may be less convenient for the physician. The expected treatment benefits should not be outweighed by difficulties in complying with the suggested therapeutic programme. For example geographical considerations may be important in the decision whether to use a parenteral or oral remittive agent in rheumatoid arthritis. It is obviously most important that the relative risks and benefits of treatment be most carefully explained to the patient, with supporting written information provided. Except in life-threatening or rapidly crippling conditions it is wise to allow the patient a waiting period for discussion with family members and the referring physician rather than to try to persuade a reluctant patient to accept your suggestions. One phrase we find useful in this circumstance is 'If I or one of my family members had your disease, and another rheumatologist were to suggest this course of treatment, I would have no hesitation in accepting it'.

In any management plan for patients with rheumatological disease a careful assessment must be made of the relative architectural, mechanical, and inflammatory components. All of these must be seen in the context of the patient's general health. Generally speaking, architectural problems may benefit from surgical solutions and secondary mechanical problems, such as muscle weakness, are susceptible to physical therapy. In rheumatoid arthritis, seriously disabling problems in the feet and ankles can be prevented by the provision of adequate transverse and longitudinal supports for the metatarsal arches, combined with suitable footwear. Surgery does not necessarily pose any increased risk in elderly patients, if risk factors such as electrolyte imbalance, heart failure, or arrhythmia can be corrected or treated. It is most important to determine that surgical correction of the musculoskeletal problem, in, for example, a painful unstable knee, will improve the patient's quality of life. There may be little point in giving a patient a perfectly stable, painless joint if exercise distance cannot be significantly increased due to cardiac, respiratory, or peripheral vascular problems. Under these circumstances it would be more practical simply to provide an external brace. Similarly, age can be an important consideration in any plan for staged orthopaedic surgery accompanied by the necessary long periods of rehabilitation between procedures. Such a treatment plan, which may extend over several months or years, may be entirely sensible for a patient in their 40s or 50s whose life expectancy is 20 or 30 years. It could hardly be thought to improve the overall quality of life for someone of 80 whose life expectancy is much less.

If someone's life expectancy is very short because of, say, terminal malignant disease, we may choose to prescribe steroids for their rheumatoid arthritis, but we have on occasion been surprised at the length of the terminal illness, which has allowed serious steroid complications to arise.

Conclusions

Finally, it should be remembered that the ability to recognize a seriously ill patient before a specific diagnosis is as important for rheumatologists as for all other physicians. While this remains part of the art of medicine, it is also a reason for the triggering of appropriate perceptive questions, even to patients who may at first sight appear to have minor disease.

Chapter References

Davidson, R.J. and Goleman, D.J. (1977). The role of attention in meditation and hypnosis: a psychobiologic perspective on transformations of consciousness. *International Journal of Clinical Experimental Hypnosis*, **15**, 291–308.

Kellgren, J.H. (1977). The anatomical source of back pain. *Rheumatology and Rehabilitation*, **16**, 7–12.

Mark, V.H., Irvin, F.R., and Hackett, T.P. (1960). Clinical aspects of stereotactic thalomotomy in the human. 1. The treatment of chronic severe pain. *Archives of Neurology*, **3**, 351–67.

Schon, D.A. (1987). *Educating the reflective practitioner: toward a new design for teaching and learning in the professions*. Jossey-Bass, San Francisco.

Zborowski, M. (1952). Cultural components in responses to pain. *Journal of Social Issues*, **8**, 17–30.

1.1.2 Children and adolescents

Ross E. Petty

Introduction

Musculoskeletal complaints in childhood are common, but usually short-lived, and most are ultimately inconsequential (Apley 1976). It may be difficult, however, to differentiate definitively such problems from those associated with long-term illness. An appreciation of the spectrum of rheumatic disorders in childhood, their frequency, and the age-related factors that may affect their presentation or course is essential for the physician caring for children or adolescents. A full discussion of childhood rheumatic diseases is beyond the scope of this chapter and summaries of many of the important entities are provided in later chapters or in other sources (Cassidy and Petty 1995; Southwood and Malleson 1993).

Rheumatic diseases in childhood differ considerably from those in adults. In some instances the diseases themselves are unique to the young; in all instances, age-related factors are important determinants of clinical presentation, differential diagnosis, and management. It is the purpose of this chapter to consider some aspects of the epidemiology of childhood rheumatic diseases and the clinical characteristics of the growing child that may affect the presentation, diagnosis, and management of such diseases.

Epidemiology

An understanding of the relative frequencies of the rheumatic diseases in childhood, and their distribution through the age range, facilitates their recognition and diagnosis. There is a paucity of valid data on the incidence and prevalence of rheumatic diseases in childhood. This is due, in part, to the absence of definitive diagnostic criteria for these diseases, or to the existence of competing criteria.

Definitions and criteria

There are two major sets of criteria for the diagnosis of chronic childhood arthritis, those of the American College of Rheumatology (**ACR**) (Brewer *et al*. 1977) and those of the European League Against Rheumatism (**EULAR**) (Table 1) (European League Against Rheumatism 1977). The major difference between the two is the inclusion of the seronegative spondylarthropathies in the EULAR definition of juvenile chronic arthritis, and their exclusion from the ACR definition of juvenile rheumatoid arthritis. The EULAR classification restricts the term juvenile rheumatoid arthritis to children with IgM rheumatoid factor. There is little doubt that the IgM rheumatoid factor-positive girl with polyarthritis resembles the adult with seropositive rheumatoid arthritis; the other arthritides defined by the ACR criteria (oligoarticular-onset and systemic-onset juvenile rheumatoid arthritis) are clinically distinct. Oligoarticular-onset arthritis is essentially a uniquely childhood disease. Systemic-onset disease may have an adult counterpart, so-called adult-onset Still's disease (Elkon *et al*. 1982).

	ACR Criteria: JRA[a]	EULAR Criteria: JCA[b]
Age at onset	Under 16 years	Under 16 years
Duration	6 weeks or more	3 months or more
Onset types	JRA: Pauciarticular, <4 joints Polyarticular, >4 joints	Still's disease: Pauciarticular, <5 joints Polyarticular, >4 joints (rheumatoid factor negative)
	Systemic arthritis with characteristic fever	Systemic arthritis with characteristic fever
Other groups	None	JRA: rheumatoid factor positive polyarthritis Ankylosing spondylitis Psoriatic arthritis
Exclusions	Other types of arthritis	

[a]JRA, Juvenile rheumatoid arthritis (Brewer *et al*. 1977)
[b]JCA, Juvenile chronic arthritis (EULAR 1977)

Table 1 Comparison of the ACR and EULAR criteria for the diagnosis of chronic arthritis of childhood

Recently, a third approach to classification of childhood arthritis has been proposed by a committee of the International League Against Rheumatism (**ILAR**) in co-operation with the World Health Organization (**WHO**) (Fink 1995). This classification (Table 2) is proposed in the hope that it will allow the identification of homogeneous groups of patients suitable for genetic, immunological, and epidemiological study. It recognizes that some children will not be classifiable at the present time and it is designed to be revised after testing and validation. It is recommended that at present this classification be reserved for use in research settings and that the ACR or EULAR criteria be retained until proof of the validity of the ILAR/WHO approach has been shown. For consistency with the other chapters in this book, the EULAR terminology for chronic arthritis occurring in children will be used throughout this chapter except where specifically indicated.

Oligoarthritis (<5 joints affected during first 6 months of disease)

Polyarthritis (>4 affected during first 6 months of disease)
- Rheumatoid factor negative
- Rheumatoid factor positive

Extended oligoarthritis (oligoarticular onset but accumulation of >4 affected joints during first year of illness)

Systemic arthritis (arthritis with typical fever, rash) (definite if arthritis is present, probable in the absence of arthritis)

Psoriatic arthritis (arthritis with psoriasis, or arthritis with a family history of psoriasis, and dactylitis or nail abnormalities)

Enthesitis-related arthritis (arthritis with enthesitis)

(Fink 1995)

Table 2 Proposed WHO/ILAR classification of arthritides of childhood

Two other sets of criteria deserve mention. The criteria for the seronegative enthesitis and arthritis (SEA) syndrome (Table 3) (Rosenberg and Petty 1982) identify children who, although they do not currently fulfil criteria for the diagnosis of ankylosing spondylitis, are likely to do so in the future (Cabral *et al*. 1992a; Burgos-Vargas and Clark 1989). These criteria facilitate differentiation of such children from those with other types of arthritis. The so-called 'Vancouver criteria' (Southwood *et al*. 1989) were developed to facilitate the early identification of children who have or will develop psoriatic arthritis (Table 4). Although both of these sets of criteria have gained widespread use, they have not been formally validated.

Age less than 16 years

Seronegativity (absence of rheumatoid factor and antinuclear antibody)

Enthesitis

Arthritis or arthralgia

(Rosenberg and Petty 1982)

Table 3 Criteria for the diagnosis of seronegative enthesopathy, arthropathy (SEA) syndrome

Definite psoriatic arthritis:
- arthritis with typical psoriatic rash, or
- arthritis with three of four criteria:
  - dactylitis
  - nail pitting or onycholysis
  - psoriasis-like rash (atypical in location or form)
  - family history of psoriasis (first or second degree relatives)

Probable psoriatic arthritis:
- arthritis with two of four criteria

(Southwood *et al*. 1989)

Table 4 'Vancouver criteria' proposed for the diagnosis of juvenile psoriatic arthritis

The only other diagnostic criteria that have been developed for childhood rheumatic diseases are those for acute rheumatic fever (Special Writing Group of the American Heart Association 1992) (Table 5) and those for Kawasaki disease (Table 6) (Sekiguchi *et al*. 1985). The application of the criteria for Kawasaki disease has been useful in identifying most children who have this childhood vasculitis but the criteria lack the benefit of rigorous evaluation and could include children with other exanthematous illnesses. Furthermore, their strict application would exclude children whose later disease course indicates that they had Kawasaki disease.

Major manifestations	Minor manifestations
Carditis	Fever
Polyarthritis	Arthralgia
Chorea	Previous rheumatic carditis
Subcutaneous nodules	Prolonged PR interval
Erythema marginatum	Increased erythrocyte sedimentation rate or C-reactive protein

Diagnosis requires the presence of two major criteria, or one major and two minor criteria, and evidence of a preceding streptococcal infection (Special Writing Group of the American Heart Association 1992)

Table 5 Criteria for the diagnosis of acute rheumatic fever

Fever	duration 5 days or more
Conjunctivitis	bilateral, bulbar, non-suppurative
Lymph node enlargement	cervical, non-purulent, >1.5 cm
Rash	polymorphous, no vesicles or crusts
Changes in lips and mucosa	dry, red, vertically cracked lips, or strawberry tongue, or diffuse oropharyngeal erythema
Changes in extremities	erythema of palms or soles, or indurative oedema of hands, feet, or desquamation of tips of fingers

Diagnosis requires the presence of five of six criteria. Recommendations of the Japan Mucocutaneous Lymph Node Syndrome Research Committee 1984 (Sekiguchi *et al*. 1985)

Table 6 Criteria for the diagnosis of Kawasaki disease

Criteria for the diagnosis of other chronic rheumatic diseases of childhood either do not exist, have been adopted in whole or part from criteria for the classification of diseases in adults, as for systemic lupus erythematosus (Tan *et al*. 1982) or dermatomyositis (Bohan and Peter 1975), or have not been validated in clinical practice.

Frequencies of childhood rheumatic diseases

It is frequently perceived that rheumatic diseases are rare in childhood and adolescence. As a result, consideration of such a diagnosis may be inappropriately delayed. In fact, a significant fraction of adults with rheumatic diseases have onset of their disease in childhood (Table 7). Thus, up to one-fifth of all patients with systemic lupus erythematosus or dermatomyositis/polymyositis have onset of their disease before age 17 years. Onset of disease in childhood occurs in from 5 to 10 per cent of individuals with one of the more common chronic arthropathies such as ankylosing spondylitis and rheumatoid arthritis. It is likely that these estimates represent minima because, particularly in the diseases such as oligoarticular-onset juvenile chronic arthritis, many children have remission of their disease in childhood.

Adult rheumatic disease	Annual rate per 100 000	Percentage with childhood onset
Spondylarthropathies	130–1000	10
Rheumatoid arthritis (juvenile)	40	5
Systemic lupus erythematosus	6	18
Dermatomyositis/polymyositis	0.8	20
Scleroderma	0.4	3
Vasculitides	0.2	10
Gout	20	0

Table 7 Incidences of connective tissue diseases

Some rheumatic diseases occur almost exclusively in those under 16 years of age. Included in this category are oligoarticular-onset juvenile chronic arthritis with uveitis, systemic-onset juvenile chronic arthritis, acute rheumatic fever, and, among the vasculitides, Henoch-Schönlein purpura and Kawasaki disease. In contrast, certain rheumatic diseases almost never occur in childhood. Included in this group are gout, calcium pyrophosphate-deposition disease, polymyalgia rheumata, and primary osteoarthritis.

The actual frequencies of each of the types of arthritis in a community is difficult to ascertain. A study from Finland (Table 8) (Kunnamo *et al*. 1986) indicates that acute, presumably postviral arthritis of the hip or knee accounts for three-quarters of all children with arthritis; chronic childhood arthritis accounts for approximately one-fifth. Children with other rheumatic diseases were not included in this study. The overall impression is, however, that the annual incidence of significant joint inflammation or infection involves approximately 1 in 1000 children and that chronic arthritis occurs in 1 in 5000 children per year.

Diagnostic group	Percentage of total	Incidence per 100 000 children
Transient synovitis of the hip	47.8	51.9
Acute transient synovitis	23.5	25.8
Juvenile chronic arthritis	16.8	18.2
Septic arthritis	6.2	6.7
Reactive arthritis	5.0	5.4

Adapted from Kunnamo *et al*. (1986)

Table 8 Incidence of arthritis in children in a Finnish community

The composition of paediatric rheumatology clinic populations has been described in a study by Rosenberg (Rosenberg 1990), and in national registries of children seen in paediatric rheumatology clinics (Bowyer *et al*. 1996; Malleson *et al*. 1996; Symmons *et al*. 1996). In Rosenberg's study, 46 per cent of children with an identifiable rheumatic disease had juvenile rheumatoid arthritis (ACR criteria), 31 per cent had a spondylarthropathy, 18 per cent had a connective tissue disease (systemic lupus erythematosus, dermatomyositis, Henoch-Schönlein purpura, Kawasaki disease, other vasculitis), and 5 per cent had a variety of other conditions. From these data, an annual incidence of juvenile rheumatoid arthritis of 8 in 100 000 and a prevalence of 40 in 100 000 were estimated.

Age at presentation

Many of the rheumatic diseases of childhood have characteristic ages at onset, (Fig. 1) and consideration of the age of the child at the onset of symptoms can be of value in diagnosis. For this purpose, four age groups can be considered. Three rare syndromes make their appearance in the neonatal period; neonatal lupus syndromes, neonatal-onset multisystem inflammatory disease—NOMID (also called chronic infantile neurological cutaneous and articular syndrome—CINCA), and infantile sarcoidosis. In addition, septic arthritis may occur in the neonate. In early childhood (from 1 to 4 years), oligoarticular-onset juvenile chronic arthritis, juvenile psoriatic arthritis, Kawasaki disease, and septic arthritis are most frequent. In midchildhood (7 to 11 years), polyarticular-onset juvenile chronic arthritis and ankylosing spondylitis begin to be seen and juvenile dermatomyositis, Henoch-Schönlein purpura, and polyarteritis nodosa have their peak frequencies. In late childhood and the teenage years, juvenile ankylosing spondylitis (in boys) and systemic lupus erythematosus (in girls) show a marked increase and the rare vasculitis syndromes, such as Wegener's granulomatosis and Takayasu's arteritis, occur.

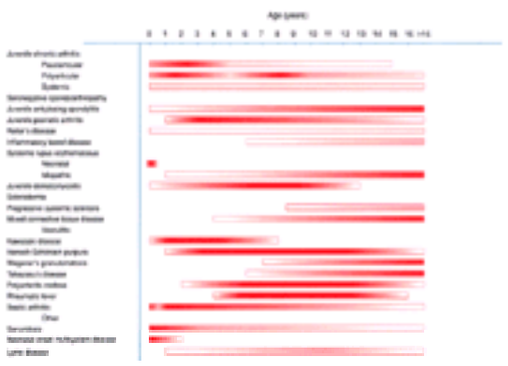


Fig. 1 Characteristic ages of onset of rheumatic diseases of childhood.

Influence of gender

Consideration of the sex of the child may be of help in formulating the differential diagnosis (Fig. 2). The most striking differential sex ratios are seen in the predominance of juvenile ankylosing spondylitis in boys and of systemic lupus erythematosus in girls. In systemic lupus the sex ratio is closer to unity in the very young, however (Cassidy *et al*. 1977). Oligoarticular juvenile chronic arthritis, especially when accompanied by uveitis, is much more common in girls than in boys, whereas systemic-onset juvenile chronic arthritis occurs with equal frequency in both sexes. In general, the spondylarthropathies and the vasculitides are more common in boys than girls.

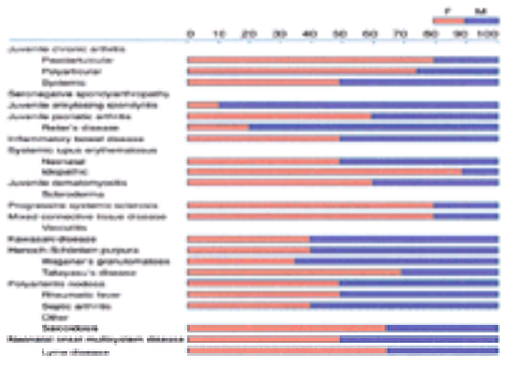


Fig. 2 Sex differences in rheumatological diseases of childhood.

Growth and development

Important age-related characteristics of the patient affect the physician's approach to rheumatic diseases in childhood. Some of the most important are briefly discussed below.

Psychosocial and educational considerations

Chronic illness can have a major impact on the quality of life of the child and the family (Quirk and Young 1990). It is often useful for the health-care team to meet with the family as a whole to acknowledge this fact and to educate other family members about the health problem faced by the child. The physician must establish a professional relationship not only with the child but also with the parents, who usually provide most of the information needed to make a diagnosis, and upon whom implementation of treatment depends. Equally importantly, the physician must recognize the primacy of the patient, even the very young patient, and every reasonable attempt must be made to incorporate the child into the process of obtaining a history and explaining the findings of the physical examination and laboratory investigations. Throughout childhood the roles of child and parent change, and the adolescent must be expected to assume increasing responsibility for attendance at medical appointments, taking prescribed drugs, and complying with prescribed exercise programmes. The older teenager should be given the opportunity to visit the physician without the parent.

The school is one of the child's most important environments; it is to the child what the workplace is to the adult. Attention to school performance, both academic and social, is critical to the development of a complete picture of the child's health, and in children with functional musculoskeletal or abdominal pain it is often the key to understanding the cause of the distress. The co-operation of the teachers and counsellors in facilitating compliance with medications that are taken during school hours, in modifying physical education programmes, and in ensuring the optimal participation of the child with a rheumatic disease in all classroom activities is a vital component of the child's long-term management.

A comprehensive team approach will help address the broad areas of concern to the older child and adolescent, including self-esteem, sexuality, and educational and vocational goals. The timing of transition from a paediatric to an adult health-care setting varies with patient, physician, and availability of appropriate facilities. In general, children are seen in paediatric clinics until they complete secondary education, enter college or university, or join the work-force. A transition clinic in which the expertise of both paediatric and adult rheumatologists are available is an ideal way to facilitate successful entry into the adult medical care setting. From the psychosocial aspect, a 21-year-old patient with a 15-year history of a chronic disease is very different from a 21-year-old patient in whom a similar disease has just begun. Little is known about the best ways to provide care for adults with childhood onset of chronic illnesses. The transition clinic concept will help provide this much needed information.

Motor development

Recognition of the range of normal motor development in the young child is essential for interpretation of findings on physical examination and functional assessment. The major motor milestones are shown in Table 9. Delay in such achievement may occur in a child with any chronic disease, and may be accentuated in the child with disease affecting the musculoskeletal system, the function of which is the basis for the estimation of developmental stages. Although developmental delay is common in chronically ill children, milestones are eventually achieved. There is considerable normal variation in the age at which individual developmental milestones are achieved and the ages in Table 9 should be used as guidelines only (Vaughan and Litt 1992).

Age (months)	Milestone achieved by this age
7	Sits independently
10	Crawls
12	Cruises, walks with one hand held
15	Walks alone; crawls up stairs
18	Runs stiffly; walks up stairs with one hand held
24	Runs well; walks up and down stairs one step at a time
30	Jumps
36	Walks up stairs alternative feet; rides tricycle
48	Hops on one foot
60	Skips

Table 9 Gross motor developmental milestones

Physical growth

Chronic inflammatory disease of the musculoskeletal system, and sometimes the side-effects of its therapy, have profound effects on physical growth. Careful documentation of height and weight should be made at each encounter and compared to growth curves for healthy children of the same sex and age (Fig. 3 and Fig. 4) (Tanner and Whitehouse 1976). In so doing, it is important to consider the parents' height and weight. Sequential measurements ensure not only that the child's height and weight are appropriate for age and sex, but also that the velocity of growth is normal. A growth curve of a child with severe polyarticular juvenile chronic arthritis is shown in Fig. 5. When the disease was controlled, growth velocity increased and a final height close to the expected value was achieved. This is not always the case, however, and severe stunting may result from disease or from prolonged, high dosage of glucocorticoids. Growth failure becomes a serious issue for the adolescent whose peers are growing at a normal rate. Attempts to accelerate growth by giving nutritional supplements or growth hormone have met with limited

success, and feasibility and cost often preclude their use (Butenandt 1979).

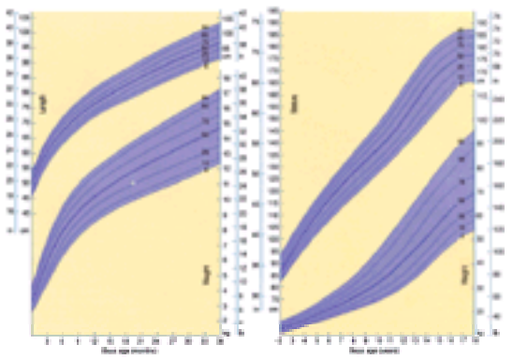


Fig. 3 Physical growth in boys. (a) From birth to 36 months. (b) From 2 to 18 years. Adapted from National Center for Health Statistics Growth Charts (USA) 1986.

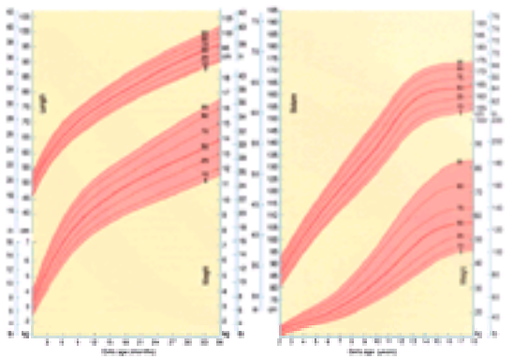


Fig. 4 Physical growth in girls. (a) From birth to 36 months. (b) From 2 to 18 years. Adapted from National Center for Health Statistics Growth Charts (USA) 1976.

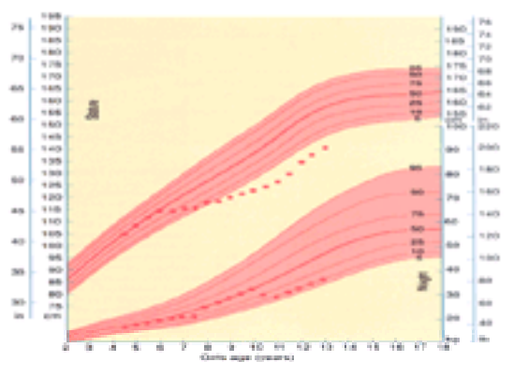


Fig. 5 Growth record of child with polyarticular juvenile chronic arthritis of onset age 5.5 years and severe active disease until age 8 years. The disease was controlled with corticosteroids age 8 to 10 years. The disease was in remission age 11 to 13 years.

Sexual maturation

Delay in sexual maturation in the child with a chronic rheumatic disease is usually a manifestation of inflammatory disease, rather than as an intrinsic abnormality of sexual development. It is useful to document the stages of development of secondary sexual characteristics as shown in Table 10 and Table 11. In boys the genitalia may reach Tanner Stage 2 at any time after the age of 9 years, but need not do so until 14 or 15 years of age. Complete genital maturity (Tanner Stage 5) may occur before the age of 13, but may not occur until 18 years of age (Marshall and Tanner 1986). Axillary hair develops approximately a year after the appearance of pubic hair, and facial hair begins to grow in males a year after the appearance of axillary hair. Menarche, the time of the first menstrual period, occurs at approximately 12.5 to 13 years of age, followed, after a year or more, by the development of the breasts and appearance of pubic and axillary hair (Marshall and Tanner 1986). Delay or irregularity of menstruation is characteristic of adolescent girls who have active inflammatory disease (Fraser *et al*. 1988).

Stage	Pubic hair	Penis	Testes
1	None	Preadolescent	Preadolescent
2	Scant, long, straight	Slight enlargement	Slight enlargement
3	Darker, starts to curl	Longer	Larger
4	Coarse, curly	Broader, glans enlarges	Larger, scrotum darker
5	Adult, spreads to thighs	Adult	Adult

Table 10 Sexual maturation in adolescent boys

Stage	Pubic hair	Breasts
1	Preadolescent	Preadolescent
2	Scant, straight, medial margin of labia	Small mound; areolar diameter increased
3	Darker, starts to curl	Breasts and areolas enlarged; no contour separation
4	Coarse, curly	Areola and papilla form secondary mound
5	Adult; spreads to thighs	Mature, nipple projects; areola part of breast contour

Table 11 Sexual maturation in adolescent girls

In addition to indicating the presence of chronic inflammatory disease, delay in sexual maturation is sometimes a source of severe, often unexpressed, anxiety for the adolescent. The physician or nurse should raise these issues with the adolescent patient if they do not initiate the discussion. The adolescent should understand the reason for the delay in sexual development, and be assured that it is almost certain that normal sexual development will eventually occur. Concerns about fertility, and in girls, the ability to deliver a baby should also be recognized and discussed, sometimes with the advice of an obstetrician. Questions of inheritance of the rheumatic disease are often raised and require answers as permitted by the limited information available.

Diagnostic approach

History

In paediatric rheumatology, perhaps more than in most other branches of medicine, the history and physical examination are the foundations of the diagnostic approach; the laboratory provides limited help. It is important to obtain the historical information from both the care-giver (parent or other adult) and the child. Second-hand historical information, even if provided by the mother, may be less reliable than direct information from the care-giver. Important, too, is information offered by the child. Even a young child may be capable of giving a more accurate history of the symptom than the parent. It is, after all, the child, not the parent, who has the disease. The process of getting historical information from the child may be a time-consuming one, but the reward is often worth the time invested.

Physical examination

Rheumatic diseases are systemic disorders, and the physical examination must include a general physical examination as well as an examination of the musculoskeletal system. In order to gain the most information from the physical examination, the child should be as comfortable as possible. The child's modesty and dignity should be preserved at all times and appropriate examination gowns, shorts, and private changing space provided. The child's permission to examine him or her should be sought informally. Reassurance should be given that the physician will not persist if examination causes pain, and, if possible, examination of the painful area should be left to the end. The experienced physician will identify the 'examinable moment' when the child has gained sufficient confidence in the examiner and acceptance of the environment that useful, trauma-free examination is possible. In small children, examination may begin by observing the child at play or moving about the room. Examination of the joints may begin with the child in the mother's lap. As the child gains confidence in the procedure, the general and musculoskeletal examinations can be completed on the examining table.

Evaluation of the musculoskeletal system must take into account age-related variations from normal, some of which are commented on below.

Range of motion

In full-term new-borns, the elbows, knees, and hips do not fully extend. In the young child, and even the adolescent, hyperextension at the knees and elbows is often greater than in the adult. In other ways, however, normal adult range is achieved by approximately 3 or 4 years of age. Asymmetry of range at any age should be considered to be abnormal.

Muscle strength

Evaluation of muscle strength in the child is difficult. Not only are there obvious age-related differences in muscle strength, but it may be difficult to gain the child's co-operation in order to examine individual muscle groups. In very young children, the examiner must rely on function rather than strength *per se*. Myometry can be used in older children, but repeated evaluation by the same well-trained examiner (usually a physiotherapist), using a standard muscle grading system, is probably the best way of evaluating muscle strength over time.

Gait development

Cruising (walking while holding on to a hand or to furniture) develops by 12 months of age, and independent walking should occur by 15 months. Until the age of 3 years, the child has a gait pattern that differs from that in an adult; the rate of walking is faster, the stance is more broad-based in relation to pelvic width, the knee may not be fully extended, and the ankle may be plantar flexed at foot-strike (Sutherland *et al*. 1980). The ability to climb stairs one at a time develops at about 2 years of age; by 3 years of age the child can usually go up and down stairs, alternating feet. Walking on tiptoe is not abnormal for a child who has just begun to walk. This pattern should disappear by the age of 2 years, however. If it persists beyond this age, causes such as spasticity, muscle weakness, or tethered cord should be considered.

Structural variations from normal

The appearance of flat feet is usually normal in the young child. The distribution of fat and the paucity of muscle development in the infant and young child cause the feet to look flat. This appearance is common in children up to about 5 or 6 years of age. In the absence of symptoms, no investigation or treatment are indicated.

The alignment of a child's legs is a frequent source of parental concern, and age-dependent characteristics should be identified (Salenius and Vankka 1975). From birth until 2 years of age, it is normal to have symmetrical genu varus. From 2 to 5 years of age, mild genu valgum may occur. At no time should the degree of valgus or varus exceed 10°, and it should be bilaterally symmetrical. Age-related normal ranges of varus and valgus are shown in Fig. 6.

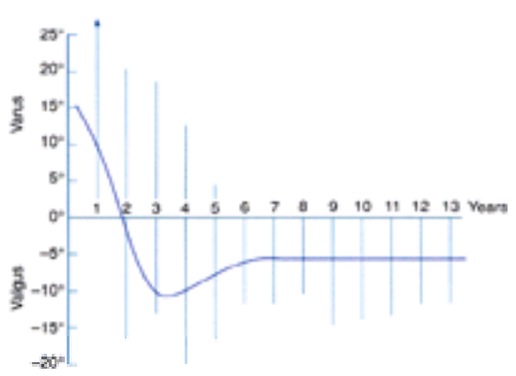


Fig. 6 The femoral–tibial angle (measurement of varus and vulgus at the knee) in normal children. Vertical lines indicate range. Solid line indicates mean (Salenius and Vankka 1975).

Functional assessment

Age-related, functional assessment can be ascertained with some knowledge of developmental stages, and can be documented more precisely with tools developed specifically for this purpose. The development of the juvenile arthritis functional assessment scale (JAFAR) (Lovell *et al*. 1989) has permitted a semiquantitative approach to the measurement of disability in children with chronic arthritis. This assessment can be administered in a few minutes in the outpatient setting.

Laboratory investigations

Laboratory evaluation of the child with a rheumatic disease serves three functions; to provide or exclude evidence of inflammation, to provide evidence of diagnostic significance, and to exclude non-rheumatic diseases (Table 12).

Indicators of inflammation
White blood-cell count, differential, platelet count, erythrocyte sedimentation rate
Tests of diagnostic significance
Antinuclear antibody
Antinuclear antibody specificities if antinuclear antibody is present
Muscle enzymes (if dermatomyositis is suspected)
Antineutrophil cytoplasmic antibodies (if vasculitis is suspected)

Table 12 Laboratory evaluation of the child with musculoskeletal pain or a suspected rheumatic disease

In screening for evidence of inflammation, the white blood cell count and differential, haemoglobin, platelet count, and the erythrocyte sedimentation rate usually suffice. In most instances of chronic inflammatory disease one or more of these is abnormal, although children with monoarticular juvenile chronic arthritis may have normal test results. Discordance between the white blood-cell count or the platelet count and the erythrocyte sedimentation rate (i.e. a low platelet count with an elevated sedimentation rate) may be suggestive of an underlying malignancy such as leukaemia or neuroblastoma. In some disorders, such as Kawasaki disease, platelet counts may exceed one million. Decreased albumin is consistent with inflammation of many causes but if it is marked, and articular disease is mild or moderate, inflammatory bowel disease should be considered. In the child with a fever of unknown origin and an elevated sedimentation rate, but no abnormalities detected on physical examination, inflammatory bowel disease, polyarteritis nodosa, or occult malignancy should be considered.

Tests that have diagnostic specificity are few in number. Although antinuclear antibody is present in many children with a wide variety of rheumatic disease, it occurs in non-rheumatic disease as well (Cabral *et al*. 1992b), but it is absent in children with the seronegative spondylarthropathies. If clinically indicated, the antigenic specificity of the antinuclear antibody should be determined; thus, for example, antinuclear antibody directed to double-stranded DNA strongly suggests a diagnosis of systemic lupus erythematosus.

Studies of the sensitivity and specificity of testing for rheumatoid factor in the child have shown unequivocally that it has little value as a diagnostic test (Eichenfield *et al*. 1986). Rheumatoid factor is usually absent in children with rheumatic disease, except in those with polyarticular juvenile rheumatoid arthritis where its presence is required for diagnosis by the EULAR classification (although not by the ACR criteria), and where it is associated especially with the presence of nodules and erosions.

Some laboratory investigations are useful in excluding rather than diagnosing disease. An unequivocally normal white blood-cell count, differential platelet count, and erythrocyte sedimentation rate make the possibility of an inflammatory, infectious, or malignant disease very unlikely. A normal radiograph of the bones and soft tissues of a symptomatic joint effectively excludes major trauma, and if the symptom is of long standing, inflammation or infection as well. A normal technetium bone scan virtually excludes primary bone or joint disease in a child with undiagnosed musculoskeletal pain.

Specific approaches

Pattern recognition

Using knowledge of the relative frequencies of individual rheumatic diseases, the age of the child at onset of symptoms, and the sex of the child, coupled with the pattern of joint involvement and the character of extra-articular signs and symptoms, it is possible to make some informed guesses about the likely diagnosis of a specific rheumatic disease. Certain patterns of these clinical characteristics are illustrated below.

1. Acute onset of large joint monoarthritis in a young girl should be considered to be oligoarticular-onset juvenile chronic arthritis unless there is evidence of infection.
2. At any age and in either sex, a pattern of scattered, asymmetrical, large and small oligoarthritis or limited polyarthritis suggests the possibility of juvenile psoriatic arthritis.
3. Oligoarthritis of large joints of the lower extremity in an older boy suggests a seronegative spondylarthropathy, most likely ankylosing spondylitis.
4. The presence of enthesitis with arthritis indicates the likely presence of seronegative enthesitis and arthritis syndrome (Rosenberg and Petty 1982), and the probable later development of ankylosing spondylitis (Cabral *et al*. 1992a; Burgos-Vargas and Clark 1989).
5. The onset of polyarthritis in a teenage girl should suggest the possibility of systemic lupus erythematosus as well as polyarticular juvenile chronic arthritis.
6. Isolated hip joint arthritis may be caused by toxic synovitis, Legg–Perthes' disease, a slipped capital femoral epiphysis, or less likely, by chronic inflammatory arthritis, probably ankylosing spondylitis.
7. An older child or young adolescent male with low back pain may have ankylosing spondylitis but mechanical causes, infection, or malignancy are more likely.

The child with limb pain

The causes of limb pain vary with the age of the child. The very young child may not complain of or admit to having pain, but may instead alter the manner in which he or she uses a limb. Pain in the preambulatory child may be first noticed when the child is being held, dressed, or bathed. The child may be observed to refuse to use one arm or leg, or may regress in major motor milestones. For example the child capable of standing may cease to do so. The child who is an established walker may, for no apparent reason, cease walking or develop a limp or other alteration of gait (Table 13). Until fine motor movement of the hands has developed, it may be difficult to notice the functional effects of arthritis in the joints of the hands and fingers. Preference for the use of the right or left hand before the age of 2 to 4 years, when preferential use of one or other hand normally develops, suggests the possibility of an abnormality of the underused hand.

Location of pain	Cause of pain
Back	Trauma, spondylolysis or spondylolisthesis Inflammation (discitis, spondylitis) Infection (osteomyelitis) Tumours (osteoid osteoma, malignancy)
Sacroiliac joint	Trauma Septic sacroiliitis Inflammatory sacroiliitis
Hip	Transient synovitis, Legg–Perthes' disease, slipped capital femoral epiphysis Chronic synovitis (juvenile chronic arthritis) Painful muscles (dermatomyositis) Congenital dysplasia or dislocation
Knee	Hypermobility, patellar dislocation Osgood–Schlatter disease Inflammatory enthesitis Osteochondritis dissecans, torn meniscus or cartilage Chronic synovitis (juvenile chronic arthritis)
Ankle and foot	Hypermobility, trauma (including stress fracture and footwear problems) Enthesitis, chronic synovitis (juvenile chronic arthritis)

Table 13 Painful conditions that may cause a limp

The history of the onset of the pain is important. The presence of an identifiable time of onset may suggest trauma or infection; chronic arthritis most frequently has a very insidious onset, especially in the very young child. Diurnal variation in symptoms is significant, but in the infant or young child this may be difficult to identify. Precipitating or alleviating factors should be noted. The presence of associated systemic symptoms, such as fever, weight loss, rash, upper respiratory tract infection, or diarrhoeal illness, should be noted.

Limb pain may originate in bone, joint, or soft tissues. Localization of the source of the pain to one of these tissues permits appropriate investigations. Young children often localize pain poorly and the physician must carefully identify the area of discomfort by palpation and observation. The child may not admit to pain but will withdraw the limb or appear anxious when the affected part is examined. Observation of the child's facial expression during examination is very important, as is the parent's interpretation of the child's response. Leg pain may be referred from the back to the thigh, from the sacroiliac joint to the buttocks, or from the hip to the knee. Pain in the absence of tenderness suggests that it may be referred from another site.

Quantification of pain in children often requires the use of non-verbal cues. The child's behaviour is the most important indicator of the presence or absence of pain. Visual tools to assess pain such as that shown in Fig. 7 may be useful, especially in following change over time (McGrath and Unruh 1987).

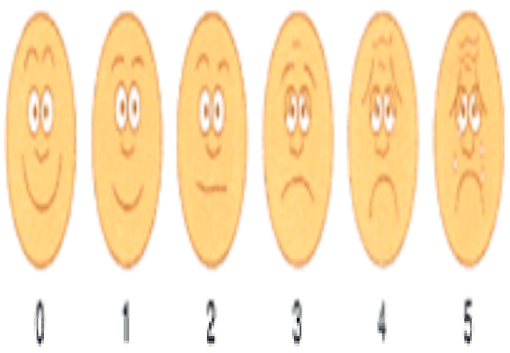


Fig. 7 Pain assessment in children—faces rating scale.

The child with back pain

Back pain is an uncommon complaint, particularly in young children, and should always be thoroughly investigated as its underlying cause may require prompt, specific therapy. Significant causes of back pain in children and adolescents are outlined in Table 14. Common causes of back pain in adults, such as overuse, disc prolapse, or degenerative joint disease, are rare to non-existent in children. In childhood, the important causes of back pain include infection (discitis, osteomyelitis), tumours (benign, such as osteoid osteoma, or malignant), mechanical causes (such as spondylolysis and spondylolisthesis, idiopathic adolescent scoliosis—Scheuermann's disease), and inflammatory causes (such as ankylosing spondylitis). The age relationships and usual sites of back pain in childhood are outlined in Table 15. Pain associated with pyelonephritis and glomerulonephritis, and osteoporosis secondary to glucocorticoid use should also be considered.

Pain in the back
- Acute trauma:
  - fracture/dislocation/haematoma
  - spondylolysis/spondylolisthesis
- Tumour:
  - benign (osteoid osteoma)
  - malignant (reticulum cell sarcoma, metastatic neuroblastoma, leukaemia, lymphoma)
- Infection:
  - discitis
  - vertebral osteomyelitis
- Osteoporosis
- Scheuermann's disease (painful adolescent kyphosis)
- Fibromyalgia

Pain in the sacroiliac joints
- Septic sacroiliitis
- Spondylarthropathy

Pain in the pelvis
- Osteomyelitis
- Tumour (usually osteogenic sarcoma or Ewing's tumour)

Table 14 Back pain in children and adolescents

Cause of back pain	Peak age (years)	Usual site of pain
Discitis	1–3	L4–5 or L3–4
Osteomyelitis	<12	Any site
Scheuermann's disease	13–17	Lower thoracic
Osteoid osteoma	10–30	Any site
Spondylolysis	10–20	L5–S1
Metastatic neuroblastoma	<10	Multiple
Inflammatory sacroiliitis	>12	Unilateral or bilateral
Osteogenic sarcoma	10–20	Ilium
Ewing's tumour	10–20	Ilium

Table 15 Age relationship of causes of back pain in children

Trauma

Trauma is seldom the cause of persisting joint pain or swelling in a child, and the diagnosis of sprain in a young child is usually incorrect. Children of all ages are subject to physical abuse, which may include injuries resulting in musculoskeletal pain. It is estimated that 30 per cent of fractures in children under the age of 3 years are the result of non-accidental injury (Holter and Friedman 1968), and that under the age of 1 year more than half of fractures result from child abuse (McClelland and Heiple 1982). It may be difficult to differentiate fractures secondary to child abuse from those secondary to accidents or to diseases such as osteogenesis imperfecta. However, the characteristic pattern of multiple fractures of different ages, especially involving the posterior ribs, femur, and skull, in combination with skin lesions (bruises, abrasions, burns) strongly suggests the diagnosis. Attention to the problem will help prevent recurrence of the abuse, and allow appropriate intervention which may include taking the child into care.

Monoarticular arthritis

Among the presentations of arthritis in childhood, monoarticular arthritis is particularly problematic. In a child with acute onset of monoarthritis, the initial differential diagnosis includes infection (septic arthritis and/or osteomyelitis), trauma (accidental or non-accidental), and malignancy (especially leukaemia and neuroblastoma). Chronic monoarthritis may represent oligoarticular juvenile chronic arthritis, a seronegative spondylarthropathy, or juvenile psoriatic arthritis. Additionally, a number of uncommon disorders may mimic chronic monoarticular inflammatory arthritis. These include intra-articular haemangiomas, osteochondritis dissecans (especially at

the knee), synovial chondromatosis, lipomatosis arborescens, villonodular synovitis, osteoid osteomas, and discoid meniscus.

In arthritis in the joints of a leg, an altered gait is a common presentation. A limp (asymmetry of gait) may results from structural asymmetry, muscle weakness, or pain. Painful causes of limps are given in Table 14 (Hensinger 1986; Petty 1994). Pain and stiffness are seldom complained of as such in very young children. Instead, non-use, altered use, or irritability are the chief manifestations.

Localized growth disturbances are characteristic of the effect of joint inflammation in the child whose epiphyses have not yet fused. In general, the bones adjacent to an inflamed joint grow more rapidly, and in consequence, the affected digit or limb is longer. If this occurs in a single knee, an inequality of leg length results.

The child with fever of unknown origin

Fever of unknown origin in a child frequently results in consultation with the paediatric rheumatologist with the question: 'Could this girl have one of the rheumatic diseases?' The answer to the question is often 'Yes, she could', but elucidation of the cause may be difficult.

The child who has a history of weeks or even months of fever without demonstration of a cause should be evaluated first to exclude the possibility of occult infection or malignancy (especially leukaemia or lymphoma). Complete blood count, bone marrow biopsy, chest radiograph, and abdominal ultrasonography are appropriate initial investigations. A technetium bone scan and gallium scan may also be indicated. Occult infection, especially osteomyelitis, should be ruled out by radiographs and bone scans. Bacterial endocarditis should be excluded by blood cultures and echocardiography.

The onset of systemic juvenile chronic arthritis may be characterized by fever and rash before the development of joint disease. It is very unusual, however, for arthritis to lag more than a few weeks behind the onset of the systemic features of this disease, and a diagnosis of systemic-onset juvenile chronic arthritis in the absence of arthritis should always be tentative. The presence of pleural or pericardial effusions, splenomegaly, anaemia, leucocytosis, and elevated erythrocyte sedimentation rate, together with a quotidian fever and characteristic evanescent macular rash, would support the possibility of this diagnosis.

Fever with weight loss, with or without gastrointestinal disturbances, erythema nodosum, or arthralgia, suggests the possibility of inflammatory bowel disease. In such instances there may be anaemia, leucocytosis, hypoalbuminaemia, and marked elevation of the erythrocyte sedimentation rate. Definitive diagnosis requires biopsy of the gastrointestinal tract.

Children with vasculitis may have only fever and elevation of indicators of inflammation without objective physical signs. Indirect evidence of vasculitis may be obtained by demonstration of antibody to neutrophil cytoplasm (ANCA), high levels of von Willebrand factor (factor VIII-related antigen), abnormalities on Doppler ultrasonography of large arteries, or with other imaging techniques such as magnetic resonance arteriography or contrast arteriography. Evidence of coronary arteritis in children with Kawasaki disease is shown by electrocardiograph and echocardiograph. Hypertension or haematuria may indicate the presence of renal arterial or glomerular disease. Deficiencies in peripheral pulses suggests the possibility of Takayasu's arteritis. In polyarteritis, the presence of a painful subcutaneous nodule, characteristically in the calf or sole of the foot, provides the opportunity for definitive excisional biopsy.

The child with muscle weakness

Muscle weakness is common in children with inflammatory musculoskeletal disease, although its detection and quantitation may be difficult because of age-related differences in normal muscle strength. As a result, it is often necessary to rely on a history of regression of major motor developmental stages. Thus the child who was an established walker may, because of weakness, stop walking, or ask to be carried. Observation of the child at play, of the manner in which he or she gets up from sitting and lying positions, and of the ability to stand on tiptoes, heels, and squat and recover from a squat all provide useful clues to the presence of muscle weakness. The child with weakness of trunk muscles, as might be seen in dermatomyositis, may have to roll over on to the side or abdomen in order to assume a sitting position from a supine position. The child with weakness of the hip girdle may be unwilling to squat, and unable to stand upright after a squat without exhibiting Gower's sign. Observation of the gait for a wide base or Trendelenburg sign may suggest weakness of the hip girdle. It may be difficult to pick up the young child with weakness of the shoulder girdle because of the instability of the upper trunk. The presence of head lag when lifting the child from the supine to sitting position suggests weakness of neck flexors. With experience and patience, the physician or physiotherapist can demonstrate muscle weakness of a grade or more by formal testing in the child over 3 or 4 years of age. Assessment of muscle strength should also include a neurological examination. Prolonged muscle weakness without muscle wasting is unlikely to be the result of organic disease.

The symmetry of muscle weakness provides an important diagnostic clue. Asymmetric weakness is not likely to result from an inflammatory myositis such as dermatomyositis, or from a primary myopathy. Weakness of one limb or muscle group is most likely to reflect a peripheral neuropathic lesion or inflammation in a major joint such as hip, knee, or ankle. Sacroiliitis can cause severe weakness of the proximal thigh and wasting of muscle. Osteoid osteoma can produce a similar effect. Symmetrical proximal weakness accompanied by muscle tenderness strongly suggests dermatomyositis, and the presence of classical cutaneous changes and elevated muscle enzymes would confirm that diagnosis and differentiate it clinically from primary myopathies.

Summary

The child with a rheumatic disease is often a diagnostic and therapeutic challenge. An effective approach to diagnosis requires a careful history, and a physical examination that takes into account normal age-related variations, which may otherwise be misleading. Treatment requires recognition that the child is rapidly growing and changing physically, psychologically, and socially. It requires recognition of the child's place as a member of a family and the impact of chronic disease on all family members. Problems related to psychosocial and educational development, challenges in school, and the possibility of limitations on employment necessitate management by a well co-ordinated team of health professionals. Attention to these features of care are all the more important because, with few exceptions, rheumatic diseases of childhood are as yet incurable, although in many instances effective control is possible and the child is able to lead a happy fulfilling childhood and adulthood.

Chapter References

Apley, J. (1976). Limb pains with no organic disease. *Clinics in Rheumatic Disease*, **2**, 487–91.

Bohan, A. and Peter, J.B. (1975). Polymyositis and dermatomyositis. *New England Journal of Medicine*, **292**, 344–7, 403–7.

Bowyer, S., Roettcher, P., and the members of the Pediatric Rheumatology Database Research Group. (1996). *Journal of Rheumatology*, **23**, 1968–74.

Brewer, E.J., Bass, J., Baum, J., Cassidy, J.T., Fink, C., Jacobs, J., Hanson, V., Levinson, J.E., Schaller, J., and Stillman, J.S. (1977). Current proposed revision of JRA criteria. *Arthritis and Rheumatism*, **20**(Suppl.), 194–9.

Burgos-Vargas, R. and Clark, P. (1989). Axial involvement in the seronegative enthesopathy and arthropathy syndrome and its progression to ankylosing spondylitis. *Journal of Rheumatology*, **16**, 192–7.

Butenandt, O. (1979). Rheumatoid arthritis and growth retardation in children: treatment with human growth hormone. *European Journal of Paediatrics*, **130**, 15–28.

Cabral, D.A, Oen, K.G., and Petty, R.E. (1992 *a*). SEA syndrome revisited: a long-term followup of children with a syndrome of seronegative enthesopathy and arthropathy. *Journal of Rheumatology*, **19**, 1282–5.

Cabral, D.A., Petty, R.E., Fung, M., and Malleson, P.N. (1992 *b*). Persistent antinuclear antibodies in children without identifiable inflammatory, rheumatic or autoimmune disease. *Pediatrics*, **89**, 441–4.

Cassidy, J.T., Sullivan, D.B., Petty, R.E., and Ragsdale, C. (1977). Lupus nephritis and encephalopathy. Prognosis in 58 children. *Arthritis and Rheumatism*, **20**(Suppl.), 315–22.

Cassidy, J.T. and Petty, R.E (1995). *Textbook of pediatric rheumatology*, (3rd edn). W.B. Saunders and Co., Philadelphia.

Eichenfield, A.H., Athreya, B.H., Doughly, R.A., and Cebul, R.D. (1986). Utility of rheumatoid factor in the diagnosis of juvenile rheumatoid arthritis. *Pediatrics*, **78**, 480–4.

Elkon, K.B., Hughes, G.R.V., Bywaters, E.G.L., *et al*. (1982). Adult-onset Still's disease. Twenty year follow-up and further studies of patients with active disease. *Arthritis and Rheumatism*, **25**, 647.

European League Against Rheumatism (EULAR) (1977). *Bulletin 4: Nomenclature and classification of arthritis in children*. National Zeitung AG Basel, Basel.

Fink, C.W. (1995). Proposal for the development of classification criteria for idiopathic arthritides of childhood. *Journal of Rheumatology*, **22**, 1566–9.

Fraser, P.A., Hoch, S., Erlandson, D., Partridge, R., and ackson, J.M. (1988). The timing of menarche in juvenile rheumatoid arthritis. *Journal of Adolescent Health Care*, **9**, 483–7.

Hensinger, R.N. (1986). Limp. *Pediatric Clinics of North America*, **33**, 1355–64.

Holter, J.C. and Friedman, S.B. (1968). Child abuse. Early case finding in the emergency department. *Pediatrics*, **42**, 128–38.

Kunnamo, I., Kallio P., and Pelkonen, P. (1986). Incidence of arthritis in urban Finnish children. *Arthritis and Rheumatism*, **29**, 1232–8.

Lovell, D.J., Howe, S., Shear, E., Hartner, S., McGirr, G., Schulte, M., and Levinson, J. (1989). Development of a disability measurement tool for juvenile rheumatoid arthritis. The juvenile arthritis functional assessment scale. *Arthritis and Rheumatism*, **32**, 1390–5.

Malleson, P.N., Fung, M.V., Rosenberg, A.M., for the Canadian Pediatric Rheumatology Association (1996). The incidence of childhood rheumatic diseases in Canada. *Journal of Rheumatology*, **23**, 1981–7.

Marshall, B.A. and Tanner, J.M. (1986). Puberty. In *Human growth, a consprehensive treatise*, (2nd edn) (eds F. Falkner and J.M. Tanner), pp. 171–209. Plenum, New York.

McClelland, C.O. and Heiple, K.G. (1982). Fractures in the first year of life. *American Journal of Diseases of Children*, **135**, 26–9.

McGrath, P.J. and Unruh, A.M. (1987). *Pain in children and adolescents*. Elsevier, New York.

Petty, R.E. (1994). Limp. In *Handbook of pediatric emergencies*, (2nd edn) (ed. G.A. Baldwin), p. 338. Little Brown, Boston.

Quirk, M.E. and Young, M.H. (1990). The impact of JRA on children, adolescents, and their families. *Arthritis Care and Research*, **3**, 36–43.

Rosenberg, A.M. (1990). Analysis of a pediatric rheumatology clinic population. *Journal of Rheumatology*, **17**, 827–30.

Rosenberg, A.M. and Petty, R.E. (1982). A syndrome of seronegative enthesopathy and arthropathy in children. *Arthritis and Rheumatism*, **25**, 1041–7.

Salenius, P. and Vankka, E. (1975). The development of the tibiofemoral angle in children. *Journal of Bone and Joint Surgery*, **57A**, 259–61.

Sekiguchi, M. *et al*. (1985). On the mucocutaneous lymph node syndrome or Kawaski disease. In *Progress in cardiology*,13, (eds P.N. Yu and J.F. Goodwin), p. 97. Lea and Febiger, Philadelphia.

Southwood, T.R. and Malleson, P.N. (eds) (1993). *Arthritis in children and adolescents*. *Baillière's Clinical Paediatrics* 1, 3. Baillière Tindall, London.

Southwood, T.R., Petty, R.E., Malleson, P.N., Delgado, E.A., Hunt, D.W., Wood, B., and Schroeder, M.-L. (1989). Psoriatic arthritis in children. *Arthritis and Rheumatism*, **32**, 1007–13.

Special Writing Group of the American Heart Association (1992). Guidelines for the diagnosis of rheumatic fever. *Journal of the American Medical Association*, **268**, 2069–73.

Sutherland, D.H., Ohlsen, R., Cooper, L., and Woo, S.-L.-Y. (1980). The development of the mature gait. *Journal of Bone and Joint Surgery*, **62A**, 336–53.

Symmons, D.P.M., Jones, M., Osborne, J., Sills, J., Southwood, T.R., Woo, P., for the British Paediatric Rheumatology Group National Diagnostic Index. (1996). *Journal of Rheumatology*, **23**, 1975–80.

Tan, E.M., Cohen, A.S., Fries, J.F., Masi, A.T., McShane, D.J., Rothfield, N.F., Schaller, J.G., Talal, N.B., and Winchester, R.J., (1982). The 1982 revised criteria for the classification of systemic lupus erythematosus. *Arthritis and Rheumatism*, **25**, 1271–7.

Tanner, J.M. and Whitehouse, R.H. (1976). Clinical longitudinal standards for height, weight, height velocity, weighty velocity, and the stages of puberty. *Archives of Diseases in Childhood*, **51**, 170–9.

Vaughan, V.C.III and Litt, I.F. (1992). Assessment of growth and development. In *Textbook of pediatrics* (14th edn), (ed. R.E. Behrman), p. 32–43. Saunders, Philadelphia.

1.1.3 The geriatric age group

Evan Calkins

Rheumatological practice among older patients is substantially different from that with persons at midlife with regard to the array of diseases, the physiological and psychological characteristics of the patients, treatment goals, and the approach to management. This chapter will first outline, briefly, aspects of the biology of ageing that underlie these differences. Second, it will consider implications of the biology of ageing on rheumatological practice in a general sense—the geriatric perspective. Third, it will review how the process of ageing influences the characteristics of some of the specific rheumatological entities and their diagnosis and management.

The biology of ageing

Ageing is a continuous, life-long process, involving sequential changes in cells, organs, and the entire organism. Rapid changes occur during early developmental life and extend through youth and adolescence. Most functions achieve their peak by the time the person reaches his/her mid-twenties. From that point on, with very few exceptions, the physiological capacity of a given organ undergoes a pattern of progressive decline, illustrated schematically in Fig. 1, gradually approaching the minimal level required for maintenance of life.

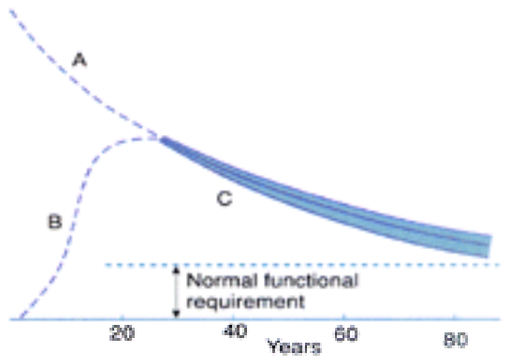


Fig. 1 Schematic depiction of the lifetime curve of physiological function and structure of many components of the body. Some start in embryonic life at negligible levels, others at high levels. Rapid changes occur in childhood and youth. Most functions reach their peak in the mid-twenties. This is followed by a gradual decline, approaching, in time, the minimal requirement for life, indicated by the dotted line. (Reprinted, with permission, from Katz *et al*. 1986.)

Figure 2 illustrates this concept by depicting the progressive structural changes occurring in the Meissner corpuscle (Cauna 1964). This is the nerve ending, located on the finger tips, which permits one to feel in one's pocket and differentiate a 1p from a 5p coin, or a dime from a penny. Numerous changes occur during embryonic development and the early years of childhood. Later changes evolve much more slowly, but they are progressive and, apparently, inexorable. One can hypothesize that, by the time one reaches the age of 40, it is difficult to distinguish a pound coin from a 10p, or a penny from a dime. Two or three decades later, the person has trouble differentiating a coin from a button.

Fig. 2 Changes in the Meissner corpuscle from embryonic life to old age. Changes occur rapidly during early development. They evolve more slowly as one ages, to a state of hypertrophy and, later, to an atrophic rudiment. (Reprinted, with permission, from Cauna 1964.)

Figure 3 depicts a more serious change—the sequential decrease in creatinine clearance as a person ages (Rowe *et al*. 1976). More recent studies have raised the question: do these changes reflect the process of ageing, itself, or the presence of intercurrent disease (Evans and Williams 1992; Lindeman *et al*. 1985). The fact that the decline occurs in approximately two-thirds of individuals, but not all, suggest that it is not a mandatory component of ageing. Still, the change is present in the

majority of older people, even in the presence of normal values for blood urea nitrogen and serum creatinine, and must be kept in mind as one prescribes drugs that are excreted by the kidney.

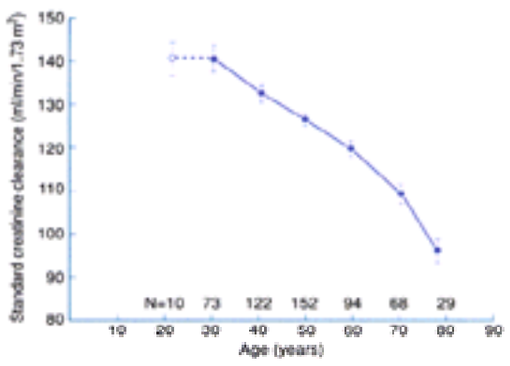


Fig. 3 The relationship of standard true creatinine clearance with age. Reprinted with permission from Rowe, J.W. *et al*. (1976). The effect of age on creatinine clearance in man: a cross sectional and longitudinal study. *Journal of Gerontology*, **31**, 155–63. Note: subsequent authors have shown that these changes occur in two-thirds of patients, but not all, and question whether they are due to age or age-related disease (Rowe *et al*. 1976).

Figure 4 depicts plasma levels of amylobarbitone sodium 4 and 24 h after administration to two groups of patients, one group aged between 20 and 40 years and the other over 65 years (Irvine *et al*. 1974). The data illustrates not only the increased plasma levels achieved in many older patients but also the variability of the response. Despite their superficially similar appearance, older people are more different, one from another, than younger individuals. This is a cardinal principle of geriatric medicine that should be kept in mind by all physicians who treat older persons.

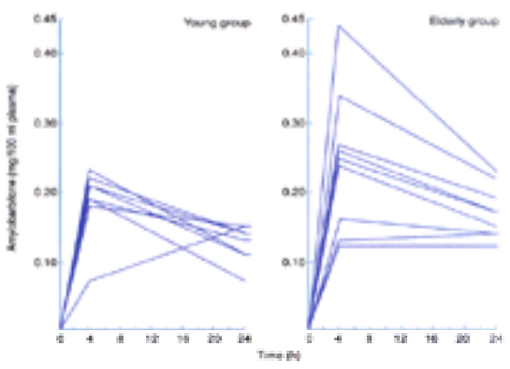


Fig. 4 Plasma levels of amylobarbitone sodium 4 and 24 h after oral administration of a standard dose to groups of subjects age 20–40 years (left panel) and over age 65 years (right panel). The drug is detoxified in the liver. Older persons show far greater variability than young subjects. (Reproduced from Irvine *et al*. 1974, with permission.)

Figure 5 was derived from a study with small animals (mice) but illustrates another important concomitant of the ageing process—the loss of homeostatic ability (Finch *et al*. 1969). In his PhD thesis project, Finch compared rectal temperatures of young and old mice following immersion in a bucket of ice water. The young mice responded by vigorous shivering and equally vigorous thrashing about and were able to sustain their body temperature. The old mice were able to muster only a feeble response and sustained deep hypothermia. Hypothermia is a common occurrence among old people, who have a reduced capacity to respond to external demands of many sorts, including infection and physical trauma, and have difficulty in sustaining homeostasis in the face of these demands. This is due, in part, to the decreased reserve capacity in the function of most organs as the person ages and also to attrition of neuroreceptors.

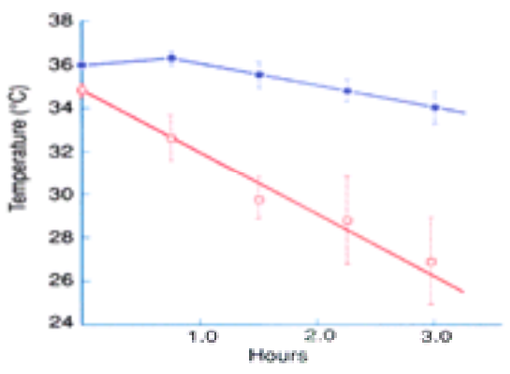


Fig. 5 The effect of age on colon temperature of mice during 3-h exposure to cold (9 to 10 °C). Seven mice per age group. Young adult mice (10 months old); senescent mice (30 months old).(Reproduced from Finch *et al*. 1969, with permission.)

While the general direction of the curve depicted in Fig. 1 is the same for different body components, the shape varies. For example muscle strength declines only gradually during the period between 24 and 40 years of age, and then begins a more rapid decline (Pendergast *et al*. 1993). Further, there is considerable individual variation within what is regarded as normal. Genetic variation may result in significant shifts in the curve in some individuals, resulting in premature death. Similarly, health habits, such as a lifetime pattern of smoking or exposure to industrial toxins, may result in more rapid declines. Temporary reversal of a downward slope may also occur but only with the greatest difficulty. For example vigorous exercise, such as achieved by marathon runners, will result in a substantial increase in the metabolic efficiency of skeletal muscle (Heath *et al*. 1981). However, once the new functional level has been achieved, progressive age-related declines still occur, despite whatever effort the runner may undertake. Although by appropriate preventative measures the decline in physiological function can be delayed to a certain point, no means has yet been achieved to add significantly to the time when declining function crosses the minimum requirement for life, leading to death (Strehler 1975).

Some age-related changes are easy to detect—greying hair, wrinkles, and decreased agility. Careful anthropometric and analytical studies have shown progressive changes in the dimension of many body components (Rossman 1977). These may be as trivial as broadening of the nose or lengthening of the ears, or significant, as in approximate doubling of total body fat, decrease in total body water, and an even greater decrease in body cell solids and bone mineral. Decreases in height are experienced by most men and women, averaging 4.9 cm in women and 2.9 cm in men. Although these losses are related primarily to decreased spinal length, there are also aged-related decreases in leg length. Narrowing of the shoulders and widening of the pelvis are also seen. Bennett, Wayne, and Bauer described progressive degenerative and proliferative changes in the articular surfaces of the knee, as persons age from the second to ninth decade (Bennett *et al*. 1942). McFarland and Dieppe pointed out that many older persons with osteoarthritis of the hands exhibit ulnar drift at the metacarpophalangeal joints, resembling the changes seen in rheumatoid arthritis (McFarland and Dieppe 1983).

While many of these changes are easily observed, the extent of a person's age-related physiological changes and decreased reserve capacity is often not apparent.

Most older people go to great lengths to conceal their disabilities from their friends, family, and society as a whole and also from their physicians. One of the reasons for this is the fear that they will be regarded as senile and a candidate for admission to a nursing home. It is only as the older person faces additional demands, such as an acute infection, that this lack of physiological reserve becomes apparent. As the organ system initially involved begins to fail, additional compensatory burdens are placed on other organ systems which may, already, be functioning at maximal capacity. This leads, in many instances, to a cascade of organ failure and to what is suddenly perceived as multisystem disease.

Taken together these losses add up to what can well be described as frailty, an almost universal characteristic of very old patients (in their 90s and older) and of many people in their 70s and 80s. This term refers not only to the decreased physiological reserve of many organs, but also, more specifically, to decreased muscle strength, balance, and overall functional capacity and to increased likelihood of falls.

An additional factor contributing to these age-related changes and, we are learning, the cause of many, is the decreased physical activity that almost always accompanies the ageing process. Inactivity itself contributes substantially to decline in function of muscles, bones, and other organs (Buchner and Wagner 1992). While cardiovascular deconditioning is the best example of this, we are just beginning to appreciate its widespread effects. For example it has recently been shown that the propensity of an older person to develop gastropathy, secondary to non-steroidal anti-inflammatory drugs, is increased in persons who are physically inactive and can be decreased as the result of physical exercise (Pahor *et al*. 1994a; Pahor *et al*. 1994b). Exercise is the most important means, to date, by which a person can influence his or her functional capacity and general fitness and, by this means, attenuate or even reverse some of the aspects of frailty of old age.

Reflecting, in part, the concepts of multiple organ failure referred to above, most persons age 75 years and older suffer from multiple chronic diseases. Gruenburg has pointed out that medical science has succeeded, at least for the time being, in preventing or treating many of the acute life threatening illnesses, such as serious infection, but has, as yet, had limited success in preventing or curing many of the chronic diseases such as osteoarthritis or macular degeneration of the eye—diseases which do not kill but linger on to collect, in multiples, as one ages (Gruenburg 1977). The presence of multiple diseases in most patients aged 75 years and older has major implications for diagnosis and treatment.

The geriatric perspective

How do these physiological changes accompanying ageing affect the practice of rheumatology? I have selected 10 areas for mention.

Age-related frequency of diseases

The influence of age, sex, and race on the incidence and frequency of specific diseases is a topic of increasing interest (Beeson 1994). These considerations are particularly relevant in the field of rheumatology. Table 1 lists some of the most common musculoskeletal and rheumatic diseases with regard to their age of onset. As the result of these age-specific relationships, the sort of patients one sees in a geriatric rheumatology consulting practice is quite different from that encountered with younger persons. Reiter's disease, Still's disease, and ankylosing spondylitis are not seen as new onset disease among older people. By contrast, osteoarthritis, polymyalgia rheumatica, giant cell arteritis, Sjögren's syndrome, and calcium pyrophosphate deposition disease occur almost exclusively in the older population. Equally important in clinical practice is the fact that musculoskeletal symptoms occurring in older persons may reflect a wide variety of conditions, many of them with serious implications, which are not usually regarded as falling within the constellation of rheumatic disease. Examples include carcinomatosis, multiple myeloma, hyperthyroidism, myxoedema, and a variety of neurological entities including Parkinson's disease and Alzheimer's disease. Differential diagnosis requires a broad understanding of many aspects of clinical medicine, including ones which fall outside the frame of reference of traditional rheumatology.

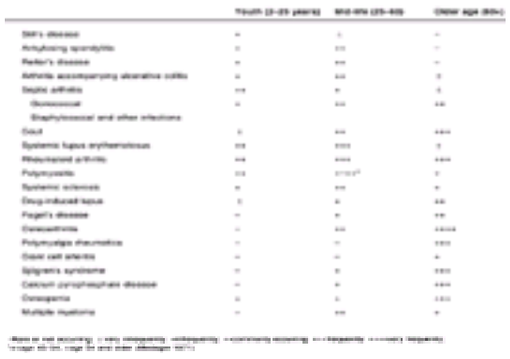


Table 1 Frequency of representative musculoskeletal diseases at various stage of life

The most frequent source of musculoskeletal complaints, in person age 70 to 79 years, is back pain, present in one study from Sweden (Bergstrom *et al*. 1986a; Bergstrom *et al*. 1986b), in 48, 55, and 32 per cent of women age 70, 75, and 79 years, respectively, and in 32, 20, and 28 per cent of men at similar ages. In a study from Iowa, United States, (Lavsky-Schulan *et al*. 1985), 25 per cent of woman and 20 per cent of men, age 75 to 79, were reported as having experienced severe back pain—78 per cent of these women and 68 per cent of these men had sought the help of physicians; 44 per cent of the women and 82 per cent of the men had obtained chiropractic care; 29 per cent of the total group had been hospitalized at least once for their back pain and 8 per cent of the woman and 2 per cent of men had undergone back surgery. Clearly, back pain emerges as one of the most significant aspects of geriatric/rheumatological practice.

The next most frequent musculoskeletal condition, in this age group, is osteoarthritis (Lawrence *et al*. 1989). In the Swedish study mentioned above, at age 70 years 12 per cent of the sample (19 per cent of the men and 6 per cent of the women) were judged, on clinical grounds, to have osteoarthritis of the knees. Surprisingly, these figures declined among those who survived to age 75 and 79 years (5 per cent and 6 per cent respectfully). Radiographic changes, consistent with the diagnosis of osteoarthritis of the knees, were present in 20 per cent of the patients at 70 years of age (26 per cent of women and 15 per cent of men), but in only 14 per cent of those who survived to age 75 years and 13 per cent of those who survived to age 79 years. Since it seems highly unlikely that the radiological changes are reversible, this unexpected finding suggests the possibility that osteoarthritis, or concomitant conditions of which it is a marker, may be associated with a decreased life span. In view of the effect osteoarthritis has on a person's physical function and the benefits of exercise on overall health, the association is not totally illogical. A recent study has lent added support to this hypothesis (Cerhan *et al*. 1995).

In the Swedish study, the frequency of symptoms of rheumatoid arthritis, either present or as a sequel to earlier onset disease, were evidenced in 7 per cent of the 537 patients age 79 years. In three women (0.9 per cent of the total group) and three men (1.5 per cent) the disease had begun within the previous 4 years, that is after age 75 years. Almost identical frequencies were derived from the clinical examination of a subset of these patients. The fact that rheumatoid arthritis is an age-related disease afflicting an increasing percentage of patients at least up to age 79 years, with new-onset disease occurring as late as age 75 to 79, seems well established (Mikkelsen *et al*. 1967; Linos *et al*. 1980).

Multiple diseases

The symptoms of rheumatic diseases are, themselves, relatively non-specific. When a person suffers from a single disease the product of these symptoms gains specificity and one reaches the correct diagnosis relatively easily. When a person has multiple diseases, several of which may be characterized by certain of these non-specific manifestations, or none, the pattern provides a special challenge to the careful clinician. For example because a patient presents with the classic stigmata of long-standing rheumatoid arthritis does not necessarily mean that the symptoms of which the patient currently complains are due to that disorder. An example was a 65-year-old man with long-standing, deforming rheumatoid arthritis who was referred to the author by a primary care physician because of failure to respond to gold. History indicated the presence of severe constitutional symptoms, with sweats and weight loss, which the primary care physician had attributed to increased activity of the rheumatoid process. However, the patient did not complain of morning stiffness and, on physical examination, the joints did not display the extent of active synovitis one would have expected in a patient with constitutional symptoms of that degree. Further studies disclosed the presence of Hodgkin's disease. Another example was a 62-year-old man with extensive radiological evidence of cervical spondylosis and numbness and tingling in the upper extremities. Further work-up disclosed the presence of pernicious anaemia with significant improvement in the neurological manifestations following administration of vitamin B_{12}. For this reason, in an older patient referred for a musculoskeletal disorder it is important to conduct a complete history and physical examination, to explore the

possibility that the symptoms may not reflect the musculoskeletal disease which is initially apparent but may be due to a less obvious concomitant disorder. The examination should include a rectal examination for carcinoma of the prostate, breast examination, and neurological examination.

Use of drugs

The use of drugs in older patients with rheumatic diseases holds promise of great benefit but also great hazard due to the omnipresent threat of adverse drug reactions, often serious and sometimes fatal (Girdwood 1974; Beers and Ouslander 1989; Griffin 1991). Fries *et al*. have stated, with regard to the non-steroidal antirheumatic drugs (non-steroidal anti-inflammatory drugs), 'gastropathy associated with non-steroidal anti-inflammatory drug use probably represents the most frequent drug side-effect in the United States' (Fries *et al*. 1989). Reasons for the increased side-effects of drugs in older patients include changes in pharmacokinetics, pharmacodynamics, and the issue of drug interaction—the effect that one drug may have on the pharmacokinetics of a second, simultaneously-administered agent (Furst 1988; Montamat *et al*. 1989; Feely and Coakley 1990; Gurwitz and Avorn 1991). In the field of rheumatology, the increasing tendency to prescribe multiple antirheumatic and anti-inflammatory drugs to a given patient provides the potential for serious interactions of this sort. These problems become compounded in the person who suffers, simultaneously, from a range of additional chronic diseases, including, perhaps, hypertension, diabetes mellitus, gout, deep venous thrombosis, peptic ulcer disease, and a urinary tract infection.

Several guidelines may prove helpful (Table 2). A careful drug history should be obtained in all older patients, identifying drugs prescribed by each of the different physicians the patient is seeing, and also over-the-counter medications. It is important to limit, to a minimum, the number of drugs a patient is expected to take. In some patients carefully designed exercise, local heat or cold, patient education, and, perhaps, a local injection will avoid utilization of a systemic agent. Few patients can adhere, accurately, to the requirements for taking more than three or four different drugs at one time. Many older persons, especially those who are seeing several different physicians at the same time, will be expected to comply with instructions to take eight or nine drugs, each according to its own time schedule! Simplifying the number of drugs will make life much easier for the patient, decrease the chance of non-compliance or a compliance error, and decrease the likelihood of an adverse drug reaction.

Obtain a list of all current medications, including over the counter preparations, and document any adverse reactions that may have occurred.
Limit the number of drugs. Use non-pharmacological treatment when appropriate
Identify mode of excretion, detoxification, and drug interactions of all drugs you use.
Reduce dosage for older patients. Start treatment with even lower dosage. Instruct patients about possible side-effects.
Obtain appropriate pretreatment laboratory data to assess risk of therapy and provide baseline for identification of possible future toxicity.
Define specific goals for each drug used. Discontinue drug if goal is not achieved.
If toxic effects to drug A ensue, it is better to shift to another agent (B), than to add a drug to offset the toxic effects of drug A.

Table 2 Guidelines for safe administration of drugs in older patients

A physician should be fully informed concerning the mode of excretion or catabolism of each drug he or she is prescribing, and also the potential for interactions. It is wise to limit the number of different drugs a given physician selects to include in his or her therapeutic armamentarium, and to select for a given patient agents with a maximal ratio of benefit to risk, especially in patients who are physiologically compromised in the first place.

Start treatment with a low dose. For most older patients, the maintenance dose will be lower than it would be for younger individuals. At the onset of therapy it is well to prescribe the first few doses at an even lower level, perhaps one-quarter or one-half of the expected maintenance dose. The patient should be instructed to discontinue the drug immediately and telephone the office if side-effects ensue. Base-line laboratory data, relevant to the area in which problems might occur, should be obtained prior to starting a new drug. These values should be repeated at appropriate intervals.

For any drug, the risk of adverse effects varies with each individual patient. The definition of patients at increased risk, in response to a number of agents, is becoming clearer. For example for non-steroidal anti-inflammatory drugs there is a growing consensus that the chance of gastropathy is significantly enhanced if the patient has a history of previous peptic ulcer, previous upper gastrointestinal bleeding of any cause, and in patients with multiple chronic diseases (Llewellyn and Pritchard 1988; Janssen *et al*. 1994). Care should be taken in prescribing warfarin for patients with a predilection for falling.

It is useful to define the specific goal one hopes to be achieved through use of a given agent. If toxic effects ensue, it is much better to shift to a different agent than to add yet another drug, designed to decrease the severity of the adverse drug effects. The entire drug and treatment programme should be reviewed periodically. Drugs that have not been proven to be effective should be discontinued.

Laboratory data

Carefully selected laboratory data is an essential part of pharmacological therapy for three reasons:

1. by helping to assess the risk of toxicity in given patients;
2. to identify persons in whom the drug should not be used;
3. in monitoring patients for potential side-effects.

Laboratory data also constitutes an important component of diagnosis. A great deal of data on this topic is provided elsewhere in this book. Because of the possibility that a given patient may have several, simultaneous, unrelated clinical disorders, a carefully selected group of laboratory determinations should be carried out on all new patients with arthritis, except, perhaps, in its most localized form. While the list will vary in different clinics, Table 3 reflects our policy in this regard.

Obtain always or almost always
Complete blood cell count and differential
Biochemical screen, including blood urea nitrogen, creatinine, glucose, electrolytes, calcium, phosphorus, alkaline phosphatase, uric acid, albumin, total serum protein, and liver enzymes
Urinalysis
Radiograph of hands and other involved joints
Obtain if appropriate
Antinuclear antibody (ANA) screen
If ANA screen is positive, obtain ANA titre, with pattern, anti-dsDNA, and other antinuclear and anticytoplasmic antibodies as appropriate
Rheumatoid factor
Chest radiograph
T_4 thyroid-stimulating hormone
Synovial fluid examination for culture, crystals, and cell count
Other diagnostic studies: Schirmer's test, electromyography, bone scan, MRI, CT scan, arthroscopy, dual energy X-ray absorptiometry, biopsy of skin, muscle, kidney, lacrimal gland, etc.

Table 3 Baseline laboratory studies in the diagnosis of musculoskeletal disease

Established values for most blood chemistries and special studies conducted in rheumatology are the same for older persons as for younger individuals. This is not true, however, for the erythrocyte sedimentation rate, which will often exhibit values of 25 to 30 mm/1 h in an older person without demonstrable rheumatic or inflammatory disease (Fig. 6). Similarly, both the rheumatoid factor and antinuclear antibody test, in low dilution, will be exhibited by many apparently normal older

persons in the absence of any demonstrable inflammatory or rheumatic disease (Hallgren *et al*. 1973).

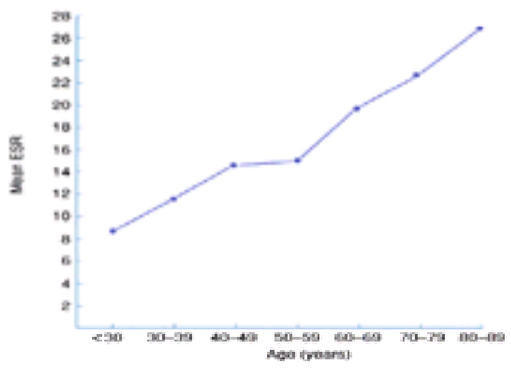


Fig. 6 Increase in mean erythrocyte sedimentation rate with advancing age. Wintrobe method, measured in millimetres per hour. (Reproduced from Hayes and Stinson 1976, with permission.)

Exercise

As noted earlier, muscle function undergoes a progressive, age-related decline. Conspicuous among the changes accompanying this decline are alterations in length–tension relationships (Pendergast *et al*. 1993). In youth, muscles act something like elastic bands—the more one stretches them, the greater the resistance. As one ages, however, this relationship changes and maximum strength or resistance is achieved somewhat short of the point of the maximum length. This results in an effective shortening of the muscle, contributing to the muscle aches experienced by many older people, especially those with arthritis (Fisher *et al*. 1991*b*).

Muscle exercises are of two sorts, aerobic and anaerobic. Aerobic exercises, often recommended to patients in later life, consist primarily of flexibility exercises. These provide entertainment and some improvement in flexibility and balance, but do nothing to improve strength. Anaerobic exercises, however, involving vigorous muscle contraction against resistance, do result in increased strength and in an alteration of the strength–tension relationship toward normal. Exercises of this sort can be conducted, safely, under supervision, among older persons, even those in their late 90s, and result in significant increases in strength (Fisher *et al*. 1991*a*; Fiaterone and Evans 1993; Pyka *et al*. 1994). Unfortunately, if the exercise is confined to a relatively short period, 3 to 6 months for example, the strength will usually return to its original level over ensuing months. To be effective, therefore, a somewhat less intensive programme of continued exercise should be maintained.

These considerations have significant relevance to the field of rheumatology. Arthritis, either generalized or localized to a given joint such as a knee or hip, almost inevitably leads to decreased overall physical activity, with significant consequences in terms of overall muscle strength, cardiovascular function, the likelihood of falls, and other components of frailty. However, these declines in muscle function in patients with osteoarthritis have the capacity to be reversed by well designed anaerobic exercise (Fisher *et al*. 1991*b*).

Since exercise is maximally effective at the specific muscle length at which it is conducted, it should be carried out at several different muscle lengths for the greatest benefit. For example the attachment of the quadriceps muscle is such that it is at its greatest length when the hip is fully extended (i.e. the patient is lying flat). Therefore, quadriceps strengthening exercises should be carried out with the patient both in the sitting and lying position with, ideally, an intermediate position added. A given programme of exercise will result in a slight increase in muscle aching, analogous to that experienced by normal people after a 'workout'. The exercise should not be conducted to the point of actual pain, which is not only uncomfortable for the patient but actually inhibits muscle function and the effectiveness of the programme.

Unfortunately, exercises of this sort are extremely boring. Few patients will actually carry them out by themselves at home. We have found that a series of exercise classes, including approximately 12 patients at a time, and combining exercise with patient education and social support, are enthusiastically received by many patients. Twenty classes, 1 1/4 h each, conducted over the course of 10 weeks, have led to significant enhancement in overall function, balance, and psychological well being (Gunther, J., Brown, D., Karuza, J., and Calkins, E., in press).

Diseases of the joints and muscles not only yield pain and functional limitation, they also assume importance because they tend to limit the ability of the person to participate in normal physical activities such as housework, chores, climbing stairs, and sports. Thus, by appropriate treatment of musculoskeletal problems which inhibit these normal patterns of exercise, the rheumatologist can make an important contribution to the patient's overall health, happiness, and, possibly, longevity.

Prevention

One of the characteristics of frailty or fragility is that it is easier to prevent damage than to repair it. This is particularly true for the geriatric population (German and Fried 1989). Despite progressive functional losses in many organ systems, many people, age 80 years and older, are able to function quite well until some relatively minor episode—an attack of influenza or a fall—triggers a cascade of events with high probability of leading to the need for extensive (and expensive) medical care, losses of functional capacity, increased need for support services, possible institutionalization, and, in some, death.

Thus, a significant portion of the attention given to medical care of older people should be devoted to careful consideration of the full range of measures that will decrease the likelihood of intercurrent illness and accidents, such as falls. In addition to exercises, these measures include use of a cane as an aid to balance as well as to ease pain in affected joints, avoidance of drugs (such as the benzodiazepines) which are known to predispose to falls, and appropriate safety measures at home, such as avoidance of throw rugs and provision of good illumination on stairs. Assurance that these measures are being properly addressed is an appropriate component of rheumatological practice with older patients.

The best current example of the importance of prevention in the field of musculoskeletal disease is osteoporosis (Odell and Heath 1993). In an article entitled 'Health of the nation and osteoporosis', Dixon (Dixon 1992) pointed out that prevention of osteoporosis ideally fulfils the three basic criteria for prioritizing preventive care:

1. the burden of unprevented disease is great;
2. opportunities for prevention exist;
3. practical targets can be set.

In 1989, a report of The Royal College of Surgeons pointed out that patients with hip fractures occupied 20 per cent of orthopaedic beds; 80 per cent were woman over 65 years. The direct hospital costs were £160m at 1977 to 88 prices. Approximately 20 per cent of patients with hip fractures are dead in 6 months and, of the remainder, approximately half lose independence. Patients who are thin, smokers, with a family history of severe osteoporosis, and minimal calcium intake during adolescence are at greatest risk. Initiation of oestrogen replacement at the earliest indication of menopause, especially in high risk patients, is highly effective in preventing this bone loss. Supplementation of calcium intake, to 1500 mg/day, together with small amounts of vitamin D (400 international units either daily or 3 times a week) further contributes to maintenance of bone mineralization. Exercise, especially of a weight bearing variety, will also contribute to retaining calcium in the bones and, possibly, some degree of increased mineral content.

The opportunity to address late-life osteoporosis has recently received an important impetus through two developments. The first is the wide availability of equipment for dual-energy X-ray absorbtiometry, which provides an excellent, relatively low cost means of assessing the mineral content of both vertebral and long bones. In addition to assisting in the identification of early osteoporosis, this technique provides the first readily available means of following the course of patients and determining the effectiveness of a given regimen.

The second advance is the increased appreciation of the effectiveness of two agents in actually reversing osteoporosis—calcitonin and the bisphosphonates. The latter are proving to be the most practical. These agents are synthetic analogues of inorganic pyrophosphate. Like pyrophosphate, they have a high affinity for hydroxyapatite but are resistant to metabolism by endogenous phosphatase. They act both by inhibiting bone resorption and also by inhibiting bone mineralization.

The agent whose use initially demonstrated the potential benefit of this form of therapy, etidronate, exhibits both actions at approximately the same dosage. In clinical practice, the inhibitory effect on bone mineralization is decreased, however, by the administration of the drug in a cyclical fashion, for 2 weeks every 4 months (Watts *et al*. 1990; Harris *et al*. 1993). Even so, there is some evidence suggesting that after 2 years of therapy the effect of the drug on diminishing mineralization of bone may become apparent. It is questionable whether use of the drug should be continued for periods of 3 years or more.

A more recently developed agent, alendronate, is about 1000 times more potent, in terms of inhibiting bone resorption, than etidronate and provides effective inhibition of bone resorption at doses that do not affect mineralization. The agent can be taken daily, on a continuous basis. It is recommended to be taken in a dose of 10 mg/day, one-half hour before breakfast. This avoids the complexity of cyclical therapy and provides the potential opportunity for continuous usage over a long period of time. A recent, large-scale, controlled study showed that daily administration of the drug increased the bone mass of the spine, hip, and total body and reduced the risk of vertebral fractures, the progression of vertebral deformities, and the height loss in both men and women with osteoporosis (Liebermann *et al*. 1995). Although experience has not yet been gained in long-term use of this drug, initial reports suggest that this agent, and others currently being developed, have a high likelihood of yielding a major impact on the frequency and consequences of osteoporosis (Sambrook 1995). It has been shown that administration of etidronate will prevent the loss of bone mineral density in patients receiving long-term prednisone therapy (Mulder and Struys 1994), and one would anticipate that alendronate would also be effective in this regard.

The one period in life when bones readily absorb dietary calcium lies during the 3 or 4 years immediately following puberty (Lloyd *et al*. 1993). There is a direct association between calcium intake of children during this period and the bone mineral mass which they will carry the rest of their lives. As Dixon has observed, 'good bones last a lifetime' (Dixon 1992). It is, therefore, recommended that all children receive diets, appropriately supplemented, which will yield an intake of approximately 1500 mg of calcium daily. This is one of several examples of the fact that patterns of life during childhood have a direct relationship to the way one responds to the challenge of ageing.

Psychological factors

The most common and, probably, most devastating loss, often but not always associated with advanced age, is that of cognitive capacity. The extent of this loss is not readily identified. Older patients are very good at concealing it. Physicians seldom include objective assessment of cognitive capacity as part of their physical examination and frequently overlook losses which have a significant effect on the person's ability to comply with medication regimens and rehabilitation guidelines. Assessments of cognitive capacity can easily be performed through a series of simple questions requiring not more than 10 min (Folstein *et al*. 1975). While Alzheimer's disease and multi-infarct dementia are the most common causes of cognitive loss at this age, there are a number of other causes, some of which are reversible, at least on a short-term basis (Rabins *et al*. 1984). These include depression, delirium, (often secondary to infection or inappropriate medication), hyperthyroidism or myxoedema, hypercalcaemia, and vitamin B12 deficiency. Depression has been shown to be present in approximately 20 per cent of patients in an ambulatory care practice (Blazer 1989). It, too, is often overlooked by physicians. Good objective means of assessment have been designed (Burnam *et al*. 1988) but are rarely put into effective use in a primary care practice. The pain and limitation of musculoskeletal diseases, often preventing patients from participating in life-long activities and interests, produce, in many patients, a sense of discouragement and depression. For this reason it is important that the rheumatologist become skilled in differentiating these symptoms from those of true depression, which is readily amenable to appropriate pharmacological therapy. Chronic anxiety and, not infrequently, panic attacks are other psychological problems experienced by older patients, frequently overlooked by physicians, and readily amenable to appropriate therapy (Lindesay *et al*. 1989; Brown *et al*. 1991).

Older patients also experience serious social losses. These include loss of a job, decreased income, death of friends, possible death of a spouse, children moving away, and possible relocation from the house one has occupied for 40 or 50 years. Despite these losses, essentially all older patients adhere, fiercely, to two goals: ability to live independently in the community and retention of a measure of control. They are aided by two assets—experience and social skills they have acquired since childhood. It is fascinating to see how different people use these two resources to fend off or compensate for the often staggering social and physical losses, sometimes maintaining their proud independence into their early 90s. Others, endowed with a less flexible, resourceful, and determined psychological makeup, crumple at an early age in the face of far less formidable losses and obstacles.

It should be clear from earlier portions of this chapter that the way one responds to the challenge of musculoskeletal disease has a great deal to do with the outcome. Therefore, attention to the psychological and social aspects of a patient emerge as major considerations in a rheumatological practice with older people.

Nutrition

Maintenance of adequate nutrition, in terms of intake of protein, calories, vitamins, calcium, and trace metals is an important component of the ability to maintain independence (Morley *et al*. 1986; Lipsitz 1992). While protein calorie malnutrition is seldom included among the discharge diagnoses of hospitalized patients, it is present, to a significant degree, in a high percentage of hospitalized persons, especially older people, and is a major factor contributing to their inability to respond to acute illness, especially infection (Bienia *et al*. 1982).

Systematic observation of the patient's weight is an important component of ambulatory care practice. A loss of weight, of 8 to 10 per cent over the course of 6 months, is an indication of serious disease or inadequate food intake and should stimulate thorough assessment. While losses of this sort may reflect coincidental chronic illness, older people with rheumatic diseases, especially those living alone, may face difficulties in gaining access to the grocery store, transporting food to their homes, proper preparation of meals, and the process of eating itself. Meals on Wheels and similar programmes assist substantially by providing one good meal a day and a small amount of social support.

Environmental considerations

In addition to efforts to improve the patient's functional capacity, the ability to maintain independence can be enhanced through appropriate attention to the availability and usefulness of the wide range of functional aids that are currently available. These include various types of wheelchairs, canes, and other devices to assist ambulation, and special equipment to aid in eating, dressing, hearing, seeing, and attending to personal needs (Wasson *et al*. 1990). Modification of the physical environment by elevated toilet seats, appropriate ramps, and introduction of community-based resources for social support are also important. While few physicians have the time to assume direct responsibility for instructing patients concerning the use of these resources, the physician, in our culture, is the person to whom the patient turns as the ultimate source of help. Physicians should be aware of the importance of these aspects of care and assume some responsibility for seeing that these needs are met.

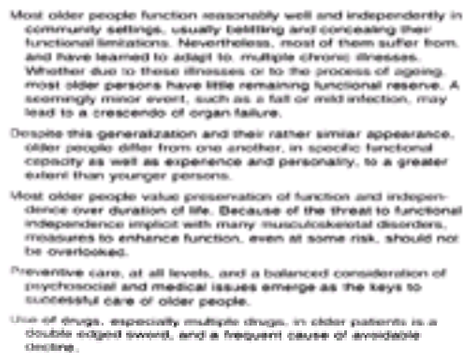
Organizing comprehensive care for the older patient with a rheumatic disease

Thus, the goal of therapy in persons of advanced age, especially 80 years and older, is substantially different from that in a younger person or someone at midlife. In a 30-year-old person with early rheumatoid arthritis, for example, the goal is to do everything one can to achieve a remission, avoiding, if possible, long-term commitment to an agent such as continuing prednisone therapy, which would have a deleterious impact over the decades to come. In a person age 80 years with a similar disease, the goal is to restore the person, in the shortest possible period of time, to the previously established lifestyle. This may involve, for example, early institution of low-dose prednisone therapy, with relatively little concern as to the effect of this agent over the course of a decade or more to come. For a person whose ability to continue independent life is compromised by a painful hip with serious limitation in motion, one might, in consultation with the patient and family, recommend a joint replacement, realizing that the operative risk, in a person of this age, will be significantly enhanced.

Thus, for older persons with arthritis, the programme of management transcends what is, traditionally, the province and training of the physician. It can only be carried out through a team approach—an interdisciplinary pattern of management in which various aspects of care and support are entrusted to people with special skills and experience (Halstead 1976; Rubenstein *et al*. 1994). An approach of this sort is utilized in many, probably most, centres for comprehensive care of patients with arthritis, and is essential in geriatrics. It involves recognition by the physician that he or she can not be the sole provider of health care nor even, necessarily, the 'person in charge'. This approach to management involves interdisciplinary skills rarely if ever taught in medical school or fellowship training. To be effective, it also requires a comprehensive assessment of the range of problems confronted by an individual patient. This has led to the development of a variety of assessment instruments or tools, which have become essential component of geriatric care (Lawton and Broady 1969; Katz *et al*. 1970; Reuben and Siu 1990; McHorney *et al*. 1994).

Finally, and particularly important, is the topic of patient education (Winfield 1989). As Wood has pointed out 'We cannot expect major reduction in the scale of problems for which some help from the health and social services will be needed. The strategic goal should be towards policies of enablement, seeking ways of

maximizing the ability of old people to handle their own health and other problems' (Wood 1992). Measures to achieve this objective, through patient education and mutual support groups, become important components of comprehensive rheumatological practice. Information concerning various aspects of arthritis, appropriate for patients and families, is included in the publication *Arthritis today*, available from the Arthritis Foundation, 1314 Spring Street NW, Atlanta, GA 30309, telephone (404) 872–7100. Other national organizations may provide similar support. Important points to remember regarding special characteristics of older patients and their management are summarized in Table 4.


Most older people function reasonably well and independently in community settings, usually belittling and concealing their functional limitations. Nevertheless, most of them suffer from, and have learned to adapt to, multiple chronic illnesses. Whether due to these illnesses or to the process of ageing, most older persons have little remaining functional reserve. A seemingly minor event, such as a fall or mild infection, may lead to a crescendo of organ failure.

Despite this generalization and their rather similar appearance, older people differ from one another, in specific functional capacity as well as experience and personality, to a greater extent than younger persons.

Most older people value preservation of function and independence over duration of life. Because of the threat to functional independence implicit with many musculoskeletal disorders, measures to enhance function, even at some risk, should not be overlooked.

Preventive care, at all levels, and a balanced consideration of psychosocial and medical issues emerge as the keys to successful care of older people.

Use of drugs, especially multiple drugs, in older patients is a double-edged sword, and a frequent cause of avoidable distress.

Table 4 Important points to remember: Special characteristics of older patients

The geriatric perspective on specific rheumatic diseases

Osteoarthritis

Osteoarthritis is a typical geriatric disorder. Although changes in histological appearance of cartilage and adjacent bone can be seen as early as the third decade, clinical manifestations rarely appear before the fifth or sixth decade. Details concerning the epidemiology, pathology, and clinical manifestations of osteoarthritis are presented elsewhere in this book. Numerous studies have described the physical and biochemical changes of cartilage associated with ageing and with osteoarthritis. Unfortunately, except for elucidating a few genetically controlled abnormalities, these studies have, to date, produced relatively little information concerning the pathogenesis and means of prevention of common forms of osteoarthritis. In this author's opinion, much more attention needs to be paid to the role of stresses on joints, secondary to the way people, especially children, use their legs. Studies on effectiveness of measures to correct abnormal gait in children, through orthotics and other means, on the development of osteoarthritis in late life are difficult to carry out, due to the element of time, but might pay important dividends. Such studies would be consistent with a theme I have tried to present in this chapter—that ageing is a life-long process and that patterns established during childhood have a great deal to do with the way people respond to the challenge of ageing.

Clinical aspects

One of the striking features of this disorder, of particular relevance in practice with older patients, is the lack of concordance between radiological findings and clinical symptomatology. Many patients with severe radiological changes indicative of osteoarthritis of the spine, for example, exhibit no symptoms. Conversely, other patients will complain of severe disabling back pain without radiological evidence of osteoarthritis or any other identifiable structural disorder.

The reasons for this lack of concordance are not fully understood. There are wide discrepancies in pain perception among individual persons (Lichtenberg *et al*. 1986). While this undoubtedly reflects, to some degree, physiological differences, it is inescapable that the major differences are psychological in nature. Musculoskeletal symptoms are strongly influenced by psychological factors such as attitudes towards pain and disability and the sense of control over one's disease (Keefe *et al*. 1987). Failure to recognize this aspect of rheumatological practice and to respond to it appropriately accounts for a considerable measure of the frustration experienced by many primary care physicians in attempting to treat patients with osteoarthritis.

The lack of concordance between radiological evidence of osteoarthritis and symptoms is important for another reason. The existence of radiological evidence of osteoarthritis, whether in the neck, spine, hips, or knees, and pain generally referable to that area should not be taken as evidence that the pain is due to the osteoarthritis rather than a concomitant problem of an entirely different sort, such as carcinoma of the pancreas or prostate, multiple myeloma, dissecting aortic aneurysm, or a fracture of an adjacent bone.

In older patients with osteoarthritis there is a general agreement that acetaminophen in good doses is the safest agent for relief of pain and is effective in many patients. Two 650 mg tablets three times a day can be tolerated by most patients. Higher doses may result in a 'spaced out' sensation. The possibility of nephrotoxicity has recently been emphasized. This is much more likely to be seen in patients with already-established renal disease than in otherwise normal persons. Serum estimations of creatinine and blood urea nitrogen should be obtained during long-term treatment with acetophenetidin. The hazards accompanying use of the non-steroidal anti-inflammatory drugs in older patients has already been discussed. These drugs should be used with the greatest caution or omitted altogether in patients who present high risk for toxic side-effects. Because of the hazards of addiction, codeine and other narcotics should not be used in patients with osteoarthritis or other chronic forms of arthritis, except under unusual circumstances and for short periods of time. Although propoxyphene has been widely criticized and is regarded, by some, as inappropriate for use in older persons, we have found that 75 mg of propoxyphene once daily, perhaps in the evening, in addition to acetaminophen or aspirin, is helpful in many older patients. Higher doses of propoxyphene may be necessary in some patients.

The concepts concerning design and implementation of exercise programmes, discussed earlier, are particularly applicable and effective in patients with osteoarthritis. So, too, are earlier statements on the use of assistive aids, environmental support, the issue of multiple disease, and the concepts of interdisciplinary team care. Details on surgical management are presented elsewhere in this book.

Rheumatoid arthritis

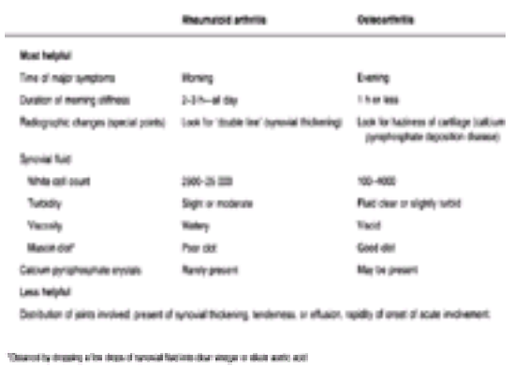
As pointed out earlier, very few demographic studies on the incidence or characteristics of rheumatoid arthritis in persons 70 years of age and older are available, and the studies that have begun to appear contain little information on the clinical characteristics of patients in this age group. Therefore, our understanding of the clinical features of this disease in older persons is based on observations of patients who present themselves or are referred to the rheumatologist's office or hospital, and are subject to wide variations in patient selection. Persons with this disorder fall into two general categories—those whose disease commenced many years before and now present with sequelae of the disease and previous modes of therapy, and those with new onset disease. Little information is available concerning the characteristics of persons in the former group. One gains the impression that far fewer of them exists than one would have anticipated from incidence rates over the decades, consistent with other data concerning the negative impact of rheumatoid arthritis, or its treatment, on life expectancy (Allebeck 1982; Pincus and Callahan 1986).

A number of clinical studies have suggested that new onset disease, in older persons, is different from that in people at midlife (Deal *et al*. 1985; Healey 1986). Approximately two-thirds of the cases exhibit a gradual onset, occurring over the course of weeks or months, and constitutional manifestations are mild. In a smaller group of patients the arthritis develops much more rapidly, often over the course of a few days. These patients may experience severe constitutional symptoms, including fever and even chills. They appear to have a greater chance for spontaneous remission than do patients in the gradual-onset group. Although many older persons with new onset rheumatoid arthritis will have a positive titre for rheumatoid factor, one must bear in mind that rheumatoid factor in low titre will be seen in many older persons without any evidence of rheumatic or inflammatory disease (Hallgren *et al*. 1973). In our experience, the number of older persons with new onset disease who exhibit the clinical features of seropositive rheumatoid arthritis is rather small.

In view of the extensive coverage of the pathogenesis, course, manifestations, and treatment of rheumatoid arthritis in other sections of this volume, this discussion will focus on selected areas of special relevance in the older patient.

Differential diagnosis of rheumatoid arthritis and osteoarthritis

This differential has been presented, for years, to entering medical students as one of the pedestals of rheumatology. Among older people, however, the differential is far less easy. It is also important in that antirheumatic therapy, such as methotrexate or low dose prednisone, appropriate for use in patients with rheumatoid arthritis, is not indicated in osteoarthritis and will be of little value. For this reason, uncertainty concerning this diagnosis is one of the chief reasons for referrals from the primary care physician to the geriatric rheumatologist. Table 5 summarizes points that we have found to be helpful.



	Rheumatoid arthritis	Osteoarthritis
Most helpful		
Time of major symptoms	Morning	Evening
Duration of morning stiffness	2–3 h—all day	1 h or less
Radiographic changes (special joints)	Look for 'double line' (synovial thickening)	Look for fuzziness of cartilage (calcium pyrophosphate deposition disease)
Synovial fluid		
White cell count	2000–25 000	100–4000
Turbidity	Slight or moderate	Fluid clear or slightly turbid
Viscosity	Watery	Viscid
Mucin clot*	Poor clot	Good clot
Calcium pyrophosphate crystals	Rarely present	May be present
Less helpful		
Distribution of joints involved, present of synovial thickening, tenderness, or effusion, rapidity of onset of acute involvement		

*Obtained by dropping a few drops of synovial fluid into dilute vinegar or dilute acetic acid

Table 5 Differential diagnosis of rheumatoid arthritis and osteoarthritis in an older patient

Speed of onset

In theory, this should be a fundamental part of the differentiation. Osteoarthritis is a chronic disease with, usually, a gradual onset while, especially in older persons, rheumatoid arthritis may exhibit a rapid onset, over the course of 1 week. However, older patients with osteoarthritis will often exhibit sudden onset of swelling, tenderness, and effusion in a given joint, especially a knee, following relatively minor trauma. In some instances this will be due to haemarthrosis due to the 'nipping' of a synovial lip by an osteoarthritic spur.

Distribution of involved joints

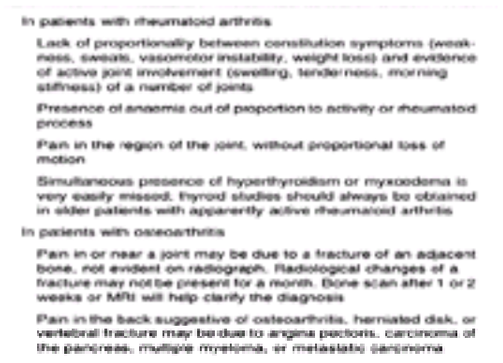
This differential, too, may be blurred. Osteoarthritis frequently involves the proximal interphalangeal joints and this may over-shadow changes in the distal interphalangeal joints. Occasionally, involvement will be seen in the metacarpophalangeal joints as well, even including a mild ulnar drift (MacFarlane and Dieppe 1983). While synovial thickening, accompanied by warmth and tenderness, is a characteristic feature of rheumatoid arthritis, these changes are often seen in patients with osteoarthritis, especially in the knees. It has been speculated that this may reflect immunological reaction to breakdown products of cartilage. Alternatively, this may be due to calcium pyrophosphate deposition disease, a frequent concomitant of osteoarthritis.

The time of day when symptoms are most severe, and the presence or absence of morning stiffness, are both useful in differentiating these entities. Further help can be derived from the radiograph. The differential points are exhibited with special clarity in hand films, due to the multiplicity of joints and the closeness of the joints to the radiographic film. The bony and cartilage changes of the two disorders are described elsewhere in this book. Older patients with rheumatoid arthritis will frequently exhibit radiological changes of osteoarthritis as well, either as a simultaneous, coincidental disease or secondary to the rheumatoid process. Soft tissue changes may also be of assistance. If one carefully examines a film with the aid of a bright light, one is often able to detect, adjacent to a proximal interphalangeal joint, a 'double line' indicating the presence of fusiform synovial thickening characteristic of rheumatoid arthritis. One should also look carefully for evidence of calcium pyrophosinate deposition disease, a common concomitant of osteoarthritis but not of rheumatoid arthritis.

Synovial fluid analysis remains an important component of differential diagnosis. There will be patients in whom the diagnosis of rheumatoid arthritis or osteoarthritis is so clear that fluid analysis will not be necessary. However, at least one analysis should be conducted on all patients in whom fluid is available. Details concerning synovial fluid examination and its use in differential diagnosis are presented elsewhere in this book. Attention should be paid to appearance, viscosity, cell count, culture, and a search for crystals. If a single joint shows evidence of greater inflammation than is present in other joints, one should suspect the presence of septic arthritis. Many patients with osteoarthritis have underlying or concomitant calcium pyrophosphate deposition disease. Therapy with colchicine, although not as an effective as in gout, is helpful in some patients. Treatment of either rheumatoid arthritis or osteoarthritis is by no means easy. To attempt it in the absence of a clear diagnosis is difficult, and every effort should be made to establish the diagnosis.

Treatment of the older patient with rheumatoid arthritis

In patients with rheumatoid arthritis, as in those with osteoarthritis, it is important to search for the possible presence of concomitant disease. Key questions that should be asked in this regard are outlined in Table 6.



In patients with rheumatoid arthritis
Lack of proportionality between constitution symptoms (weakness, sweats, vasomotor instability, weight loss) and evidence of active joint involvement (swelling, tenderness, morning stiffness) of a number of joints
Presence of anaemia out of proportion to activity or rheumatoid process
Pain in the region of the joint, without proportional loss of motion
Simultaneous presence of hyperthyroidism or myxoedema is very easily missed. thyroid studies should always be obtained in older patients with apparently active rheumatoid arthritis
In patients with osteoarthritis
Pain in or near a joint may be due to a fracture of an adjacent bone, not evident on radiograph. Radiological changes of a fracture may not be present for a month. Bone scan after 1 or 2 weeks or MRI will help clarify the diagnosis
Pain in the back suggestive of osteoarthritis, herniated disk, or vertebral fracture may be due to angina pectoris, carcinoma of the pancreas, multiple myeloma, or metastatic carcinoma

Table 6 Clues to the presence of concomitant disease complicating rheumatoid arthritis or osteoarthritis

The physician should be particularly attentive to non-pharmacological therapy (Podgorski and Edmonds 1985). Use of any drug in older persons is fraught with greater risk than in those at midlife. Non-pharmacological therapy has few complications and, if carefully supervised by the physician, is remarkably effective. By involving a commitment on the part of the patient, non-pharmacological therapy can contribute, substantially, to the sense of control a patient has over his or her disease.

Rheumatoid arthritis is a constitutional illness. Patients should get sufficient rest to avoid fatigue. The physician should spell this out in detail. For example, the patient might rest in bed for 2 h after lunch, undertake light activities at home during the rest of the day, but be given permission to undertake one activity outside the home, such as a shopping trip or going out to dinner with friends. The concept of rest also includes use of appropriate splints for actively involved joints, always accompanied by periods of range of motion exercises.

In the older patient with rheumatoid arthritis who fails to respond to a comprehensive programme of rest, exercises, and salicylates or non-steroidal anti-inflammatory drugs, methotrexate (Wolfe and Cathey 1991) and low doses of prednisone are, in our view, the agents of choice. Methotrexate is particularly effective in patients who exhibit marked synovial thickening. Prior to initiating treatment, a complete blood cell count, liver function tests, and a chest radiograph must be obtained. The

complete blood cell count should be repeated every month; liver function tests every 3 months. An initial dose of 7.5 mg of methotrexate per week may be increased to 15 mg if necessary. Patients may experience bothersome upper gastrointestinal symptoms, which subside with decreased dosage. If gastrointestinal symptoms persist, or the patient does not achieve a satisfactory response, methotrexate may be administered intramuscularly at the same dosage levels. The dosage should be decreased, in any case, once an initial response has been achieved. All patients receiving methotrexate should also be given folic acid, 1 mg daily or 5 mg each week. This will not inhibit the effectiveness of the methotrexate.

In contrast to use of low-dose prednisone in patients with lupus erythematosus or giant cell arteritis, the initial dose in patients with rheumatoid arthritis should be fairly close to what would be an appropriate maintenance level in that particular patient, that is approximately 8 mg/day, given in the morning. The dose should be titrated against symptomatic response in 1 mg aliquots so as to achieve a reasonably satisfactory response at a minimal dosage level. Concomitant calcium and vitamin D should be given and, for patients at high risk for the development of osteoporosis, etidronate or better, alendronate. Methotrexate may be given simultaneously with prednisone.

I have not included sulfasalazine, plaquenil, or penicillamine in this list. The reason is that, in my view, these agents are less effective than methotrexate or low-dose prednisone. In the older patient, delay in initiating full rehabilitation brings the serious hazard of social isolation, discouragement, and the development of decubitus ulcers and other sequelae of inanition. The benefit/hazard ratio of either methotrexate or prednisone in proper doses, accompanied by a full, well-organized, conservative programme, will, in my view, be greater than that associated with the other agents.

Polymyalgia rheumatica and temporal arthritis

Polymyalgia rheumatica and temporal arthritis are among the entities that generate a reasonable flow of consultations from the primary care physician to the rheumatologist. Three aspects account for most of these consultations. The first is the fact that a considerable number of patients, perhaps half of those referred to me in my own consultation practice, with classic clinical characteristics of polymyalgia rheumatica including rapid response to low-dose prednisone, turn out, after follow-up for a year or so, to have either seronegative rheumatoid arthritis or, less frequently, systemic lupus erythematosus. Dixon has pointed out the diagnostic importance of the precise timing of the initial response to prednisone (Dixon 1978). Patients with seronegative rheumatoid arthritis, masquerading as polymyalgia rheumatica, will exhibit a good clinical response to prednisone, in a dose of 10 or 15 mg/day, but this usually occurs on the second or third day after therapy; patients with polymyalgia rheumatica typically manifest this improvement on the first day after starting prednisone. The nature of this response, together with the discordance between the constitutional symptoms, especially muscle stiffness and weakness and objective evidence of inflammation of the peripheral joints, in patients with polymyalgia rheumatica, provide good grounds for initial differentiation. However, even in these patients, by the end of a year or so of prednisone therapy, a number will begin to exhibit sufficient involvement of the peripheral joints, including proximal interphalangeal and/or metacarpophalangeal joints, accompanied by early radiological changes, to justify the diagnosis of seronegative rheumatoid arthritis. Healey and Sheets have emphasized the close relationship of these two entities (Healey and Sheets 1988). A problem is created by the fact that, by the time a year has gone by, one is fairly well committed to long-term prednisone therapy in a patient in whom one might not have selected this treatment if the correct diagnosis had been made at the outset. An alternative approach, that of treating patients with polymyalgia rheumatica with a non-steroidal antirheumatic drug from the onset, would avoid this problem. However, not all patients will respond and there is the added hazard of the toxic effect of non-steroidal anti-inflammatory drugs in older patients.

A second problem, occurring in patients in whom the diagnosis of polymyalgia rheumatica was correctly made, is inability to withdraw steroid therapy after 1 or even 2 years without a significant flair up of the symptoms of polymyalgia. This occurs in approximately one-third to one-half of the patients who are referred to me with this condition. For these patients, a very small dose of prednisone, perhaps 2 or 3 mg/day, often proves to be effective and a long-term commitment at this dosage level seems unavoidable. It has been suggested that this problem would be less apt to occur if the initial dose were established at the lowest possible effective level, perhaps 8 to 9 mg/day.

A third problem is presented by the patient who has temporal arteritis suppressed by the initial high doses of prednisone, in whom the symptoms recur as the prednisone dose is decreased to a level of, perhaps, 30 mg/day. In these patients, addition of methotrexate, 7.5 to 15 mg/week, may permit further reduction of prednisone dosage. There are also patients in whom it is possible to lower the dose of prednisone to 8 or 9 mg/day, while further decreases are accompanied by flair up of the manifestations of temporal arteritis and continued therapy at this dosage level is unavoidable.

Remitting seronegative symmetrical synovitis with pitting oedema syndrome (RS $_3$ PE)

McCarty *et al*. have described a syndrome which they have termed remitting seronegative symmetrical synovitis with pitting (o)edema, the first letters of which form the abbreviated title RS$_3$ PE syndrome (McCarty *et al*. 1985). Occurring in men twice as frequently as in women, with an average age of 75 years, the syndrome is characterized by the sudden onset, sometimes within hours, of a symmetrical polysynovitis manifested in extreme cases by tenderness and swelling of the peripheral joints, especially wrists, and flexordigitorium tendons over the volar surface of the wrist and forearm, accompanied by pitting oedema and, sometimes, redness and warmth of the dorsum of the hands and feet and, in some instances, pretibial areas. The condition is accompanied by marked morning stiffness, lasting up to 6 h. Carpal tunnel syndrome may be present. Constitutional symptoms, such as fever and fatigue are infrequent. Radiographs reveal no evidence of joint destruction, either initially or over a follow-up of up to 5 years. While synovitis may be present, the involvement focuses primarily in tendon sheaths. Rheumatoid factor, by latex fixation test, is absent. Low titres of antinuclear antibodies may be present. Erythrocyte sedimentation rate is often elevated, sometimes as high as 93 mm/1 h. HLA typing reveals a disproportionate number of patients with HLA-B7.

By 1994, 31 cases of this syndrome had been described (Russell *et al*. 1990; Olivo *et al*. 1994). Many patients become quite incapacitated by the symptoms, sometimes requiring hospitalization. Various treatment regimens have been utilized, including non-steroidal anti-inflammatory drugs, hydroxychloroquine, gold, and prednisone. Prednisone, in doses comparable to those employs in polymyalgia rheumatica, yielded prompt resolution of the symptoms in several cases, and low-dose prednisone therapy appears to be the optimal form of treatment. Regardless of therapy, all patients entered a substantial remission within a year, with complete subsidence of the inflammatory manifestations but, in some, a small reduction in range of motion in the hands and wrists. While many rheumatologists regard this syndrome as a specific entity, Healey has postulated that this entity, seronegative rheumatoid arthritis, and polymyalgia rheumatica may represent different phases or manifestations of a common disorder (Healey 1990).

Systemic lupus erythematosus

Systemic lupus erythematosus occurs primarily in young adults, with progressive declines occurring throughout late adult life and a frequency close to zero among the very old. However, the disease does occur in the latter age group and provides a particular diagnostic challenge for two reasons. First, the rarity of the entity, in contrast to the increasing frequency of other conditions with somewhat similar manifestations (such as rheumatoid arthritis, Sjögren's syndrome, and carcinomatosis) places it low in the index of suspicion of physicians caring for persons in this age group. Second, the clinical manifestations of lupus, among persons at late midlife, appear to be significantly different from the 'textbook description' of this disease, which has been based primarily on studies of patients at a younger age.

While there is general agreement that these differences do exist, the published studies on this point do not yet achieve total agreement concerning the details. Review of data published by Dimant *et al*. (216 younger patients/16 older), Maddison (93 younger/19 older), Cervera (1910 younger/90 older), and a meta-analysis of nine studies conducted by Ward and Polisson illustrate this point (Dimant *et al*. 1979; Maddison 1987; Cervera 1993; Ward and Polisson 1989). All authors agree that patients with onset of the disease at late midlife have a increased frequency of pulmonary manifestations. The three articles which comment on Raynaud's phenomena (Dimant *et al*. 1979; Maddison 1987; Cervera 1993) state that the frequency is reduced among persons at late midlife but Ward and Pollison found no difference. The same three articles contained comments on the frequency of Sjögren's-like manifestations, and found them to be increased among those with onset at late midlife. Madison and Dimant found the frequency of photosensitivity to be increased, but Cervera found it to be decreased, and Ward and Polisson found no difference. The two studies which commented on neurological manifestations (Dimant *et al*. 1979; Ward and Polisson 1989) both found them to be decreased in frequency in late midlife.

With regard to serological manifestations of later-onset systemic lupus erythematosus, Maddisons' findings will serve as an example (Table 7). However, in Ward and Polisson's meta analysis, anti-Ro antibodies were encountered in only 16 per cent of later onset patients; anti-La in 6 per cent. Anti-nRNP and anti-Sm were present in 5 per cent each. There is general agreement that further studies are needed to document these clinical and serological differences, and these should be based on defined populations rather than referred cases and include more data on patients age 70 years and older. To date, most authors agree that the differences between early and late-onset systemic lupus erythematosus are more likely to represent progressive age-related changes rather than two distinct syndromes.

Serological feature	Patients (%)	
	Late onset (*n*=19)	Younger onset (*n*=93)
Antinuclear antibody	95	95
Anti-DNA	53	65
Anti-Ro (SSA)	84	29 $p<0.0001$
Anti-La (SSB)	63	9 $p<0.0001$
Anti-nRNP	0	37 $p<0.001$
Anti-Sm	0	6
Rheumatoid factor	63	40
Hypocomplementaemia	53	57

Reproduced (with slight modification) from Maddison (1987), with permission.

Table 7 Serological features of systemic lupus erythematosus with onset in youth and midlife as compared with elder-onset disease (61–89 years)

From the clinical point of view, the occasional presence of systemic lupus erythematosus in patients age 70 years and older assumes importance for two reasons. Firstly, the condition does occur and deserves to be correctly diagnosed and appropriately treated. Secondly, for every old patient one sees who actually has systemic lupus erythematosus, there will be many who are referred by primary care physicians for the possibility of this disorder. The explanation lies in the frequent occurrence of rash, alopecia, chronic renal disease, polyarthritis, epilepsy, and neuropsychiatric disorders suggestive of a clinical diagnosis of systemic lupus erythematosus but due, instead, to a combination of four or five independent chronic diseases. The problem is confounded by the fact that a considerable number of older persons, without any known immunological disorder, will exhibit a positive antinuclear antibody test (Hallgren *et al*. 1973). Obviously, to present to a person with an over-supply of chronic diseases of this sort the added burden of corticosteroid or immunosuppressive therapy, in the mistaken impression that he or she is suffering from systemic lupus erythematosus, is a serious abrogation of the Hippocratic oath to do no harm.

Little objective information has as yet been obtained concerning treatment of persons with older-onset systemic lupus erythematosus. Initial treatment with high doses of prednisone is clearly indicated, with tapering in accordance with clinical manifestations. Recent information on the effectiveness of the bisphosphonates in inhibiting the osteopenic effect of long-term prednisone therapy provides an important new arena of preventive care in these patients. Although one prefers not to use cytotoxic drugs in this age group, the adverse prognosis of many patients with older-onset systemic lupus erythematosus must be kept in mind (Reveille *et al*. 1990). In this study, there was no relationship between the dosage of prednisone at last hospitalization or the 'frequency of recent immunosuppressive therapy' and death due to infection.

Information for patients and the families of patients suffering from lupus erythematosus is provided in *Lupus news*, published by the Lupus Foundation of America, Inc., 4 Research Place, Suite 180, Rockville, MD 20850–3226, telephone (301) 670–9292.

Drug-induced lupus

In contrast to systemic lupus erythematosus, this syndrome occurs with increased frequency in patients in the older age group due, at least in part, to the large number of drugs many of these patients consume (Solenger 1988). Procainamide or hydralazine therapy result in the development of antinuclear antibodies in 50 to 70 per cent of cases, but only about 30 per cent of Caucasians who exhibit these antibodies develop symptoms. The syndrome is very rare among blacks. The time of appearance of the symptoms after commencing procainamide therapy depends, at least to some degree, on the genetically controlled activity of the *N*-acetotransferase system in the liver (Woodley *et al*. 1978). Patients with reduced activity of this enzyme (slow acetylators) have a higher frequency of drug induced lupus than rapid acetylators, and develop the clinical manifestations after a shorter period of drug therapy (an average of 12 months as compared to 48 months). Clinical characteristics of drug-induced lupus include symmetrical polyarticular arthritis, often accompanied by striking morning stiffness, pleuritis, pericarditis, and, in some, parenchymal involvement of the lung. Thus, the manifestations resemble those of older-onset systemic lupus erythematosus. Rash, alopecia, and renal involvement are rarely seen. Although patients may exhibit haemolytic anaemia, thrombocytopenia and leucopenia are rarely encountered. Most patients with procainamide induced lupus are males; the majority of those with hydralazine induced lupus are females. Essentially all patients with this syndrome will exhibit positive antinuclear antibody reactions with titres of 1:640 or higher. Antibodies to n-DNA, RNP, Sm, Ro, and La are rarely seen. If the clinical manifestations do not subside within 2 months of discontinuing the offending drug, one must consider the possibility that the drug has proved to be an inciting agent for the development of systemic lupus erythematosus.

Sjögren's syndrome

Sicca symptoms occur in approximately 40 per cent of patients age 65 years and older, with increased frequency as the patient moves towards the older years (Strickland *et al*. 1987). In most cases this is due to age-related atrophy of the salivary glands or as a result of medications that have anticholinergic properties. Many of these patients will have a positive Schirmer's test. Sjögren's syndrome differs from sicca syndrome in the aged in that the dryness of eyes and mouth is accompanied by any of a wide variety of systemic manifestations, such as hoarseness, difficulty in swallowing, pneumonitis, interstitial lung disease, achlorhydria, nephritis, peripheral neuropathy, and a variety of psychiatric disorders—see Chapter 5.10. In view of the propensity of older people to suffer from several chronic diseases, it is easy for a physician to overlook the possibility that these manifestations might be due to a single syndrome. Thus, the disease undoubtedly occurs more frequently than is recognized. The presence of the disorder can be substantiated by a lip biopsy and demonstration of focal lymphocytic infiltration in the minor salivary glands. A positive test for anti-Ro(SSA) antibodies, present in 57 per cent of cases, provides added confirmation.

Although patients with this disorder have an increased likelihood for the development of lymphoma, and also renal involvement, most patients show good survival rates but face highly troublesome symptoms. In the United States, the Sjögren's Syndrome Foundation provides extremely effective support services, including a monthly publication, *The moisture seekers*, circulation of information materials through the internet, national and regional symposia, and a support network for patients and families. This is a disorder in which a positive attitude and attention to detail on the part of the patient can make a big difference to the patient's overall effectiveness, happiness, and independence. The foundation can be reached at 333 North Broadway, Hericho, New York, telephone (516) 933–6365, fax (516) 933–6368.

Systemic sclerosis

Systemic sclerosis (scleroderma) is a relatively uncommon rheumatic disease with onset occurring primarily at midlife. Of 63 patients with either diffuse or limited disease, described by Maddison *et al*., 55 had their onset between ages 20 and 60 years; only one developed the disease after age 60 (Maddison *et al*. 1993). However, since the patients usually live for a number of years, with the mean duration of disease for the CREST syndrome in one series (Aeschlimann *et al*. 1989) of 26 years, one should expect to see a number of patients with this disorder in a rheumatological practice with older patients.

When systemic sclerosis occurs in old persons, the close resemblance of its manifestations, especially changes in facial expression, to those of normal ageing is such that the diagnosis is easily overlooked (Holti and Schuster 1981). Obtaining an antinuclear antibody test in persons suspected of having this disorder will almost always lead to the diagnosis, since this test is positive in essentially all patients. The antitopoisomerase antibody (anti-Scl-70), present in between 20 and 40 per cent of the patients, will provide added confirmation, if present. The diagnosis is confirmed by skin biopsy. For further details concerning scleroderma and its treatment, the reader is referred to Chapter 5.8.1 of this book.

Soft tissue rheumatism

Articular structures themselves represent only one component of the musculoskeletal system. In older patients, disorders of the periarticular tissues (capsule, tendinous insertions, and bursas) and other tissues remote from the joints, such as muscle and fascial sheaths, probably account for a larger share of musculoskeletal symptoms than the classic articular diseases. Age-related decreases in muscle function and atrophy of tendons and muscle contribute, substantially, to the frequency of these disorders. The range of disorders includes acute and chronic tendinitis, bursitis (at numerous sites), enthesitis (inflammation of the tendinous insertions into the bony periosteum), muscle tears and strains, fasciitis, and involvement of paraspinal tissues secondary to chronic strain from poor posture and/or spondylolisthesis. Even intervertebral disc disease might be included in this list since the intervertebral disc, in life, is not a solid or semisolid structure but is a soft

sack, filled with a liquid gel, within which floats a firmer structure or nucleus.

While these periarticular disorders are frequently accompanied by decreased function of the adjacent joint, and are referred to by patient as 'arthritis,' they can be distinguished from involvement of the joint itself by careful history and physical examination, and, in most instances, will respond well to appropriate treatment.

Characteristically, the pain from periarticular involvement does not occur at rest but is exacerbated by motion of the adjacent joint, especially if it is resisted. Pain due to involvement of periarticular tissues of the shoulder is often exacerbated when the patient lies on the affected shoulder, thus interfering with sleep. In many instances, the pain due to these soft tissue lesions is perceived not in the specific area of involvement but in a more diffuse area, distal to the joint. This so-called referred pain is due to shared innervation of many connective tissues with specific areas of the skin.

The physical examination is important. For example for involvement of the shoulder region, one first seeks to identify, by careful palpation, diffuse swelling and tenderness of the synovial tissues. The presence of articular involvement is further confirmed by demonstration of decreased range of motion. This is best elicited by placing one hand firmly on the clavicle, trapezius, and scapula and, with the other hand, assessing the extent of abduction and internal and external rotation of the shoulder joint. In a normal individual, each of these motions should be approximately 80 to 90 per cent.

In patients who exhibit a full range of motion and the absence of diffuse synovial thickening one can be virtually certain that the symptoms arise from the periarticular tissues. Careful palpation of each area, seeking localized tenderness, will usually permit identification of the specific problem. Confirmation of bicipital tendinitis can be confirmed by eliciting pain over the tendon while the patient attempts to externally rotate the arm against resistance.

Treatment consists of rest, local application of heat or ice (whichever proves more effective), careful use of non-steroidal anti-inflammatory drugs and, if necessary, local injection of the bursa or tendon sheaths. It is our practice to infuse 20 mg of methylprednisolone acetate (Depo-Medrol) plus 2 ml of 1 per cent lidocaine, through a 21 gauge needle. Guidance for techniques which will help ensure that the material is injection within the bursa or tendon sheaths, rather than within the tendon itself, periosteum, or non-specifically in surrounding connective tissues, is provided in the references (Botstein 1990; Dorman and Ravin 1991; Neustadt 1991; Doherty *et al*. 1992).

In the management of patients with involvement of the periarticular structures surrounding the shoulder, attention should also be given to instructing the patient in pendulum exercises with, perhaps, a 1 kg weight held in the hand, for 5 to 10 min each day, in order to avoid the superimposed problem of a frozen shoulder. Guidance in the management of periarticular tissues in the elbow, wrist, hip, knee, neck, and back and additional information relevant to the shoulder is provided in the references cited above.

The topic of fibromyalgia, also covered in Chapter 5.14, is of special interest and importance. Although many rheumatologists remain unconvinced of the importance or even existence of this syndrome, others encounter patients meeting the criteria for this disease in a growing proportion of their practice. Many aspects of this syndrome overlap with another frequent and puzzling disorder, chronic fatigue syndrome (Buchwald and Garrity 1994). A new and promising approach to the treatment of latter disorder (Boh-Holaigah *et al*. 1995) may provide a new approach to the issue of fibromyalgia as well. Information for patients with fibromyalgia is provided in the *Fibromyalgia network newsletter*, obtainable form the Fibromyalgia Network, PO Box 31750, Tucson, AZ 85751, telephone (602) 290–5508.

Summary

Musculoskeletal symptoms provide the most common cause of morbidity among patients age 65 years and older (National Center for Health Statistics 1986). This fact, plus the increasing size of this population, renders an understanding of the biology of ageing and the unique characteristics of care for older persons of increasing relevance to the practice of rheumatology. This chapter has attempted to outline the major elements of the 'geriatric perspective', and discussed the application of these principles to several rheumatic diseases of special importance among older patients.

Chapter References

Aeschlimann, A., *et al*. (1989). Anti-Scl-70 antibodies detected by immunoblotting in progressive systemic sclerosis: specificity and clinical correlations. *Annals of the Rheumatic Diseases*, **48**, 992–7.

Allebeck, P. (1982). Increased mortality in rheumatoid arthritis. *Scandinavian Journal of Rheumatology*, **11**, 81–6.

Beers, M.H. and Ouslander, J.G. (1989). Risk factors in geriatric drug prescribing. A practical guide to avoiding problems. *Drugs*, **37**, 105–12.

Beeson, P.B. (1994). Age and sex associations of 40 autoimmune diseases. *American Journal of Medicine*, **96**, 457–62.

Bennett, G.A., Waine, H., and Bauer, W. (1942). *Changes in the knee joint at various ages*. The Commonwealth Fund, New York.

Bergstrom, G., Bjelle, A., Sorensen, L.B., Sundh, V., and Svanborg, A. (1986 *a*). Prevalence of rheumatoid arthritis, osteoarthritis, chondrocalcinosis and gouty arthritis at age 79. *Journal of Rheumatology*, **13**, 527–34.

Bergstrom, G., Bjelle A., Sundh, V., and Svanborg, A. (1986 *b*). Joint disorders at ages 70, 75 and 79 years—a cross sectional comparison. *British Journal of Rheumatology*, **25**, 333–41.

Bienia, R., Ratcliff, S., Barbour, G.L., and Kummer, M. (1982). Malnutrition in the hospitalized geriatric patient. *Journal of the American Geriatric Society*, **30**, 433–6.

Blazer, D.G. (1989). Depression in the elderly. *New England Journal of Medicine*, **320**, 164–6.

Botstein, G.R. (1990). Soft tissue rheumatism of the upper extremities: Diagnosis and management. *Geriatrics*, **45**, 30–43

Boh-Holaigah, I., Rowe, P.C., Kan, J., and Calkins, H. (1995). The relationship between neurally medicated hypotension and the chronic fatigue syndrome. *Journal of the American Medical Association*, **24**, 961–7.

Brown, C.S., Rakel, R.E., Wells, B.G., Downs, J.M., and Akiskal, H.S. (1991). A practical update on anxiety disorders and their pharmacological treatment. *Archives of Internal Medicine*, **151**, 873–84.

Buchner, D.M. and Wagner, E.H. (1992). Preventing frail health. *Clinical Geriatric Medicine*, **8**, 1–17.

Buchwald, D. and Garrity, D. (1994). Comparison of patients with chronic fatigue syndrome, fibrylmialgia and multiple chemical sensitivities. *Achieves of Internal Medicine*, **154**, 2049–53.

Burnam, M.A., Wells, K.B., Leake, B., and Landsvert, J. (1988). Development of a brief screening instrument for detecting depressive disorders. *Medical Care*, **26**, 775–89.

Cauna, N. (1964). Aging on receptor organs of human dermis. In *Advances in biology of skin* (5th edn) (ed. W. Montagna), pp. 73–85. Pergamon Press, Oxford.

Cerhan, J.R., Wallace, R.B., El-Khoury, G.Y., Moore, T., and Long, C.R. (1995). Decreased survival with increased prevalence of full-body radiologically defined osteoarthritis in women. *American Journal of Epidemiology*, **141**, 225–34.

Cervera, R. *et al*. (1993). Systemic lupus erythematosus: Clinical and immunologic patterns of disease expression in a cohort of 1000 patients. *Medicine*, **72**, 113–24.

Deal, C.L. *et al*. (1985). The clinical features of elderly-onset rheumatoid arthritis. A comparison with younger-onset disease of similar duration. *Arthritis and Rheumatism*, **28**, 987–94.

Dimant, J. *et al*. (1979). Systemic lupus erythematosus in the older age group: computer analysis. *Journal of the American Geriatrics Society*, **27**, 58–61.

Dixon, A. St.J. (1978). The diagnosis of polymyalgia rheumatica. In *Reports on rheumatic diseases. Collected reports 1959–1977* (ed. Hawkins and H.L.F. Currey), pp. 61–3. Geriatrics and Rheumatism Council, London.

Dixon, A. St.J. (1992). Health of the nation and osteoporosis. *Annals of the Rheumatic Diseases*, **51**, 914–18.

Doherty, M., Hazleman, B.L., Hutton, C.W., Maddison, P.J., and Perry, J.D. (1992). *Rheumatology examination and injections techniques*. WB Saunders, London.

Dorman, T.A. and Ravin, T.H. (1991). *Diagnosis and injection techniques in orthopedic medicine*. Williams and Wilkins, Baltimore.

Evans, J.G. and Williams, T.F. (1992). Introduction. In *Oxford textbook of geriatric medicine* (ed. J.G. Evans and T.R. Williams), pp. XV-XVII. Oxford University Press, Oxford.

Feely, J. and Coakley, D. (1990). Altered pharmacodynamics in the elderly. *Clinical Geriatric Medicine*, **6**, 269–83.

Fiaterone, M.A. and Evans, W.J. (1993). Etiology and reversibility of muscle dysfunction in the aged. *Journal of Gastroenterology*, **48**, 77–83.

Finch, C.E., Foster, J.R., and Mirsky, A.E. (1969). Ageing and the regulation of cell activities during exposure to cold. *Journal of General Physiology*, **54**, 690–712.

Fisher, N.M., Pendergast, D.R., and Calkins, E. (1991*a*). Muscle Rehabilitation in impaired nursing home residents. *Archives of Physical Medicine and Rehabilitation*, **72**, 181–5.

Fisher, N.M., Pendergast, G.R., Gresham, G.E., and Calkins, E. (1991 *b*). Muscle rehabilitation: Its effect on muscular and functional performance of patients with knee osteoarthritis. *Archives of Physical Medicine and Rehabilitation*, **72**, 367–74.

Folstein, M.F., Folstein, S., and McHugh, P.R. (1975). Mini-mental state. A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, **12**, 189–98.

Fries, J.F., Miller, S.R., Spitz, P.W., Williams, C.A., Hubert, H.B., and Bloch, D.A. (1989). Toward an epidemiology of gastropathy associated with nonsteroidal anti-inflammatory drug use. *Gastroenterology*, **96**, 647–55.

Furst, D.E. (1988). Clinically important interactions of non-steroidal anti-inflammatory drugs with other medications. *Journal of Rheumatology*, **15** (suppl. 17), 58–62.

German, P.S. and Fried, L.P. (1989). Prevention and the elderly. public health issues and strategies. *Annual Review of Public Health*, **10**, 319–32.

Girdwood, R.H. (1974). Death after taking medicaments. *British Medical Journal*, **1**, 501–4.

Griffin, M.R., Piper, J.M., Daugherty, J.R., Snowden, M., and Ray, W.A. (1991). Nonsteroidal anti-inflammatory drug use and increased risk for peptic ulcer disease in elderly person. *Annals of Internal Medicine*, **114**, 257–63.

Gruenberg, E.M. (1977). The failures of success. *Milbank Memorial Fund Quarterly—Health and Society*, **55**, 3–24.

Gurwitz, J.H. and Avorn, J. (1991). The ambiguous relation between aging and adverse drug reactions. *Annals of Internal Medicine*, **114**, 956–66.

Hallgren, H.M., Buckley, C.E. III, Gilbertsen, V.A., and Yunis, E.J. (1973). Lymphocyte phytohemagglutinin responsiveness, immunoglobulin and autoantibodies in aging humans. *Journal of Immunology*, **111**, 1101–7.

Halstead, L.S. (1976). Team care in chronic illness: A critical review of the literature of the past 25 years. *Archives of Physical Medicine and Rehabilitation*, **57**, 501–11.

Harris, S.T., Watts, N.B., Jackson, R.D., *et al*. (1993). Four-year study of intermittent cyclic etidronate treatment of post-menopausal osteoporosis: 3 years of blinded therapy followed by one year of open therapy. *American Journal of Medicine*, **95**, 557–67.

Hayes, C.S. and Stinson, I.N. (1976). Erythrocyte sedimentation rate and age. *Archives of Ophthalmology*, **94**, 939–40.

Healey, L.A. (1986). Rheumatoid arthritis in the elderly. *Clinics in Rheumatic Diseases*, **12**, 173–9.

Healey, L.A. (1990). RS_3 PE Syndrome. Letter to the Editor. *Journal of Rheumatology*, **17**, 414.

Healey, L.A. and Sheets, P.K. (1988). The relation of polymyalgia rheumatica to rheumatoid arthritis. *Journal of Rheumatology*, **15**, 750–2.

Heath, G.W., Hagberg, J.M., Ehsani, A.A., and Holloszy, J.O. (1981). A physiological comparison of young and older endurance athletes. *Journal of Applied Physiology*, **51**, 634–40.

Holti, G. and Schuster, S. (1981). Scleroderma in the elderly (letter). *British Medical Journal*, **282**, 1400.

Irvine, R.E., Grove, J., Taseland, P.A., and Traunce, J.R. (1974). The effect of age on the hydroxylation of amylobarbitone sodium in man. *British Journal of Clinical Pharmacology*, **1**, 41–3.

Janssen, M., Dijkman, B.A.C., Lamers, C.B.H.W., Zwonderman, A.H., and Vandenbroucke, J.P. (1994). A gastroscopic survey of the predictive value of risk factors for non-steroidal anti-inflammatory drug associated ulcer disease in rheumatoid arthritis patients. *British Journal of Rheumatology*, **33**, 449–54.

Katz, S., Downs, T.D., Cash, H.R., and Grotz, R.C. (1970). Progress in development of the index of ADL. *Gerontologist*, **10**, 20–30.

Katz, P., Dube, D., Calkins, E. (1986). Aging and disease. In *The practice of geriatrics* (ed. E. Calkins, P.J. Davis, and A.B. Ford), pp. 1–6. WB Saunders, Philadelphia.

Keefe, F.J. *et al*. (1987). Pain coping strategies in osteoarthritis patients. *Journal of Consulting and Clinical Psychology*, **55**, 208–12.

Larson, E.B., Reifler, B.V., Sumi, S.M., Canfield, C.G., and Chinn, N.M. (1985). Diagnostic evaluation of 200 elderly outpatients with suspected dementia. *Journal of Gerontology*, **40**, 536–43.

Lavsky-Shulan, M., Wallace, R.B., Kohout, F.J., Lemke, J.H., Morris, M.C., and Smith, I.M. (1985). Prevalence and functional correlates of low back pain in the elderly: the Iowa 65+ Rural Health Study. *Journal of the American Geriatric Society*, **33**, 23–8.

Lawrence, R.C. *et al*. (1989). Estimates of the prevalence of selected arthritic and musculoskeletal diseases in the United States. *Journal of Rheumatology*, **16**, 427–41.

Lawton, M.P. and Broady, E.M. (1969). Assessment of older people. Self monitoring and instrumental activities of daily living. *Gerontologist*, **9**, 179–86.

Lichtenberg, P.A., Swensen, C.H., and Skehan, M.W. (1986). Further investigation of the role of personality, lifestyle and arthritic severity in predicting pain. *Journal of Psychosomatic Research*, **30**, 327–37.

Liebermann, U.A., Weiss, S.R., Broll, J., Minne, H.W., *et al*. (1995). Effect of oral alendronate on bone mineral density and the incidence of fractures in post-menopausal osteoporosis. *New England Journal of Medicine*, **333**, 1437–43.

Lindeman, R.D., Tobin J., and Shock, N.W. (1985). Longitudinal studies on the rate of decline in renal function with age. *Journal of American Geriatric Society*, **33**, 278–85.

Lindesay, J., Briggs, K., and Murphy, E. (1989). The Guy's/Age Concern survey: Prevalence rates of cognitive impairment, depression and anxiety in an urban elderly community. *British Journal of Psychiatry*, **155**, 317–29.

Linos, A., Worthington, J.W., O'Fallon, W.N., and Kurland, L.T. (1980). The epidemiology of rheumatoid arthritis in Rochester, MN. A study of incidence, prevalence, and morbidity. *American Journal of Epidemiology*, **111**, 87–98.

Lipsitz, D.A. (1992). Nutrition and ageing. In *Oxford textbook of geriatric medicine* (ed. J.G. Evans and T.F. Williams), pp. 119–27. Oxford University Press, Oxford.

Llewellyn, J.G. and Pritchard, M.H. (1988). Influence of age and disease state in nonsteroidal antiinflammatory drug associated gastric bleeding. *Journal of Rheumatology*, **15**, 691–4.

Lloyd, T. *et al*. (1993). Calcium supplementation and bone mineral density in adolescent girls. *Journal of the American Medical Association*, **270**, 841–4.

McCarty, D.J., O'Duffy, J.D., Pearson, L., and Hunter, J.B. (1985). Remitting seronegative symmetrical synovitis with pitting edema: R.S. $_3$P.E. syndrome. *Journal of the American Medical Association*, **254**, 2763–7.

MacFarlane, D.G. and Dieppe, P.A. (1983). Pseudo-rheumatoid deformity in elderly osteoarthritic hands. *Journal of Rheumatology*, **10**, 489–90.

McHorney, C.A., Ware, J.E., Lu, J.F., and Sherbourne, C.D. (1994). The MOS 36-item short-form Health Survey (SF-36) III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. *Medical Care*, **32**, 40–66.

Maddison, P.J. (1987). Systemic lupus erythematosus in the elderly. *Journal of Rheumatology*, **14** (suppl. 13), 182–7.

Maddison, P.J. *et al*. (1993). Connective tissue disease and autoantibodies in the kindred of 63 patients with systemic sclerosis. *Medicine*, **72**, 103–72.

Mikkelsen, W.M., Dodge, H.J., Duff, I.F., and Kato H. (1967). Estimates of the prevalence of rheumatoid arthritis in the population of Tecumseh, MI. 1959–60. *Journal of Chronic Diseases*, **20**.

351–69.

Montamat, S.C., Cusack, B.J., and Vestal, R.E. (1989). Management of drug therapy in the elderly. *New England Journal of Medicine*, **321**, 303–9.

Morley, J.E., Silver, A.J., Fiatarone, M., and Mooradian, A.D. (1986). Geriatric grand rounds: nutrition and the elderly. University of California, Los Angeles. (Clinical Conference). *Journal of the American Geriatric Society*, **34**, 823–32.

Mulder, H. and Struys, A. (1994). Intermittent cyclical etidronate in the prevention of corticosteroid-induced bone loss. *British Journal Rheumatology*, **33**, 348–50.

National Center for Health Statistics (1986). *Current estimates from the national health survey*. United States. Vital and Health Statistics, Series 10: 164 October 1987.

Neustadt, D.H. (1991). Local corticosteroid injection therapy in soft tissue rheumatic conditions of the hand and wrist. Editorial. *Arthritis and Rheumatism*, **34**, 923–6.

Odell, W.D. and Heath, H.III (1993). Osteoporosis: pathophysiology, prevention, diagnosis and treatment. *Disease A Month*, **39**, 789–867.

Olivo, M.D., D'Amore, M., Lacava, R., Rossi, M.G., Gareri, P., Fiorentini, C., and Mattace, R. (1994). Benign edematous polysynovitis in the elderly (RS $_3$PE syndrome). *Clinical and Experimental Rheumatology*, **12**, 669–73.

Pahor, M., Guralink, J.M., Salive, M.E., Chrischilles, E.A., Manto. A., and Wallace, R.B. (1994 *a*). Disability and severe gastrointestinal hemorrhage. A prospective study of community dwelling older persons. *Journal of the American Geriatric Society*, **42**, 816–25.

Pahor, M., Guralnik, J.M., Salive, M.E., Chrischilles, E.A., Brown, S.L., and Wallace, R.B. (1994 *b*). Physical activity and risk of severe gastrointestinal hemorrhage in older persons. *Journal of the American Medical Association*, **272**, 595–9.

Pendergast, D.R., Fisher, N.M., and Calkins, E. (1993). Cardiovascular, neuromuscular and metabolic alterations with age leading to frailty. *Journal of Gerontology*, **48**, 61–7.

Pincus, T. and Calahan, L.F. (1986). Taking mortality and rheumatoid arthritis seriously—predictive markers, socio-economic status, and co-morbidity. *Journal of Rheumatology*, **13**, 841–5.

Podgorski, M. and Edmonds, J. (1985). Non-pharmacologic treatment of patients with rheumatoid arthritis. *Medical Journal of Australia*, **25**, 511–6.

Pyka, G., Lindenberger, E., Charette, S., and Marcus, R. (1994). Muscle strength and fiber adaptations to a year-long resistance training program in elderly men and woman. *Journal of Gerontology*, **49**, M22–M25.

Rabins, P., Merchant, A., and Nestradt, G. (1984). Criteria for diagnosing reversible dementia caused by depression. validation by 2 year follow-up. *British Journal of Psychiatry*, **144**, 488–92.

Reuben, D.B. and Siu, A.L. (1990). An objective measurement of physical function of elderly outpatients: The physical performance test. *Journal of the American Geriatric Society*, **38**, 1105–12.

Reveille, J.D., Bartolucci, A., and Alarcon, G.S. (1990). Progress in systemic lupus erythematosus. Negative impact of increasing age at onset, black race, and thrombocytopenia as well as causes of death. *Arthritis and Rheumatism*, **33**, 37–48.

Rossman, I. (1977). Anatomic and body composition changes with aging. In *Handbook of the biology of aging* (ed. C.E. Finch and L. Hayflick), pp. 189–221. Van Nostrand, Reinhold Company, New York.

Rowe, J.W., Andres, R., Tobin, J.D., Norris, A.H., and Shock, N.W. (1976). The effect of age on creatinine clearance in man: A cross-sectional and longitudinal study. *Journal of Gerontology*, **31**, 155–63.

Rubenstein, L.Z., Josephson, K.R., Wieland, G.D., English, P.A., Sayre, J.A., and Kane, R.L. (1994). Effectiveness of a geriatric evaluation unit: A randomized clinical trial. *New England Journal of Medicine*, **311**, 1664–70.

Russell, E.B., Hunter, J.B., Pearson, L., and McCarty, D.J. (1990). Remitting, seronegative, symmetrical synovitis with pitting edema—13 additional cases. *Journal of Rheumatology*, **17**, 633–9.

Sambrook, P.N. (1995). Editorial: The treatment of post-menopausal osteoporosis. *New England Journal of Medicine*, **333**, 1495–6.

Solenger, A.M. (1988). Drug-related lupus. *Clinics in Rheumatic Diseases*, **1**, 187–202.

Strehler, B.L. (1975). Implications of aging research for society. In *Biology of aging and development* (ed. G.T. Thorbecke), pp. 3–9. Plenum Press, New York.

Strickland, R.W., Tesar, J.T., Berne, H.H., Hobbs, B.R., Lewis, D.M., and Welton, R.C. (1987). The frequency of the Sicca syndrome in the elderly population. *Journal of Rheumatology*, **14**, 766–71.

Ward, M.M. and Polisson, R.P. (1989). A meta analysis of the clinical manifestations of older-onset systemic lupus erythematosus. *Arthritis and Rheumatism*, **32**, 1226–1332.

Wasson, J.H., Gale, V., McDoanld, R., and Liang, M.H. (1990). The prescription of assistive devices for the elderly: Practical considerations. *Journal of General Internal Medicine*, **5**, 46–54.

Watts, N.B. *et al*. (1990). Intermittent cyclical etidronate of postmenopausal osteoporosis. *New England Journal of Medicine*, **323**, 73–9. See also Comments in *New England Journal of Medicine*, **323**, 124–5 and Comments in *New England Journal of Medicine*, **323**, 1633–5.

Winfield, J.B. (1989). Arthritis patient education. Efficacy, implementation and financing. The ACR/AHPA/AF/NAAB Task Force on Arthritis Patient Education. *Arthritis and Rheumatism*, **32**, 1330–3.

Wolfe, F. and Cathey, M.A. (1991). The effect of age on methotrexate efficacy and toxicity. *Journal of Rheumatology*, **18**, 973–7.

Wood, P.H.N. (1992). Epidemiology of joint and connective tissue disorders. In *Oxford textbook of geriatric medicine* (ed. J.G. Evans and T.R. Williams), pp. 351–6. Oxford University Press, Oxford.

Woodley, R.L., *et al*. (1978). Effects of acetylator phenotype on the rate at which procainamide induces antinuclear antibodies and lupus syndrome. *New England Journal of Medicine*, **298**, 1157–60.

1.1.4 Principles of examination

Elisabeth Paice

General principles

The examination of the musculoskeletal system can be considered under two headings: (i) the systematic screening of virtually all the accessible joints of the body to identify abnormalities and establish their distribution; (ii) the more detailed analysis of an abnormal joint or group of joints. The screening is intended to be brief but comprehensive and efficient. It can be accomplished in a few minutes and should be part of the routine examination taught to all medical students.

The patient is best examined lying comfortably on a couch or bed for most of the examination, but sitting up for the shoulders and neck, and standing for the final stages of examining the feet and the movements of the back. Walking should be observed, either at the start in the outpatient clinic or at the end of the examination if the patient is already undressed and in bed. In a very young child, much of the examination can be carried out while he or she sits on the parent's lap.

The examination starts most logically with the hands, the calling cards of many of the rheumatic diseases, then moves up the arms to the joints of the shoulder girdle and the jaw, down the spine from the neck to the coccyx and on to the legs, finishing with the toes.

The same basic pattern governs the examination of the most of the joints: inspection, palpation, and establishing the range of movement. This pattern is altered for some parts, for example the back, in the interests of efficiency.

Inspection (Table 1)

Colour
Skin or other local changes
Muscle wasting
Swelling
Deformity

Table 1 Inspection of joints

Where a joint is paired, both should be exposed to allow comparison of the two sides. Redness overlying a joint is an important indicator of acute inflammation, as in gout or septic arthritis. Any skin rash, subcutaneous nodules, cysts, scars, or evidence of local infection are noted. Muscle wasting is often associated with joint disease, but may be masked by joint swelling. Swelling or deformity should be sought. Discrepancies in limb or digit length may indicate growth problems secondary to inflammation, as in juvenile chronic arthritis.

Palpation (Table 2)

Temperature
Tenderness
Swelling:
Bone
Soft tissue
Fluid
Crepitus

Table 2 Palpation of joints

Temperature change is assessed relative to the same joint on the other side, or to surrounding normal skin. The back of the hand is rapidly moved between the areas to be compared. Errors may occur if one side has been kept warmer by a bandage, glove, or splint.

Tenderness is elicited with firm pressure along the joint margins, and over tendons and ligaments. This should be carried out with the examiner's eyes on those of the patient, rather than on the joint, so as to pick up the first signs of discomfort.

Swelling noted on inspection is palpated, to answer the question: 'Is this swelling bone, soft tissue, or fluid?' Bony swelling is especially a feature of osteoarthritis. Other causes are new bone formation with psoriatic arthritis, Charcot's joints, and callus formation after a fracture. Soft tissue swelling about a joint is most often synovial. The consistency of synovial swelling is more doughy than bouncy, like foam rubber. Where synovial fluid and synovial swelling coexist, the fluid can be pushed away or aspirated, allowing the synovium to be palpated. Fluid is recognizable for its incompressibility. Palpating a lax effusion is easy, because the fluid can be milked from one part of the joint to another. A tense effusion can be recognized by its 'bouncy' quality, like a firm rubber ball.

Joint noises are frequently better felt than heard. Loud cracks and snaps, particularly on pulling the fingers or rotating the ankles, may be quite normal. Pulling the fingers creates a vacuum and the popping sound represents the sudden development of a gas cavity in the joint fluid. The snap on rotating an ankle is related to the slipping of one tendon over another. Similar harmless snapping and crunching noises occur when the head is extended and rotated from side to side, or the shoulders are braced back and rotated.

Velvet crepitations are too soft to be audible, but may be felt. They occur when the joint contains small particles of proteinaceous material, most typically in rheumatoid arthritis. Crepitus in the joint of a patient with osteoarthritis feels like a dull, coarse crunching, and is caused by irregularities in the cartilage.

Eburnation crepitus is both heard and felt, occurring when the cartilage has been destroyed and the two bony surfaces are in contact. The name means 'turning to ivory' and the sound is thought to be similar to that of ivory grinding on ivory. Most often it emanates from the hip or the knee, and then care is needed to work out which is the affected joint as the vibrations are well transmitted up or down the femur. To demonstrate eburnation crepitus the examiner lays a hand on the joint and asks the patient to move it. Passive movement may fail to elicit the sign because the joint surfaces are not so closely opposed when the muscles are relaxed.

Range of movement

The patient is usually first asked to move the joint actively. If a full range of active movement can be carried out without discomfort, there is rarely any need to proceed to passive movement. If the patient is unable to carry out the full range of active movement, then the passive movements will be most informative. In general, where active and passive restriction are the same, limitation of movement will reflect one of the following: inflammation of a joint; contracture of the tendons or ligaments surrounding the joint; destruction of bone or cartilage in the joint. Restriction of active but not passive movement indicates rupture or inflammation of a tendon, muscle weakness, or failure of the nerve supply to the muscle.

The 'normal' range of movement must be interpreted with caution. There is much variation in joint mobility with age and race, and from individual to individual. The figures commonly given are those considered to be at the lower end of the usual range. An excessive range of movement in otherwise normal joints is referred to as the hypermobility syndrome. An excessive range in diseased joints reflects damage to the articular surfaces or the capsule and ligaments, leading to instability.

Recording measurements

In order to keep accurate records of a patient's progress, especially in the context of a clinical trial, reproducible measurements are needed. The following have proved reliable for serial observations by the same examiner, although significant interobserver variation occurs.

Ritchie index

This is a method of recording activity in rheumatoid arthritis by means of grading the tenderness to firm pressure or pain on movement in a standard selection of joints.

Ring size

This is useful for recording the degree of swelling in the proximal interphalangeal joints. The smallest size that will slip easily over the joint is recorded.

Goniometer

For accurate measurement of angles of movement or degrees of deformity a goniometer is necessary. It is useful to have a long-limbed one for large joints and a smaller one for joints of hands and feet. The limbs of the goniometer should be lined up along the long axis of the joint, the hinge on the joint line, and the maximum angles of flexion and extension recorded. The position of the joint in the extended anatomical position is taken as zero, and further extension is recorded as a minus value. A spirit-level inclinometer is also available for measuring angles of spinal movement.

Grip strength

A special sphygmomanometer with its own manometer is most convenient, though a small cuff attached to a standard mercury manometer is also satisfactory. The cuff is inflated to 30 mmHg and the patient squeezes it with one hand. The maximum reading achieved in three attempts is recorded.

The systematic survey

The whole patient (Table 3)

Dress: unkempt or neat

Shoes: high fashion or slippers

Getting out of chair: use of arms, help

Walking aids: stick, frame

First steps: gelling phenomenon

Gait: antalgic, waddling

Posture: erect, kyphotic, hangdog

Undressing: trouble with buttons, jacket

Table 3 Observing the whole patient

Rheumatic diseases are often multisystem in their effects, and many systemic diseases will present with rheumatic complaints. The discussion of the examination of the other systems in the body is outside the scope of this chapter, but it should be remembered that the examination of the musculoskeletal system is only one aspect of the careful and comprehensive history-taking and general physical examination that makes up the assessment of the patient with rheumatic symptoms.

Posture and gait

It is instructive to observe how the patient gets off a chair and starts moving after a period of inactivity. Gelling, or stiffness after inactivity, is a common feature of many arthritic conditions. Posture may give a clue to disorders of the back, and gait to problems in the lower limbs. Watching the patient undress is a valuable opportunity for observing whether there are difficulties. In a paediatric setting, quiet observation of the child at play can be extremely valuable in discovering whether a limb is being protected.

The hands

Screening

Inspect the hands, back and front. Palpate each of the finger joints in turn, one hand at a time, moving from the distal row proximally, including the thumbs, checking for tenderness or swelling. Ask the patient to make the hands into fists and straighten them out again. Offer the patient two fingers to grip each side and assess grip strength.

Inspection (Table 4)

Nails
Pitting
Onycholysis
Splinter haemorrhages
Shaggy cuticles
Nailfold infarcts
Periungual erythema
Dilated capillary loops
Skin
Rash of psoriasis, dermatomyositis
Vitiligo
Purpura
Raynaud's colour changes
Infarcts
Trophic
Calcinosis
Waxy thickening
Nodules
Palmar erythema
Muscles
Wasting
Tendons
Swelling of sheath
Rupture
Displacement
Joints
Swelling
Deformity

Table 4 Inspecting the hands

Nails

The nails may show the pitting or lifting from the nail bed (onycholysis) typical of psoriasis, nail-fold infarcts typical of vasculitis, or the shaggy cuticles and periungual erythema of dermatomyositis. Dilated capillary loops in the nail fold can be seen with the naked eye, though more easily with a magnifying glass, and tend to accompany certain rheumatic diseases, for example systemic sclerosis and dermatomyositis.

Skin

A rash on the backs of the hands may be due to psoriasis or to dermatomyositis, best distinguished from each other by the distribution. Dermatomyositis tends to affect the extensor surfaces of the joints in a neat and symmetrical fashion while psoriasis is more likely to be distributed at random across the hand. Palmar erythema is common in patients with connective tissue diseases, and is frequently different in nature from the palmar erythema of pregnancy or liver disease, looking more mottled and indeed more like palmar livedo reticularis than erythema. This appearance seems to be associated with systemic vasculitis.

Muscle wasting

This is common whenever joints are inflamed, but attention should be paid to any significant patterns of muscle wasting, such as the wasting of the thenar eminence in carpal tunnel syndrome, which spares only the adductor pollicis, visible as a band parallel to the wrist. Wasting of opponens pollicis commonly accompanies osteoarthritis of the first carpometacarpal joint.

Tendon involvement

Tenosynovitis of the flexor tendons may be seen as a fullness in the palm. It may be differentiated from Dupuytren's contracture by the absence of skin tethering over the surface, and the presence of crepitus on flexing the fingers. To detect crepitus of the flexor tendon the examiner places the index and middle fingers over the course of the tendon, with the thumb on the back of the hand exerting firm pressure. The patient is then asked to make a fist, and crepitus or nodularity may be felt as the tendon moves within its sheath. The procedure is repeated for each of the flexor tendons in turn.

Trigger finger occurs when the finger locks in flexion, but can be passively (albeit painfully) straightened. This is caused by a nodule on the tendon passing through a stricture in the tendon sheath. The weaker extensor muscles are unable to pull the nodule back through the same obstruction. The nodule is usually to be found in the palm just proximal to the metacarpal head, but will only be detected during active movement of the affected digit. Tendon involvement may also be recognized in the fingers by pinching the thickened soft tissues at the palmar surface of the base of each finger.

Swelling

The hands may look generally puffy with no localization of this over joints. This is especially common in mixed connective tissue disease or early systemic sclerosis, where the hands may resemble a bunch of sausages, and also occurs with systemic lupus erythematosus, dermatomyositis, polymyalgia rheumatica, reflex sympathetic dystrophy, and with early rheumatoid arthritis.

Where only one or two fingers have a diffuse, cylindrical swelling, not centred on a joint, we speak of a 'sausage' digit. These may represent swelling of the flexor tendon, as in the seronegative spondylarthropathies, diffuse soft tissue inflammation as in gout, or infection as in leprosy or tuberculosis.

Synovial swelling in the proximal interphalangeal joints is sometimes called spindling because of the fusiform appearance. Osteoarthritis causes a swelling of these joints that is bony and irregular, occasionally with effusions.

Distribution

The distribution of joint swelling gives an immediate clue to the diagnosis. Swelling predominantly of the terminal (distal) interphalangeal joints indicates osteoarthritis, psoriatic arthritis, or occasionally gout. Osteoarthritis also predominantly affects the first carpometacarpal joint, an articulation spared by most other arthropathies. The joint looks to be squared, because of osteophyte formation, wasting of the surrounding muscles, and adduction of the thumb. Rheumatoid arthritis has a predilection for the second and third metacarpophalangeal joints, as well as the proximal interphalangeal joints.

Deformities

Note deformities such as swan neck, boutonnière, ulnar drift, and Z thumbs. Palmar subluxation of the metacarpophalangeal joints is often mistaken for synovial swelling, but the exposed metacarpal heads are rounded and bony on palpation. Palmar subluxation is easily confirmed by running a finger along the back of the patient's hand and fingers to assess whether the phalanges are on the same plane as the metacarpals. Wherever deformities are found, assess whether they are

reversible, either actively or passively.

Range of movement

The patient should be able to make a fist, burying the fingertips in the palm, a movement that requires 90° of flexion in each of the three phalangeal joints. Hypermobility is indicated by more than 90° extension in the metacarpophalangeal joints and ability to approximate thumb to forearm, passively.

The wrist

Screening

Inspect both wrists, front and back, and palpate the joint line for tenderness or swelling. Ask the patient to press both palms together and bring the elbows out at right angles in the 'prayer position'. Then reverse the position with the backs of the hands together to demonstrate palmar flexion (Fig. 1, Fig. 2). If the hands, wrists, or elbows are deformed, this test cannot be carried out. In this case, assess the range of active and passive movement at each wrist in turn.

Fig. 1 The 'prayer position', demonstrating wrist dorsiflexion. (photography inthis andoither Figures in this chapter by courtesy of Eric Leung, The Photography and Illustration Centre, University College London.)

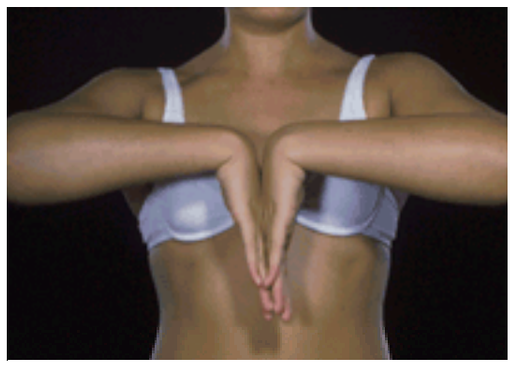


Fig. 2 The reversed 'prayer position', demonstrating wrist palmar flexion

Swelling

In rheumatoid arthritis, exuberant soft tissue swelling may be seen, originating either from the wrist joint itself or from the synovial tendon sheaths, and spreading both sides of the extensor pollucis retinaculum. An undulating swelling of the ulnar border of the wrist is a particular feature of rheumatoid arthritis. The undulations are the result of synovial swelling pushing its way through the fibres of the extensor retinaculum. This appearance is sometimes called the caput ulnae, and is associated with an increased risk of rupture of the fourth and fifth extensor pollucis tendons. Synovial proliferation is less often visible on the palmar aspect of the joint, although sometimes a visible fullness extends under the flexor retinaculum and into the palm.

Swelling of the wrist joint is associated with compression of the median nerve in the carpal tunnel. Particular attention should be paid to looking for evidence of this complication.

Inflammation of the abductor pollicis longus tendon at the wrist (de Quervain's tenosynovitis) is best diagnosed by inspecting the wrists from the radial side, with the palms together and any watches or bracelets removed. Unilateral swelling may then easily be seen, extending from the first carpometacarpal joint proximally. The diagnosis can be confirmed using Finkelstein's test. In this the patient is asked to bring the thumb across the palm and clasp the fingers around it. The examiner holds the clenched fist and tweaks it sharply in an ulnar direction, putting a sudden but not excessive pull on the inflamed tendon. This elicits a sharp pain, and should be done with care, watching the patient's face for signs of discomfort.

Repetitive strain injuries often affect the muscles of the wrist, giving a diffuse swelling of the muscles of the forearm, with crepitus on moving the fingers.

Deformity

Deformity of the wrist joint usually takes the form of palmar subluxation, owing to the stronger pull of the flexor than the extensor muscles of the forearm. Prominence of the lower end of the ulna is also common, because synovial proliferation at the inferior radioulnar joint weakens the ligaments and allows the bone to ride up in a deformity that may be either fixed or mobile. When the bone is mobile it is often painful, giving rise to the 'piano-key' sign, whereby pressure on the bone leads to the production of an audible protest from the patient.

Range of movement

The 'prayer position' normally allows 90° of passive palmar and dorsiflexion. The active range is 20° less in each direction.

The elbow

Screening

Ask the patient to fold the arms across the chest. Compare the elbow joints in this position. Inspect the extensor surface of each. Ask the patient to extend and flex the elbows, then with the elbows flexed and held against the sides, to show in turn the palms and the backs of the hands.

Inspection

Inspecting the elbows is a fruitful source of clues to the diagnosis in many rheumatic diseases. Rheumatoid nodules, gouty tophi, tendon xanthomata, olecranon bursitis, and the rash of psoriasis or of dermatomyositis may all be detected by a glance at the extensor surface of the elbow, just distal to the olecranon. The joints

are most easily compared with the elbows flexed and the arms folded.

Palpation

Palpation for synovial proliferation or joint effusion is best done in the groove between the lateral epicondyle and the olecranon process. The radioulnar joints are frequently involved in rheumatoid arthritis, leading to loss of pronation and supination. The examiner places one thumb on the radial head and the other on the inferior radioulnar joint at the wrist, then passively pronates and supinates the patient's forearm, feeling for crepitus in one or both joints.

Tenderness of the lateral or medial epicondyles and the muscles inserting into them is found in 'tennis' or 'golfer's elbow', respectively. To confirm, the patient is asked to clench the fist and to extend or flex the wrist against resistance. These manoeuvres should cause pain at the affected muscle insertion.

Range of movement

Flexion is usually limited by the interposition of soft tissues, and should be about 150°. Extension beyond –10° implies hypermobility. Pronation and supination take place at the superior and inferior radioulnar joints and must be done with the elbows flexed to exclude movement at the shoulder joint. The hands can normally be fully supinated and pronated.

The shoulder girdle

The three shoulder-girdle joints are best considered together, as all three work together in movements of the shoulder, and pain from all three is felt around the shoulder and upper arm. Acromioclavicular pain is usually felt at the tip of the shoulder. Pain in the sternoclavicular and glenohumeral joints is most often felt over the deltoid muscle and midhumerus. The manubriosternal joint is often swollen and sometimes painful in rheumatoid arthritis and in seronegative spondylarthritis.

Screening

Inspect the shoulder girdles both from the front and the back. Palpate the sternoclavicular and acromioclavicular joints for tenderness or swelling, then the glenohumeral joints starting from the anterior joint line, across the subacromial bursa, and on to the posterior joint line. Ask the patient to raise the arms sideways above the head, clasp the hands behind the neck with elbows well back, then clasp them behind the back as far up as possible (Fig. 3, Fig. 4).

Fig. 3 Clasping the hands behind the neck to demonstrate external rotation.

Fig. 4 Raising the hands as high as possible behind the back. The opposite scapula can just be reached.

A common source of error lies in what seems to be the attraction of symmetry. The patient with one stiff shoulder will often respond to a request to raise both arms by lifting both to the limit of the range of the abnormal shoulder, thereby giving the impression that both are stiff. It is worth encouraging the patient to make sure that the maximum range of each joint is being demonstrated.

The sternoclavicular joint

The sternoclavicular joint is often neglected, despite the fact that it is frequently affected in a variety of arthropathies, including polymyalgia rheumatica, rheumatoid arthritis, the seronegative spondylarthropathies, and septic arthritis. Pain is rarely localized to the joint, even when there is obvious swelling and redness, but is referred to the upper arm. Palpation should include the inner end of the clavicle and the articular surface of the manubrium, directly below it. Pain can best be elicited by asking the patient to brace the shoulders forward and back, or to shrug the shoulders up and down, movements which do not involve the glenohumeral joint. If these movements are painful but external rotation of the shoulder is not, there is strong evidence of sternoclavicular disease.

The acromioclavicular joint

The acromioclavicular joint is near the surface and can easily be seen and palpated if swollen. Passive adduction of the arm across the chest will often cause pain if the joint is inflamed. There may be a painful arc on abduction above 90°.

The shoulder (glenohumeral) joint

Inspection

Wasting of muscles should be noted both from the front and the back.

Swelling of the shoulder joint is not always obvious. A shoulder effusion in a patient with rheumatoid arthritis may give that shoulder a more normal-looking contour. Rupture of the shoulder capsule may lead to fluid tracking down to the upper arm.

A dislocated shoulder may appear square and dropped, while in rheumatoid arthritis the shoulders appear both square and raised, owing to upward subluxation of the

humeral heads as a result of dysfunction of the rotator cuff and muscle contracture.

Palpation

A shoulder effusion is best felt by placing the thumb in front and the fingers behind the joint and fluctuating the fluid back and forth across the shoulder. Tenderness may be elicited over the whole line of the joint in capsulitis or inflammatory arthritis, or may be localized where there is an isolated lesion of the tendon. Inflammation of the long head of biceps gives tenderness in the bicipital groove anteriorly, while supraspinatus tendinitis causes tenderness in the subacromial bursa.

Rupture of the long head of biceps is best appreciated by asking the patient to contract the muscle against resistance. The affected muscle will form a ball (the 'Popeye sign').

Movements of the shoulder girdle

Although individual arcs of movement may be attributed to each of the three joints of the shoulder girdle, in practice they work smoothly together when the arm is raised. The sternoclavicular joint is the single point of bony connection between the arm and the trunk, and allows considerable movement of the clavicle, which in turn acts as a strut to add freedom of movement to the arm. The acromioclavicular joint has less movement, but allows for rotation of the scapula on the thorax.

Pure glenohumeral movement can be isolated from movement of the scapula across the thorax by fixing the inferior angle of the scapula with one hand and passively abducting the arm with the other (Fig. 5). Normally the scapula starts to move at 30° of abduction, but if it is fixed at least 90° of glenohumeral movement is possible. Passive external rotation is another reliable measure of pure glenohumeral movement. The last few degrees of abduction may be lost in patients with a painful neck lesion.

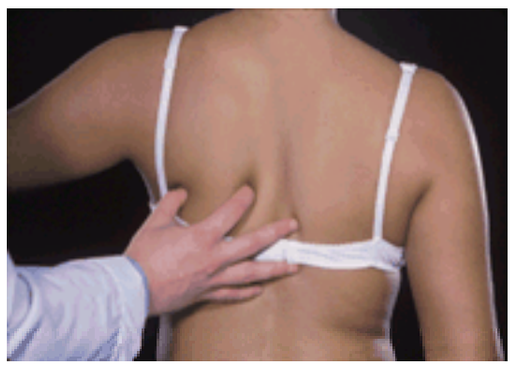


Fig. 5 Fixing the inferior angle of the scapula, while passively abducting the arm, to demonstrate true glenohumeral movement.

Passive movements should always be checked if the active range is incomplete. Loss of active and passive range in equal measure suggests intracapsular disease such as arthritis or adhesive capsulitis ('frozen shoulder'). Loss of active more than passive movement, accompanied by pain, suggests a lesion of tendon or rotator cuff. Limitation or weakness of active movement with a full and pain-free passive range suggests muscular or neurological disease.

The painful arc

During abduction of the arm, pain may be experienced in an arc, above or below which movement is pain free. The most common cause is supra spinatus tendinitis. In this condition pain occurs somewhere in the arc from 60 to 100° as the greater tuberosity impinges on the acromion, pinching the supraspinatus tendon. The pain may be severe enough to prevent abduction. The patient is often able to elevate the arm fully by forward flexion, but if asked to take it down sideways, will complain of pain as the upper limit of the painful arc is reached. A painful arc above 90° occurs in acromioclavicular disease.

Resisted movements

Specific tendon lesions can be identified by using resisted movements to elicit pain. Pain on resisted abduction suggests a supraspinatus lesion. Pain on resisted external rotation suggests an infraspinatus lesion. Pain on resisted internal rotation suggests a subscapularis lesion (Fig. 6).

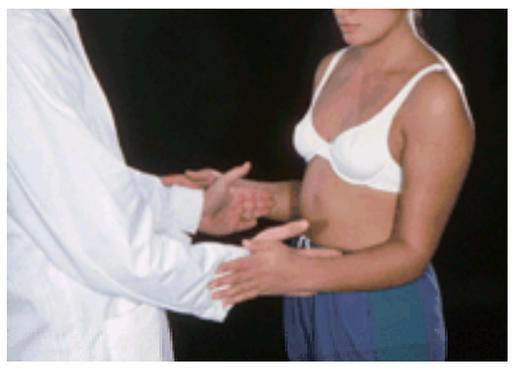


Fig. 6 Testing resisted internal rotation of the shoulder.

Range of movement

Movement at the shoulder comprises forward flexion to 180°, abduction to 180°, extension to 45°, and rotation to 70° internally and externally. Movement at the sternoclavicular joint includes protraction and retraction of the clavicle in an arc of 60°, and elevation of the clavicle to 60°.

The temporomandibular joint

Screening

Palpate the joints for tenderness or swelling and ask the patient to open the mouth wide.

These joints are most frequently involved in rheumatoid arthritis but other inflammatory arthropathies may also affect them. The complaint is of pain on chewing or yawning and sometimes of pain in the ear. The joints rarely look abnormal, but there is local tenderness and pain on opening the jaws. Normally it is possible for the fully open jaws to accommodate the middle three digits of the patient's hand, held vertically between the teeth. Restricted mouth opening from disease of the temporomandibular joints, may be differentiated from that caused by scleroderma by inspecting the lips, which will be drawn tight in scleroderma.

Micrognathia occurs as a result of temporomandibular joint inflammation in childhood, as in juvenile chronic arthritis (juvenile rheumatoid arthritis). Premature closure

of the epiphysis leads to failure of normal development of the mandible.

The cervical spine

Screening

Inspect posture: palpate the spinous processes for tenderness and alteration of the normal curve; ask the patient to flex, extend, rotate, and laterally flex the neck to left and right. It is important to remember that the detailed examination of the neck should not be considered complete until a neurological assessment has also been made.

Inspection

Neck posture may indicate underlying disease. A poking chin with dorsal kyphosis often accompanies cervical spondylosis. Acute torticollis can be recognized as the patient enters the room, one sternocleidomastoid muscle in spasm and the head tilted to the side and slightly rotated. The patient with rheumatoid arthritis will often have lost neck height and there may be a lateral asymmetry if there has been softening and erosion of bone.

Palpation

Palpation of the neck should include the occipital ridge, the spinous processes, and the paravertebral muscles overlying the facet joints. The occipital ridge is frequently tender in cervical spondylosis, especially over the greater occipital nerve. The spinous processes normally form a lordotic arc, with C2 and C7 being prominent at either end of the curve. Loss of lordosis or even a reversed curve can be appreciated by palpation. Palpation of the paravertebral muscles may reveal spasm in any case of neck pain, but especially after a whiplash injury or in cases of headache and neck pain related to stress. Through the paravertebral muscles, the facet joints can be felt, and tenderness or swelling elicited. Each pair of facet joints should be palpated in turn.

Range of movement

The patient should be able to touch the chin to the chest, tip the head back until the forehead and nose are parallel to the ceiling, and rotate in each direction to 70°. Lateral flexion is normally 45° each side, and this movement is lost early in ankylosing spondylitis. Passive neck movements are not a part of the routine examination, as neck instability is a feature of the rheumatic diseases.

Lhermitte's sign is more correctly a symptom. The patient complains of paraesthesia down the body on flexing the neck. It is sometimes present in cases of compression of the cervical cord, especially in rheumatoid arthritis. The examiner should not attempt to elicit this sign by forcible neck flexion.

The thoracic spine

Screening

Inspect the thoracic spine for swellings or deformities; palpate the spinous processes in turn; assess chest expansion. It is important to remember that the detailed examination of the thoracic spine should not be considered complete until a neurological assessment, including all four limbs, has also been made.

Inspection

This will reveal postural abnormality such as kyphosis or scoliosis. Destruction or collapse of vertebrae may lead to a sharply angled kyphosis, or angular gibbus, most often seen in tuberculosis. Kyphosis in adolescents suggests Scheuermann's disease. Scoliosis may be developmental, in which case there is usually some rotation of the ribs, which can be exaggerated by asking the patient to bend forward so that the rotation is more easily appreciated. Scoliosis resulting from muscle spasm becomes less obvious on forward flexion.

Palpation

The spinous processes and paravertebral muscles are next palpated, especially for tenderness. Recent collapse of a vertebra, malignant deposits, or vertebral osteomyelitis all produce local or 'point' tenderness. If direct pressure over the spinous process is insufficient to elicit this, it is worth trying percussion, laying two fingers over each process in turn and striking these with the fist, gently at first, then more firmly if no discomfort is elicited. Tenderness of the upper thoracic vertebrae and the paravertebral muscles is a common feature of thoracic spondylosis.

Range of movement

Rotation is best assessed with the patient seated to fix the pelvis and is usually 45°. Flexion is most easily measured with a tape measure. Distraction on flexion between marks made at C7 and T12 is normally at least 2.5 cm. Costovertebral movement is measured by chest expansion at the level of the nipples, usually more than 7 cm.

The lumbar spine

Screening

Inspect the standing patient from behind; palpate for tenderness or muscle spasm; ask them to touch the toes, lean back, and bend from side to side. With four fingers along the lumbar spine, assess whether movement is taking place at the lumbar spine rather than the hips. It is important to remember that the detailed examination of the lumbar spine should not be considered complete until a neurological assessment of the lower limbs has been made. This may conveniently be done as an intrinsic part of the examination of the back, which takes place in three positions as indicated in Table 5 and below.

Table 5 Examination of the lumbar spine

Examination

The patient standing

The examiner inspects the back for abnormality of posture or deformity, such as scoliosis or kyphosis, then palpates the paravertebral muscles for evidence of spasm. Palpation for tenderness is best left until the patient is lying down.

The patient is then asked to carry out the movements outlined in Table 5. Forward flexion and extension are relatively more restricted in the patient with a mechanical back pain, while lateral flexion is the first to be lost in ankylosing spondylitis. Scheber's test is a useful method of measuring lumbar flexion, especially in ankylosing spondylitis. Using a tape measure, the examiner makes two marks, one 5 cm below the level of the sacroiliac dimples and one 10 cm above. The patient bends forward and the examiner measures the distance between the marks. Distraction of at least 5 cm is normal.

Asking the patient to take a few steps on the heels and toes is a crude but useful test of power in the legs and feet, and at the same time the buttocks can be inspected for any evidence of a Trendelenburg sign (dipping down of the pelvis when the ipsilateral foot is raised from the ground) suggesting muscle weakness.

The patient lying face up

Straight-leg raise (Lasègue's sign)
The patient lies supine while the examiner grasps one of the ankles and raises it from the couch, keeping the knee straight with a hand on the knee, and watching the patient's face for any sign of discomfort (Fig. 7). Normally 80 to 90° can be attained without discomfort, depending on the tightness of the hamstrings. Irritation of the dural sleeve of any of the nerve roots contributing to the sciatic nerve will result in pain when the sciatic nerve is put on the stretch, and a much reduced straight-leg raise on the affected side. The usual cause of this is a prolapsed lumbar disc protruding posterolaterally. Bilateral reduction in straight-leg raising implies a central disc prolapse. To distinguish between sciatic irritation and hip disease, the leg is raised with the knee flexed. This will now be painless in sciatic irritation, but just as uncomfortable in hip disease.

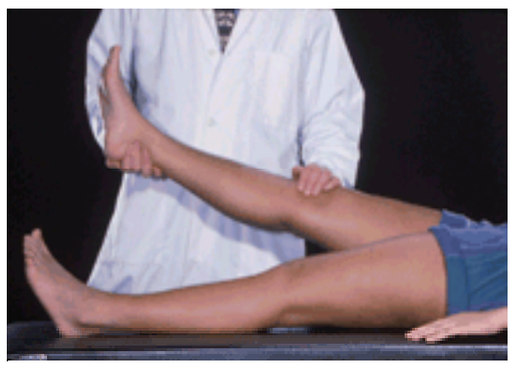


Fig. 7 The straight leg raise.

The sciatic stretch test may be useful in distinguishing sciatic irritation from tight hamstrings if the straight-leg raise is reduced. Once the limit of straight-leg raising is established, the leg is lowered a fraction until the pain is relieved. Now the examiner dorsiflexes the foot. If the pain was due to sciatic irritation, it returns.

After the straight-leg raise, the examiner tests the knee, ankle, and plantar reflexes. Testing power in each of the major muscle groups of the legs, especially the hip muscles, may cause difficulties if the patient is in pain. It may be necessary to reassess the patient after adequate control of the pain. Sensation to light touch and pinprick is next elicited, following a dermatomal pattern over the anterior aspects of both legs and the soles of the feet.

The patient lying face down

It is important to have the couch flat, with pillows removed to avoid painful hyperextension. A pillow under the patient's abdomen will usually increase comfort and render the examination easier. The examiner palpates each spinous process in turn, gently at first, then increasing the pressure by using one hand laid on the other. Then the examiner palpates for tenderness in the paravertebral muscles, the posterior iliac crests, and over the sacrum and coccyx. Coccygeal tenderness may also be felt by a rectal examination. Point tenderness is a feature of local infection or malignancy. More diffuse tenderness is common in patients with mechanical back pain.

The femoral stretch test is the parallel of the straight leg-raising test, putting the femoral nerve roots on the stretch. The examiner takes the prone patient's ankle in one hand, flexing the knee to 90°, and raising the bent leg to extend the hip (Fig. 8). Pain during this manoeuvre may indicate a lesion of a high lumbar disc or hip disease.

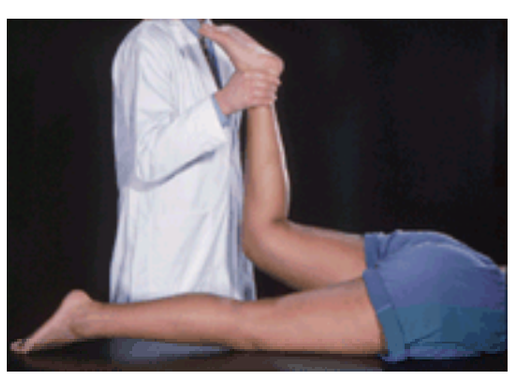


Fig. 8 The femoral stretch test.

Loss of sensation around the saddle area is an important clue to the presence of cauda equina lesions and should always be tested for. Buttock tone is assessed by asking the patient to pinch the cheeks of the buttocks together as hard as possible, and palpating the two sides. A rectal examination is helpful where there is any hint of a lesion of the cauda equina, to assess sphincter tone and sensation, in addition to its importance as part of the general assessment of the patient.

The sacroiliac joints

Screening

Inspect and palpate the joints from behind (Fig. 9), and stress them by pressure on the sacrum; roll the patient on one side and press again over the uppermost iliac crest.

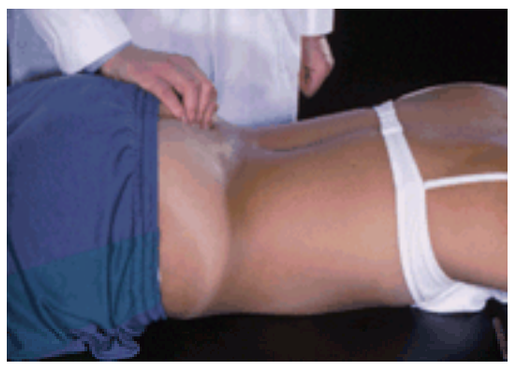


Fig. 9 Palpating over the sacroiliac joint.

Inspection and palpation

Inspecting the joints from behind may reveal evidence of swelling or deformity, especially in tuberculous infection. Palpation is best done with the four fingers of the hand along the length of the joint. Nodules of tender fibrofatty tissue often overlie the joints, and are of no significance except that they are often found in patients with chronic back pain. Bony fusion of the joints in ankylosing spondylitis obliterates the joint line. The synovial portion of the joint lies too deep to be felt under normal circumstances.

Range of movement

Moving the sacroiliac joints is not easy, as there is only a small amount of rotational glide, but they may be stressed by a variety of manoeuvres. The simplest method is to press firmly on the sacrum while the patient lies on the front, then on the pelvic brim as he or she lies on the back and on the side (Fig. 10, Fig. 11). If the sacroiliac joints are inflamed, these manoeuvres may elicit pain. Instability of the sacroiliac joints, especially postpartum, may cause pain, which is aggravated when the patient is asked to stand on one leg. Sometimes the excessive movement can be appreciated by placing one finger on each side of the joint as the patient makes the manoeuvre.

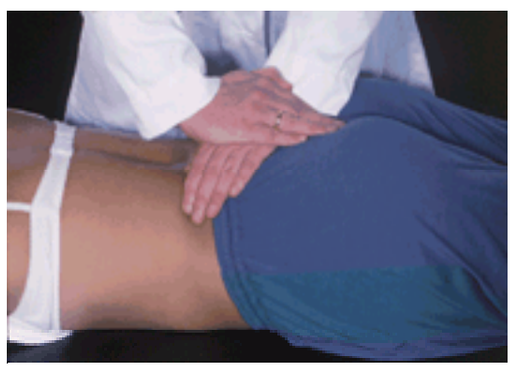


Fig. 10 Stressing the sacroiliac joints from the back.

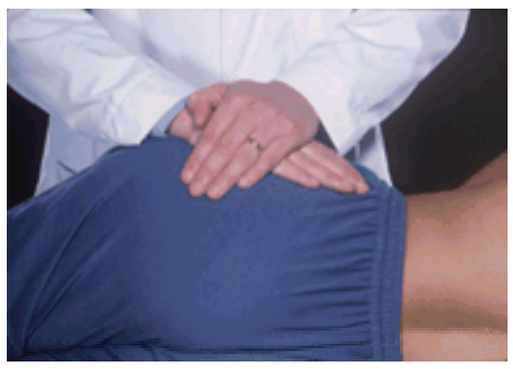


Fig. 11 Stressing the sacroiliac joints from the side.

The hip

Screening

Inspect leg length and position. Grasp the leg at knee and ankle, with the knee flexed; now flex the hip as far as it will go. Return the hip to 90° of flexion and test internal and external rotation using the foot as a pointer and the knee as a pivot (Fig. 12). Straighten the leg, grasp the opposite iliac crest and test abduction and adduction with the pelvis thus flexed (Fig. 13).

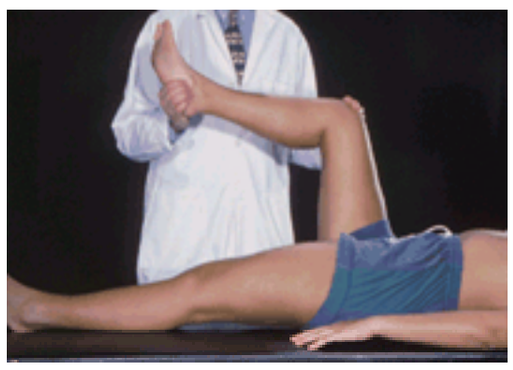


Fig. 12 Testing hip rotation using the foot as a pointer.

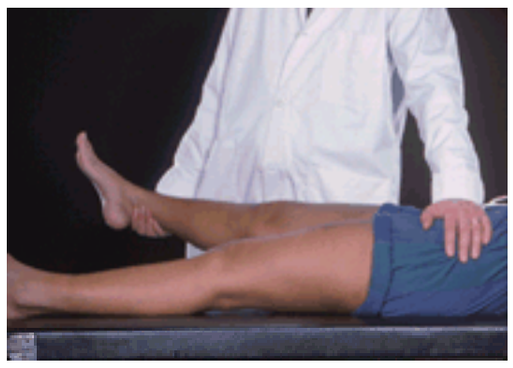


Fig. 13 Abducting the hip, with the pelvis fixed.

Inspection

The gait is particularly important in the assessment of the lower limb joints. The patient with a fixed flexion deformity of the hips throws the spine into an exaggerated lordosis to compensate. The patient with weak hip muscles walks with a waddle, and the one with a painful hip walks in such a way as to spend the least time possible on the painful leg, leaning to the opposite side, often with a walking-stick. This is called the antalgic gait.

Inspection of the hips should always include comparison of leg lengths and attention to the position of the leg. The leg that is short and externally rotated suggests a fractured neck of femur, or a legacy of juvenile chronic arthritis. Failure to observe the back of the hips may mean that the evidence of previous surgery is missed.

Palpation

Palpation of the hip is difficult because the joint is so deep, although occasionally an effusion can be detected by fluctuation just beneath the inguinal ligament. A psoas abscess may cause a very painful, stiff hip with spasm that cannot be overcome without anaesthesia. The psoas abscess will eventually point as a mass below the inguinal ligament. The greater trochanteric bursa lies over the greater trochanter and if it is inflamed, tenderness can easily be elicited with the patient lying on the unaffected side.

Range of movement

Flexion, 110°; extension, –30°; abduction, 50°; adduction, 30°; internal and external rotation, each 45°.

A fixed flexion deformity can be demonstrated with Thomas's test. The patient lies supine and one knee and hip are flexed until any lumbar lordosis has been obliterated. Flexion of the other hip during this manoeuvre indicates a fixed flexion deformity of that hip.

A catch for the inexperienced is the patient with a painful knee that is held in flexion. Attempts to straighten the knee in the bed are unsuccessful because of pain, and yet no other abnormality is seen in the knee. This patient probably has disease of the hip, with a fixed flexion deformity and pain referred to the knee. If the examiner were to straighten the knee while maintaining a flexed hip, a full range of movement would be found in most cases.

Rotation may also be tested in extension, by rolling the leg. Where both hips have a limited range of abduction, it is useful to record the maximum distance between the two medial malleoli. Extension of the hip is best measured with the patient lying on the side or prone.

The knee

Screening

Inspect for swelling, abnormal alignment, or quadriceps wasting. Palpate for tenderness, effusion, popliteal cyst. Flex and extend. Test ligament stability.

Inspection

Knee swellings can be recognized by fullness of the suprapatellar pouch and obliteration of the hollows either side of the patella. Quadriceps wasting may be masked by swelling of the knee and measurement of quadriceps bulk should be made above the upper limit of the supra patellar bursa. Deformities may be varus or valgus, forward or backward slip.

Palpation

P>Tenderness of the medial collateral ligaments occurs early in osteoarthritis, followed by the development of a tender bony ridge as osteophytes develop.

A knee effusion can be demonstrated in one of three ways, depending on the quantity of fluid present, as follows.

1. A small effusion can be milked from one side of the patella to the other in the 'bulge sign'. Any fluid in the medial side of the joint is swept firmly upward and laterally into the suprapatellar pouch. Then the back of the hand presses firmly and sharply on the lateral side of the joint. When a small effusion is present, a bulge will appear on the medial side of the joint as the fluid returns.
2. A moderate effusion is best detected by eliciting the 'patellar tap' (Fig. 14). One hand is placed on the suprapatellar pouch, the finger and thumb exerting side-to-side pressure, squeezing any fluid in the pouch to the retropatellar space. The patella is then sharply depressed with the fingers of the other hand, so that it floats through the fluid to strike the lower end of the femur. This sharp tap will not be felt if the undersurface of the patella is covered with synovial pannus, which acts as a blanket to muffle the impact. A tap may be felt in the absence of fluid if intra-articular fat is squeezed behind the patella.

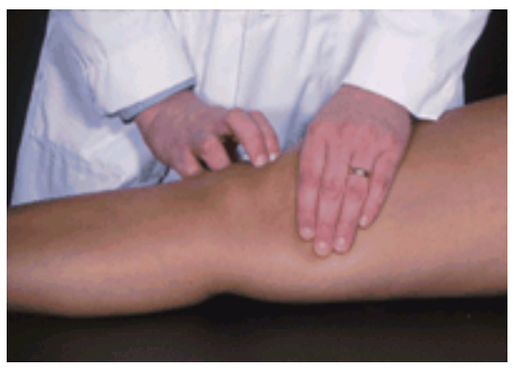


Fig. 14 Eliciting a patellar tap.

3. A large, tense effusion can be confirmed by fluctuation across the joint, from one side of the patella to the other.

A cyst of the calf or popliteal fossa is often easier to feel than see, and has the consistency of a firm rubber ball.

The patellar compression test

The patient lies supine with the quadriceps relaxed. The examiner presses the patella firmly distally between finger and thumb. The patient is then asked to push the knee hard into the bed. If this causes pain or crepitus it suggests retropatellar disease, such as chondromalacia patellae in the young or osteoarthritis in the elderly.

Testing stability

To test the cruciate ligaments the examiner grasps the leg just below the knee, with both hands, and exerts pressure first anteriorly, then posteriorly. To test the medial and lateral collaterals the examiner raises the patient's leg, supported with the fingers of one hand under the knee, holding it slightly flexed, and the other under the ankle (Fig. 15). Then pressure is exerted with the heel of each hand, producing a stress across the knee joint, centred on the ligament to be tested.

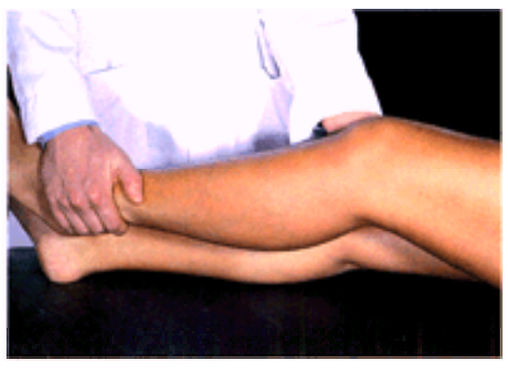


Fig. 15 Testing for lateral collateral ligament stability at the knee.

McMurray's sign

This is a test for cartilage tears. The patient lies supine and the examiner grasps the knee in one hand and the ankle in the other. The examiner then flexes the knee fully, internally rotates the lower leg as far as it will go, and slowly straightens the knee. Then the test is repeated with external rotation. If the hand over the knee detects a clunk on straightening, accompanied by wincing or other expression of pain from the patient, the test is positive.

Range of movement

Extension, 0°; flexion, 135°. Hyperextension can be tested by holding the knee firmly on the bed and lifting the ankle. Extension beyond –10° is abnormal.

The ankle joint

Screening

Inspect for swelling or deformity. Palpate the anterior joint line for tenderness, swelling. Assess dorsiflexion and plantar flexion.

Swelling can be detected anteriorly, between the malleoli, where it has the appearance and feeling of a 'spare tyre', quite unlike the diffuse swelling of ankle oedema. Swelling behind the lateral malleolus suggests peroneal tendinitis. Synovitis of the extensor tendons results in swelling both above and below the extensor retinaculum.

Range of movement

Dorsiflexion, –20°; plantar flexion, 70°.

The subtalar (talocalcaneal) joint

Screening

Inspect for varus or valgus deformity of the hindfoot. Grasp the heel and assess inversion and eversion (Fig. 16).

Fig. 16 Testing subtalar movement.

Swelling is seen below the malleoli. This joint is prone to symptomless involvement, especially in rheumatoid arthritis and juvenile chronic arthritis. Talipes valgus deformity usually results, and is best appreciated from behind, with the patient standing.

Range of movement

Inversion, 25°; eversion, 15°.

The heel and foot

Screening

Inspect both the dorsal and plantar aspects of the foot, and the back of the heel. Palpate the Achilles tendon insertion, the plantar fascia insertion (Fig. 17), and the metatarsophalangeal joints. Apply torsion to the midfoot to test movement in the midtarsal joints.

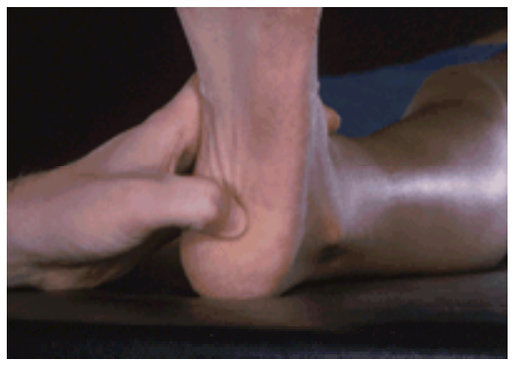


Fig. 17 Eliciting tenderness over the plantar fascia insertion.

Painful feet may be caused by vascular or neurological lesions. The peripheral pulses and sensation should always be assessed, and search made for ulcers, sinuses, or skin infarcts. Occasionally a prolapsed lumbar disc will present with foot pain without back pain or sciatica.

The Achilles tendon

This is inspected from behind, either with the patient standing or lying prone. Thickening, tenderness, and rupture can be detected. With rupture, the patient may have difficulty in standing on tiptoe at first, though fibrosis later occurs, rendering the movement possible. The Achilles tendon is a common site for traumatic bursas related to pressure from shoes, rheumatoid nodules, gouty tophi, and tendon xanthomata.

Heel

Heel pain is commonly due to plantar fasciitis. The heel can be inspected from below with the patient lying down. It is unusual to see swelling or discoloration, though the affected side often feels firmer. Tenderness is elicited by pressure directly upwards from the centre of the heel and by pressure from side to side at the level of the anterior margin of the calcaneum.

Midfoot

The naviculocuneiform joint is subject to osteoarthritis and is often found to show bony enlargement.

The sole

Callosities under the metatarsal heads develop when the metatarsophalangeal joints are subluxated. The rash of psoriasis or keratoderma blennorrhagica may be detected on the soles. Perforating foot ulcers may be a sign of systemic disease.

The arches

Flattening of the transverse or longitudinal arches is best seen with the patient standing. Pes cavus should also be assessed while standing.

The toes

Inspect the toes for deformities such as hallux valgus, clawing, overlapping toes, undersized toes. In rheumatoid arthritis, fibular drift of the toes is common. Clawed toes press on the uppers of shoes and callosities form over the proximal interphalangeal joints. 'Sausage toes' suggest Reiter's disease, gout, or psoriatic arthropathy. A bursa between the joints will spread the toes. Diffuse swelling over the metatarsophalangeal joints may be easier to feel than see; tenderness may be elicited by palpating each joint individually. Beware of squeezing the forefoot, as this can cause severe pain if the joints are inflamed.

Further reading

Bland, J.H. (1987). *Rheumatism of the cervical spine*. W.B. Saunders, London.

Cyriax, J. (1982). *Textbook of orthopaedic medicine*, Vol. I. *Diagnosis of soft tissue lesions*. Baillière Tindall, London.

Dieppe, P.A., Bacon, P.A., Bamji, A.N., and Watt, I. (1986) The clinical evaluation of the rheumatic diseases. In *Atlas of clinical rheumatology*, pp. 1.2–1.16. Gower Medical Publishing, London.

Edgar, M.A. and Park, W.M. (1974). Induced pain patterns on passive straight-leg-raising in lower lumbar disc protrusion. *Journal of Bone and Joint Surgery*, **56B**, 658.

Finkelstein, H. (1930). Stenosing tenovaginitis at the radial styloid process. *Journal of Bone and Joint Surgery*, **12**, 509.

Grieve, G. (1971). The hip. *Physiotherapy*, **57**, 212.

Klenerman, L. (1982). Examination of the foot. In *The foot and its disorders* (2nd edn) (ed. L. Klenerman). Blackwell Scientific Publications, Boston.

Larson, R.L. and Grana, W.A. (eds) (1993). *The knee: form, function, pathology and treatment*. W.B. Saunders, Philadelphia.

Magee, D.J. (1987). *Orthopedic physical assessment*. WB Saunders, Philadelphia.

Moll, J. and Wright, V. (1987). Measurement of spinal movements. In *The lumbar spine and back pain*. (3rd edn) (ed. M.I.V. Jayson), pp. 215–34. Churchill Livingstone, Edinburgh.

Ritchie, D.M., Boyle, J.A., McInnes, J.M., *et al*. (1968). Clinical studies with an articular index for the assessment of joint tenderness in patients with RA. *Quarterly Journal of Medicine*, **37**, 393–406.

Rudge, S.R., Swannell, A.J., Rose, D.H., and Todd, J.H. (1982). The clinical assessment of sacro-iliac joint involvement in ankylosing spondylitis. *Rheumatology Rehabilitation*, **51**, 15.

Schumacher, R.H. and Gall, E.P. (1988). *Rheumatoid arthritis: an illustrated guide to pathology, diagnosis and management*. J.B. Lippincott, Philadelphia.

Stickland, A. (1984). Examination of the knee joint. *Physiotherapy*, **70**, 144.

Thompson, P.W., Silman, A.J., Kirwan, J.R., and Currey, H.L.F. (1987). Articular indices of joint inflammation in rheumatoid arthritis. *Arthritis and Rheumatism*, **30**, 618–23.

Yocum, L.A. (1983). Assessing the shoulder: history, physical examination, differential diagnosis, and special tests used. *Clinical Sports Medicine*, **2**, 281.

1.1.5 Outcomes assessment in rheumatology

Oliver Sangha, Gerold Stucki and Matthew H. Liang

Introduction

Rheumatic disorders vary in their clinical expression, yet each has a major effect on function and health status. The major goal in the management of rheumatic disorders is to control or cure the disease and to preserve and control function and health status. This requires standardized assessment of organ morphology and function, as well as health status. Health-care professionals are accustomed to the measurement of organ pathology but if effective treatment is to improve, the assessment of health status is critical. There are several possible approaches to assess health status in a more standardized way: (i) the judgement of a health professional, (ii) performing standardized activities by the patient, and (iii) self-report of patients to standardized questionnaires.

In 1949, the American Rheumatism Association (now the American College of Rheumatology) Functional Classes were developed (Steinbrocker *et al*. 1949). Since then, hundreds of *ad hoc*, non-standardized, assessments of activities of daily living using one of the three approaches have been used. Beginning in the early 1980s, sophisticated, validated, and reproducible questionnaires to measure health status became available.

The literature on the psychometrics and properties of these instruments and their application to clinical trials is now commonplace, whereas their use in clinical practice is not established and the reasons why under study (Deyo 1982; Deyo and Carter 1992; Golden 1992).

This chapter reviews the patient-outcomes measurement of function, health status, and quality of life in the rheumatic and musculoskeletal diseases, conceptual limitations to their use, and guidelines for the selection of measures for specific applications.

Definitions and conceptual framework

The terms used in this chapter which describe the impact of rheumatic conditions on the individual are defined:

'Impairment' is defined as any loss or abnormality of psychological, physiological, or anatomical structure or function. It refers to the level of an individual organ or an organ system. Altered organ morphology or 'damage' may cause organ dysfunction. Impairment is concerned with abnormalities of body structure and appearance, and with organ or system function resulting from any cause. For instance, patients with synovitis usually demonstrate a limited range of motion. It is important to distinguish morphological changes and organ function from functional disability. A patient with a limited range of motion in joints may or may not be limited in activities of daily living (Ferraz *et al*. 1990).

'Disability' is the physical and psychological functional limitation caused by an impairment which is described by an individual when there is a discrepancy between one's capacity and an actual or perceived need for a specific function. A patient's expectations, motivation, the social support system, and the actual demands imposed by the physical limitations are critical determinants of this perception. Function changes over the course of people's development as to their capacity to do and what they wish or need to do.

In children and adolescents, rapid change and maturation of physical, cognitive, behavioural, emotional, and psychological function are the rule, whereas in adult life, those capacities are stable but life circumstances are changing and physical abilities may be declining gradually.

Physical function is dependent on physical integrity of the joints and neuromotor system to perform tasks needed for daily living, to care for oneself, to do recreational activities, to work, etc. When these activities of the daily living concept are extended beyond activities in the home and community, the term 'instrumental activities of daily living' is used. These include using a telephone, shopping for groceries, etc. Disability in instrumental activities of daily living suggests a need for special services.

'Health status' or 'quality of life', an ephemeral concept, embodies the dimensions of physical, social, and emotional function. If these concepts are attributed to health, the term health-related quality of life is used. Although there seems to be a general agreement on the dimensions which are included, there is considerable variation in the terminology and interpretation of this concept. Most correspond to the World Health Organization definition of health as a state of 'complete physical, mental and social well-being' (WHO 1958). Health status instruments usually include the dimensions of physical function, social function, emotional function, pain, and perception of well being.

Characteristics of a health status measure

The attributes of any quantitative measure are validity, reliability, responsiveness, and practical usefulness (Liang and Jette 1981).

'Validity' refers to whether an instrument measures what it is supposed to measure. Ideally, one would compare a measure with a gold standard, for example, comparing a suspicious nodule on a chest radiograph with a biopsy showing cancer (criterion validity).

For health status no reference or gold standard exists to judge the validity of a particular instrument. Instead one assesses the extent to which a measure is consistent with a theoretical concept (construct) concerning the phenomenon of interest (construct validity). Face validity (it 'looks like' it measures what it intends to measure) or content validity (it represents the domain of interest) are other techniques to strengthen the validity of a construct.

'Reliability' is the extent to which a measurement yields the same result on repeated administration of the questionnaire under the same circumstances (reproducibility). If scores of a health status instrument have little random error, they are considered reliable.

Validity and reliability are minimal criteria when one wants to differentiate individuals at one point in time. However, when used to evaluate changes over time, an instrument needs to be able to capture clinically meaningful changes. 'Sensitivity' denotes the capacity of a measure to show any change whether it is meaningful or not. This can be done by statistical techniques such as the standardized response mean (Liang *et al*. 1985), the effect size (Kazis *et al*. 1989), or Guyatt's responsiveness statistic (Guyatt *et al*. 1987). 'Responsiveness', on the other hand, is the capacity to show a change that is clinically meaningful to the patient and/or the physician. Responsiveness of a measure, is the criterion which ultimately determines the usefulness of any outcome measure in the evaluation of chronic conditions but is the measurement criterion least established for health status instruments. Finally, one needs to assess 'the practical utility' of a health status instrument for a given setting. In practice and research applications, the time needed to complete a questionnaire should be no more than 10 to 15 minutes to ensure compliance. In general, self-administered questionnaires are more practical than instruments requiring a trained interviewer. However, in multicultural populations or

where literacy levels are variable, a standardized interview might be the only way to obtain reliable information.

In the following we describe commonly used instruments which have been evaluated for their metric properties.

Health status and quality of life measurement

The instruments for measuring health status or quality of life cover a variety of dimensions of health, including physical, social, and emotional functioning. Some instruments are general health status measures which have been applied to rheumatic disorders. Others are measures developed for rheumatoid arthritis, osteoarthritis, the spondylarthropathies, and low back. Two rheumatoid arthritis instruments assess patient satisfaction with their function (Pincus *et al*. 1983; Meenan *et al*. 1992) and only one is an individualized measure of health status in rheumatoid arthritis (Tugwell *et al*. 1987) (Table 1).

General
Sickness Impact Profile (SIP) (Bergner *et al*. 1976)
Quality of Well-Being Index (QWB) (Balaban *et al*.)
Nottingham Health Profile (NHP) (McDowell *et al*. 1978)
Medical Outcome Study Short Form 36 (SF-36) (Ware and Sherbourne 1992)
Arthritis specific
Health Assessment Questionnaire (HAQ) (Fries *et al*. 1980)
Arthritis Impact Measurement Scales (AIMS) (Meenan *et al*. 1980)
McMaster Toronto Arthritis Patient Preference Disability Questionnaire (MACTAR) (Tugwell *et al*. 1987)
Lee Functional Status Index (Lee *et al*. 1973)
Toronto Functional Capacity Questionnaire (TFCQ) (Helewa *et al*. 1982)

Table 1 Selected health status measures in systemic rheumatic disease

General health status instruments measure multiple aspects of health, including physical function, social function, and pain and are suitable for the comparison of health status across multiple diseases or the value of competing clinical programmes. Generic health status instruments are useful in the evaluation of subjects with multiple chronic conditions, since they can detect changes arising from different organ systems. This is of particular interest when interventions can have effects or adverse effects on several organ systems.

Disease-specific instruments are useful for measuring clinically important changes in response to treatments (Patrick and Deyo 1989). Since these instruments include the elements most relevant to a particular disease, they are usually more sensitive to subtle improvements in health status. Disease-specific instruments are reliable as traditional measures of improvement in clinical status, such as anthropometric approaches (25-metre walk time) or laboratory tests (erythrocyte sedimentation rate). Health status instruments are interchangeable in their ability to measure major clinically significant improvement, but have varying ability to demonstrate changes in subdimensions such as social and global function (Liang *et al*. 1985; Bombardier *et al*. 1986). Measures of function or health status predict mortality and utilization of health services (McNevitt *et al*. 1986; Mitchell *et al*. 1986). A functional questionnaire is an economical and efficient technique for finding subjects with rheumatic and musculoskeletal diseases and has been applied in developing countries for evaluating community burden (Liang *et al*. 1981).

Generic health status measures

The Sickness Impact Profile (**SIP**) (Bergner *et al*. 1976) is a widely used general health status instrument containing 136 items that have true or false answers. Scores use predetermined weights based on rater panel estimates of relative severity of the dysfunction. The categories of ambulation, body care, and mobility are aggregated into a physical dimension, and four categories (emotional behaviour, social interaction, alertness behaviour, and communication) into a psychosocial dimension (Table 2).

Instrument	Dimensions covered	Mode of administration	Time to complete
Sickness Impact Profile (SIP) (Bergner *et al*. 1976)	Physical function, psychosocial, work, sleep and rest, eating, home management, recreation and pastimes	Self-administered or interview	~30 min
Quality of Well-Being (QWB) (Balaban *et al*. 1986)	Mobility, physical activity, social activity	Interview	~20 min
McMaster Health Index Questionnaire (MHIQ) (Chambers *et al*. 1982)	Physical function, social function, emotional function	Self-administered	~15–20 min

Table 2 Generic health status measures

The remaining categories are work, sleep and rest, eating, home management, and recreation and pastimes. The Sickness Impact Profile is available as a self-administered questionnaire or interview. It takes up to 30 min to complete as an interview. The SIP has been used in a number of studies on arthritis. The instrument demonstrates change in groups of arthritis patients, but is relatively insensitive to changes in individual patients (Deyo and Inui 1984; Liang *et al*. 1985; Stucki *et al*. 1995).

The Quality of Well-Being Index (**QWB**) and an earlier version, the Index of Well-Being, (Kaplan *et al*. 1976) assess mobility, physical activity, and social activity. An interviewer asks whether the patient was limited performing certain tasks during the last 6 days because of illness. Scoring for particular functions is based on preference weights derived from the normal population. These have been validated in patients with rheumatoid arthritis (Balaban *et al*. 1986). The interview requires a trained assessor and takes about 20 min. The instrument has been validated, but was not as sensitive to change as other measures in a clinical trial in rheumatoid arthritis (Bombardier *et al*. 1986). Major limitations are the complexity of the instrument and the requirement of a specially trained interviewer.

The McMaster Health Index Questionnaire (**MHIQ**) (Chambers *et al*. 1982) evaluates the quality of life in patients with rheumatoid diseases but has had limited application. A physical function index covers physical activities, self-care activities, mobility, communication, and global physical activity. A social index combines general well being, work performance, material welfare, support, participation with friends and family, and global social function. The emotional index measures feelings about personal relationships, self-esteem, the future, critical life events, and global emotional function. The MHIQ has 59 self-administered questions and takes 15 to 20 min to complete.

Short forms

One disadvantage of the instruments described above is that most require at least 15 to 20 min to complete. This becomes burdensome if general health status is measured along with disease-specific measures in a study, and may be a problem in daily practice. When studied, shorter measures appear to retain the psychometric properties of the longer instruments (Table 3).

Instrument	Dimensions covered	Mode of administration	Time to complete
Medical Outcome Study Short Form 36 (SF-36) (Ware and Sherbourne 1992)	Physical health, mental health, social functioning, role functioning, general health, vitality	Self-administered or interview or telephone interview	~10 min
Nottingham Health Profile (NHP) (McDowell et al. 1978)	Physical mobility, pain, emotional reaction, energy level, sleep, social interaction	Self-administered	~10 min

Table 3 Generic health status (short questionnaires)

The Medical Outcome Study Short Form 36 (**SF-36**) (Ware and Sherbourne 1992) comes from a larger battery of questions administered in the Medical Outcomes Study. The SF-36 includes eight multi-item scales containing 2 to 10 items each and a single item to assess health transition. The scales cover the dimensions of physical health, mental health, social functioning, role functioning, general health, and vitality. Forms cover a week or a month. The use of subscales is encouraged and the questionnaire can be self-administered or interviewer administered. The SF-36 is the most widely used general health status instrument and has been translated into many languages. The instrument is suitable for subjects aged 14 years and older and takes approximately 10 min to complete. Studies show excellent psychometric properties and there seems to be good sensitivity to change in patients with rheumatic conditions compared with longer instruments (Katz *et al*. 1990b).

The SF-36 allows scoring of the eight subscales and the construction of two summary scales, the physical component summary (**PCS**) and the mental component summary (**MCS**) scales. Futher evaluation of these summary scales provided the foundation for the construction of an instrument that is much shorter than the SF-36 (Ware *et al*. 1996). This new short form, the SF-12, uses 12 items from the SF-36, and demonstrates satisfactory reproducibility of the PCS and MCS with correlation coefficients of 0.93 to 0.97 on cross validation with the whole instrument. The new short form is likely to perform well enough for monitoring general populations; however, it does not allow scoring of individual SF-36 subscales such as bodily pain or social functioning.

While the SF-36 and SF-12 are commonplace in clinical and health services research, a previous version, the SF-20, is losing importance. The SF-20 was also developed and tested as part of the Medical Outcomes Survey (Stewart *et al*. 1988). Six health concepts are represented (physical, role, and social functioning, mental health, health perceptions, and pain) with between one and six items for each dimension. The instrument was developed in response to the problem that larger instruments taking up to 45 min to complete were impractical in clinical settings, and the poorer precision and reliability of single-item measures. The SF-20 is potentially less sensitive than the SF-36.

The Nottingham Health Profile (**NHP**) (McDowell *et al*. 1978) and its predecessor the Nottingham Health Index (**NHI**) (Hunt *et al*. 1985) assess perceived physical, social, and emotional health with 38 items answered yes or no. It uses weighted scores from panels' judgement about the severity of individual items. The NHP covers physical mobility, pain, emotional reaction, energy level, sleep, and social isolation and can provide dimension-specific scores.

The NHP is designed to represent rather severe problems, giving individuals with minor difficulties little room to improve. The NHP's reliability and validity have been assessed in rheumatoid arthritis, osteoarthritis, and in patients undergoing hip replacement. The instrument is self-administered and takes approximately 10 min to complete.

Short health status forms offer the promise of decreasing respondent burden in clinical research and integrating health status measurement into clinical practice. However, these measures have few items relevant to upper extremity function and therefore should be supplemented with disease-specific questions for use in rheumatoid arthritis. While shorter instruments are desirable, sensitivity is probably greater for the longer instruments.

Condition-specific measures

Rheumatoid arthritis

Disability as seen in rheumatoid arthritis is characterized by multiple joint involvement, chronicity, variable symptoms of disability, and occasional constitutional symptoms such as fatigue, fever, and weight loss. Rheumatoid disability is affected by the person's age and social role. In middle age, polyarthritis may affect career development, raising a family, family relationships, or return to work when children are grown up. Polyarthritis in older persons accentuates the ageing process and accelerates physical dependency.

Functional disability is the major consequence of disease in patients with rheumatoid arthritis. Relevant outcomes in rheumatoid arthritis include the evaluation of physical function, pain, treatment side-effects, quality of life, and cost-effectiveness (Table 4).

Instrument	Dimensions covered	Mode of administration	Time to complete
Arthritis Impact Measurement Scales (AIMS) (Meenan *et al*. 1980)	Mobility, physical activity, dexterity, social role functioning, social activity, activities of daily living, depression, anxiety, pain	Self-administered	~20 min
Health Assessment Questionnaire (HAQ) (Fries *et al*. 1980)	Physical disability (dressing and grooming, arising, eating, walking, hygiene, grip, activities), pain	Self-administered	~10 min
McMaster Toronto Arthritis Patient Preference Disability Questionnaire (MACTAR) (Tugwell *et al*. 1987)	Open-ended self-identified functional priorities, closed-ended physical activities	Interview	~10 min
Functional Status Index (FSI) (Jette 1980)	Activities of daily living (basic activities, gross mobility, personal care, home chores, interpersonal activities)	Interview	20–30 min
Toronto Functional Capacity Questionnaire (TFCQ) (Reisine *et al*. 1982)	Function in personal care, activities of upper extremities, mobility, leisure activities	Interview	~20 min
Kaleé Index (Kaleé *et al*. 1971)	Mobility, ambulation, self-care tasks	Clinician observation	~15 min
Lee Functional Status Index (Lee *et al*. 1973)	Physical function	Self-administered or interview	~10 min
Convery Polyarticular Disability Index (Convery *et al*. 1967)	Mobility, activities of daily living	Self-administered	~15 min

Table 4 Arthritis-specific measures—rheumatoid arthritis

The Arthritis Impact Measurement Scales (**AIMS**), a health status or quality of life instrument developed for rheumatoid arthritis, has 48 multiple-choice questions with nine subscales measuring mobility, physical activity, dexterity, social role functioning, social activity, activities of daily living, depression, anxiety, and pain (Meenan *et al*. 1980). The possible range of scores on each subscale is 0 to 10; subscale results are averaged to obtain a global score. The questionnaire is self-administered and takes 15 to 20 min to complete. AIMS 2 (Meenan *et al*. 1992) is an improved and expanded version which has additional scales measuring arm function, work, and support from family and friends, and a problem attribution section with yes/no questions. The scales also include three new questions to assess satisfaction with current level of functioning, specific impact of arthritis on an individual's health status, and prioritization of three areas where patients would most likely achieve improvement. Reliability and internal validity of the AIMS 2 have been demonstrated. Although the sensitivity to change has not been tested for AIMS 2, it is likely to be the same as or more sensitive than the original version.

Wallston *et al*. (Wallston *et al*. 1989) have shortened the original AIMS from 48 to 18 items by selecting two questions for each of the nine subscales. The abbreviated AIMS is slightly less reliable but has comparable validity to the whole instrument.

The Health Assessment Questionnaire (**HAQ**) (Fries *et al*. 1980) asks questions about physical and psychological disability, pain, global severity, employment, income, cost of medical care, and drug-related side-effects. The portion about disability and pain is composed of 24 questions on activities of daily living and mobility

including a single pain scale. The HAQ takes approximately 5 min to complete. The short form yields a summated index between 0 and 3 on a continuous scale.

The HAQ is probably the most widely used health status instrument in rheumatoid arthritis and has been tested extensively for validity and reliability and captures clinically meaningful changes over time but may be insensitive to early or advanced disability. It is included in the United States National Health and Nutrition Examination Survey (**NHANES**).

The Modified Health Assessment Questionnaire (**M-HAQ**) (Pincus *et al*. 1983) reduces the HAQ disability subscales to 8 questions and is reliable, valid, and sensitive to clinically meaningful change and is a useful alternative to the parent instrument.

The McMaster Toronto Arthritis Patient Preference Disability Questionnaire (**MACTAR**) (Tugwell *et al*. 1987) asks patients in a semi-structured interview to specify their own key functional activities. The five activities that rank highest are evaluated. On reassessment, patients are asked if their previously indicated limitations have improved, worsened, or remained the same. The MACTAR takes about 10 min to complete.

The Functional Status Index (**FSI**) (Jette 1980) measures pain, limitations, and dependence in performing 18 activities of daily living grouped as hand activities, gross mobility, personal care, home chores, and interpersonal activities. Each activity is rated between 0 and 4 on each dimension. Studies on patients with rheumatoid arthritis show validity and a high degree of interobserver reliability. However, the FSI has had little usage as judged by publications.

The Toronto Functional Capacity Questionnaire (**TFCQ**) (Helewa *et al*. 1982) assesses function in personal care, activities using the upper extremities, mobility, and leisure activities. The TFCQ takes approximately 20 min to complete and requires an interviewer. Scoring includes weighting based on panels of occupational and physical therapists and rheumatologists. The instrument is valid and responsive in clinical trials (Bombardier *et al*. 1986).

The Convery Polyarticular Disability Index (Convery *et al*. 1977) uses 16 items to rate functional impairment in polyarticular arthritis. The index is based on 9 mobility items and 7 activities of daily living. The questionnaire requires approximately 15 min to complete and is reliable and valid. There are no studies on its responsiveness.

The American College of Rheumatology (**ACR**) Revised Criteria for Classification of Global Functional Status in Rheumatoid Arthritis (Hochberg *et al*. 1992) rates patients from complete independence to limited abilities in performing usual self-care, vocational, and avocational activities. The criteria expand the original Rheumatoid Arthritis Classification of Functional Capacity (Steinbrocker *et al*. 1949). As an ordinal scale it is useful for stratifying or describing patients but less useful for showing change.

Unlike the questionnaires above, the Keitel Index (Keitel *et al*. 1971) consists of 24 standardized tasks which are rated by a trained examiner. The index tests motions of axial and peripheral joints focusing on mobility, ambulation, and self-care tasks. The complete test takes approximately 15 min to complete. Validity and reliability of the instrument, and the responsiveness in a clinical trial have been demonstrated.

The Lee Functional Status Index (Lee *et al*. 1973) includes 17 activities to assess the functional ability of the upper and lower extremities and the axial joints. The instrument has good psychometric properties and can be self-administered or interviewer based and takes less than 10 min to complete.

Childhood arthritis

In children, polyarthritis affects musculoskeletal, psychological, and social growth and development. It may interfere with the attainment of educational goals. In adolescence and young adulthood it may interfere with the acquisition of job skills, with emancipation from parents, achieving economic independence, interacting with peers, self-esteem, body image, and finding a partner.

The evaluation of physical function and health status in young patients requires measures different than those used to assess health status of adults with rheumatic diseases. Health status in children is conceptualized as achieving normal development and participating fully in developmentally appropriate physical, psychological, and social activities. Pain reporting is different in children than adults with a tendency to report less pain. Self-report is not considered reliable in children whose developmental or chronological age is below 4 years and necessitates a proxy or performance testing.

In children with juvenile rheumatoid arthritis, research has focused on the assessment of gross and fine motor skills and the evaluation of disability and pain. (Table 5)

Instrument	Dimensions covered	Age	Mode of administration	Time to complete
Development				
Hoskins and Squires Test for Gross Motor and Reflex Development (Hoskins and Squires 1973)	Gross motor, reflexes	0–5 years	Clinician observation	30–60 min
Denver Development and Screening Test (DDST) (Frankenburg *et al*. 1970; 1971)	Personal-social, fine motor adaptive and gross motor skills, language	1 month–6 years	Clinician observation	15–30 min
Newington Children's Hospital Juvenile Rheumatoid Arthritis Evaluation (Rhodes *et al*. 1988)	Motor tasks	1–6 years	Clinician observation	30–60 min
Function				
Adapted Arthritis Impact Measurement Scales for Children (AIMS) (Coulton *et al*. 1987)	Physical disability (dressing and grooming, arising, eating, walking, hygiene, grip activities), pain	All ages	Self-administered (parent or child)	~ 10 min
Juvenile Arthritis Functional Assessment Scale (JAFAS) (Lovell *et al*. 1988)	Physical activities	1–18 years	Clinician observation	~ 10 min
Juvenile Arthritis Functional Assessment Report for Children (JAFAR-C) and parents (JAFAR-P) (Howe *et al*. 1991)	Physical activities	1–18 years	Self-administered (child or parent)	~ 10 min
World Health Organization Study Scale (WHO) (Ellertsen *et al*. 1979)	Mobility, physical activity, role functioning, self-care	0–13 years	Interview and parent report	

Table 5 Arthritis-specific measures—childhood arthritis

Measures of developmental status

The Hoskins and Squires Test for Gross Motor and Reflex Development (Hoskins and Squires 1973) asks an infant to perform 60 voluntary physical tasks that are rated by a clinician on a scale from 1 to 4. The 'motor age' is calculated by finding the highest level where the subject can perform at least 50 per cent of the required voluntary skills. The assessment covers children from birth to 5 years with normal values derived from the literature. The instrument has been tested for interrater reliability and validity.

The Denver Development and Screening Test (**DDST**) (Frankenburg *et al*. 1970; Frankenburg *et al*. 1971) consists of a questionnaire to parents and a simple observational screening test for children. The test covers personal–social, fine motor-adaptive, language, and gross motor skills and is applicable to children aged between 1 month and 6 years. Test scores are used to assess whether a child has achieved age standardized milestones.

The Newington Children's Hospital Juvenile Rheumatoid Arthritis Evaluation (Rhodes *et al*. 1988) measures quality of motor performance against chronological age at which the task is typically mastered on a scale of 0 to 4. The index assumes that central nervous system development is normal and that tasks are sequentially learned. Fifty-eight motor tasks are evaluated and weighted equally. The instrument is designed to assess children aged from 1 to 6 years. Measures that evaluate developmental status measure impairment rather than functional disability and are restricted to children under 6 years of age.

Health status measures

The Adapted Arthritis Impact Measurement Scales for Children (Coulton *et al*. 1987) apply components of the AIMS to children with juvenile arthritis. The instrument includes scales covering physical disability and pain. Each scale contains four or five items with response options scored on a 0 to 10 scale. The pain scale is most reliable in children with active and inactive arthritis, and the physical activity and dexterity scales have reasonable reliability. The AIMS for children is administered by a health professional to the parents of an affected child.

The Childhood Health Assessment Questionnaire (**CHAQ**) (Singh *et al*. 1994) is based on the adult HAQ (Fries *et al*. 1980). The questionnaire is either parent or self-administered and can be used for all age groups. The CHAQ covers two domains, pain and disability. The latter assesses function in dressing and grooming, arising, eating, walking, hygiene, grip, and activities. Each item is rated on a four-point scale with the highest scoring question determining the score for the functional area. The CHAQ has been tested for validity and reliability and can be completed in less than 10 min.

The Juvenile Arthritis Functional Assessment Scale (**JAFAS**) (Lovell *et al*. 1989) is a series of timed tests using items from the AIMS, HAQ, and MHIQ and a consensus process among paediatric occupational and physical therapists experienced with patients with juvenile rheumatoid arthritis. The instrument has 10 items that differentiate children with juvenile rheumatoid arthritis and age-matched controls. The test is administered by a health professional, takes approximately 10 min to complete, and requires special equipment.

The Juvenile Arthritis Functional Assessment Report (**JAFAR**) is a self-reported version of the JAFAS (Howe *et al*. 1991). There are two questionnaires. The JAFAR-C is administered to children, the JAFAR-P to their parents. Patient and parent reports correlate highly with each other and with objective assessment by a physical/occupational therapist. Questionnaire scores were independent of the child's age.

JAFAS and JAFAR appear to be a convenient method to assess disability in patients with juvenile rheumatoid arthritis. Both instruments are reliable and valid, but there is no published data on their sensitivity. JAFAS and JAFAR cannot be administered to children under 6 years of age.

The Rand Health Insurance Study Scale (**HIS**) (Eisen *et al*. 1979; Eisen *et al*. 1980) assesses health status in children. The domains covered are mobility, physical activity, role functioning, and self-care. The Rand scale is administered by an interviewer who asks the parents questions. The items on social relations and parental concerns are relevant only to children aged 5 to 13 years. However, the remaining questions pertain to children from birth to 13 years. The instrument has good psychometric properties and was used in a study of the impact of childhood rheumatic diseases on the family (McCormick *et al*. 1986).

Osteoarthritis

Osteoarthritis disability, with its involvement of the small joints of the hand, the weight-bearing joints, and/or the axial joints is characterized by onset in middle or old age, lack of systemic symptoms, and mono- or pauci-articular involvement. Osteoarthritis disability unfolds slowly, paralleling the ageing process. Its impact on function occurs at a time when expectations and physical demands are less. Several studies have show that the radiographic appearance of osteoarthritis does not correspond with symptoms (Lawrence *et al*. 1966).

Patient-centred measures of osteoarthritis have concentrated on a careful assessment of pain and functional status. These include measures for evaluating specific interventions such as surgery, osteoarthritis-specific measures, and measures of generic health status.

Generic health status instruments (see Table 1 and Table 2) covering dimensions of social function, emotional function, role function, pain, and physical function are relevant to patients with chronic osteoarthritis (Table 6).

Instrument	Dimensions covered	Mode of administration	Time to complete
Western Ontario McMaster University (WOMAC) Osteoarthritis Index (Bellamy *et al*. 1988)	Physical disability, pain, stiffness	Self-administered	~10 min
Lequesne Index (Lequesne *et al*. 1987)	Pain, discomfort, activities of daily living, walking distance (separate scale for hip and knee joint)	Interview	~10 min
Hip Rating Questionnaire (Johanson *et al*. 1992)	Pain, function, walking	Self-administered	~10 min
Hip Arthroplasty Outcome Evaluation Questionnaire (Katz *et al*. 1995)	Pain, function	Self-administered	~10 min

Table 6 Arthritis-specific measures—osteoarthritis

Measures to evaluate osteoarthritis-specific interventions

Since the advent of total joint arthroplasty surgery, rating schemes have been used to evaluate the success of surgery (Larson 1963; Harris 1969) but until recently were unstandardized, untested for validity and reliability, and based solely on the surgeons assessment or anthropometric criteria, disregarding the patient's judgement.

It is not surprising, therefore, that whether a surgical intervention is judged as a success is dependent on the rating scale employed (Andersson 1972). In recognition of these limitations orthopaedic societies have led the way in developing psychometrically sound, standardized instruments in shifting over to groups of related joints and to the evaluation of all interventions rather than only joint arthroplasty (Liang *et al*. 1991; Johanson *et al*. 1992; Katz *et al*. 1995; Daltroy *et al*. 1996).

The Hip Rating Questionnaire (Johanson *et al*. 1992) is a standardized instrument to assess the outcome of total hip replacement. The questionnaire is one of the very few rating scales in orthopaedic surgery that has been tested for validity, reliability, and sensitivity to change. The domains of global impact of arthritic pain, walking, and function are represented in the self-administered questionnaire that can be completed in about 10 min.

Osteoarthritis-specific measures

Standardized health status measures have been validated for the use in patients with osteoarthritis of the joints of the lower extremities and all assess pain and physical function. The Arthritis Impact Measurement Scales (**AIMS**) (Meenan *et al*. 1980) and the Stanford Health Assessment Questionnaire (**HAQ**) (Fries *et al*. 1980) have been applied as outcome measures in pharmacological, surgical, and rehabilitative interventions in osteoarthritis.

The Lequesne Index (Lequesne *et al*. 1987) and the Western Ontario and McMaster Universities Osteoarthritis Index (**WOMAC**) (Bellamy *et al*. 1988) were exclusively developed for patients with osteoarthritis. The Lequesne Index contains two scales, one for the hip and one for the knee. Each scale covers pain and discomfort, activities of daily living, and maximum walking distance. The hip scale includes a question on sexual unction.

The Lequesne Index has been widely used in clinical and epidemiological studies of osteoarthritis and is tested for validity and reliability. A time period covered by the items is not defined and may affect its sensitivity. The hip or the knee scale can be completed in less than 10 min.

The WOMAC is a 24-item questionnaire focusing on the domains of pain, stiffness, and physical disability. The items were selected from the AIMS and HAQ and supplemented by questions felt to be important by 100 patients with hip and knee osteoarthritis. The WOMAC is tested for validity, reliability, and sensitivity to change, and has been used as the principal outcome measure in studies of surgical, physiotherapy, and pharmacological interventions. The instrument is self-administered and takes approximately 10 min to complete. Two versions of the WOMAC are available, a version utilizing five-point Likert response categories and a version scoring responses on a 10 cm visual analogue scale. Both scales produce comparable results.

Multiple studies that have used either instrument as key outcome measure in osteoarthritis research, and both scales have been recommended as suitable measures for testing slow acting drugs in osteoarthritis (Lequesne *et al*. 1994).

Spondylarthropathies

Disability in spondylarthropathy is characterized by stiffness or restrictive movement of the spine, with occasional involvement of the peripheral joints, most commonly the shoulder, hip, or knee. The disability is characterized by male predominance, onset in young and middle-aged adults, and by the episodic nature of the peripheral arthritis. Disability in spondylarthropathies is generally compatible with good function except when it involves weight-bearing joints in which case mobility may be affected. Spondylitis in young adult males may affect self-esteem, body image, and leisure.

Measurements used to evaluate the efficacy of treatments in ankylosing spondylitis include spinal and chest movement, duration and severity of morning stiffness, and quality of sleep. Health status indices such as the HAQ or AIMS are not readily applicable to ankylosing spondylitis since the items which might be affected by spinal mobility and symptoms are sparse (Table 7)

Instrument	Dimensions covered	Mode of administration	Time to complete
Health Assessment Questionnaire for the Spondylarthropathies (S-HAQ) (Daltroy *et al*. 1988)	Physical disability (dressing and grooming, arising, eating, walking, hygiene, grip activities), pain	Self-administered	~10 min
Dougados Functional Index (Dougados *et al*. 1988)	Physical function	Self-administered	~10 min
Leeds Disability Questionnaire (Abbott *et al*. 1994)	Pain, mobility, bending, reaching, neck movement, posture	Self-administered	~10 min
Bath Ankylosing Spondylitis Functional Index (Calin *et al*. 1995)	Physical function, activities of daily living	Self-administered	<10 min

Table 7 Arthritis-specific measures—spondylarthropathies

Daltroy *et al*. (1988) developed a functional status measure (**S-HAQ**) for patients with spondylitis by adding five items to the HAQ, to cover the activities identified to be most problematic in a survey of 300 British patients with ankylosing spondylitis. The instrument is valid and reliable.

Dougados *et al*. (1988) developed a 20-item functional index for the assessment of ankylosing spondylitis (Dougados Index). An item inventory was derived from a consensus of three rheumatologists and this was reduced by statistical procedures. The Dougados Index is valid and reliable and shows sufficient responsiveness (Dougados *et al*. 1990).

HAQ, AIMS, S-HAQ, and the Dougados Index have been compared to test their ability to capture physical limitations in patients with spondylarthropathies, demonstrating that the S-HAQ had the highest correlations with physical measures, particularly in patients with long-standing disease (Kuzis and Ward 1994).

The Leeds Disability Questionnaire (Abbott *et al*. 1994) assesses disability in ankylosing spondylitis, inquiring about four areas of function: mobility, bending down, reaching up and neck movements, and postures. The 16-item self-administered questionnaire is valid, reliable, and sensitive to change.

The Bath Ankylosing Spondylitis Functional Index (**BASFI**) (Calin *et al*. 1995) is a 10-item self-administered questionnaire to assess function and activities of daily living in patients with ankylosing spondylitis. The questionnaire takes less then 10 min to complete and has been tested for validity, reliability, and sensitivity to change.

Low back pain and other miscellaneous measures

A number of questionnaires have been specifically designed for people with back pain of mechanical or structural origin. Probably, the most widely used are the Oswestry Low Back Pain Disability Questionnaire (Fairbank *et al*. 1980), the Roland Disability Questionnaire derived from the Sickness Impact Profile (Roland and Morris 1983), and the Million Visual Analogue Scale (Million *et al*. 1982). Other measures have been reported (Mooney *et al*. 1976; Lankhorst *et al*. 1982; Lehmann *et al*. 1983; Evans and Kagan 1986; Lawlis *et al*. 1989; Greenough and Fraser 1992; Coste *et al*. 1993; Kopec *et al*. 1995; Daltroy *et al*. 1996).

The dimensions measured in all these measures include pain, functional, and pain-limiting activities. The instruments whose indicators of reliability and validity have been reported include the Roland Disability Questionnaire (Roland and Morris 1983), Oswestry Low Back Pain Disability Questionnaire (Fairbank *et al*. 1980), Million Visual Analogue Scale (Million *et al*. 1982), Dallas Pain Questionnaire (Lawlis *et al*. 1989), North American Spine Society (**NASS**) Lumbar Spine Outcome Assessment Instrument (Daltroy *et al*. 1996), Quebec Back Pain Disability Scale (Kopec *et al*. 1995), and EIFEL Questionnaire (Coste *et al*. 1993). The responsiveness of the Roland Disability Questionnaire was at least as good as that of the full SIP (Deyo 1986). Responsiveness has not been examined for the remaining scales, apart from demonstrating changes in the total score over time (the Oswestry and Million instruments).

Other outcome questionnaires have been developed for carpal tunnel syndrome (Levine *et al*. 1993; Katz *et al*. 1994a), fatigue in systemic lupus erythematosus (Krupp *et al*. 1989), shoulder pain (Roach *et al*. 1991), and fibromyalgia (Burckhardt *et al*. 1991).

Limitations of outcomes measures

Multidimensional instruments or subscales of these instruments have floors or ceilings in their capacity to detect change and this relates to the subject's baseline score and the number of items or gradations of change that address one end or the other of the continuum of the function observed (Stucki *et al*. 1996).

Function is multideterminate. Psychometrically sound instruments assume that function can be measured in all patients with the same instrument. However, patients' function is relative to their age, sex, motivation, social supports, priorities, and goals, and to their needs. Function is relative and, therefore, a small change in an individual's function may make a lot of difference.

The small change may be totally adequate for the person's needs, yet may not be statistically significant or captured by a questionnaire. Individualized quality of life and functional measures have been developed to capture patient priorities better, but whether the greater resources required to administer these measures results in sufficiently improved validity or responsiveness remains largely unexplored.

The measurement of specific functions in questionnaires is too coarse for monitoring patients closely. For example, function of hand and fingers may be assessed by a question on difficulty with fastening buttons or doing zippers. More subtle change would be assessed more appropriately by impairment measures such as grip and pinch strength and standardized dexterity measures. Self-reported symptoms and function do not necessarily correlate with objective impairments (Spiegel *et al*. 1985; Katz *et al*. 1990a; Liang *et al*. 1991; Davis *et al*. 1992) underscoring the fact that self-reported information complements measures of impairment and vice versa.

From a clinical perspective, not all activities of daily living are equal for a given patient, and the technology of eliciting preference weights is far from perfect (Thompson *et al*. 1982; Thompson *et al*. 1984; Katz *et al*. 1994b). Preferences of the sick and anxious do not remain constant during the vicissitudes of rheumatic illnesses, which are characterized by chronicity and an unpredictable waxing and waning course. Values change with time, and experiences of illness or personal circumstances change because people learn, adjust, or accommodate over the course of illness.

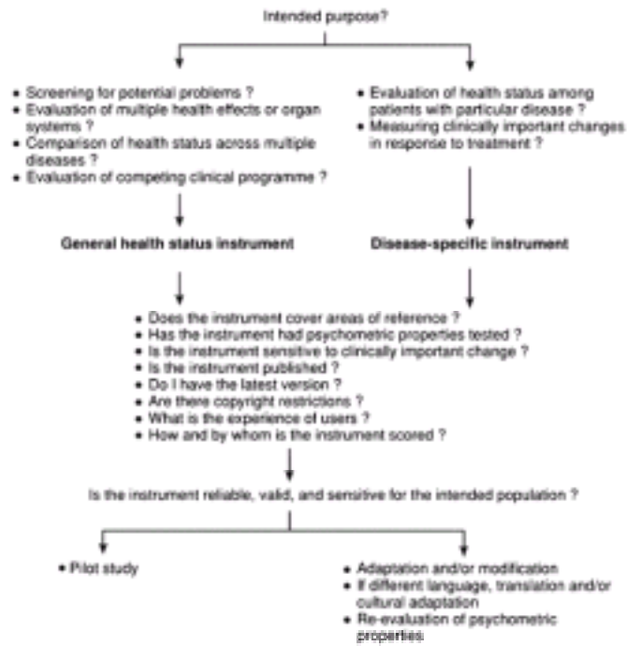
Selection of appropriate measures

For research applications, one should only use instruments with demonstrated psychometric properties which have been published. See Box 1 for the selection of

appropriate measures.

Published instruments are frequently revised and improved, and it is important to get the latest version and permission if protected by copyright. Strictly speaking, the validity and reliability of an instrument are characteristic of the instrument for a specific population and should be re-evaluated in a new population. This may not always be possible but, at a minimum, individual items should be inspected carefully to assess face validity and to make sure that all relevant outcomes and potential adverse consequences are included. The scale should cover the range of severity and the magnitude of the changes expected. A small pilot on individuals that are representative of the ones to be studied can be extremely informative. When a functional or health status measure is used in circumstances where language or cultural setting of the original instrument is different, a comprehensive procedure must be used to ensure its reliability and validity. Guidelines have been proposed (Guillemin *et al*. 1993).

Box 1 Selection of appropriate health status measures



The goal of medical care is to do no harm, to relieve pain and suffering, and to improve and maintain the patient's function, and in some instances, to prevent disability or improve self-image. Thus, any evaluation of an intervention, whether it is medication, surgery, or a programme with elements of these, should examine if these goals are achieved and to what extent.

In selecting measures for clinical trials in rheumatic conditions one needs an understanding of the range of benefits and adverse effects that might be expected and to get the best measure for each of the positive and negative consequences. Traditional anthropometric measures should be supplemented by self-reported measures of physical function and health status. The battery should include a general health status measure if the results are to be used in health policy. If available, a disease-specific health status measure such as the Arthritis Impact Measurement Scale or the Health Assessment Questionnaire should also be used.

Some interventions relieve symptoms and improve function promptly, whereas others may take time or require that a patient be rehabilitated from some stable level of function. Similarly, known and unknown adverse consequences can occur immediately or after a period of time. For a medical intervention designed to treat a root cause or primary mechanism of joint destruction, such as synovitis, one should see evidence that inflammation has been controlled. To the degree that psychosocial dysfunction results from the rheumatic condition, one would expect improvement of these parameters. However, prolonged psychological symptoms or innate traits are not likely to be helped by attention to synovitis alone. In fact, a discrepancy between one's perceived function and objective signs of disease is often a clue that there is something else going on which needs attention.

General health status measures are needed in studies used in health policy in which decision makers must allocate resources to different conditions. The use of such scales in a clinical trial can help the investigator relate the findings to other diseases and the policy maker to understand the trade-offs in resource allocation. The use of a general health status measure may not capture the specific outcomes seen in a disorder.

Arthritis-specific scales usually provide better coverage of dimensions of health thought to be important for that condition. The advantage of using an arthritis-specific scale is that clinicians may have a better feeling for what a change on the scale means.

The evaluation should include a measure of whether or not a change has occurred and if the change was important to the patient, and whether the treatment was worthwhile considering all the positive effects on disease and negative effects. Studies show that the importance of change can be judged differently by the patient and the physician (Redelmeier *et al*. 1993). We would recommend two simple questions: 'Have you experienced a change in your condition?' (and, if so) 'Has this been an important change?'

The Sickness Impact Profile, Index of Well-Being, Functional Status Index, Arthritis Impact Measurement Scale, and Modified Health Assessment Questionnaire have been compared using joint replacement surgery as an assay system to study their relative measurement sensitivity (Liang *et al*. 1985). The instruments correlate highly with one another and showed change. Of the five instruments, the Arthritis Impact Measurement Scale, Functional Status Index, and Sickness Impact Profile were equally efficient in detecting improvement in mobility, but the Health Assessment Questionnaire and Index of Well-Being were about one-half as efficient as the other three instruments. For pain evaluation, the Arthritis Impact Measurement Scale was more sensitive than the Health Assessment Questionnaire. The Index of Well-Being and Sickness Impact Profile do not have a pain subscale. With regard to social function, the Sickness Impact Profile, Index of Well-Being, and Health Assessment Questionnaire were more sensitive than the Arthritis Impact Measurement Scale. For overall or global function, the Sickness Impact Profile, Arthritis Impact Measurement Scale, and Index of Well-Being were the most sensitive.

In practice, the Index of Well-Being is a difficult questionnaire to administer, somewhat artificial for patients, and requires trained interviewers. However, it has the advantage of being a ratio scale with a true zero point which, though not often relevant in rheumatic disease, allows for the calculation of quality-adjusted life–years in cost–benefit studies. In a controlled drug trial in rheumatoid arthritis, it displayed the smallest change (Bombardier *et al*. 1986). The omission of pain as a dimension makes it less desirable as a single instrument in rheumatic disorders since pain is a central concern for patients. The Sickness Impact Profile is much easier to understand by patients, is self-administered, and has been used successfully in several rheumatic disorders. It results in one number. Its dichotomous response categories are less responsive than multi-item responses. It contains questions on continence and communication which are not relevant to rheumatic disorders.

Ultimately, if patient care and outcome are to be improved, attention to function must be incorporated into practice. Improved outcome will probably come from improved understanding of basic pathophysiology of the disease, improved treatment and understanding of the natural course of functional decline, and the critical points in which intervention might make a difference. To know a functional problem exists is only the start, and one needs to have improved understanding of when and what can be done to improve function.

For routine patient care the most desirable features of function or health status are severity, their utility in systematic problem identification (screening), and in following a patient and deciding whether that patient is better or worse (sensitivity and responsiveness). The instruments that have the widest usage in clinical settings are the HAQ and AIMS or the SF-36. To these we would add a transition question alluded to above.

Increasingly, policy makers are looking to outcome measurement as a way to assess the quality of care delivered by individual physicians, institutions, or systems of care, and to reduce ineffective practice variation. Randomized trials evaluating whether the collection of functional status information aids decision-making or improves outcomes have been disappointingly negative (McVey *et al*. 1989; Rubenstein *et al*. 1989; Kazis *et al*. 1990; Calkins *et al*. 1994), and indicate that much work remains to be done before the technology of functional status assessment is successfully transferred to improved patient care.

The reason for lack of improvement in functional status in these studies may have been the insensitivity of the functional measures, inadequate training of the

physicians to deal with functional problems, the timing of the information fed to the physician, or the intractability and multifactorial nature of disability.

Summary

The evaluation of clinical status in rheumatic and musculoskeletal diseases is not complete without an assessment of its impact in the patient's terms. Measures evolving over the last 20 years have made the measurement of these so-called 'soft' outcomes harder and more practical and should make research and patient care more relevant to patient's concerns.

Chapter References

Abbott, C.A., Helliwell, P.S., and Chamberlain, M.A. (1994). Functional assessment in ankylosing spondylitis: evaluation of a new self-administered questionnaire and correlation with anthropometric variables. *British Journal of Rheumatology*, **33**, 1060–6.

Andersson, G. (1972). Hip assessment: comparison of nine different methods. *Journal of Bone and Joint Diseases*, **54B**, 621–5.

Balaban, D.J., Sagi, P.C., Goldfarb, N.L., and Nettler, S. (1986). Weights for scoring the quality of well-being (QWB) instrument among rheumatoid arthritis: a comparison to general population weights. *Medical Care*, **24**, 973–80.

Bellamy, N., Buchanan, W.W., Goldsmith, C.H., Campbell, J.J., and Stitt, L. (1988). Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. *Journal of Rheumatology*, **15**, 1833–40.

Bergner, M., Babbitt, R.A., Pollard, W.E., Martin, D.P., and Gilson, B.S. (1976). The Sickness Impact Profile: validation of a health status measure. *Medical Care*, **14**, 57–67.

Bombardier, C., Ware, J., Russell, I.J., Larson, M., Chalmers, A., Read, J.L., and the Auranofin Cooperating Group (1986). Auranofin therapy and quality of life in patients with rheumatoid arthritis: results of a multicenter trial. *American Journal of Medicine*, **81**, 565–78.

Burckhardt, C.S., Clark, S.R., and Bennett, R.M. (1991). The fibromyalgia impact questionnaire: development and validation. *Journal of Rheumatology*, **18**, 728–33.

Calin, A., Garrett, S., Whitelock, H., Kennedy, L.G., O'Hea, J., Mallorie, P., and Jenkinson, T. (1995). A new approach to defining functional ability in ankylosing spondylitis: the development of the Bath Ankylosing Spondylitis Functional Index. *Journal of Rheumatology*, **21**, 2281–5.

Calkins, D.R., *et al*. (1994). Functional disability screening of ambulatory patients: a randomized controlled trial in a hospital-based group practice. *Journal of General Internal Medicine*, **9**, 590–2.

Chambers, L.W., MacDonald, L.A., Tugwell, P., Buchanan, W.W., and Kraag, G. (1982). The McMaster Health Index Questionnaire as a measure of quality of life for patients with rheumatoid disease, *Journal of Rheumatology*, **9**, 780–4.

Convery, F.R., Minteer, M.A., Amiel, D., and Connett, K.L. (1977). Polyarticular disability: a functional assessment. *Archives of Physical Medicine and Rehabilitation*, **58**, 494–9.

Coste, J., LeParc, J.M., Berge, E., Delecouillerie, G., and Paolaggi, J.B. (1993). Adaptation and validation in French of a Disability Rating Scale for Low Back Pain Patients. (The EIFEL Questionnaire). *Revue Rheumatism (English Edition)*, **60**, 295–301.

Coulton, C.J., Zborowsky, E., Lipton, J., and Newman, A.J. (1987). Assessment of the reliability and validity of the Arthritis Impact Measurement Scales for children with juvenile arthritis. *Arthritis and Rheumatism*, **30**, 819–24.

Daltroy, L.H., Larson, M.G., Roberts, W.N., and Liang, M.H. (1988). A modification of the Health Assessment Questionnaire for the spondyloarthropathies. *Journal of Rheumatology*, **17**, 946–50.

Daltroy, L.H., Cats-Baril, W.L., Katz, J.N., Fossel, A.H., and Liang, M.H. (1996). The North American Spine Society (NASS) Lumbar Spine Outcome Assessment Instrument: reliability and validity tests. *Spine*, **21**, 741–9.

Davis, M.A., Ettinger, W.H., Neuhaus, J.M., Barclay, J.D., and Segal, M.R. (1992). Correlates of knee pain among US adults with and without radiographic knee osteoarthritis. *Journal of Rheumatology*, **19**, 1943–9.

Deyo, R.A. (1986). Comparative validity of the Sickness Impact Profile and shorter scales for functional assessment in low-back pain. *Spine*, **11**, 951–4.

Deyo, R.A. and Carter, W.B. (1992). Strategies for improving and expanding the application of health status measures in clinical settings: a researcher–developer viewpoint. *Medical Care*, **30** (Suppl.), MS176–86.

Deyo, R.A. and Inui, T.S. (1984). Towards clinical applications of health status measures: sensitivity of scales to clinical important changes. *Health Services Research*, **19**, 275–89.

Deyo, R.A., Inui, T.S., Leiniger, J., and Overman, S. (1982). Physical and psychosocial function in rheumatoid arthritis. Clinical use of a self-administered health status instrument. *Archives of Internal Medicine*, **142**, 879–82.

Dougados, M., Gueguen, A., Nakache, J.-P., Nguyen, M., Mery, C., and Amor, B. (1988). Evaluation of a functional index and an articular index in ankylosing spondylitis. *Journal of Rheumatology*, **15**, 302–7.

Dougados, M., Gueguen, A., Nakache, J.-P., Nguyen, M., and Amor, B. (1990). Evaluation of a functional index for patients with ankylosing spondylitis. *Journal of Rheumatology*, **17**, 1254–5.

Eisen, M., Ware, J.E., Jr., and Donald, D. (1979). Measuring components of children's health status. *Medical Care*, **17**, 902–21.

Eisen, M., Donald, C.A., Ware, J.E., Jr., and Brook, R.H. (1980). *Conceptualization and measurement of health for children in the health insurance study*. Report Number R-2313-HEW. Rand Corporation, Santa Monica, CA.

Evans, J.H. and Kagan, A. (1986). The development of a functional rating scale to measure the treatment outcome in chronic spinal patients. *Spine*, **11**, 277–81.

Fairbank, J.C.T., Mbaot, J.C., Davies, J.B., and O'Brien, J.P. (1980). The Oswestry Low Back Pain Disability Questionnaire. *Physiotherapy*, **66**, 271–3.

Ferraz, B.M., Magalhaes Oliveira, L., Araujo, P.M.P., Atra, E., and Walter, S.D. (1990). EPM-ROM scale: an evaluation instrument to be used in rheumatoid arthritis trials. *Clinical and Experimental Rheumatology*, **8**, 491–4.

Frankenburg, W.K., Dodds, J., and Fandal, A. (1970). *The Denver Developmental Screening Test Manual*. Denver University Colorado Press, Denver, CO.

Frankenburg, W.K., Camp, B.W., and Van Natta, P.A. (1971). Validity of the DDST. *Child Development*, **42**, 475–85.

Fries, J.F., Spitz, P., Kraines, R.G., and Holman H.R. (1980). Measurement of patient outcome in arthritis. *Arthritis and Rheumatism*, **23**, 137–45.

Golden, W.E. (1992). Health status measurement. Implementation strategies. *Medical Care*, **30** (Suppl.), MS187–95.

Greenough, C.G. and Fraser, R.D. (1992). Assessment of Outcome in patients with low back pain. *Spine*, **17**, 36–41.

Guillemin, F., Bombardier, C., and Beaton, D. (1993). Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *Journal of Clinical Epidemiology*, **46**, 1417–32.

Guyatt, G., Walter, S., and Norman, G. (1987). Measuring change over time: assessing the usefulness of evaluative instruments. *Journal of Chronic Diseases*, **40**, 171–8.

Harris, W.H. (1969). Traumatic arthritis of the hip after dislocation and acetabular fractures: treatment by mold arthroplasty. An end-result study using a new method of result evaluation. *Journal of Bone and Joint Surgery*, **51A**, 737–55.

Helewa, A., Goldsmith, C.H., and Smyth, H.A. (1982). Independent measurement of functional capacity in rheumatoid arthritis. *Journal of Rheumatology*, **9**, 794–7.

Hochberg, M.C., Chang, R.W., Dwosh, I., Lindsey, S., Pincus, T., and Wolfe, F. (1992). The American College of Rheumatology 1991 revised criteria for the classification of global functional status in rheumatoid arthritis. *Arthritis and Rheumatism*, **35**, 498–502.

Hoskins, T.A. and Squires, J.E. (1973). Developmental assessment: a test for gross motor and reflex development. *Physical Therapy*, **53**, 117–25.

Howe, S., Levinson, J., Shear, E., Hanner, S., McGirr, G., Schulte, M., and Lovell, D. (1991). Development of a disability measurement tool for juvenile rheumatoid arthritis. *Arthritis and Rheumatism*, **34**, 873–80.

Hunt, S.M., NcEwen, J., and McKenna, S.P. (1985). Measuring health status: a new tool for clinicians and epidemiologists. *Journal of the Royal College of General Practitioners*, **35**, 185–8.

Jette, A.M. (1980). Functional Status Index: reliability of a chronic disease evaluation instrument. *Archives of Physical Medicine and Rehabilitation*, **61**, 395–401.

Johanson, M.A., Charlson, M.E., Szatrowski, T.P., and Ranawat, C.S. (1992). A self-administered hip rating questionnaire for assessment of outcome after total hip replacement. *Journal of Bone and Joint Surgery*, **74A**, 587–97.

Kaplan, R.M., Bush, J.W., and Berry, C.C. (1976). Health status: types of validity for an index of well-being. *Health Services Research*, **11**, 478–507.

Katz, J.N. *et al*. (1990*a*). Carpal tunnel syndrome: diagnostic utility of the history and physical examination findings. *Annals of Internal Medicine*, **27**, 1495–8.

Katz, J.N., Larson, M.G., Phillips, C.B., Fossel, A.H., and Liang, M.H. (1990 *b*). Comparative measurement sensitivity of short and longer health status measures. *Medical Care*, **28**, 632–42.

Katz, J.N., Gelberman, R.H., Wright, E.A., Lew, R.A., and Liang, M.H. (1994 *a*). Responsiveness of self-reported and objective measures of disease severity in carpal tunnel syndrome. *Medical Care*, **32**, 1127–33.

Katz, J.N., Wright, E.A., Liang, M.H., and Cleary, P.D. (1994 *b*). Differences between men and women undergoing orthopedic surgery for osteoarthritis. *Arthritis and Rheumatism*, **37**, 687–94.

Katz, J.N. *et al*. (1995). The validity and reliability of a Total Hip Arthroplasty Outcome Evaluation Questionnaire. *Journal of Bone and Joint Surgery*, **77A**, 1528–34.

Kazis, L.E., Anderson, J.J., and Meenan, R.F. (1989). Effect sizes for interpreting changes in health status. *Medical Care*, **27**, 178–89.

Kazis, L.E., Callahan, L.F., Meenan, R.F., and Pincus, T. (1990). Health status reports in the care of patients with rheumatoid arthritis. *Journal of Clinical Epidemiology*, **43**, 1243–53.

Keitel, W., Hoffman, H.L., and Weber, G. (1971). Ermittlung der prozentualen Funktionsminderung der Gelenke durch einen Bewegungsfunktionstest in der Rheumatologie. *Deutsches Gesundheitswesen*, **26**, 1901–2.

Kopec, J.A. *et al*. (1995). The Quebec Back Pain Disability Scale. Measurement properties. *Spine*, **20**, 341–52.

Krupp, L.B., LaRocca, N.G., Muir-Nash, J., and Steinberg, A.D. (1989). The Fatigue Severity Scale: application to patients with multiple sclerosis and systemic lupus erythematosus. *Archives of Neurology*, **46**, 1121–3.

Kuzis, S. and Ward, M.M. (1994). Measurement of functional disability in ankylosing spondylitis: a comparison of four self-report questionnaires. *Arthritis and Rheumatism*, **9** (Suppl.), S226.

Lankhorst, G.J., Van de Stadt, R.J., Vogelaar, T.W., Van der Korst, J.K., and Prevo, A.J.H. (1982). Objectivity and repeatability of measurements in low back pain. *Scandinavian Journal of Rehabilitative Medicine*, **14**, 21–6.

Larson, C.B. (1963). Rating scale for hip disabilities. *Clinical Orthopedics and Related Research*, **31**, 85–93.

Lawlis, G.F., Cuencas, R,, Selby, D., and McCoy, C.E. (1989). The development of the Dallas Pain Questionnaire. An assessment of the impact of spinal pain on behavior. *Spine*, **14**, 511–16.

Lawrence, J.S., Breamer, J.M., and Brier, F. (1966). Osteoarthrosis: prevalence in the population and relationship between symptoms and X-ray changes. *Annals of the Rheumatic Diseases*, **25**, 1–23.

Lee, P., Jasani, M.K., Dick, W.C., and Buchanan, W.W (1973). Evaluation of a functional index in rheumatoid arthritis. *Scandinavian Journal of Rheumatology*, **2**, 71–7.

Lehmann, T.R., Brand, R.A., and Gorman, T.W.O. (1983). A low back rating scale. *Spine*, **8**, 308–15.

Lequesne, M.G., Mery, C., Samson, M., and Gerard, P. (1987). Indexes of severity for OA of the hip and knee. *Scandinavian Journal of Rheumatology*, Suppl. **65**, 85–9.

Lequesne, M.G., Brandt, K., Bellamy, N., Moskowitz, R., Menkes, C.J., Pelletier, J.P., and Altman, R. (1994). Guidelines for testing slow acting drugs in osteoarthritis. *Journal of Rheumatology*, **21** (Suppl. 41), 65–73.

Levine, D.W. *et al*. (1993). Development and validation of symptom severity and functional status scales for carpal tunnel syndrome. *Journal of Bone and Joint Surgery*, **75A**, 1585–92.

Liang, M.H. and Jette, A.M. (1981). Measuring functional ability in chronic arthritis: a critical review. *Arthritis and Rheumatism*, **24**, 80–6.

Liang, M.H. *et al*. (1981). Evaluation of a pilot program for rheumatic disability in an urban community. *Arthritis and Rheumatism*, **24**, 937–43.

Liang, M.H., Larson, M.G., Cullen, K.E., and Schwartz, J.A. (1985). Comparative measurement efficiency and sensitivity of five health status instruments for arthritis research. *Arthritis and Rheumatism*, **28**, 542–7.

Liang, M.H., Katz, J.N., Phillips, C.B., Sledge, C.B., and Cats-Baril, W. (1991). The total hip arthroplasty outcome form of the American Academy of Orthopedic Surgeons. Results of a nominal group process. The American Academy of Orthopedic Surgeons Task Force on Outcomes Studies. *Journal of Bone and Joint Surgery*, **73A**, 639–46.

Lovell, D.H. *et al*. (1989). Development of a disability measurement tool for juvenile rheumatoid arthritis. *Arthritis and Rheumatism*, **32**, 1390–5.

McCormick, M.C., Stemmler, M.M., and Athreya, B.H. (1986). The impact of childhood rheumatic diseases on the family. *Arthritis and Rheumatism*, **29**, 872–9.

McDowell, I.M., Martini, C.J.M., and Waugh, W. (1978). A method for self-assessment of disability before and after hip replacement operations. *British Medical Journal*, **2**, 857–9.

McNevitt, M.C., Yelin, E.H., Henke, C.J., and Epstein, W.V. (1986). Risk factors for hospitalization and surgery for rheumatoid arthritis: implications for capitated medical payments. *Annals of Internal Medicine*, **105**, 421–8.

McVey, L.J., Becker, P.M., Saltz, C.G., Feussner, J.R., and Cohen, H.J. (1989). Effect of a geriatric consultation on functional status of elderly hospitalized patients, a randomized, controlled trial. *Annals of Internal Medicine*, **110**, 79–84.

Meenan, R.F., Gertman, P.M., and Mason, J.H. (1980). Measuring health status in arthritis: the Arthritis Impact Measurement Scales. *Arthritis and Rheumatism*, **23**, 146–52.

Meenan, R.F., Mason, J.H., Anderson, J.J., Guccione, A.A., and Kazis, L.E. (1992). AIMS 2. *Arthritis and Rheumatism*, **35**, 1–10.

Million, R., Hall, W., Nilsen, K.H., Baker, R.D., and Jayson M.I.V. (1982). Assessment of the progress of the back-pain patient. *Spine*, **7**, 204–12.

Mitchell, D.M., Spitz, P.W., Young, D.Y., Bloch, D.A., McShane, D.J., and Fries, J.F. (1986). Survival, prognosis, and causes of death in rheumatoid arthritis. *Arthritis and Rheumatism*, **29**, 706–14.

Mooney, V., Cairns, D., and Robertson, J. (1976). A system for evaluating and treating chronic back disability. *Western Journal of Medicine*, **124**, 370–6.

Patrick, D.L. and Deyo, R.A. (1989). Generic and disease-specific measures in assessing health status and quality of life. *Medical Care*, **27**, S217–32.

Pincus, T., Summey, J.A., Soraci, S.A, Jr., Wallston, K.A., and Hummon, N.P. (1983). Assessment of patient satisfaction in activities of daily living using a modified Stanford Health Assessment Questionnaire. *Arthritis and Rheumatism*, **26**, 1346–53.

Redelmeier, D.A., Rozin, P., and Kahneman, D. (1993). Understanding patient's decisions and emotional decisions. *Journal of the American Medical Association*, **270**, 72–6.

Rhodes, V.J., Pumphrey, K.F., and Zemel, L. (1988). Development of a functional assessment tool for children with juvenile rheumatoid arthritis. *Arthritis and Rheumatism*, **31** (Suppl. 4), S151.

Roach, K.E., Budiman-Mak, E., Songsiridej, N., and Lertratanakul, Y. (1991). Development of a shoulder pain and disability index. *Arthritis Care and Research*, **4**, 143–9.

Roland, M. and Morris, R. (1983). A study of the natural history of back pain. Part I: development of a reliable and sensitive measure of disability in low back pain. *Spine*, **8**, 141–4.

Rubenstein, L.V., Calkins, D.R., Young, R.T., *et al*. (1989). Improving patient function: a randomized trial of functional disability screening. *Annals of Internal Medicine*, **111**, 836–42.

Singh, G., Athreya, B.H., Fries, J.F., and Goldsmith, D.P. (1994). Measurement of health status in children with juvenile rheumatoid arthritis. *Arthritis and Rheumatism*, **37**, 1761–9.

Spiegel, J.S., Hirshfield, M.S., and Spiegel, T.M. (1985). Evaluating self-care activities: comparison of a self-reported questionnaire with an occupational therapist interview. *British Journal of Rheumatology*, **24**, 357–61.

Steinbrocker, O., Traeger, C.H., and Battman, R.C. (1949). Therapeutic criteria in rheumatoid arthritis. *Journal of the American Medical Association*, **140**, 659–62.

Stewart, A.L., Hays, R.D., and Ware, J.E. (1988). The MOS short-form general health survey. *Medical Care*, **26**, 724–35.

Stucki, G., Stucki, S., Brühlmann, P., and Michel, B.A. (1995). Ceiling effects of the Health Assessment Questionnaire and its modified version in some ambulatory rheumatoid arthritis patients. *Annals of the Rheumatic Diseases*, **54**, 461–5

Stucki, G., Daltroy, L.H., Katz, J.N., Johannesson, M., and Liang, M.H. (1996). Interpretation of change scores in ordinal clinical scales and health status measures: the whole may not equal the sum of the parts. *Journal of Clinical Epidemiology*, **49**, 711–7.

Thompson, M.S., Read, J.L., and Liang, M.H. (1982). Willingness-to-pay concepts for societal diseases in health. In *Values and long term care* (ed. R.L. Kane and R.A. Kane), pp. 103–5. DC Health, Lexington, MA.

Thompson, M.S., Read, J.L., and Liang, M.H. (1984). Feasibility of willingness-to-pay measurement in chronic arthritis. *Medical Decision Making*, **4**, 195–215.

Tugwell, P., Bombardier, C., Buchanan, W.W., Goldsmith, C.H., and Grace, E. (1987). The MACTAR Questionnaire—an individualized functional priority approach for assessing improvement in physical disability in clinical trials in rheumatoid arthritis. *Journal of Rheumatology*, **14**, 446–51.

Wallston, K.A., Brown, G.K., Stein, M.J., and Dobbins, C.J. (1989). Comparing the short and long versions of the Arthritis Impact Measurement Scales. *Journal of Rheumatology*, **16**, 1105–9.

Ware, J.E., Jr. and Sherbourne, C.D. (1992). The MOS 36-item short-form health survey (SF-36). A. Conceptual framework and item selection. *Medical Care*, **30**, 473–83.

Ware, J.E., Kosinski, M., and Keller, S.D. (1996). A 12-item short form health survey. Construction of scales and preliminary tests of reliability and validity. *Medical Care*, **34**, 220–33.

WHO (1958). Annex 1. Constitution of the World Health Organization. In *The first ten years of the World Health Organization*, p. 459. WHO, Geneva.

1.1.6 Psychological aspects of rheumatic disease

S. P. Newman

Introduction

The search for cures for many rheumatic conditions continues but progress has been limited. Treatment tends to revolve around attempting to manage and control the symptoms, and reduce the disabling consequences of the illness. The main thrust of treatment has been pharmacological or surgical (e.g. joint replacement), but also includes a range of non-medical interventions such as physiotherapy, pain control, weight-loss regimens, and self-management courses.

The shift from cure for individuals with a rheumatological condition mirrors the general shift in industrial societies from acute care to the management and treatment of chronic illness. This has posed a considerable challenge to health-care professionals and has led to an increasing focus on the psychological and social aspects surrounding chronic illnesses. Most chronic rheumatological conditions require major psychological adaptation when individuals face a lifetime of painful and increasingly disabling illness without prospect of a cure. They have to take regular medication and face the possibility of major surgery some time in their lives. Not only do individuals have to come to terms with the symptoms and treatment; they also have to adapt to altered life plans, reduced employment prospects, and uncertainty about the future course of the disease and its impact on their lives.

When individuals with rheumatoid arthritis are assessed about their concerns, they are most troubled about the symptoms of pain and fatigue but also have specific worries about current and future problems of mobility and disability (Burckhardt *et al*. 1993). Those with systemic lupus erythematosus have similar complaints about the symptoms of fatigue and pain, and are also concerned about uncertainty, lack of control, and appearance (Shortall *et al*. 1995).

Impact of rheumatological disorders on the individual

Introduction

Terminology continues to cast a cloud over the understanding of this domain of research. Various different and often overlapping terms are commonly used to refer to the more general consequences of rheumatological diseases on the individual. These terms include health status, functional status, impairment, disability, handicap, quality of life, activities of daily living. The term quality of life has been used to accommodate all of these dimensions but has also on occasions been used to refer to a more limited domain such as mood state and psychiatric disturbance. In this chapter 'quality of life' will be used to refer to composite measures of functional ability, mood and psychiatric state, and lifestyle. In those studies where composite measures of the quality of life have been used, significant associations have been found with levels of disability (see, for example, Bendtsen and Hornquist 1993).

The majority of studies of chronic disease in general and rheumatological disorders in particular have, however, tended to examine specific aspects of quality of life: physical function, symptoms, social and psychological well-being (Fitzpatrick *et al*. 1992). The focus on different areas of functioning enables the specific impact of the illness on different areas of life to be examined. The divisions of quality of life into symptoms, physical functioning, social functioning, and psychological well-being will be used in this chapter, although it must be recognized that this does impose arbitrary boundaries in relation to the realities of chronic illness for the individual (Newman *et al*. 1996).

Symptoms

Pain

Pain is customarily ranked as the most important symptom for adults with rheumatoid arthritis (Gibson and Clark 1985) and concerns about being free of pain are patients' primary concerns (Deyo 1988). The extent of pain is found to be related to general measures of physical disability. Using the functional classes (Steinbrocker *et al*. 1949), Parker *et al*. (1988a) found that patients classified as functional class III ('functional capacity quite limited') reported more pain on a visual analogue scale of pain intensity and that a greater area of their body was painful, in contrast to those from classes I and II. Clinical measures of disease activity or damage have, in general, not been found to predict the extent of pain reported by individuals with rheumatological conditions. In the study by Parker *et al*. (1988a), clinical variables such as erythrocyte sedimentation rate, Ritchie articular index, grip strength, walking speed, and morning stiffness were only able to account for 3 per cent of the variance in pain scores. In contrast, sociodemographic factors such as age and income were reasonable predictors of pain. Other studies, discussed further below, emphasize the importance of psychological variables such as 'self-efficacy' (Buescher *et al*. 1991) and coping (Newman *et al*. 1990) as mediators of reports of the extent of pain.

Merely describing the extent of pain does not enable the differing qualities of pain experience to be understood. The McGill pain questionnaire tests 78 qualities of pain and is widely used as an assessment of the quality of pain. Melzack (1975) reported that the majority of patients with arthritis used words like aching, exhausting, and rhythmic to describe their pain and more than a third mentioned gnawing, annoying, and constant. Other research has suggested that such patients use terms such as throbbing and burning, but do not use terms such as scalding, drilling, or *cutting* (Wagstaff *et al*. 1985). Pain quality also appears to be dependent on the nature of the activity that individuals are engaged in. In rheumatological illness, pain at rest and pain while moving appear to be qualitatively different. Pain at rest tends to be described as throbbing or aching, whereas shooting and spreading better describe pain on movement (Papageorgiou and Badley 1989).

Stiffness

In contrast to pain, stiffness has not been subjected to much research. Little is known about whether it is a symptom that can be clearly distinguished from pain by most individuals with rheumatological conditions. The subjective qualities of stiffness are also less well understood than those of pain. In one study of 100 patients, spontaneous descriptions of joints included the term stiffness, which they described as limitations in their movement (Rhind *et al*. 1980).

One characteristic of stiffness in rheumatoid arthritis is its temporal nature. Those individuals who are able to adjust their timetable to accommodate early-morning stiffness may be more able to accommodate to the problems of stiffness in arthritis. The ability to be flexible in work has been shown to be important and for many who

have early-morning stiffness rescheduling their day would be one way of dealing with this aspect of stiffness.

Fatigue

Fatigue is a primary symptom of rheumatoid arthritis and in some studies was considered by sufferers as their most severe symptom (Fitzpatrick *et al*. 1992). It tends to fluctuate with the characteristics of the disease as well as increase in times of emotional distress (Tack 1990). As with pain, when patients with rheumatoid arthritis are asked in detail they claim they can identify different types of fatigue (Papageorgiou 1994). Fatigue is also a major symptom of systemic lupus erythematosus and in some studies was the most problematical aspect of the illness (Shortall *et al*. 1995).

In both rheumatoid arthritis and systemic lupus erythematosus, fatigue is associated with sleep difficulties, and in both it also appears to be associated with mood problems and in particular depression (Basia 1995; McKinley *et al*. 1995; Shortall *et al*. 1995). In addition, fatigue in rheumatoid arthritis was also associated with functional disability, pain, and depression as well as with a lower haematocrit (Basia 1995). Determining the causal direction of these influences is difficult, but McKinley *et al*. (1995) have provided data to suggest that the experience of fatigue in systemic lupus erythematosus is best understood by the impact of the disease mediated through sleep and emotional difficulties.

Physical functioning—disability and handicap

Prevalence of disability

Population surveys of disability arising from musculoskeletal disorders produce estimates ranging from 5 per cent up to 8 per cent of individuals experiencing substantial disability (Badley 1992), with arthritis identified as the most common cause of disability (Martin *et al*. 1988). Most surveys show an age-related increase in disability (e.g. Badley and Tennant 1993).

In population studies that have specifically examined rheumatic conditions, differences in the numbers and level of disability in different types of rheumatic disease have been found. Badley and Tennant (1993) found that osteoarthritis and back disorders were the most common cause of disability in rheumatic disease, but that those reporting rheumatoid arthritis had the most current symptoms, made use of medical services the most, and had the highest level of disability.

Reporting disability

Many techniques for assessing disability rely on self-reports of individual capability without obtaining an independent observed assessment. Self-reports may be coloured by an individual's psychological state. For example, individuals who are depressed may underestimate their abilities. Studies have shown that self-reports of functional performance are related to other factors such as mental health and also health perceptions (Spiegel *et al*. 1988). Self-perceptions of health appear in turn to be related to demographic factors such as education as well as physical activity and depression (Guccione *et al*. 1995).

Among young children who have juvenile rheumatoid arthritis a problem exists as to whether their information correlates with that obtained from a parent. In essence this is a problem in all areas of information where more than one informant is used. It is emphasized within juvenile rheumatoid arthritis because of the need to obtain an accurate history and concerns that children may not be accurate. Doherty *et al*. (1993), using a formal questionnaire (the Child Health Assessment Questionnaire), found a relatively good concordance between assessments of their disability and pain made by young individuals with rheumatoid arthritis and those made by their parents.

Predicting disability

What factors influence or determine disability in rheumatic diseases is an important question, as the answers may guide clinicians to potentially useful areas of intervention. There are large individual differences in the health status of people with rheumatic disease. In particular, individuals with rheumatological illnesses that appear to be similar by clinical criteria often exhibit different levels of health status. This has led to the argument that one needs to incorporate both biological and psychosocial factors when looking at predictors of disability. Studies that have used both clinical and psychosocial variables have been able to achieve much greater levels of prediction of health status than those using disease measures alone. One study (Lorish *et al*. 1991) examined and followed 155 patients with rheumatoid arthritis and examined the association of clinical and psychosocial with physical disability. Cross-sectional step-wise regression analysis showed that the two sets of variables combined accounted for 54 per cent of the variance at baseline and 35 per cent at 12-month follow-up. The psychological measure of helplessness in this study added significantly to the ability to predict disability independent of disease severity.

A number of follow-up studies that have investigated predictors of disability have used the Health assessment questionnaire as the outcome measure. Leigh and Fries (1992) made an 8-year follow-up study and Eberhardt and Fex (1995) a 5-year follow-up study on individuals with early rheumatoid arthritis. Both these studies found the level of disability at outset of the study to be the most important predictor of future disability, as had been reported previously (Wolfe and Cathey 1991). The sex of the patient was also a significant predictor of disability, in that women were found to be more disabled (see also Thompson and Pegley 1991). What factors are included in the assessment of disability appears to influence the likelihood of finding sex differences in disability in rheumatic disease. While studies have demonstrated a gender difference on the Health assessment questionnaire, the Groningen activity restriction scale appears to show smaller sex differences (Doeglas *et al*. 1995a).

Other sociodemographic factors are important predictors of disability. Leigh and Fries (1992) found that older patients and those who were single at the time of the first assessment (never married, divorced or widowed) showed increasing disability over time. The latter finding has been related to the role and importance of social supports as buffers against the stresses of rheumatic disease and is discussed further below.

Level of pain, global health status, and number of working hours assessed at the outset were also predictive of disability 8 years later (Leigh and Fries 1992). The investigators offer two explanations to account for this last finding. First, and most obvious, is that individuals with lower levels of disability would be more likely to be in work. The second explanation is more psychological and interprets the increase in disability through the effects of unemployment on self-esteem and depression and the health-protective effects of employment. Further longitudinal study is required to see which of these explanations is the more accurate.

Higher levels of education were associated with lower levels of disability in the study by Eberhardt and Fex (1995), a well-documented finding on disability in rheumatoid arthritis also (Pincus and Callahan 1985). One interpretation of this finding is that higher levels of education are associated with more 'white collar' work and higher incomes, and as a result better work and social conditions. Improved economic resources and social conditions may lead to reduced disablement. Education may therefore be a marker of the importance of social factors on the course of disability in rheumatoid arthritis.

Social well-being

Employment and work

Rheumatoid arthritis has a dramatic impact on employment, with over half of individuals with rheumatoid arthritis who worked before the onset of the disease stopping work within 10 years of diagnosis (Yelin *et al*. 1980). Workforce participation of individuals with arthritis and limitations of movement is 20 to 33 per cent lower than that of individuals without arthritis. Participation in the workforce appears to be declining for arthritic men in the United States, in particular for those between the ages of 55 and 64 years, where from the 1970s to 1980s participation declined from 54 per cent to 38 per cent (Yelin 1992). In contrast, younger women with arthritis and limitation have increased their participation in the workforce. While both these trends are also apparent in the general population, the decline in men's participation is accentuated for those with arthritis and limitations.

Work performance and occupational status are affected by rheumatoid arthritis even soon after it has been diagnosed. Doeglas *et al*. (1995b) examined a group of patients in Holland who had a mean time since diagnosis of 2 years; they found that 90 per cent had either given up work or had made changes in their work because of their rheumatoid arthritis.

The decline in workforce participation of individuals with arthritis has a considerable impact on their income. In one study in the United States, men with rheumatoid arthritis had 48 per cent and women 27 per cent of the income of those without the disease (Mitchell *et al*. 1988). The nature of work appears to be an important

predictor of the likelihood of retaining employment. Those in professional or managerial occupations are more likely to stay in employment after the onset of rheumatoid arthritis (Callahan *et al*. 1992). The absence of paid employment for individuals with rheumatoid arthritis appears to have an independent effect on symptoms and mood. Fifield *et al*. (1991) assessed 723 patients attending clinics for rheumatoid arthritis and found that those without paid employment reported higher levels of pain and depression, even when confounding factors such as disease severity were controlled.

Whether individuals with rheumatoid arthritis will continue to work appears to be best predicted by the social characteristics of work. Yelin *et al*. (1980) found that the four best variables measuring the social characteristics of work, which included being able to control the pace of work and being self-employed, had greater explanatory power than the four best medical items (1.8 times greater) in predicting continuing work. The most important issue appears to be the flexibility of, and control over, work enabling individuals with rheumatoid arthritis to control the pace and timing of their activities at work. Doeglas *et al*. (1995b) found an association between educational level and remaining in employment despite rheumatoid arthritis. They attribute this to the nature of work performed by the more highly educated. In their study 32 per cent on the non-manual workers as opposed to 80 per cent of the manual workers with rheumatoid arthritis had left their jobs.

The problem of employment for individuals who have juvenile rheumatoid arthritis revolves around making the transition between the relatively sheltered world of education to the world of work. Programmes have been devised to assist this transition. One report of a programme in Cincinnati indicates a high level of training and education as well as high levels of employment for individuals with juvenile rheumatoid arthritis (White and Shear 1992). In the sample of 242, the majority of whom had juvenile rheumatoid arthritis, 27 per cent had completed 4 years of college education, which was well above that achieved by the general population (9 per cent), and only 6 per cent were unemployed.

The fulfilling of domestic roles and activities is particularly pertinent in considering a disease such as rheumatoid arthritis with its preponderance in women. It is not surprising that the many women with rheumatoid arthritis experienced limitations in cleaning the house (73 per cent), in laundry work (65 per cent), and shopping (61 per cent) (Reisine *et al*. 1987). More importantly, rheumatoid arthritis appears to affect the ability to perform the nurturing roles such as giving attention and support to other household members and maintaining family ties with others outside the household such as relatives and friends. (Allaire *et al*. 1991) examined the impact of rheumatoid arthritis and family and personal factors on work disability in a study that included a comparison group. Women with mild rheumatoid arthritis spent as much time on household work than the control group, but they appeared to accomplish less in that time. In contrast, women with more severe rheumatoid arthritis (Health assessment questionnaire score of over 1) did less housework and spent less time on housework than those in the mild and non-rheumatoid arthritis groups, but this reduction was compensated for by their families, who spent on average 7 h more time per week on household chores.

Family relationships

Rheumatoid arthritis

Studies on the frequency of divorce in arthritis are inconclusive (Anderson *et al*. 1985; Fitzpatrick 1993). Early studies suggested an increased frequency of divorce in rheumatoid arthritis but later studies have tended to find similar rates of divorce to those of the general population (Hull 1988; Hawley *et al*. 1991).

It is clear, however, that difficulties in relationships do arise when one partner has arthritis. In some cases these appear to result from a decreased interest in sexual relations (Reisine *et al*. 1987). Other issues that appear to create some difficulties in relationships include problems of dependency, shared problems of isolation, and reduced income (Locker 1983; Williams 1987). Specific symptoms may also lead indirectly to problems within relationships. Individuals with pain have increased levels of irritability and bad temper. These may in turn put increased pressure on their relationship.

There is, however, little apparent loss of the quality of close relationships when assessed from the perspective of the person with rheumatoid arthritis. Fitzpatrick *et al*. (1988) found little or no difference in how individuals with rheumatoid arthritis rate the quality of their close relationship. They also found no reduction in the quality of the relationship in those who had more severe disease.

Social relationships

Social relationships appear to be more at risk than close relationships in arthritis. Individuals with rheumatoid arthritis reduce their social contacts with others (Deyo *et al*. 1982). Over time as disability increases, individuals with arthritis have fewer opportunities for contact with friends and acquaintances and also derive less satisfaction with these relationships (Fitzpatrick *et al*. 1988). With further progression of the disease and reduced mobility, social contacts become even more difficult and social withdrawal may arise in some individuals. The social withdrawal may not, however, simply be attributable to increasing disability. Some individuals may withdraw socially in order to avoid the stigma and embarrassment associated with rheumatoid arthritis (Williams and Wood 1988 ; Locker 1989).

Psychological well-being

Prevalence of depression

The reported prevalence of clinical depression in rheumatoid arthritis ranges between 8 and 22 per cent in different studies (Creed 1990). While these proportions are higher than those found in the general population, similar rates were found among individuals with other chronic illnesses (Frank *et al*. 1988; Murphy *et al*. 1988; DeVellis 1993). Thus the increased levels of depression that occur in rheumatoid arthritis are not specific to that disease but appear to reflect the general effects of a chronic illness.

Hawley and Wolfe (1993) compared the incidence of depression and depressed mood in a large group of individuals with a variety of rheumatic conditions. They found no evidence to support the notion that patients with rheumatoid arthritis, in comparison to individuals with other rheumatological conditions, have a particular propensity to develop depression (Hawley and Wolfe 1993).

Factors associated with depression and low mood

A number of studies have attempted to identify the factors that lead to depression. While a simple interpretation might be that those with more severe disease are more likely to be depressed or have low mood, no direct relation between markers of the severity of arthritis and clinical depression or depressed mood has been found (McFarlane and Brooks 1988; Newman *et al*. 1989; Creed 1990; Blalock and DeVellis 1992). The most important predictor of depression and depressed mood in adults with rheumatoid arthritis and in juvenile arthritis appears to be disability (Newman *et al*. 1989; David *et al*. 1994). The research therefore suggests that two of the consequences of arthritis, depression and disability, are associated with each other but that neither is linked with clinical markers of the severity or activity of the disease. These associations serve to emphasize that how individuals respond to and experience their arthritis may be very different from what would be expected by trying to interpret the personal consequences from clinical measures of the disease.

In studies of individuals with arthritis, pain is commonly associated with depression and/or distress (e.g. Hawley and Wolfe 1988; Smedstad *et al*. 1995). It is commonly thought that the direction of this association is of pain causing the mental distress, although it could well be the opposite, with individuals who are distressed reporting higher levels of pain. Smedstad *et al*. (1995) attempted to distinguish between these two alternatives and concluded that mental distress is secondary to pain.

The impact of rheumatoid arthritis and its symptoms is apparently mediated through a number of psychological and social factors. Notable amongst these are coping responses and social supports, which are discussed below, but other important factors include social isolation and economic resources (Newman *et al*. 1989).

An important consideration is how low mood or depression may affect other aspects of the quality of life. Depressed mood is generally associated with dramatically reduced perceptions of quality of life (Sensky and Catalan 1992) and individuals with high levels of depression may restrict their social activities and become isolated. High levels of depression may also affect the individual's perceptions of their functioning and thus their self-reports of their functional status may be worse (Spiegel *et al*. 1988; Brooks *et al*. 1990; Guccione *et al*. 1995). These interactions between mood and other measures underlie the importance of studies that are conducted longitudinally and can examine causal direction between variables.

Impact of psychological variables on health status and well-being

A simple model of arthritic disease and its consequences would suggest a linear relation leading from disease to disablement and psychological well-being. The research findings suggest that this simple model is untenable and that the consequences of arthritis are dependent upon how individuals interpret their illness, how others respond to their needs, and aspects of their environment. A number of psychological concepts have been used in attempting to account for the mediation between disease and its consequences, and it is to these that we now turn. In order to place the psychological concepts into context a schematic model is presented in Fig. 1. This model illustrates the mediating position occupied by psychological and social factors in our understanding of the consequences of arthritis. The research relating to some of these central concepts is discussed below.

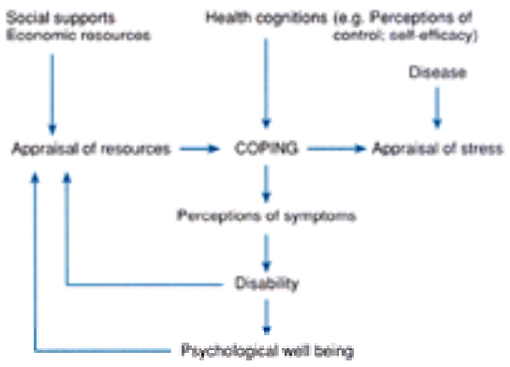


Fig. 1 How psychological and social variables influence outcome in rheumatic disease.

Coping

The term coping is used to refer to what individuals attempt to do to limit the impact of a stressor such as arthritis. Research on the way in which individuals cope with rheumatoid arthritis has been growing in recent years (see Newman and Revenson 1993). This body of work may be divided into studies that have examined coping with rheumatoid arthritis in general and those that have focused on coping with specific symptoms of rheumatoid arthritis, such as pain.

In studies that have looked at coping with arthritis in general, the most popular inventory used to assess coping in rheumatoid arthritis is the Ways of Coping Scale (Folkman and Lazarus 1980) and the Ways of Coping Scale—revised (Folkman and Lazarus 1988), although the questionnaire is frequently revised for the purposes of assessing arthritis. The questionnaire has eight subscales divided between two broad classifications of coping: problem-focused coping (e.g. 'Come up with a couple of solutions to the problem') and emotion-focused coping (e.g. 'I am changing or growing as a person in a good way'). Three studies have used variants of the Ways of Coping Scale and found that patients with rheumatoid arthritis who use the coping strategy of cognitive restructuring (broadly defined as 'attempts to seek new meaning in their situation') had better psychological well-being (Felton *et al*. 1984; Parker *et al*. 1988b; Manne and Zautra 1989). In addition, in one study, seeking information about the illness was also associated with improved psychological well-being (Manne and Zautra 1989). The importance of information seeking should be viewed against the background of the increasing popularization of medical knowledge coupled with the easy access to information about medical issues. Many individuals with arthritis have some knowledge of arthritis before they are diagnosed and/or become knowledgeable early in their illness. The value of information seeking by someone who is knowledgeable may therefore be reduced for some individuals with arthritis.

In these studies (Felton *et al*. 1984; Parker *et al*. 1988b; Manne and Zautra 1989) those who used the coping strategy of wishful thinking (e.g. 'Wish that the situation would go away or somehow be over with') tended to have lower psychological well-being. These studies all addressed arthritis in general. Long and Sangster (1993), rather than asking patients how they coped with arthritis in general, asked them how they coped with the most stressful event they had faced over the past month. Thus, the stressor was defined by the patient, and was not necessarily related to the arthritis. As in the studies mentioned above, wishful thinking was associated with poorer adaptation.

Using a questionnaire designed specifically to assess coping with rheumatoid arthritis (the London coping with rheumatoid arthritis questionnaire), Newman *et al*. (1990) used a cluster analysis to divide 158 patients into distinct groups according to how they attempted to cope with their disease. The groups defined in this way did not differ on clinical measures of disease. One group, which tended to be the most open and active in the manner in which they attempted to deal with the stresses of arthritis, had significantly lower scores on measures of pain, stiffness, and disability, as well as higher levels of psychological well-being. The findings suggest that how individuals attempt to deal with their arthritis, in terms of actions and thoughts, has an impact on psychological well-being, symptom reporting, and disability (see Fig. 1).

Some studies have examined the ways in which individuals with arthritis cope with the dominant symptom of the disease, pain. A number of measures of coping with pain have been developed and those used in arthritis include the Coping Strategies Questionnaire (Rosenstiel and Keefe 1983) and the Vanderbilt pain management inventory (Brown and Nicassio 1987), and the Vanderbilt multidimensional pain coping inventory (Smith and Wallston 1993).

Specific coping strategies appear to influence the impact of pain on psychological well-being. Perceptions of control over the pain lead to higher levels of psychological well-being and lower levels of disability (Beckham *et al*. 1991). Keefe *et al*. (1989) found that high scores on the 'catastrophising scale' of the Coping Strategies Questionnaire were associated with poorer outcome on measures of disability and psychological well-being 6 months later. The Vanderbilt pain management inventory has its origins in research on chronic pain and was designed to assess both thoughts and behavioural strategies for coping with pain. Longitudinal research with this instrument has demonstrated that active coping strategies lead to more favourable outcomes and passive strategies unfavourable outcomes; passive coping during periods of high pain may be particularly detrimental to psychological well-being (Brown and Nicassio 1987; Brown *et al*. 1989).

Taken together these studies suggest that the psychological responses to arthritis and its symptoms are important determinants of resulting health status. Both the disease and the psychological processes of trying to deal with arthritis are not static. This implies an approach to understanding the consequences of arthritis which accepts that both the disease and the individual's attempts to cope with the stresses of the disease change over time. The research in this area also opens up the possibility that psychological interventions to alter coping behaviour may be able to reduce the impact of arthritis.

Social comparison

Individuals' attempts to cope with the stresses of their arthritis take place in a social context. Part of the processes associated with coping involves making comparisons with others in order to provide a reference point to evaluate one's relative position. Social comparison constitutes a part of the coping mechanism used by some individuals. It is particularly important when the means to change the situation are not available (Wills 1987). In the context of arthritis, individuals compare themselves with others who have arthritis. This may involve a comparison with others in a worse state and in this manner they are able to judge their state as being 'relatively' good; the net effect of this comparison is to preserve their psychological well-being. The making of social comparisons has been studied in arthritis. The findings suggest that individuals with rheumatoid arthritis do predominantly tend to make comparisons with those in a worse state and that this type of comparison is associated with better psychological well-being (Affleck *et al*. 1988; DeVellis *et al*. 1990).

Social support

The process of dealing with the stresses caused by arthritis involves others in the social world not only as points of comparison but also as potential providers of both practical help and emotional support. The term 'social support' refers to the process by which interpersonal relationships promote well-being and protect people from a decline in health. In healthy individuals, social support reduces morbidity and mortality (Berkman 1985), particularly at times when they are facing stressful life circumstances (Cohen 1988), and improves mental health (Kessler and McLeod 1985).

Eliciting social support is one coping strategy that many individuals use. This is dependent upon having people in the environment from whom support may be elicited. In most societies there are natural groups and individuals who provide social support. These include family, friends, health-care professionals, and patients' groups. These individuals and groups may provide emotional support by offering a receptive ear when individuals are upset as well as practical support such as assistance

and advice about how to deal with a problem.

A large number of studies show the importance of social support for the psychological well-being of individuals with arthritis. At the simplest level, married people with rheumatoid arthritis show slower progression in their disability than those without partners (Ward and Leigh 1993). An interpretation of these findings is that a partner who provides emotional and practical support to those with rheumatoid arthritis acts to protect them against resulting disability.

Other studies have shown that individuals with rheumatoid arthritis who have more social contacts have lower levels of depression (Newman *et al*. 1989). Most studies of social support, however, suggest that the number of individuals with whom one has contact is less important than the belief that one has adequate support available. These studies have demonstrated that those individuals with rheumatoid arthritis who perceive greater social support exhibit greater self-esteem (Fitzpatrick *et al*. 1988) and life satisfaction (Smith *et al*. 1991), are better adjusted (Affleck *et al*. 1988), show less depression (Fitzpatrick *et al*. 1991; Doeglas *et al*. 1994), and cope better with their arthritis (Manne and Zautra 1990). The important role of the individuals' judgement of their social support parallel findings in both epidemiological studies and studies in other illnesses. It further emphasizes the mediating role of psychological factors in determining the impact of processes in the world on an individual.

It is important to recognize that the provision of support implies shared understandings of need, what is required, when it is required, and an ability to negotiate the provision of support at a time when the provider is ready to offer it (Newman *et al*. 1996). This process takes place within a reciprocal relationship in which the individual with rheumatoid arthritis is also a support provider. Given these factors it is not surprising that some social support may have negative effects (Revenson *et al*. 1991). In some cases the negative effects may be intended. The relation between negative social support and coping efficacy was well demonstrated in a study by Manne and Zautra (1989), who examined the impact of husbands' criticism on their wives' coping and found a direct relation between spouses' critical comments and wives' coping and adaptation to their rheumatoid arthritis. Patients with critical spouses tended to engage in more wishful thinking, which led to poorer psychological outcomes. The authors argue that spouses' criticism 'may encourage ineffective and even harmful coping strategies'. Patients who perceived their spouses as supportive engaged more in cognitive restructuring and information seeking, strategies that proved to be more adaptive in dealing with the stresses of arthritis. (Kraaimaat *et al*. 1995) report similar findings but found that spousal criticism led to anxiety in men and both anxiety and depression in women.

Some researchers have looked at the impact of different types of social support and found that not all elements of support lead to improved health status or psychological well-being in individuals with arthritis. Taal *et al*. (1993) found that while emotional support was not positively related to health status, practical support was. Doeglas *et al*. (1994) found that receiving greater amounts of emotional support around problems had a negative relation to some aspects of psychological well-being. These findings suggest that to understand fully the impact of social support in arthritis further research needs to examine the nature of the support provided and who provides it. In addition it is likely that the timing of the support is important as well as whether it is requested.

Control, self-efficacy, and perceived competence

Aside from specific attempts to deal with the stresses of arthritis and the important role that others in the environment can play through social support, individuals hold more general beliefs about their health and what they are able to do about the course of their arthritis (see Fig. 1). Some of these more general beliefs are discussed next.

Beliefs about whether and who may influence the course of health and illness in the future have been widely studied. This concept, known as health locus of control, has three dimensions. The internal health locus of control (**IHLC**) is the belief that the individual has an influence over future health and illness; the chance health locus of control (CHLC) is the belief that fate and chance will determine future health and illness, and 'powerful other' health locus of control (POHLC) is the belief that others such as health-care professionals will wield an important influence on future health and illness (Wallston *et al*. 1978). This concept usefully identifies that individuals with arthritis who join self-help groups have higher IHLC scores (Volle *et al*. 1990). While this is not a surprising finding it raises the important methodological issue that studies of self-help groups will not necessarily be representative of all individuals with arthritis.

Early cross-sectional research suggested that beliefs that one had control over future illness and health (IHLC) had positive effects on psychological well-being in rheumatoid and osteoarthritis (Wallston 1993). Later studies have shown the issue about control to be more complex. A study that distinguished between control over treatment and control over symptoms found that personal control over treatment was associated with higher psychological well-being while perceptions of greater control over symptoms by health-care providers were associated with lower psychological well-being (Affleck *et al*. 1987). A later study by the same group (Tennen *et al*. 1992) showed that patients who believed they had control over their pain (IHLC) at the beginning of an intensive daily study reported less pain. More interestingly, it appeared that when individuals believed they could control their pain but were thwarted by higher levels of pain than expected, they become distressed. These studies clearly demonstrate that perceptions of personal control over arthritis do not necessarily lead to improved outcomes.

Perceived competence and self-efficacy are terms used to describe individuals' beliefs that they have the skills to be able to deal effectively with issues in the environment. It has been argued that these beliefs act as a mediator between measures of disease and adaptation. One study (Smith *et al*. 1991) found that, in cross-sectional and longitudinal analyses, levels of competence did act as a mediator between levels of disease, perceptions of control and social support on the one hand, and life satisfaction and depression on the other.

Intervention studies

Studies on psychological factors have demonstrated how they mediate between the disease on the one hand and outcome on the other. These investigations have led to the development of intervention studies, which attempt to modify individuals' understanding, beliefs, coping styles, and social supports in order to influence psychological well-being and health status. These studies customarily involve a range of factors, which are discussed next.

Information

Information about arthritis is a small but significant component that appears integral to most intervention studies. Lorig *et al*. (1987) reviewed the literature in 1987 and reported on 34 educational studies, of which 32 (94 per cent) found increases in knowledge among the participants. The difficulty has been to observe a translation of knowledge into behaviour. There is little evidence to support the view that the provision of general information alone has any benefits on individuals with arthritis (DeVellis and Blalock 1993).

Patient education

Most patient education programmes have a number of different components. Most include skills taught by other health-care professionals such as physiotherapists and nurses, with the aim of developing self-management skills. Most well-known amongst this form of intervention in arthritis is the arthritis self-management programme (**ASMP**) developed by Lorig and colleagues (Lorig 1986). The ASMP uses 'lay leaders' in groups of patients with arthritis to deliver a programme in which education and skills training are core components. Although the effects of this intervention have not always been consistent, early research suggested that the programme is effective in reducing levels of pain and depression and increasing physical activity (DeVellis and Blalock 1993; Lorig *et al*. 1993). These benefits were apparent 4 years after the intervention. Most striking was the research which suggested that these changes were not the result of specific self-management activity but rather individuals' beliefs (self-efficacy) that they could cope with the consequences of arthritis.

Of importance are the potential economic benefits of such a programme. The ASMP reduced the number of visits to the physician, and Lorig *et al*. (1993) estimated the 4-year savings of the programme to be US$701.68 per patient with rheumatoid arthritis and US$189.24 per patient with osteoarthritis. These savings, if projected to the entire population of patients with these arthritides in the United States, would result in savings of over US$13 billion for rheumatoid arthritis and over US$19 billion for osteoarthritis.

Another investigation of patient education has also produced some positive results. Lindroth *et al*. (1995) made a 5-year follow-up study on a control group and an intervention group of patients with rheumatoid and osteoarthritis. The intervention group had received six sessions of 2.5 h of education that focused on medical and treatment aspects, pain management, stress management, self-awareness and communication skills, and exercise and work. Knowledge was higher in the intervention group and, in contrast to that group, the control group reported significant increased disability and pain. As with the ASMP, there were no differences in behaviour (exercise and joint protection) between the two groups at 5 years.

Social support interventions

Many of the psychological and educational interventions have ignored the social context and have been directed solely at the individual with rheumatoid arthritis. Despite this focus, by bringing individuals with arthritis into contact with others, social support may have been an influence on the outcome of the intervention (DeVellis and Blalock 1993). Other studies have used a social support intervention and directed their attention to the individual with rheumatoid or osteoarthritis (Shearn and Fireman 1985).

At the minimal level of intervention, Weinberger *et al*. (1986) and Weinberger *et al*. (1989) used a telephone contact with patients with osteoarthritis and found improved levels of functional status along with increased levels of social support. In a randomized control study they found that sustained telephone contact with a non-professional was more effective than interventions by physicians at outpatient appointments. In a more detailed analysis to determine what factors could account for the improved functional status, the investigators used telephone calls to discuss symptoms, provide advice, impart information, reduce non-compliance, and improve morale. They found no effects of the telephone contacts on social support, satisfaction with care, morale or compliance (Weinberger *et al*. 1991). They suggest that this may be due to the telephone calls having a large advice component, which may have led to them being perceived as an extension of medical treatment.

Conclusion

Research on the psychological factors associated with arthritis is still in its infancy. The work has shown that attempts to predict disability and psychological well-being in individuals with arthritis need to incorporate how individuals perceive their arthritis, how they try to deal with the stresses it creates, as well as factors in their social environments. This area of research and its resulting interventions are likely to lead to important findings in the future.

Chapter References

Affleck, G., Tennen, H., Pfeiffer, C., and Fifield, J. (1987). Appraisals of control and predictability in adapting to a chronic disease. *Journal of Personality and Social Psychology*, **53**, 273–9.

Affleck, G., Pfeiffer, C., Tennen, H., and Fifield, J. (1988). Social support and psychosocial adjustment to rheumatoid arthritis. *Arthritis Care and Research*, **1**, 71–7.

Allaire, S. H., Meenan, R. F., and Anderson, J. F. (1991). The impact of rheumatoid arthritis on the household work performance of women. *Arthritis and Rheumatism*, **34**, 669–78.

Anderson, K., Bradley, L., Young, L., McDaniel, L., and Wise, C. (1985). Rheumatoid arthritis: review of psychological factors related to etiology, effects and treatment. *Psychological Bulletin*, **98**, 358–7.

Badley, E. (1992). The impact of musculoskeletal disorders on the Canadian population. *Journal of Rheumatology*, **19**, 337–40.

Badley, E. M. and Tennant, A. (1993). Impact of disablement due to rheumatic disorders in a British population: estimates of severity and prevalence from the Calderdale Rheumatic Disablement Survey. *Annals of the Rheumatic Diseases*, **52**, 6–13.

Basia, L. (1995). Comparison of self reported fatigue in rheumatoid arthritis and controls. *Journal of Rheumatology*, **22**, 639–43.

Beckham, J. C. *et al*. (1991). Pain coping strategies in rheumatoid arthritis. *Behaviour Therapy*, **22**, 113–24.

Bendtsen, P. and Hornquist, J. O. (1993). Severity of rheumatoid arthritis, function and quality of life: subgroup comparisons. *Clinical and Experimental Rheumatology*, **11**, 495–502.

Berkman, L. F. (1985). The relationship of social networks and social support to morbidity and mortality. In *Social support and health* (ed. S. Cohen and S. L. Syme), pp. 243–62. Academic Press, New York.

Blalock, S. J. and DeVellis, R. F. (1992). Rheumatoid arthritis and depression: an overview. *Bulletin of the Rheumatic Diseases*, **41**, 6–8.

Brooks, B., Jordan, J., Divinew, G., Smith, K., and Neelon, F. (1990). The impact of psychologic factors on measurement of functional status. *Medical Care*, **28**, 793–804.

Brown, G. K. and Nicassio, P. M. (1987). Development of a questionnaire for the assessment of active and passive coping strategies in chronic pain patients. *Pain*, **31**, 53–63.

Brown, G. K. *et al*. (1989). Pain coping strategies and depression in rheumatoid arthritis. *Journal of Consultative Clinical Psychology*, **57**, 652–7.

Buescher, K. L. *et al*. (1991). Relationship of self efficacy to pain behaviour. *Journal of Rheumatology*, **18**, 968–72.

Burckhardt, C. S., Archenholtz, B., and Bjelle A. (1993). Quality of life of women with systemic lupus erythematosus: a comparison with women with rheumatoid arthritis. *Journal of Rheumatology*, ***2***, 977–81.

Callahan, L., Bloch, D., and Pincus, T. (1992). Identification of work disability in rheumatoid arthritis. *Journal of Clinical Epidemiology*, **45**, 127–38.

Cohen, S. (1988). Psychosocial models of the role of social support in the etiology of physical disease. *Health Psychology*, **7**, 269–97.

Creed, F. (1990). Psychological disorders in rheumatoid arthritis: a growing consensus. *Annals of the Rheumatic Diseases*, **49**, 808–12.

David, J. *et al*. (1994). The functional and psychological outcomes of juvenile chronic arthritis in young adulthood. *British Journal of Rheumatology*, **33**, 876–81.

DeVellis, B. M. (1993). Depression in rheumatological diseases. In *Psychological aspects of rheumatic*disease, Vol. 7 (2), *Baillière's clinical rheumatology* (ed. S. Newman and M. Shipley), pp. 241–58. Baillière Tindall, London.

DeVellis, R. and Blalock, S. (1993). Psychological and educational interventions to reduce arthritis disability. In *Psychological aspects of rheumatic disease*, Vol. 7 (2), *Baillière's clinical rheumatology* (ed. S. Newman and M. Shipley), pp. 397–416. Baillière Tindall, London.

DeVellis, R., Holt, K., and Renner, B. (1990). The relationship of social comparison to rheumatoid arthritis symptoms and affect. *Basic and Applied Social Psychology*, **11**,1–18.

Deyo, R. A. (1988). Measuring the quality of life of patients with rheumatoid arthritis. In *Quality of life: assessment and application* (ed. S. R. Walker and R. M. Rosser), p. 205. MTP, Lancaster.

Deyo, R. A., Inui, T., Leininger, J., and Overman, S. (1982). Physical and psychosocial function in rheumatoid arthritis. *Archives of Internal Medicine*, **142**, 879–82.

Doeglas, D. *et al*. (1994). Social support, social disability and psychological well being in rheumatoid arthritis. *Arthritis Care and Research*, **7**, 10–15.

Doeglas, D. *et al*. (1995*a*). The assessment of functional status in rheumatoid arthritis: a cross cultural, longitudinal comparison of the Health Assessment Questionnaire and the Groningen Activity Restriction Scale. *Journal of Rheumatology*, **22**, 1834–43.

Doeglas, D. *et al*. (1995*b*). Work disability in early rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **54**, 455–60

Doherty, E., Yanni, G., Conroy, R. M., and Bresnihan, B. (1993). A comparison of child and parent ratings of disability and pain in juvenile chronic arthritis. *Journal of Rheumatology*, **20**, 1563–6.

Eberhardt, K. and Fex, E. (1995). Functional impairment and disability in early rheumatoid arthritis—development over 5 years. *Journal of Rheumatology*, **22**, 1037–42.

Felton, B. *et al*. (1984). Stress and coping in the explanation of psychological adjustment among chronically ill adults. *Social Science and Medicine*, **18**, 889–98.

Fifield, J., Reisine, S. T., and Grady, K. (1991). Work disability and the experience of pain and depression in rheumatoid arthritis. *Social Science and Medicine*, **33**, 579–85.

Fitzpatrick, R. (1993). The measurement of health status and quality of life in rheumatoid arthritis. In *Psychological aspects of rheumatic disease*, Vol. 7 (2), *Baillière's clinical rheumatology* (ed. S. Newman and M. Shipley), pp. 297–317. Baillière Tindall, London.

Fitzpatrick, R., Newman, S., Lamb, R., and Shipley, M. (1988). Social relationships and psychological well-being in rheumatoid arthritis. *Social Science and Medicine*, **27**, 399–403.

Fitzpatrick, R., Newman, S., Archer, R., and Shipley, M. (1991). Social support, disability and depression: a longitudinal study of RA. *Social Science and Medicine*, **33**, 605–11.

Fitzpatrick, R., Ziebland, S., Jenkinson, C., and Mowat, A. (1992). A generic health status instrument in the assessment of rheumatoid arthritis. *British Journal of Rheumatology*, **31**, 87–90.

Folkman, S. and Lazarus, R. (1980). An analysis of coping in a middle-aged community sample. *Journal of Health and Social Behaviour*, **21**, 219–39.

Folkman, S. and. Lazarus, R. S. (1988). *Manual for the ways of coping questionnaire*. Consulting Psychologists Press, Palo Alto, CA.

Frank, R. *et al*. (1988). Depression in rheumatoid arthritis. *Journal of Rheumatology*, **15**, 920–5.

Gibson, T. and Clark, B. (1985). Use of simple analgesics in rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **44**, 27–9.

Guccione, A. *et al*. (1995). The correlates of health perceptions in rheumatoid arthritis. *Journal of Rheumatology*, **22**, 432–9.

Hawley, D. J. and Wolfe, F. (1988). Anxiety and depression in patients with rheumatoid arthritis. *Journal of Rheumatology*, **15**,932–41.

Hawley, D. J. and Wolfe, F. (1993). Depression is not more common in rheumatoid arthritis: a 10-year longitudinal study of 6 153 patients with rheumatic disease. *Journal of Rheumatology*, **20**, 2025–31.

Hawley, D. J. *et al*. (1991). Marital status in rheumatoid arthritis and other rheumatic disorders: a study of 7,293 patients. *Journal of Rheumatology*, **18**, 654–60.

Hull, R. G. (1988). Outcome in juvenile arthritis. *British Journal of Rheumatology*, **27**, 66–71.

Keefe, F. J. *et al*. (1989). Coping with rheumatoid arthritis pain: catastrophising as a maladaptive strategy. *Pain*, **37**, 51–6.

Kessler, R. C. and McLeod, J. D. (1985). Social support and mental health in community samples. In *Social support and health* (ed. S. Cohen and S. L. Syme), pp. 219–40. Academic Press, New York.

Kraaimaat, F. *et al*. (1995). Association of social support and the spouse's reaction with psychological distress in male and female patients with rheumatoid arthritis. *Journal of Rheumatology*, **22**, 644–8.

Leigh, P. J. and Fries, J. F. (1992). Predictors of disability in a longitudinal sample of patients with rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **51**, 581–7.

Lindroth, Y. *et al*. (1995). A 5-year follow-up of a controlled trial of an arthritis education programme. *British Journal of Rheumatology*, **34**, 647–52.

Locker, D. (1983). *Disability and disadvantage: the consequences of chronic illness*. Tavistock, London.

Locker, D. (1989). Coping with disability and handicap. In *Disablement in the community* (ed. D. Patrick, and H. Peach), pp. 176–96. Oxford University Press.

Long, B. C. and Sangster, J. L. (1993). Dispositional optimism/pessimism and coping strategies: predictors of psychosocial adjustment of rheumatoid and osteoarthritis patients. *Journal of Applied Social Psychology*, **23**, 1069–91.

Lorig, K. (1986). Development and dissemination of an arthritis patient education course. *Family and Community Health*, **9**,23–32.

Lorig, K., Konkol, L., and Gonzolez, V. (1987). Arthritis patient education: a review of the literature. *Patient Education and Counselling*, **10**, 207–52.

Lorig, K., Mazonen, P., and Holman, H. (1993). Evidence suggesting that health education for self-management in patients with chronic arthritis has sustained health benefits while reducing health care costs. *Arthritis and Rheumatism*, **36**, 439–46.

Lorish, C. D., Abraham, N., and Austin, J. E. A. (1991). Disease and psychosocial factors related to physical functioning in rheumatoid arthritis. *Journal of Rheumatology*, **18**, 1150–7.

McFarlane, A. C. and Brooks, P. (1988). An analysis of the relationship between psychological morbidity and disease activity in rheumatoid arthritis. *Journal of Rheumatology*, **15**, 926–31.

McKinley, P., Ouellette, S. C., and Winkel, G. (1995). The contributions of disease activity, sleep patterns, and depression to fatigue in systemic lupus erythematosus. *Arthritis and Rheumatism*, **38**, 826–34.

Manne, S. and Zautra, A. (1989). Spouse criticism and support: their association with coping and psychological adjustment among women with rheumatoid arthritis. *Journal of Personal and Social Psychology*, **56**, 608–17.

Manne, S. and Zautra, A. (1990). Couples coping with chronic illness: women with rheumatoid arthritis and their healthy husbands. *Journal of Behavioural Medicine*, **13**, 327–43.

Martin, J., Meltzer, H., and Elliott, D. (1988). *The prevalence of disability among adults*. HMSO, London

Melzack, R. (1975). The McGill Pain Questionnaire—major properties and scoring methods. *Pain*, **1**, 277–99.

Mitchell, J., Burkhauser, R., and Pincus, T. (1988). The importance of age, education and comorbidity in the substantial earnings losses of individuals with symmetric polyarthritis. *Arthritis and Rheumatism*, **31**, 348–57.

Murphy, S., *et al*. (1988). Psychiatric disorder and illness behaviour in rheumatoid arthritis. *British Journal of Rheumatology*, **27**, 357–63.

Newman, S. and Revenson, T. A. (1993). Coping with rheumatoid arthritis. In *Psychological aspects of rheumatic disease*, Vol. 7 (2), *Baillière's Clinical Rheumatology* (ed. S. Newman and M. Shipley), pp. 259–80. Baillière Tindall, London.

Newman, S. P., Fitzpatrick, R., Lamb, R., and Shipley, M. (1989). The origins of depressed mood in rheumatoid arthritis. *Journal of Rheumatology*, **16**, 740–4.

Newman, S. *et al*. (1990). Patterns of coping in rheumatoid arthritis. *Psychology and Health*, **4**, 187–200.

Newman, S., Fitzpatrick, R., Revenson, T. A., Skevington, S., and Williams, G. (1996). *Understanding rheumatoid arthritis*. Routledge, London.

Papageorgiou, A. (1994). Fatigue in rheumatoid arthritis. Unpublished M.Sc. Dissertation, Faculty of Medicine, University of Manchester.

Papageorgiou, A. C. and Badley, E. M. (1989). The quality of pain in arthritis: the words patients use to describe overall pain and pain in individual joints at rest and on movement. *Journal of Rheumatology*, **16**, 106.

Parker, J. *et al*. (1988*a*). Pain in rheumatoid arthritis: relationship to demographic medical and psychological factors. *Journal of Rheumatology*, **15**, 433.

Parker, J. *et al*. (1988*b*). Coping strategies in rheumatoid arthritis. *Journal of Rheumatology*, **15**, 1376–83.

Pincus, T. and Callahan, L. F. C. (1985). Formal education as a marker for increased mortality and morbidity in rheumatoid arthritis. *Journal of Chronic Diseases*, **38**, 973–84.

Reisine, S., Goodenow, C., and Grady, K. (1987). The impact of rheumatoid arthritis on the homemaker. *Social Science and Medicine*, **25**, 89–96.

Revenson, T. A., Schiaffino, K. M., Majerovitz, S. D., and Gibofsky, A. (1991). Social support as a double-edged sword: the relation of positive and problematic support to depression among rheumatoid arthritis patients. *Social Science and Medicine*, **33**, 801–13.

Rhind, V. M., Bird, H. A., and Wright, V. (1980). Comparison of clinical assessments of disease activity in rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **39**, 135–7.

Rosenstiel, A. K. and Keefe, F. J. (1983). The use of coping strategies in chronic low back pain patients: relationship to patients characteristics and current adjustment. *Pain*, **17**, 33–40.

Sensky, T. and Catalan, J. (1992). Asking patients about their treatment. *British Medical Journal*, **305**, 1109–10.

Shearn, M. A. and Fireman, B. H. (1985). Stress management and mutual support groups in rheumatoid arthritis. *American Journal of Medicine*, **78**, 771–5.

Shortall, E., Isenberg, D., and Newman, S. P. (1995). Factors associated with mood and mood disorders in SLE. *Lupus*, **4**, 272–9.

Smedstad, L. *et al*. (1995). The relationship between pain and sociodemographic variables, anxiety, and depressive symptoms in rheumatoid arthritis. *Journal of Rheumatology*, **22**, 514–20.

Smith, C. A. and Wallston, K. (1993). Validation of a multidimensional pain coping inventory. (*Proceedings*) 28th National Scientific Meeting of the Arthritis Health Professions Association, San Antonio, CA.

Smith, C. A., Dobbins, C. J., and Wallston, K. A. (1991). The mediational role of perceived competence in psychological adjustment to rheumatoid arthritis. *Journal of Applied Social Psychology*, **21**, 1218–47.

Spiegel, J. *et al*. (1988). What are we measuring? An examination of self-reported functional status measures. *Arthritis and Rheumatism*, **31**, 721–8.

Steinbrocker, O., Traeger, C., and Batterman, R. (1949). Therapeutic criteria in rheumatoid arthritis. *Journal of the American Medical Association*, **140**, 659–62.

Taal, E., Rasker, J. J., Seydel, E. R., Weigman, O. *et al*. (1993). Health status, adherence with health recommendations, self-efficacy and social support in patients with rheumatoid arthritis. Special Issue: psychosocial aspects of rheumatic diseases. *Patient Education and Counseling*, **20**, 63–76

Tack, B. (1990). Fatigue in rheumatoid arthritis. *Arthritis Care and Research*, **3**, 65–70.

Tennen, H. *et al*. (1992). Perceiving control, construing benefits, and daily processes in rheumatoid arthritis. *Canadian Journal of the Behavioral Sciences*, **24**, 186–203.

Thompson, P. W. and Pegley, F. S. (1991). A comparison of disability measured by the Stanford Health Assessment Questionnaire disability scales (HAQ) in male and female rheumatoid outpatients. *British Journal of Rheumatology*, **30**, 298–300.

Volle, B., Wiedenbusch, S., and Lohaus, A. (1990). Psychological correlates of self help group membership in patients with rheumatic diseases. *Psychotherapie, Psychosomatik Medizinische Psychologie*, **40**, 230–7.

Wagstaff, S., Smith, O. V, and Wood, P. H. N. (1985). Verbal pain descriptors used by patients with arthritis. *Annals of the Rheumatic Diseases*, **44**, 262–5.

Wallston, K. (1993). Psychological control and its impact in the management of rheumatological disorders. In *Psychological aspects of rheumatic disease*, Vol. 7 (2), *Baillière's Clinical Rheumatology* (ed. S. Newman and M. Shipley), pp. 281–95. Baillière Tindall, London.

Wallston, K. A., Wallston, B., and DeVellis, R. F. (1978). Development of the multidimensional health locus of control scale (MHLC). *Health Education Monographs*, **6**, 160–70.

Ward, M. and Leigh, P. (1993). Marital status and the progression of functional disability in patients with rheumatoid arthritis. *Arthritis and Rheumatism*, **36**, 581–92.

Weinberger, M., Hiner, S. L., and Tierney, W. M. (1986). Improving functional status in arthritis: the effect of social support. *Social Science and Medicine*, **23**, 899–904.

Weinberger, M. *et al*. (1989). Can the provision of information to patients with osteoarthritis improve functional status: a randomised controlled trial. *Arthritis and Rheumatism*, **32**,1577–83.

Weinberger, M. *et al*. (1991). The impact of increased contact on psychosocial outcomes in patients with osteoarthritis: a randomised controlled trial. *Journal of Rheumatology*, **18**, 849–54.

White, P. and Shear, E. S. (1992). Transition/job readiness for adolescents with juvenile arthritis and other chronic illness. *Journal of Rheumatology*, **19**, 23–7.

Williams, G. H. (1987). Disablement and the social context of daily activity. *International Disability Studies*, **9**, 97–102.

Williams, G. H. and Wood, P. H. N. (1988). Coming to terms with chronic illness: the negotiation of autonomy in rheumatoid arthritis. *International Disability Studies*, **10**, 128–32.

Wills, T. A. (1987). Downward comparison as a coping mechanism. In *Coping with negative life events* (ed. C. R. Snyder and C. E. Ford), pp. 243–68. Plenum, New York.

Wolfe, F. and Cathey, M. (1991). The assessment and prediction of functional disability in rheumatoid arthritis. *Journal of Rheumatology*, **18**, 1298–306.

Yelin, E. (1992). Arthritis: the cumulative impact of a common chronic condition. *Arthritis and Rheumatism*, **35**, 489–97.

Yelin, E. *et al*. (1980). Work disability in rheumatoid arthritis: effects of disease, social and work factors. *Annals of Internal Medicine*, **93**, 551–6.

1.1.7 Growth and skeletal maturation

Daniel J. Lovell and Patricia Woo

Introduction

Growth, like beauty, is a multifaceted concept varying greatly from person to person. For some, 'growth' is a continuum that involves changes in body size and form, physiological function, and biological maturation. Some of the changes in this continuum include embryogenesis, compensatory growth, and wound healing. For others, 'growth' is the process of increasing intellectual and emotional sophistication which for many is a lifelong pursuit but for all extends much beyond the traditional paediatric age. Growth, historically, has been an area of intense interest to paediatricians but only fairly recently have individuals involved in paediatrics and child health come to realize that growth is a basic science peculiar to their art (Falkner and Tanner 1986). This is a science that incorporates, among others, the disciplines of anatomy, physiology, biochemistry, and biology, and is no longer only observational. Molecular biology has provided the ability to synthesize a variety of proteins (e.g. growth hormone, insulin-like growth factors, releasing hormones) which can profoundly change human growth processes. These growth-modifying agents have stimulated an already intense interest in the physiology, psychology, ethics, and economics of poor growth.

In this chapter, an overview of concepts of early growth, growth and maturation measurement, analysis of growth data, growth rates in various developmental stages, and factors influencing growth (especially the role of chronic inflammation in growth) will be developed. More detailed information can be found in Falkner and Tanner (Falkner and Tanner 1986).

Prenatal concepts

The variation of mature body weights in different species of mammals is enormous. It is therefore remarkable that all species start life as a single cell of approximately the same size (1000 μm^3) and that at the time of implantation in the uterus the majority of the cells are directed towards supporting nutrient flow from the mother to the fetus and only a small number of cells are precursors for organogenesis (Reeds and Fiorotto 1990).

Two important, but related, concepts are necessary to understand the process of development—critical periods and canalization. The phenomenon of critical periods in development was first described by Stockard (Stockard 1921). Critical periods occur at times of rapid and important change in a system. A system may be an embryological cell, an organ system, or an entire animal. But most importantly, during this time of rapid change, a decision is made that cannot be repeated or reversed at a later time. The most striking examples come from embryology. There is only a brief period during development in which an individual can be influenced to develop more than the normal number of digits. No later interference can alter the number produced except by amputation. Moreover, many critical periods occur during development and as they proceed, the process places the organism in a deeper and deeper canal whose boundaries limit further choices—'canalization' (Waddington 1962).

A more detailed description of prenatal development is beyond the scope of this chapter but the reader is referred to Chapter 2.3, Chapter 2.4, and Chapter 2.6 for detailed discussions of the development of bone, cartilage, and muscle.

Measurement of growth

When seen from a distance with uninitiated eyes, the problem of measurement of physical growth seems relatively simple. When reviewing the literature of 'auxological anthropometry', that is the systemized art of measuring and taking observations of somatic growth in humans, one begins to develop the feeling that this discipline is plagued by the human equivalent of the Heisenberg Uncertainty Principle! In this chapter, a clinical approach will be presented that will alert the reader to commonly encountered problems in the measurement of growth in children. It is certainly true that the measurement of somatic growth involves a close, fine dance between the subject, instrument, and observer in an interactive situation that the observer is constantly attempting to standardize so as to measure the only important change—that of the subject. At this time, sufficiently sophisticated instruments are available to make the hardware part of the dance quite a reliable partner. To complicate the situation in paediatrics further, the subject is in a constant mode of change from measurement to measurement; in addition, in many clinical situations the identity of the observer may change. Efforts should be made to minimize variation in the observer especially if growth parameters are to be utilized as part of a research study. The record for longevity of observer stability to which all subsequent researchers have aspired is that of the Harpenden Growth Study initiated in 1949 by Professor J.M. Tanner and continued for 25 years. One of the results of this dogged dedication to anthropometric measurements was the realization of measurement instruments that are both ergonomically designed and accurate. These anthropometric instruments are widely accepted for their accuracy, consistency, and ease of use.

A long list of standardized measures have been developed by auxological anthropometrists but only the basic measurements that might be used in standard clinical or growth research studies reported in the rheumatological literature will be included in this chapter. Figure 1 outlines a small subset of the bodily landmarks and standardized surface measurements that are utilized by anthropometrists.

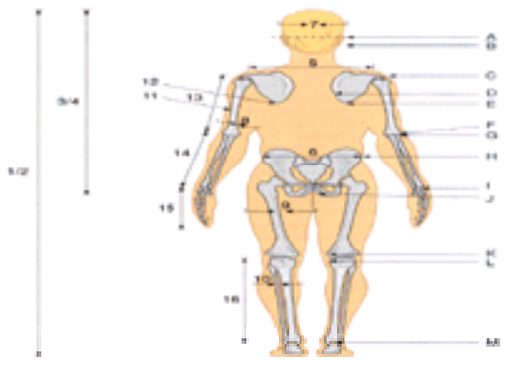


Fig. 1 Body landmarks and standard surface markers utilized in anthropometric measurements in humans (Tanner 1978). *Measurements*: 1/2 stature-recumbent length; 3/4 sitting height (crown–rump length); 5, biacrominal diameter; 6, biiliac diameter; 7, head circumference; 8, extended upper-arm circumference; 9,

upper-thigh circumference; 10, maximum calf circumference; 11, triceps skinfold; 12, subscapular skinfold; 13, upper-arm length; 14, forearm length; 15, hand length; 16, tibial length. *Landmarks:* A, Frankfurt plane; B, mastoid process; C, lateral border of the acromium; D, medial border of the scapula; E, inferior angle of the scapula; F, head of the radius; G, medial and lateral humeral epicondyles; H, iliac crest; I, distal end of the radius; J, gluteal fold; K, medial and lateral femoral epicondyles; L, medial aspect of the femorotibial joint space; M, medial malleolus.

Height

Assessment of height is the anthropometric measurement to which most attention has been paid in terms of development of instrumentation and accuracy. In children able to stand independently, the Harpenden stadiometer is the best readily available instrument. This stadiometer combines two critically important advances in measurement technique. It has a rigid horizontal wooden headboard which moves up and down a rigid backboard on miniature rollers and has a display counter from which the observer can read the height directly. The cost of this instrument should not preclude its use in paediatric settings. However, if alternative equipment must be used, a rigid horizontal wooden headboard must be incorporated. The greatest inaccuracies in height measurement occur when using an instrument based on a flexible horizontal headboard at an approximate angle of 90° to the wall. The use of a flexible horizontal headpiece can introduce up to 2 cm of inaccuracy in height measurements (Tanner 1978). Figure 2 shows the Harpenden stadiometer in use. The technique for measuring height is also of critical importance. The subject's shoes and socks must be removed, not only because they will affect the height of the subject but also they may conceal slight raising of the heels. The subject should stand straight so that his or her heels, buttocks, and shoulders are in contact with the backboard of the stadiometer or wall. In younger children, it may be necessary to assist by pressing down the feet so that the undersides of the heels are in contact with the ground. The legs should be positioned so that the knees or the medial mallei are touching, depending on whether the patient has knock-knee or has straight legs. The shoulders should be relaxed and sloping slightly forward in the natural position. The head is held erect in the Frankfort plane (Fig. 1) and the rigid horizontal headboard is moved downwards until the backboard touches the subject's head with slight weight or pressure on the headboard to minimize the effect of hair thickness. While the subject is in this position, he is instructed to take a deep breath and stand tall while keeping his heels flat on the ground. In younger children, stretching of the spine is assisted by gentle upward pressure applied beneath the mastoid processes by the measurer. This technique also minimizes diurnal variation. Without the use of pressure applied to the mastoid processes, diurnal variation in measured height may be as much as 20 mm (Strickland and Shearin 1972) but with pressure it is normally less than 5 mm (Whitehouse *et al*. 1974). If a subject has leg length asymmetry, the overall skeletal height can be accurately measured by having the child stand on the longer leg with the shorter leg supported by a wedge or block. The sensitivity of the Harpenden stadiometer is sufficient for an accurate assessment of the actual leg inequality to be obtained by taking two measurements of height. The two measurements are based on the child standing on each leg independently with the alternate leg supported by a block. Measurement of height utilizing the shorter leg will require that the patient place the longer leg slightly flexed on the block.

Fig. 2 Demonstration of use of Harpenden stadiometer to measure standing height.

Sitting height

In those subjects in whom muscle weakness, hip contracture, and/or knee contracture prevent an accurate measurement of standing height, sitting height is an excellent alternative. Sitting height may be measured using the same equipment that is used for standing height but requires the use of a sitting platform (Fig. 3). The technique used for positioning the patient against the backboard and stretching his or her spine with gentle pressure on the mastoid processes is similar to that used for standing height. The only precaution is to ensure that the sitting platform is properly horizontal and has sufficient depth to be stable. To obtain the actual sitting height, the height of the sitting platform is subtracted from the total height of the subject sitting upon it. The technique of having the child sit on the floor with legs outstretched as an alternative for the sitting height box is not recommended because of the inaccuracies introduced by hip flexion contractures or hamstring tightness and the compensatory lumbar lordosis.

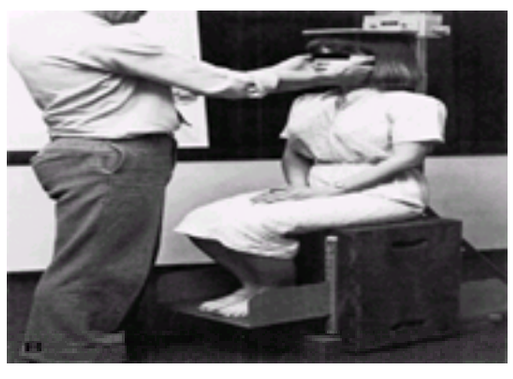


Fig. 3 Demonstration of measurement of sitting height using adaptation of standard Harpenden stadiometer and sitting platform.

Weight

Instruments accurate to 0.1 kg should be used for patients of the age range generally seen in a rheumatology clinic. Platform–balance beam instruments remain the most widely used and accepted. Digital read-out instruments which are highly accurate are available. The subject should be weighed wearing a minimum of clothing. When weighing frightened or unco-operative subjects such as young children, more accurate results are achieved by weighing the parent and the child together, then the parent alone.

Skinfold measurements

These measurements are useful to distinguish lean body mass from subcutaneous fat in growth assessment. Lange and Harpenden skinfold callipers are the preferred instruments, although many other types are available. Lange callipers are more commonly used in paediatric populations. The technique for skinfold measurement is uniform across all skinfold sites. The tissue to be measured is obtained by parting the observer's index or middle finger and thumb some 6 to 8 cm and sweeping them together over the surface of the skin to collect the subcutaneous tissue pulled away from the underlying fascia (Fig. 4). This technique is not easily mastered and is subject to error even with very skilled operators. The absolute error is approximately 5 per cent for skinfold thicknesses of 10 mm or less but 5 to 10 per cent for skinfold thicknesses greater than 10 mm (Cameron 1986).

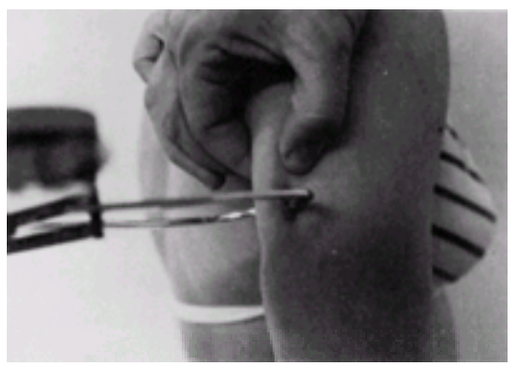


Fig. 4 Demonstration of technique for measuring skinfold thickness with callipers.

Standards for physical measurements

Growth charts have been developed to demonstrate the progressive changes in a number of growth parameters with age and sex. The range of variability is expressed either in percentiles or standard deviations from the mean. For children demonstrating normal growth patterns, percentile-based charts are entirely adequate. However, those children displaying growth patterns significantly above or below the mean are often poorly served by percentile-based growth charts. For example children with significant growth retardation are generally more than 3.5 standard deviations (SD) below the mean height for age. However, children between -2.5 and -3.5 SD have growth retardation of borderline significance. All of these children will be depicted as below the first percentile mark on a percentile-based growth chart and relative changes in growth would be difficult to determine. For this reason, growth charts giving the mean and SD for various ages are more helpful.

Currently accepted growth charts in the United States are based on an extensive cross-sectional study by the United States Public Health Service of 7000 normal children. This study measured each child on one occasion; minimal differences were seen between children residing in six different geographic areas of the United States and the heights of black and white children were closely equivalent. Growth charts for children of other ethnic backgrounds are not available at this time. Growth charts derived in the United States and United Kingdom for well-nurtured children of similar socioeconomic background show no major discrepancies, particularly in height (Kaplan 1937). Generally accepted growth charts in the United Kingdom are those derived by Tanner and Whitehouse. They are based on longitudinal growth data from British children and have the advantage of being able to discern differences in growth based on differences in timing of the adolescent growth spurt (Tanner 1978). Tables of normative values for length, weight, and head circumference by age and sex to 18 years of age are available. Figure 5 and Figure 6 show revised graphs prepared in 1995 in the United Kingdom of percentiles for stature and weight in males and females aged 2 to 19 years. In addition, the pubertal stages are included for the first time.

Fig. 5 Percentiles for stature and weight for age, boys, 2 to 19 years.

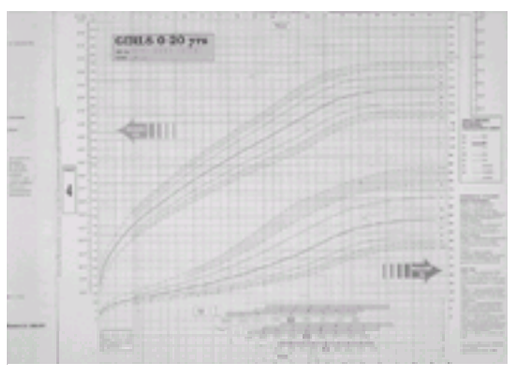


Fig. 6 Percentiles for stature and weight for age, girls, 2 to 19 years.

In those children who have contractures of the lower extremity or have weakness preventing them from standing erect, sitting height represents a reproducible and reliable alternative. Age and sex adjusted standards for sitting heights have been developed (Johnson *et al*. 1981). Standards for sitting heights for males and females are shown in Table 1 and Table 2.

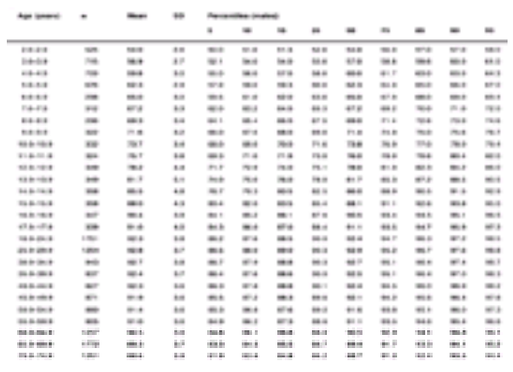


Table 1 Means, standard deviations, and percentiles of standards of sitting height (cm) by age for males (Frisancho 1990)

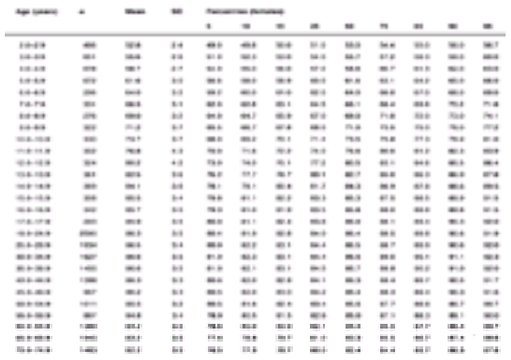


Table 2 Means, standard deviations, and percentiles of standards of sitting height (cm) by age for females (Frisancho 1990)

Standards for skinfold thicknesses have been developed. Norms vary by sex and age. Figure 7 demonstrates variation in subscapular and triceps skinfolds from birth to 19 years of age. Percentile values for triceps and subscapular skinfolds are available arranged by age and sex (Lohman *et al*. 1988; Frisancho 1990).

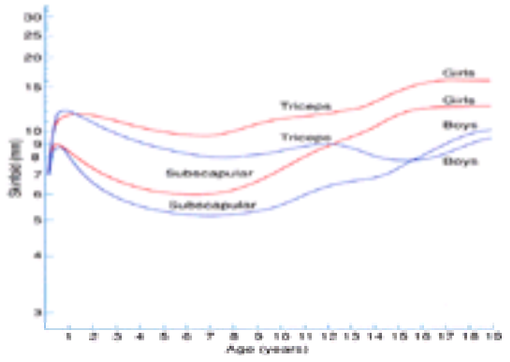


Fig. 7 Skinfold thickness by age and sex as measured by Harpenden callipers. The lines shown are the average for British children (Tanner 1978).

Approach to measurement of skeletal maturation

Several methods for quantifying skeletal maturation have been developed. These are based on the recognition of the orderly and reproducible sequences of changes in the appearance of the skeleton during childhood. Determination of the age of skeletal maturation is important in the diagnosis of a variety of endocrinopathies, assessment of the correlation between chronological age and skeletal maturation, and prediction of adult height. Skeletal maturity assessment can also be used in planning orthopaedic procedures in which the timing and the outcome of the procedure is influenced by subsequent growth, such as surgical management of leg length discrepancy or joint replacement.

Detailed discussion of the influence and overall characteristics of the process of skeletal maturations is the topic of Chapter 2.4. This section will focus only on radiographic methods used to assess skeletal maturation.

Radiographic assessment of skeletal maturity in the child is usually based on the appearance of the hand and wrist but methods are available which assess the entire hemiskeleton, knee, and ankle/foot (Hoerr *et al*. 1962; Pyle and Hoerr 1969). Results from all these methods compare well. Utilization of hand/wrist methods results in a very small radiation dose to the child, is an easily obtained and reproducible radiographic technique, and both left and right hands or wrists can be easily obtained in a single film allowing assessment of significant anharmonic skeletal growth. However, hand/wrist radiographs are unreliable during the first year of life since carpal bones are generally unossified at this stage resulting in very slight alterations in radiographic appearances. In children under 2 years of age, radiographs of the knee and/or ankle are generally used to assess skeletal maturation.

Many methods have been used to determine skeletal maturation: these include timing the onset of ossification of growth centres (Garn method) (Garn *et al*. 1967), comparison of radiographs with illustrations in a standard atlas (Greulich–Pyle method) (Greulich and Pyle 1959), or calculated maturation scores using a standardized methodology (Tanner–Whitehouse) (Tanner *et al*. 1975). There are other important influences on skeletal maturation that must be taken into account when selecting a particular methodology for clinical use. The rate of skeletal maturation is quite strikingly different between boys and girls and there are moderate differences in the rate of skeletal maturation between the black and white races. There are also regional dissimilarities between the populations used to develop standards. The importance of this becomes obvious when it is remembered that the Tanner–Whitehouse system was developed using British children as norms and the Greulich–Pyle atlas was developed using growth of American children.

The Garn method uses published tables showing the timing of onset of ossification of the growth centres throughout the body in normal children (Table 3). Skeletal age is assessed by determining which growth centres in the child under study have begun to ossify compared with published standards. With this method, assessment of radiographs is simplified, as a growth centre is either ossified or not, resulting in a relatively low interobserver variability. Separate standards have been developed for white and black children and for males and females. However, the Garn method is dependent upon the development of new growth centres which occurs only intermittently in the development process. For example, the hand and wrist centres change relatively little during the first 1 to 2 years of life and some of the growth centres in the hand have very limited predictive value because normal children will ossify a particular centre over a very wide range of ages. As a result, in younger children and in studies requiring a high degree of precision, large portions of the skeleton must be routinely surveyed, incurring greater radiation dose and cost.

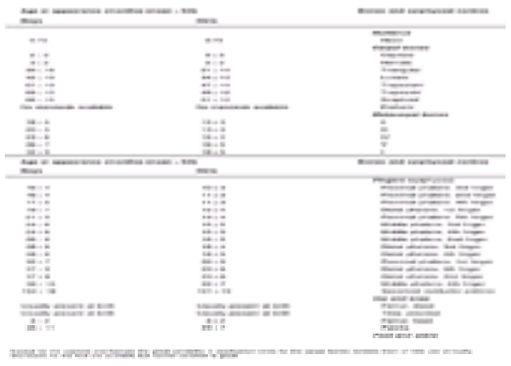


Table 3 Timing of calcification of centres of ossification in infancy and childhood to the nearest month

The most commonly used method in the United States is comparison of hand/wrist radiographs with standard illustrations in the Greulich–Pyle atlas (Greulich and Pyle 1959). The study population for which this atlas was developed was 1000 healthy, United-States-born, Caucasian children of above average socioeconomic and educational backgrounds. These children were studied radiologically longitudinally at 3-month intervals during the first year of life, at 6-month intervals until the 12th

year, and annually until maturity. The atlas consists of a series of standard radiographs representing skeletal ages ordered in increasing maturity with the intervals approximating those at which the population was studied during development of the standards. Once the closest match has been determined between the patient's radiographs and the atlas standards, the normal range for the skeletal age can be determined by using charts contained in the atlas. These provide separate values for standard deviation for boys and girls. The observed differences on the standard radiographs is usually obvious until between the ages of 10 and 14 years. At this point there are only slight changes in the appearance of the radiographic standards resulting in greater inaccuracy and interobserver variability in this age range. In general, greater weight is given to the appearance of the metacarpals, phalanges, and distal radius and ulna than to the carpal bones because of the much greater influence of systemic illness on maturation of the carpal bones.

In the United Kingdom, the Tanner–Whitehouse method is generally used. This method determines skeletal maturation based on the extent of ossification and the morphology of each of 20 bones in the hands and wrists. Each bone is evaluated individually and is assigned a particular development stage based on drawings and reproductions of radiographs. Appropriate written text is provided for each of the 20 skeletal areas (Tanner *et al*. 1975). Two separate summary scores are calculated:

1. a RUS score represents the evaluation of the distal radius and ulna and the phalanges and metacarpals of the first, third, and fifth fingers;
2. a carpal score is based on the evaluation of the carpal bones excluding the pisiform.

The RUS score and the carpal score are averaged resulting in an overall Tanner–Whitehouse score (range 0–100). Finally, skeletal age is determined by plotting the Tanner–Whitehouse score on a graph which provides the mean and a range of 3rd to 97th percentile values. Separate standards are provided for males and females. The Tanner–Whitehouse method requires much more training and experience to be used reproducibly and accurately than the Greulich–Pyle atlas method, although, when mastered, it has relatively less interobserver variability. Comparisons of the skeletal ages derived from each of the methods shows that the Tanner–Whitehouse method is generally between 2 and 12 months ahead of that assigned using the Greulich–Pyle atlas. Perhaps the most important consideration is the fact that the standards for each method are geographically distinct. The observer would be well advised to use the method that most closely coincides with the location of his or her clinic.

In those cases in which significant arthritis of the hand and wrist may make either of the methods unreliable or unworkable, a system based on assessment of anterior–posterior radiographs of the knee is available—the Roche–Wainer–Thissen method (Roche *et al*. 1975). Overall, this method has shown acceptable correlation with Greulich–Pyle standards and could be used in a child in which arthritis of the knee is either absent or unilateral.

Prediction of adult height based on radiographic assessments of skeletal maturation is possible. Garn has developed a method for predicting adult height based on published charts of skeletal maturation. Adult height is calculated by multiplying the child's height at the time of the examination by a 'multiplicity factor' for the estimated skeletal age at that point in time (Posnanzki *et al*. 1976). Bayley and Pinneau have developed an alternative method using the Greulich–Pyle atlas. Tables have been devised showing the percentage of predicted mature height at each skeletal age for children of delayed, average, and advanced skeletal maturity, with separate standards for boys and girls (Bayley and Pinneau 1952). Tanner and Whitehouse have also developed a method of predicting adult height based on a regression equation that includes current height, current chronological age, and RUS bone age plus a constant. In addition, separate tables have been developed for boys and girls which give age-specific coefficients for the factors in the regression equation (Tanner *et al*. 1975). Methods are available for both the Tanner–Whitehouse method and the Greulich–Pyle atlas method incorporating midparental heights into the equation to predict adult height. Comparison of the various methods for predicting adult height showed that the most important determinant was the radiological method used to assess skeletal maturity rather than the mathematical method used to calculate the height (Harris *et al*. 1980). The methodology used should thus be based on the geographic location of the population being studied. Whatever the method used, the accuracy of the prediction is better the closer the child is to maturity (Fig. 8) and the closer the match between skeletal age and height age. Moreover, none of these methods take into account acceleration or retardation of growth which may occur after the prediction has been made because of future illnesses or therapies.

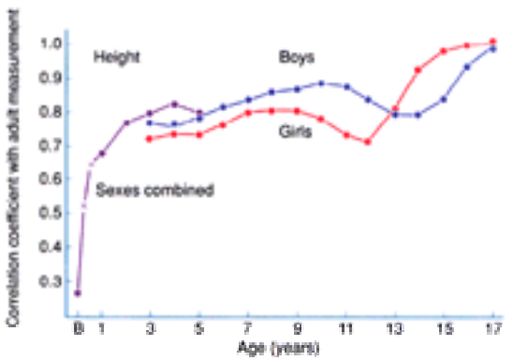
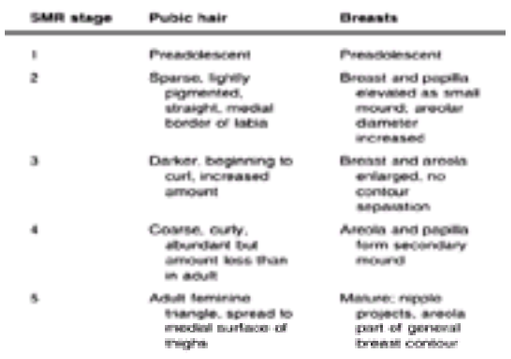


Fig. 8 Correlation between adult height and height of the same individuals in childhood (Tanner *et al*. 1975).

Assessment of pubertal development

Assessment of pubertal development has been made systematic by Tanner in such a thorough and practical fashion that it has become the standard for assessment worldwide. Assessment of the stage of pubertal development is enormously important, not only for the recognition of large changes in the size and functional structure of the reproductive organs themselves but also the major changes in hormonal and growth potential associated with the various sexual stages. Defects in hormones associated with growth at this period can be detected early if the stages of pubertal development are defined. Pubertal development is divided into five stages for both males and females—sexual maturity ratings (**SMR**). Stage I represents prepubertal status, SMR1. Early adolescence coincides with SMR2 and maturation at this stage is initiated by increased secretion of pituitary gonadotropins and growth hormone. Middle adolescence refers to the period corresponding to SMR3 and 4. It is the period of the most dramatic growth and change in body habitus. Late adolescence is associated with secondary sexual characteristics described in SMR5 which approximate adult size and appearance. Little additional linear growth is achieved during late adolescence. Table 4 and Table 5 describe the classification of the sexual maturity stages for girls and boys, respectively.



SMR stage	Pubic hair	Breasts
1	Preadolescent	Preadolescent
2	Sparse, lightly pigmented, straight, medial border of labia	Breast and papilla elevated as small mound, areolar diameter increased
3	Darker, beginning to curl, increased amount	Breast and areola enlarged, no contour separation
4	Coarse, curly, abundant but amount less than in adult	Areola and papilla form secondary mound
5	Adult feminine triangle, spread to medial surface of thighs	Mature: nipple projects, areola part of general breast contour

Table 4 Classification of physical, sexual maturity stages in girls (Tanner 1962)

SMR stage	Pubic hair	Penis	Testes
1	None	Preadolescent	Preadolescent
2	Scanty, long, slightly pigmented	Slight enlargement	Enlarged scrotum, pink texture altered
3	Darker, starts to curl, small amount	Longer	Larger
4	Resembles adult type, but less in quantity, coarse, curly	Larger, glans and breadth increase in size	Larger, scrotum dark
5	Adult distribution, spread to medial surface of thighs	Adult size	Adult size

Table 5 Classification of physical, sexual maturity stages in boys (Tanner 1962)

Despite the tremendous variety and complexity of changes in the body taking place during sexual maturation, the simplicity and accuracy of the Tanner staging method have led to its universal adoption. Numerous studies have shown that although there are secular and geographic variations in the chronological age at which these pubertal changes are attained, the overall sequence of changes is remarkably similar despite wide variations in nutritional status, genetic backgrounds, psychosocial status, and geographical location (Tanner 1978).

Analysis of growth data

The approach to the analysis of cross-sectional growth data generally used is a 'distance standard' which is a set of smooth curves giving the percentile values for a range of ages. This is the form of the standardized growth chart used by almost all paediatricians (Fig. 5 and Fig. 6), and allows determination of the actual percentile of the measured height or weight for a child of particular age and sex compared with a very large population of similar age and sex-matched normal subjects.

An alternative method is to assess growth velocity. The essential statistical point about a growth velocity measurement standard is the pair of measurements summarized by a single quantity (their difference), which is readily interpretable and for which a standard can be constructed (Fig. 9). It is of paramount importance to cross refer the observed growth velocity to some other maturational measurement. This provides much more accurate information about a child's growth than consecutive values from a single child plotted on a distance standard. A child who moves from one percentile curve to another percentile curve on distance standard may be entirely normal when the changes in dimension are compared with a velocity standard. This is especially so during a growth spurt. The interpretation of a series of measurements of height or weight will be much clearer when the velocity standard is used.

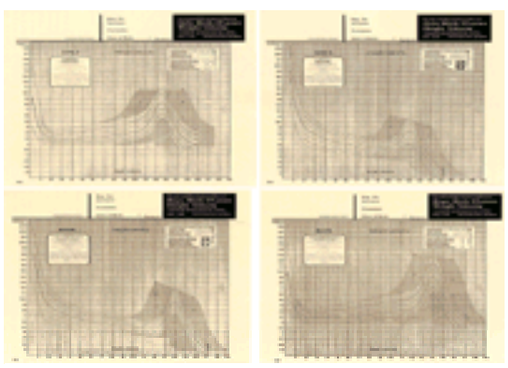


Fig. 9 (a) and (b) Examples of height and weight velocity tables used in the United Kingdom. (Reproduced with permission from Castlemead Publications.)

There are two circumstances in which growth standards are extraordinarily helpful. They are a powerful device for investigation in groups of not overtly ill children to see whether particular individuals might benefit from special medical, educational, or social care and nutritional intervention. Also growth standards, and especially those for velocity, are helpful in assessing the effects of disease and the response to treatment in children known to be ill, for example in growth hormone deficiency, adrenal hyperplasia, asthma, kidney disease, and arthritis. However, growth standards are not appropriate when growth data are to be used as an index of the health and nutrition of a population or subpopulation thought to be at risk, for example ethnic minorities. Comparison of data from the subpopulation under study with data collected from other subpopulations is not appropriate.

The effect of inflammation on growth in juvenile chronic arthritis

Disturbance of growth is a well-recognized problem for children suffering from juvenile chronic arthritis. Sir George Frederic Still referred to the general arrest of development that occurs when the disease begins before the second dentition in his first description of juvenile chronic arthritis in 1897 (Still 1941). Later Kuhns and Swain described three types of growth disorders in juvenile chronic arthritis: generalized retardation of body growth, persistence of infantile proportions, and asymmetry of growth (Kuhns and Swain 1932). Recognition of growth impairment is important because juvenile chronic arthritis is a disease which often remits by late adolescence and growth impairment may be one of the important, permanent sequelae.

Local disturbance in growth

Local disturbance in growth is seen at sites of inflammation. This can be in the form of overgrowth or undergrowth. The mechanism of epiphyseal overgrowth is not clear, but may relate to hyperaemia and growth factor release. The joints most commonly affected are the knees, followed by the ankles and subtalar joints and the wrists. There is a tendency to accelerated epiphyseal maturation, and with knee disease there is usually overgrowth, especially of the medial epicondyle of the femur, resulting in valgus deformity and unilateral increase in leg length (Fig. 10(a)). Compensatory scoliosis of the spine can result from untreated leg length discrepancy. In the wrists and ankles, increased maturation of epiphyses occurs early in the disease, leading to overcrowding and fusion of the joint (Fig. 10(b)). In addition, arthritis of the wrist often leads to growth failure of the ulnar head, causing a shortened ulna with ulnar deviation of the wrist. Because of wrist or ankle involvement, the hand or foot distally is often less well developed compared with the involved side (Fig. 10(b)). This may be due to local inflammatory factors as well as relative sparing of the affected limb in daily activities. Other joints that are underdeveloped as a result of disease are the temporomandibular and the hip joints. In both cases the fossa that articulates with the long bones are underdeveloped, probably due to lack of normal use.

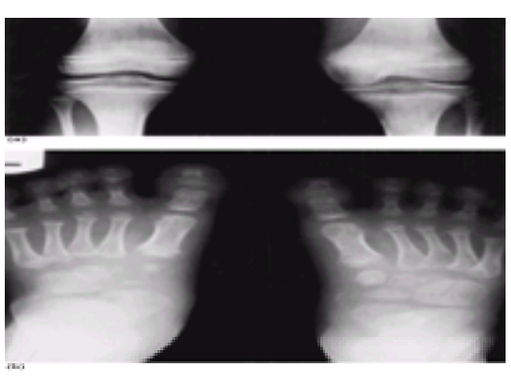


Fig. 10 (a) Radiograph of knees to show overgrowth of the medial epicondyle of the femur due to persistent arthritis of that knee. (b) Radiography of feet to show

asymmetric growth. There is acceleration of ossification in the tarsus on the affected side, and undergrowth of distal phalanges.

The spine is often involved in juvenile chronic arthritis. Common abnormalities of the cervical spine include apophyseal joint-space narrowing and eventual fusion, particularly C2–C3. As a result, undergrowth of the vertebral bodies occurs and the patient eventually has a short neck, increase in cervical lordosis, and dorsal kyphosis. Similar changes can also occur rarely in the dorsolumbar spine in severe systemic juvenile chronic arthritis. Vertebral collapse is often seen, more commonly in patients receiving corticosteroids.

Generalized disturbance in growth

In 1956, Ansell and Bywaters reported a survey of growth and bone development in 119 children aged under 14 with juvenile chronic arthritis of more than 1 year's duration. For this study the period of observation for each subject was 5 years (Ansell and Bywaters 1956). Their findings suggested that disease activity retards general growth, and indicated that corticosteroid therapy aggravated this problem (Fig. 11). Twenty of 120 height measurements from children with active disease were more than 10 per cent below the mean, whereas this was the case in only two of 70 measurements from children with inactive disease ($p < 0.01$). Provided that premature fusion of the epiphyses had not occurred, many of the younger children showed 'catch-up' or accelerated growth in remission, but not those with long-standing and severe disease. In addition, long-term treatment with cortisone was usually associated with failure to grow in height despite good disease control.

Fig. 11 (a) The shorter twin had juvenile chronic arthritis but has not been treated with corticosteroid. (b) The shorter twin had juvenile chronic arthritis and has been treated with corticosteroid.

In a more recent study in the United States by Lovell and White, subgroups of juvenile chronic arthritis were shown to affect growth differently (Lovell and White 1991). Fifty-six children with juvenile chronic arthritis of at least 1 year's duration and aged between 4 and 18, randomly chosen from a rheumatic disease clinic, were measured for height and weight. The mean heights for the systemic and polyarticular-onset groups were on the 31st and 35th centile respectively. The long-term morbidity of generalized growth delay was also reviewed from the records of 156 patients with juvenile chronic arthritis followed beyond 18 years of age. There were significantly more patients with systemic-onset and fewer with pauciarticular-onset disease in the growth failure group. There was no difference in sex, current age, or age at onset. The systemic-onset juvenile chronic arthritis subgroup was also found to have the highest growth failure rate by Bernstein *et al*. (Bernstein *et al*. 1977). In order to assess the relative effect of corticosteroids and disease activity, the authors compared the growth rate of 11 systemic juvenile chronic arthritis patients with that of 13 systemic lupus erythematosus children on equivalent doses of corticosteroid. The former had significantly more growth delay.

Corticosteroids and growth

In systemic juvenile chronic arthritis, corticosteroids are very effective in suppressing disease activity. However, it is well known that glucocorticoids inhibit growth and delay skeletal maturation, particularly when given in large doses for prolonged periods (Blodgett *et al*. 1956; and above). The form, dose, and frequency of administration of steroid are important variables. Doses of 45 mg/m^2 per day of cortisone will suppress growth in children without demonstrable endocrine abnormality, with the effect more marked in younger children (Blodgett *et al*. 1956). Ansell and Bywaters found that, in children under 5 years, the critical daily dose of cortisone above which growth was impaired was 37.5 mg, whereas in older children a daily dose of less than 50 mg of cortisone allowed normal growth (Ansell and Bywaters 1956). The synthetic glucocorticoid, prednisolone, is thought to have a more potent effect on growth; in children with asthma, growth retardation has been reported with a daily dose of 4 to 6 mg/m^2 (Falliers *et al*. 1963; Kerrebijn and Kroon 1968).

The mode of steroid administration can also effect growth differently. Following observations that the therapeutic effect of a single dose of steroid may last longer than its adrenal suppressive effect (Reichling and Kligman 1961), an alternate day corticosteroid regimen was advocated as a means of maintaining equivalent disease control with less growth retardation by Ansell and Bywaters (Ansell and Bywaters 1974). They reported continuation of growth in 28 children with juvenile chronic arthritis receiving treatment with alternate-day prednisolone in doses between 10 and 30 mg/m^2 every 48 h. However, the growth rates were less than normal. A further 21 patients were converted with some difficulty from a daily divided dose to a single morning dose alternate-day regimen to maximize night time growth hormone secretion. In most of these patients there was resumption of growth and loss of cushingoid appearance after 6 to 12 months. In a study by Loftus *et al*., an oxazolone derivative of prednisolone, deflazacort, was shown to have less effect on bone mineral depletion than prednisolone (Loftus *et al*. 1991).

Osteoporosis

Osteoporosis invariably accompanies the lack of growth in height in children with acute disease (reviewed in Woo 1994). This can be severe if the disease is uncontrolled and the relative lack of activity in children with juvenile chronic arthritis is only part of the answer. Measurement of serum calcium, vitamin D, and parathyroid hormone levels, as well as urinary excretion of calcium and phosphate, from different patient cohorts are conflicting. Measurements of bone density and bone mineral content have shown values below the 3rd centile of normal controls (Lovell *et al*. 1986; Hopp *et al*. 1991).

In a recent study in the United Kingdom, measurements of alkaline phosphatase, serum calcium, and phosphate were within normal limits compared to age and sex matched controls. Assessment of bone turnover by measuring urinary hydroxyproline and calcium/creatinine ratios have been within normal limits. However, in a very recent study, bone mineral density was seen at a number of sites in juvenile chronic arthritis patients, associated with low serum concentrations of osteocalcin and bone-specific alkaline phosphatase, suggesting reduced bone formation. Low urinary markers of bone turnover also suggested decreased resorption. Lower bone mineral density was seen with patients with active articular disease and those with a larger number of involved joints (Pepmueller 1996). Thus, these findings fulfil the biochemical and radiological definitions of osteoporosis.

Possible mechanisms of generalized growth retardation and osteoporosis

Nutrition

Several factors have been shown to be responsible for the impairment of linear growth in juvenile chronic arthritis. One important factor is the effect of prolonged periods of systemic illness with fever, leading to diminished appetite and hence poor nutrition.

The first report of nutritional deficiencies in children with arthritis was from a cohort of children in the United States (Bacon *et al*. 1988). Serum vitamin A, zinc, and vitamin D were below normal ranges for that age group. More detailed studies of nutritional requirements by several groups have also shown low protein–energy malnutrition in children with juvenile chronic arthritis (Henderson and Lovell 1989; Mortensen *et al*. 1990). Furthermore, the requirement for children with chronic inflammation may be higher than normal. Nasogastric tube feeding at night in a pilot study has shown improvement in linear growth (Lovell, unpublished

observations).

Growth hormone studies in juvenile chronic arthritis

There are several possible endocrine mechanisms for growth failure. Growth hormone may be suppressed in active chronic juvenile arthritis, or the target tissues may be somehow rendered 'less sensitive' to growth hormone by the disease with or without steroids.

Sturge *et al*. studied linear growth and the secretion of growth hormone and cortisol in 20 patients suffering from systemic juvenile chronic arthritis on four types of treatment regimen to assess the effect of corticosteroid on growth (Sturge *et al*. 1970). The height of children on prolonged daily corticosteroid therapy were below the 3rd centile, whereas children who had only received corticosteroid by an alternate-day regimen had normal growth rates.

Insulin tolerance tests were performed to assess growth hormone and cortisol secretion 10 h after corticosteroid administration. Although there was considerable variation in growth hormone response, there was no significant difference in mean levels between the four groups. In contrast, five of the six children on daily corticosteroid had subnormal mean basal plasma cortisol levels and failed to show a satisfactory rise on insulin tolerance test. Cortisol levels and response were normal in the other groups. Later, Byron *et al*. from the same unit showed the pattern of basal cortisol levels tended to parallel that for height velocity, confirming the impact of adrenal suppression on growth (Byron *et al*. 1983). More detailed studies of growth hormone secretion have been reported using insulin tolerance tests on 22 juvenile chronic arthritis patients and 20 healthy patients with a history of familial dwarfism (Butenandt *et al*. 1976). In this study, defective growth hormone secretion was demonstrated in juvenile chronic arthritis. Allen *et al*. showed increased pulse frequency of growth hormone secretion (Allen *et al*. 1991). More recently, study of 24-h growth hormone secretion in 16 United Kingdom children with juvenile chronic arthritis and severe growth retardation revealed normal pulsatile patterns of growth hormone secretion, and the mean and median levels were comparable to that of healthy children with short stature (Davies *et al*. 1994). Thus, there is unlikely to be significant growth hormone deficiency.

Insulin-like growth factor I (IGF-I) in juvenile chronic arthritis

Insulin-like growth factors (**IGF**) I and II are stimulated by growth hormone to act on target tissues and stimulate cell growth, that is they are important mediators of growth hormone action. The amino acid homology between IGF-I and -II is 62 per cent. There are two regions of homology to the proinsulin a and b chains respectively, thus conferring some cross reactivity with insulin receptors, although IGFs have their own receptors. The site of IGF-I production is mainly the liver, but it is also synthesized by many other tissues. IGF-II is more important in fetal life. These proteins are bound to carrier proteins in the circulation (the main ones being IGFBP-1 and -3). Previous reports have shown that corticosteroid therapy in childhood causes reduction in somatomedin (now known as IGF-I) activity (Philips *et al*. 1974; Elders *et al*. 1975), and subsequently there have been several studies on IGF-I in juvenile chronic arthritis.

Levels of IGF-I have been described to be reduced or normal in juvenile chronic arthritis (Takami *et al*. 1982; Bennett *et al*. 1988; Aitman *et al*. 1989; Allen *et al*. 1991). In a study of 15 systemic juvenile chronic arthritis patients and 79 normal children, Bennett *et al*. showed that serum IGF-I levels were lower than the mean normal values for age, irrespective of steroid therapy (Bennett *et al*. 1988). Furthermore, IGF-I showed a significant negative correlation with erythrocyte sedimentation rate as a measure of disease activity but no correlation with growth rate. In contrast, in the study by Aitman *et al*. of 32 children with juvenile chronic arthritis with a spectrum of growth from normal to severe retardation, serum IGF levels showed significant correlation with height ($p<0.01$) and to a lesser extent with height velocity ($p<0.05$). In this study there was no correlation between IGF-I and diet history or steroid dosage (Aitman *et al*..

More recently, low IGF-I concentrations in children with juvenile chronic arthritis have been reported (Allen *et al*. 1991). These workers also confirmed earlier reports that the disease itself is a major factor in growth retardation, independent of steroid therapy. It has also been shown that the low serum IGF-1 levels are not due to high carrier protein levels, that is IGFBP-1 is normal and IGFBP-3 is low correlating with IGF-1. Thus, there is deficient synthesis of IGF-1 in these children in spite of normal levels of growth hormone (Davies *et al*. 1997). This is in contrast to the growth failure in chronic renal failure where growth hormone 'resistance' in these patients is at least partly due to higher concentrations of IGFBP-3, reducing the amount of 'free' IGF-1 available to stimulate growth (Tonshoff *et al*. 1990).

IGF-1 synthesis is not only stimulated by growth hormone but is also affected by the nutritional state. Further, parathyroid hormone also stimulates IGF-1 and it may be low in juvenile chronic arthritis. Since IGF-1 stimulates obsteoblasts to produce osteocalcin, measurements of serum osteocalcin have shown that they are correspondingly low. However, osteoblasts are also affected by a number of other hormones, for example growth hormone and parathyroid hormone, as well as vitamin D and inflammatory cytokines. Therefore, the observed osteocalcin levels are a net effect of all these factors.

Treatment with growth hormone

There are a number of studies where growth hormone is used as a pharmacological 'anabolic' agent in growth-retarded children with juvenile chronic arthritis (Svantesson 1991; Davies *et al*. 1994; Hopp *et al*. 1995). They all show some improvement. In a comparative study with the largest cohort of patients described so far, Davies *et al*. showed significant linear growth of 3 to 5 cm during 1 year's treatment with 24 iu/m^2 per week. Further, the same group showed increase in bone density, in parallel with body surface area (Rooney *et al*., in press). The improvement, however, was abolished during flares of disease activity. The improved linear growth was maintained at 5-year follow-up (Rooney, unpublished observation).

Further prospective studies are needed to investigate the possibility of early intervention and the duration of treatment to achieve optimal final height.

Chapter References

Aitman, T.J. *et al*. (1989). Serum IGF-I levels and growth failure in juvenile chronic arthritis. *Clinical and Experimental Rheumatology*, **7**, 557–63.

Allen, R.C., Jimenez, M., and Cowell, C.T. (1991). Insulin-like growth factor and growth hormone secretion in juvenile chronic arthritis. *Annals of the Rheumatic Diseases*, **50**, 602–6.

Ansell, B.M. and Bywaters, E.G.L. (1956). Growth in Still's disease. *Annals of the Rheumatic Diseases*, **15**, 295–319.

Ansell, B.M. and Bywaters, E.G.L. (1974). Alternate-day corticosteroid therapy in juvenile chronic polyarthritis. *Journal of Rheumatology*, **1**, 176–86.

Bacon, M.C., White, P., Taiten, D. J., *et al*. (1988). Nutritional studies and growth in JRA. *Arthritis and Rheumatism*, **15**, 655–8.

Bayley, N. and Pinneau, S.R. (1952). Tables for predicting adult height from skeletal age: revised for use with the Greulich-Pyle hand standards. *Journal of Pediatrics*, **40**, 423–30.

Bennett, A.E. *et al*. (1988). Insulin-like growth factors I and II in children with systemic onset juvenile arthritis. *Journal of Rheumatology*, **15**, 655–8.

Bernstein, B.H. *et al*. (1977). Growth retardation in juvenile rheumatoid arthritis (JRA). *Arthritis and Rheumatism*, **20**, 212–16.

Blodgett, F.M. *et al*. (1956). Effects of prolonged cortisone therapy on the statural growth, skeletal maturation and metabolic status of children. *New England Journal of Medicine*, **254**, 636–41.

Butenandt, O., Eder, R., and Clados-Kelch, A. (1976). Wachstumshormonbestimmungen bei Kindern mit Rheumatoider Arthritis and Still-Syndrom. *Verhandlung der Deutschen Gesellschaft für Rheumatologie*, **4**, 47–53.

Byron, M.A., Jackson, J., and Ansell, B.M. (1983). Effect of different corticosteroid regimens on hypothalmic-pituitary-adrenal axis and growth in juvenile chronic arthritis. *Journal of the Royal Society of Medicine*, **76**, 452–7.

Cameron, N. (1986). The methods of auxological anthropometry. In *Human growth—a comprehensive treatise* (eds F. Falkner and J.M. Tanner) Vol. 3, pp. 3–46. Plenum Press, New York.

Davies, U.M., Rooney, M., Preece, M.A., Ansell, B.M., and Woo, P. (1994). Treatment of growth retardation in juvenile chronic arthritis with recombinant human growth hormone. *Journal of Rheumatology*, **21**, 153–8.

Davies, U., Jones, J., Reeve, J., *et al*. (1997). Juvenile rheumatoid arthritis: effects of disease activity and recombinant human growth hormone on insulin-like growth factor I, insulin-like growth factor binding proteins 1 and 3, and osteocalcin. *Arthritis and Rheumatism*, **40**, 332–40.

Elders, M.J. *et al*. (1975). Glucocorticoid therapy in children: effect on somatomedin secretion. *American Journal of Diseases in Children*, **129**, 1393–6.

Falkner, F. and Tanner, J.M. (eds) (1986). *Human growth—a comprehensive treatise*. Plenum Press, New York.

Falliers, C.J. *et al*. (1963). Childhood asthma and steroid therapy as influences on growth. *American Journal of Diseases of Children*, **105**, 127–37.

Frisancho, A.R. *et al*. (eds) (1990). *Anthropometric standards for the assessment of growth and nutritional status*. University of Michigan Press, Ann Arbor.

Garn, S.M., Rohman, C.G., and Silverman, F.N. (1967). Radiographic standards for postnatal ossification and tooth calcification. *Medical Radiography and Photography*, **43**, 45–52.

Greulich, W.W. and Pyle, S.L. (1959). *Radiographic atlas of skeletal development of the hand and wrist*, 2nd edn. Stanford University Press, Stanford, CA.

Harris, E.F., Weinstein, E., and Weinstein, L. (1980). Predicting adult stature: a comparison of methodologies. *Annals of Human Biology*, **7**, 225–9.

Henderson, C. J. and Lovell D.J. (1989). Assessment of protein-energy malnutrition in children and adolescents with JRA. *Arthritis Care Research*, **2**, 108–13.

Hoerr, N.L., Pyle, S.I., and Franciss, C.C. (1962). *Radiographic atlas of skeletal development of the foot and ankle. A standard reference*. Charles C. Thomas, Springfield, IL.

Hopp, R., Degan, J., Callagher, J.C., and Cassidy, J.T. (1991). Estimation of bone mineral density in children with juvenile rheumatoid arthritis. *Journal of Rheumatology*, **18**, 1235–9.

Hopp, R.J., Degan, J., Corley, K., Lindsley, C.B., and Cassidy, J.T. (1995). Evaluation of growth hormone secretion in children with JRA and short stature. *Nebraska. Medical Journal*, **80**, 52–7.

Johnson, T.R., Moore, W.M., and Jeffries, J.E. (1981). Table 1. Physical growth. In *Children are different* (eds T.R. Johnson, W.M. Moore, and J.E. Jeffries), p. 21. Ross Laboratories, Columbus.

Kaplan, S.L. (1937). Growth: normal and abnormal. In *Pediatrics* (ed. A.M. Rudolph), pp. 75–91. Appleton and Lange, Norwalk, CT.

Kerrebijn, K.F. and Kroon, J.P.M. (1968). Effect of height on corticosteroid therapy in asthmatic children. *Archives of Disease in Childhood*, **43**, 556–61.

Kuhns, J.G. and Swain, L.T. (1932). Disturbances of growth in chronic arthritis in children. *American Journal of Diseases of Children*, **43**, 1183–33.

Loftus, J., Allen, R., Hesp, R., *et al*. (1991). Randomised double-blind trial of deflazacort versus prednisolone in juvenile chronic arthritis: A relatively bone sparing effect of deflazacort. *Pediatrics*, **88**, 428–36.

Lohman, T.G., Roche, A.F., and Martorell, R. (1988). *Anthropometric standardization reference manual*. Human Kinetics Publishers, Champaign, IL.

Lovell, D.J. and White, P.H. (1991). Growth and nutrition in juvenile rheumatoid arthritis. In *Paediatric rheumatology update* (eds P. Woo, B.M. Ansell, and P.H. White), pp. 47–57. Oxford University Press, Oxford.

Lovell, D., Gregg, D., Heubi, J., and Levinson, J. (1986). Bone mineralization in JRA patients. *Arthritis and Rheumatism*, **S67**, 29.

Mortensen, A.L., Allen, J.R., and Allen, R.C. (1990). Nutritional assessment of children with juvenile chronic arthritis. *Journal of Paediatric Child Health*, **26**, 335–8.

Pepmueller, P.H., Cassidy, J.T., Allen, S.H., and Hillman, L.S. (1996). Bone mineralization and bone mineral metabolism in children with juvenile rheumatoid arthritis. *Arthritis and Rheumatism*, **39**, 746–57.

Philips, L.S., Herrington, A.C., and Daughaday, W.H. (1974). Steroid hormone effects on somatomedin. I. Somatomedin action *in vitro*. *Endocrinology*, **97**, 780–6.

Posnanzki, A.K., Kuhns, L.R., and Garn, S.M. (1976). Radiologic evaluation of maturity. In *Practical approaches to paediatric rheumatology* (ed. A.K. Poznanski), p. 291. Year Book Medical Publishers, Chicago, IL.

Pyle, S.I. and Hoerr, N.L. (1969). *A radiographic standard of reference for the growing knee*. Charles C. Thomas, Springfield, IL.

Reeds, P.J. and Fiorotto, M.L. (1990). Symposium on growth: growth in perspective. *Proceedings of the Nutrition Society*, **49**, 411–20.

Reichling, G.H. and Kligman, A.M. (1961). Alternate-day corticosteroid therapy. *Archives of Dermatology*, **83**, 980–3.

Roche, A.F., Wainer, H., and Thissen, D. (1975). *Skeletal maturity. The knee joint as a biological indicator*. Plenum Press, New York.

Still, G.F. (1941). On a form of chronic joint disease in children. *Medical-chirurgical Transactions*, **80**, 47 (reprinted in *Archives of Disease in Children*, 1941, **16**, 156–65).

Stockard, C.R. (1921). Developmental rate and structural expression. *American Journal of Anatomy*, **28**, 115–20.

Strickland, A.L. and Shearin, R.N. (1972). Diurinal height variation in children. *Paediatrics*, **80**, 1023–7.

Sturge, R.A. *et al*. (1970). Cortisol and growth hormone secretion in relation to linear growth: patients with Still's disease on different therapeutic regimens. *British Medical Journal*, **3**, 547–51.

Svantesson, H. (1991). Treatment of growth failure with growth hormone in patients with juvenile chronic arthritis. A pilot study. *Clinical and Experimental Rheumatology*, **9** (suppl. 6), 47–50.

Takami, M., Taguchi, T., and Nagashima, K. (1982). Endoimmunological studies on growth retardation in children with JRA. *Rylemachi*, **22**, 19.

Tanner, J.M. (1962). *Growth at adolescence*, 2nd edn. Blackwell Scientific Publications, Oxford.

Tanner, J.M., Whitehouse, R.H., Marshall, W.A., Healy, M.J.R., and Goldstein, H. (1975). *Assessment of skeletal maturity and prediction of adult height (TW2 method)*. Academic Press, San Francisco.

Tanner, J.M. (1978). *Fetus into man*. Harvard University Press, Cambridge, MA.

Tonshoff, B. *et al*. (1990). Growth stimulating effects of recombinant human growth hormone in children with end-stage renal disease. *Journal of Pediatrics*, **116**, 561–6.

Waddington, C.H. (1962). *New patterns in genetics and development*. Columbia University Press, New York.

Whitehouse, R.H., Tanner, J.M., and Healy, M.J.R. (1974). Diurinal variation in stature and sitting height in 12–14 year-old boys. *Annals of Human Biology*, **1**, 103–7.

Woo, P. (1994). Growth retardataion and osteoporosis in juvenile chronic arthritis. *Clinical and Experimental Rheumatology*, **12**(S-10), S87–9.

1.2.1.1 Spinal problems in adults

Andrew O. Frank

Summary

Back pain has now reached epidemic proportions in Western societies. Few with back pain have serious medical (non-spinal) conditions, significant spinal pathology, or direct involvement of nerve roots such as to require surgery. Although the causes of back pain remain unclear, the effects at work or leisure of physical stress and its consequences on discs, facet joints, and supporting soft-tissue structures are mostly implicated.

Modern management emphasizes the role of self-care, which should begin in general practice at presentation of the first episode and be reinforced by all health professionals. In the absence of root compression, bed rest should not usually be longer than 48 h. A positive approach is encouraged, acknowledging that returning to a normal life may require working through pain. More emphasis is laid on actively encouraging a return to physical fitness and other activities, including employment. Medication plays a part in facilitating these objectives. The role of tricyclic antidepressants to break pain cycles and treat the depression that may be present in long-standing pain syndromes is increasingly thought to be valuable.

Studies have shown our ability to lessen the duration and severity of individual episodes of back pain, and reduce recurrences and their cost in terms of suffering and lost work. No one should remain in pain longer than 6 weeks without being referred to a specialist service. This should include provision of information, access to manual therapists with postgraduate training (or other appropriately trained manipulators if physiotherapists are not available), education on back-strain prevention and physical fitness, pain services through clinical psychology, and consultant rheumatological or orthopaedic support. Comprehensive (and inpatient) pain and rehabilitation services should be accessible to all populations (e.g. through a regional service).

The main challenges for the future lie in detecting those individuals who are unlikely to respond well to physical and educational measures so that a major emphasis on developing coping strategies, reducing the inhibitory effects of fear, and managing social issues may reduce the number of people who will require costly, intensive, and probably inpatient rehabilitation.

Never again should it be said that 'there is nothing to be done for backache'.

Introduction

In Great Britain, between 3 to 7 million people consult their general practitioner each year for back pain and over 1 million visit hospital annually (Fig. 1). Many are likely to be out of work (Table 1), taking medication, and probably making demands on primary and secondary health care as well as the private sector (both orthodox and heterodox practitioners) (Clinical Standards Advisory Group 1994a).

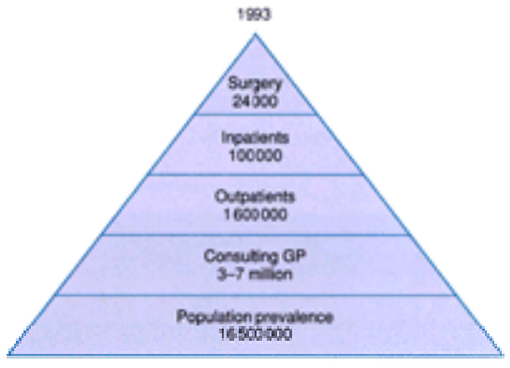


Fig. 1 Estimated annual healthcare for back pain in 1993 —from Report of a Clinical Standards Advisory Group Committee 1994 (United Kingdom)

Estimated total work incapacity:
150 million working days, 1993
Lost output:
Estimated loss (1993) £3800 million
General practitioner consultations:
Estimated 3–7 million annually
Hospital outpatient consultations:
Estimated 1.6 million annually
Hospital inpatient episodes:
Estimated 100 000

From Clinical Standards Advisory Group (1994a).

Table 1 Estimated consequences of the back-pain epidemic 1993 (United Kingdom) From Clinical Standards Advisory Group (1994a).

Definition of low back pain is difficult (Spitzer 1987). In this chapter the term 'low back pain' will imply the symptom complex where pain is localized to the lumbar spine or refers into the leg or foot, and where other specific conditions causing such pain have been excluded (hereafter referred to as low back pain). As pain from non-disc structures such as facet joints causes pain that may refer to the distal leg and foot (Mooney and Robertson 1976), the term sciatica is not used and 'leg pain' is used instead.

The Clinical Standards Advisory Group (1994 *a*) has recommended the term 'simple backache' for this group of patients. I have not adopted this term because I believe that many patients in severe distress will have difficulty accepting their doctor's advice if this term is used. Other terminology used in this chapter is shown in Table 2: it has been modified from the Quebec Task Force (Spitzer 1987), who used the period of 7 weeks to define the transition from subacute to chronic pain—assuming that healing would have taken place by 7 weeks and thus pain should have disappeared. Whether it is realistic to assume this time scale for the healing of tissues that are being continuously stressed mechanically is questionable. It must be emphasized that the precise timings suggested in this chapter are less important than the idea that the pattern of underlying spinal (dys)function will change over time. Health workers must adopt strategies that allow for this.

Acute pain: from 0–7 days duration (pain-free at onset)
Acute on chronic pain: significant exacerbation of pre-existing pain
Subacute pain: from 7 days to 3 months' duration
Chronic pain: duration over 3 months
Chronic pain syndrome: psychological and social consequences of chronic pain influencing behaviour
Intractable pain: failed conservative treatment of chronic pain
Back school: outpatient programme of treatment emphasizing educational management
Intensive rehabilitation: multiprofessional combination of skills and therapeutic modalities to minimize pain and disability
Lumbar segment: two adjacent vertebrae and their intervening soft tissues (Waddell 1982)

From Frank (1993a).

Table 2 Terminology

Most of us assume a biomechanical basis for spinal pain, although we seldom understand why some patients suffer badly and others do not. Explanations will probably be found in factors such as the elasticity of the soft-tissue, spinal-support structures, the biochemistry of the disc, and the shape and size of the spinal canal. The roles of the dorsal horn of the spinal cord and areas of the brain are discussed by Jayson (1994).

Doctors trained to intervene in well-understood disease processes may have difficulty in coping with spinal pain (Roland 1994), where there are few objective physical signs and no laboratory investigations to confirm the precise cause and severity of the complaint (Waddell 1982). The availability of MRI and CT scanning is gradually expanding our understanding of back pain (Butler *et al*. 1990), but the majority of people with minor symptoms will not have the advantage of these investigations, which may not be helpful. Consequently, many doctors have difficulty in inspiring confidence in their patient, who may therefore seek help from heterodox practitioners (Moore *et al*. 1985; Aldridge 1989).

Back pain is an extraordinarily complicated subject. Our seeming inability to help many people suffering from it is reflected in the large number of therapies available to the average district general hospital (Table 3). It is possible that medical training hinders a satisfactory approach by concentrating on the exclusion of pathology, when what patients want is understanding of their problems (Armstrong 1991), alleviation of their symptoms, and positive encouragement that activity will not harm them but be therapeutic.

Table 3 Therapeutic options

An individual's attitudes, personal strengths, and weaknesses may also play a part: e.g. in their positive or negative responses, the reactions of family and colleagues, and their ability to 'fight' to prevent recurrences. Thus a positive attitude is required from 'professional advisors' from the initial acute attack onwards. Some would argue that the two questions in back-pain management are: (i) when to take control of the pain from the patient; and (ii) when (and how) to return control to them. Thus 'person management' is a prerequisite to back-pain management, and to translating rehabilitation philosophy into back-pain practice.

Most attacks are short and self-limiting. Those remaining episodes may require additional therapy that may sometimes need to be organized through hospital. A combination of the measures outlined (Table 3) may be needed, but these more difficult problems are mostly resolved or greatly alleviated within a year. A few patients, however, do badly. They should be considered as 'failed conservative treatment' and require different strategies of management (see Intractable back pain below).

The specific management of lumbar disc disease is discussed in Chapter 5.18.1 and of the cervical spine in Chapter 5.18.2. Spinal problems in children are reviewed in Chapter 1.2.1.2.

The spine

The spine is extremely complex, constituting 139 joints, 24 discs, numerous bursas, and a vast number of ligaments and muscles comprising support structures.

The five lumbar vertebrae are linked by intervertebral discs anteriorly, two synovial facet joints posteriorly, and many supporting structures including ligaments and muscles. A segment is defined as two adjacent vertebrae and their intervening soft tissues (Waddell 1982). Pain, a subjective symptom, may arise from any of these structures, which are all innervated, and is the cardinal feature of spinal disorders.

Compartmentalization of the spine into cervical, thoracic, and lumbar regions makes sense on anatomical grounds but may hinder realistic approaches to treatment. The cervical spine has a low weight-bearing capacity. The atlantoaxial and facet joints are designed to facilitate mobility. The thoracic spine is splinted by the costovertebral joints, possibly explaining the decreased frequency of symptoms arising from the thoracic spine (Waddell 1982; see also Table 4). The lumbar spine is designed to take the weight of the upper part of the body, with less facility for movement.

Table 4 Conditions presenting to a rheumatology clinic with an interest in spinal conditions

Non-specific low back pain, that is pain to which no specific pathological (as opposed to degenerative) process can be attributed, may occur through all age groups for differing reasons. Under the age of 25, congenital factors are likely to be involved (Table 5): for example, the size and shape of the spinal canal may be a key factor in determining who will suffer symptoms from degenerative spinal disease (Porter *et al*. 1978). Disc lesions are more likely in early middle life. In later years the aftermath of previous disc disease may be seen in the form of osteoarthrosis of the facet joints. In old age, osteoarthrosis may be compounded (particularly in women) by osteoporosis. Exceptions abound: for example, adolescent discitis. Injuries caused by overuse of poorly prepared structures during sport and leisure occur at any age. Facetal osteoarthrosis rarely precedes disc disease (Butler *et al*. 1990).

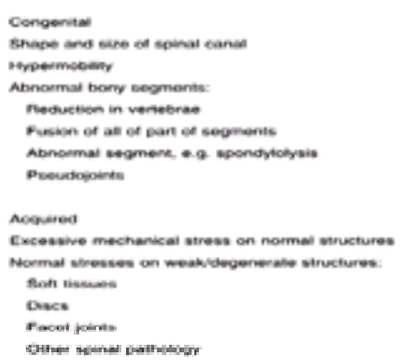


Table 5 Potential factors predisposing to spinal pain

Mechanical stress is taken predominately through the lower cervical and lower lumbar segments. It is not surprising that lesions at these two sites often coexist. It is thus important to consider the spine as a whole to ensure that advice given for one area does not aggravate another. Thus, advice to a patient with acute back pain to rest prone may be counterproductive as this sleeping posture often aggravates pain arising from the neck.

This chapter will, for brevity, concentrate on management of pain arising from the lower spine, as this problem seems to have a greater effect on the economy through sickness absence from work.

Difficulties with the evaluation of therapies for low back pain

The clinical course of low back pain creates difficulties for the investigator. Up to 85 per cent of patients with low back pain cannot be given a definitive diagnosis (Deyo 1991) because of the poor associations between symptoms, signs, imaging results, and pathological findings. In chronic back pain, the intercurrent effects of exacerbations on the underlying pain make assessment difficult, as it is hard to distinguish this from a new attack of pain. The assessment of pain is outside the scope of this review, but the pain/disability scores used in the best studies may not discriminate between the physical and psychological components of low back-pain disability (Main *et al*. 1992).

Many studies have not differentiated between acute, acute on chronic, chronic, and intractable pain. Whilst arguably failing to differentiate between nociceptive pain and pain where other factors may be of importance (e.g. pain/sickness behaviour), in practice it has been shown that some people respond to physical measures in spite of pain lasting longer than 1 month (Meade 1990) or many months (Koes *et al*. 1992). In particular, studies of chronic pain have not differentiated between previously untreated chronic pain and pain that has not responded to standard conservative therapy. For this group the term 'intractable back pain' has been suggested (Frank 1993a).

The scope and techniques involved in the many types of manipulation vary (Spitzer *et al*. 1987) and are beyond the scope of this review. In all studies, readers should take account of the method(s) of manipulation used. The study by Carette *et al*. (1991) demonstrated the importance of a placebo group in studies of low back pain. This common response may reflect not only the non-specific effects of attention and caring (Deyo 1991) but also the transfer of responsibility of care from 'self' to a trained individual (Fordyce *et al*. 1986). In addition there is the constant variation of intensity and duration of symptoms—particularly of rapid improvement for most with acute low back pain. The difficulties of assessing studies on therapy in low back pain are reviewed elsewhere (Deyo 1983; Spitzer *et al*. 1987; Koes *et al*. 1991), but, for some patients, benefit has been demonstrated in randomized controlled trials and I have concentrated (not exclusively) on reviewing such studies.

Epidemiology

Scale of the problem

The scale of the back pain epidemic (Editorial 1989a) is frightening and of enormous economic proportions (Table 1). The Clinical Standards Advisory Group (1994a) has estimated that costs to the National Health Service in the United Kingdom in 1993 were £481 million. Indirect costs were estimated to exceed £5 billion for lost work (Fig. 2) and social security benefits alone. In the United States, medical costs exceed $24 billion (Lahad *et al*. 1994) and indirect costs may exceed $30 billion (Frymoyer 1993).

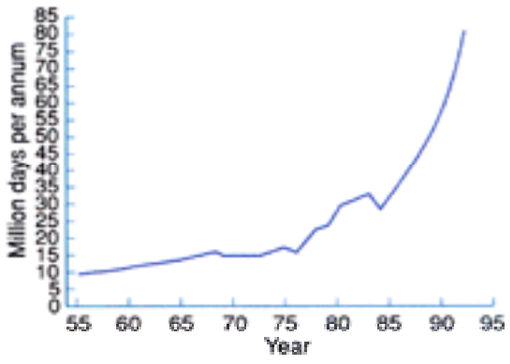


Fig. 2 Total British Sickness and Invalidity benefit for back incapacitation — from a Report of the Clinical Standards Advisory Group Committee (1994)

There are no recent data on the frequency of cervical pain: a relatively recent report did not distinguish between upper or lower back pain (Leigh and Sheetz 1989), and low back pain has been studied in more detail. The limitations of epidemiological techniques in studying this problem are discussed elsewhere (Wood and Badley 1980). In the United Kingdom it is estimated that there are 12 million general practitioner consultations annually for low back pain, resulting in 2.4 million outpatient visits (Clinical Standards Advisory Group 1994a). In 1993, approximately 100 000 people were admitted to hospital for spinal pain in England, Scotland, and Wales (Clinical Standards Advisory Group 1994a) (see Fig. 1), of which about 3 per cent may be cervical (Wells 1985).

Approximately 12000 operations are performed on lumbar discs annually in the United Kingdom (Clinical Standards Advisory Group 1994b), with less than 2000 fusions. In contrast, in the United States, approx. 200 000 surgical procedures involving lumbar disc excision are performed each year. Of these, as many as 30 000 are unsuccessful, joining the 'high-cost, high-demand, and highly emotional subset of the low back disabled' (Hanley 1992). The chances of any individual coming to back surgery are six times greater in North America than in Europe, with documented variations of nearly twofold in rates of disc excision in different parts of the United States (Keller 1992). Rates for cervical and lumbar spine surgery in hospital are increasing rapidly for both sexes in the United States (Davis 1994; Taylor *et al*. 1994). There are wide variations in rates for spinal surgery, being greatest in the United States, where there was more than five times the rate of the least, England and Scotland (Cherkin *et al*. 1994).

Conversely, low back pain is a relatively small cause of permanent, severe disability. Figures vary from 34 to 52 per 1000 people with significant morbidity from low back pain, of which only 0.2 to 4 per 1000 are severely disabled (Wood and Badley 1980). Nonetheless, consequences for employment and provision of health care are enormous. Back pain is one of the most common causes of inability to work through illness in the United Kingdom, with a total work incapacity estimated at 150 million days lost in Britain in 1993 at a cost of £5.2 billion. Such calculations of lost output and social security costs ignore the effects on individuals, their families, and of litigation.

Prevalence

Back pain is predominantly a feature of Western society. Chronic disability from the 'back pain epidemic' did not exist until legislated social support became available at the end of the nineteenth century (Allan and Waddell 1989). In the United Kingdom, the point prevalence of back pain is 14 per cent, period prevalence is 39 per cent in 1 month and 36 to 37 per cent in 1 year, and lifetime prevalence is 58 per cent (Clinical Standards Advisory Group 1994b).

The cumulative lifetime prevalence of episodes lasting two or more weeks in the United States is approx. 14 per cent, with a point prevalence of 7 per cent at any moment in time (Deyo and Tsui-Wu 1987). Of those reporting an episode of 2 weeks or more, about one-third had pain of less than 1 month's duration, about one-third had pain lasting 1 to 5 months, and the remaining one-third pain lasting 6 months or more (Deyo and Tsui-Wu 1987). The cumulative lifetime prevalence was greatest in the white population, and varied with the region of the United States and with the degree of education (greater prevalence with lesser education). The usual average age of presentation to back-pain clinics is around 40, with greatest prevalence between 45 and 64 years (Deyo and Tsui-Wu 1987; Leigh and Sheetz 1989; Clinical Standards Advisory Group 1994b).

The lifetime prevalence rate in 14-year-old Finnish schoolchildren is 30 per cent (Salminen *et al*. 1992). This suggests that lifetime prevalence data, which rely on patients' remembering symptoms from the past, may underestimate the problem.

Sex

There may be differences in spinal structure between males and females (Cooper *et al*. 1992). Back pain is equally prevalent in both sexes (Walsh *et al*. 1992a), although its clinical course may be different. Thus those requiring surgery are more frequently men (Davis 1994), possibly because disc disease is more prevalent in men (Kelsey *et al*. 1984), whilst women are more likely to attend a pain clinic (Davies *et al*. 1992). Fordyce *et al*. (1986) has noted that males may be more likely than females to incur back pain that brings them to the health-care system for help, but females may be more likely than males to have back pain that lingers into chronicity. It is unclear whether these differences relate to physical, psychological, or social factors. In men, the only socioeconomic link with back pain seems to be a manual occupation (Croft and Rigby 1994). Women with no formal educational qualification, or who lived in households in the lowest income category, are more likely to have back pain, but the reasons for this are not known (Croft and Rigby 1994).

Nearly half of all pregnant women suffer from back or posterior pelvic pain during pregnancy, often losing time from work (Ostgaard *et al*. 1994). The relation between low back pain and pregnancy, menstruation, and the menopause has been extensively investigated and reviewed. Pregnancy is a factor contributing to the problem of low back pain, with the number of abortions and a higher number of live births being associated. Hormonal changes are also important, particularly as they influence the development of osteoporosis (Svenson *et al*. 1990). Thirty-five per cent of women aged 16 to 34 years thought their back pain started in relation to pregnancy or childbirth (Mason 1994). There is currently considerable debate as to the relation between epidural analgesia and postpuerpural back pain (MacArthur *et al*. 1993). Back pain appears to occur more frequently after epidural analgesia, but the cause is thought to be postural, and the pain is not severe (Russell *et al*. 1993).

Occupation

The relation between low back pain and occupation has been recognized since at least 1705 (Allan and Waddell 1989). The 'railway spine' became recognized as being more common than backache in mariners, miners, and labourers (Allan and Waddell 1989). Jobs requiring physically heavy work, static work postures, frequent

bending and twisting, lifting and forceful movements, and repetitive work and vibrations (including those from driving vehicles; Frymoyer *et al*. 1983; Kelsey *et al*. 1984; Behrens *et al*. 1994) predispose to low back pain (Andersson 1981).

Psychological factors, such as monotony and dissatisfaction at work are implicated (Andersson 1981; Bergenudd and Nilsson 1994). Individuals who stated that they 'hardly ever' enjoyed their job tasks were 2.5 times more likely to report a back injury than those who 'almost always' enjoyed them (Bigos *et al*. 1991). Many blue-collar workers believe that work contributes to their back pain, although correlations with observed physical-work demands are not always present (Lindstrom *et al*. 1994).

There have been several studies on the frequency of back pain in hospital workers (Kaplan and Deyo 1987). Nurses are at high risk of back injury, being greatest for those doing the most physical work (aides and auxiliaries) and least for registered nurses. Indeed, nursing auxiliaries have a greater annual incidence of back pain than construction workers, garbage collectors, and truck drivers (Kaplan and Deyo 1987).

Risk of recurrent attacks/return to work

Factors influencing a return to work are sex, duration of sick leave, reported need for analgesics, pain in the dorsal and cervical regions of the spine (in addition to lumbar pain), patients' negative attitudes to their own capabilities, and after-work fatigue (Sandstrom 1986).

The recurrence rate of occupational back pain from any part of the spine in Canada was 20 per cent at 1 year and 33 per cent at 3 years. Men had a higher chance of recurrence. There was a lower risk of recurrence for thoracic symptoms than for cervical and lumbar symptoms, which had an equal risk. Drivers had the highest recurrence rate and nurses the highest average number of recurrences (Abenhaim *et al*. 1988).

In England, Troup *et al*. (1981) examined all patients returning to full-time work after sickness absence following an episode of low back pain or an accident at work. Residual pain in the leg and positive clinical signs on return to work, longer sickness absence, and two or more previous attacks were associated with recurrences. Falls were associated with a longer period of absence after the current attack and a higher rate of recurrence. Caldwell and Glanville (1980) followed a cohort of 373 patients under the age of 40 years presenting to general practitioners or hospital consultants in 1963. Only 11 per cent of patients had had a single attack, though the majority of attacks lasted fewer than 2 weeks. Only 33 per cent had had no time off work due to back pain over the 10 years, 3.3 per cent of men gave up work or took early retirement, 21 per cent changed job, 10 per cent took a reduction in working hours, and 19 per cent of men and 10 per cent of women had loss of income or potential income from their pain.

In Sweden, the pioneering work of Bergquist-Ullman (1977) showed that, in a population of workers at the Volvo plant, episodes of pain lasted less than 1 month in 35 per cent, less than 3 months in 87 per cent, and 4 per cent had pain at 1 year. The mean duration of pain was 35 days. Just over 60 per cent of this group had recurrences during the first year (mean 1.3 per patient).

A history of previous low back pain remains the single most useful predictor of future episodes of pain (Harber *et al*. 1994).

Other social factors

A high incidence of social factors has been found in back patients. Their social and economic situation is on average less good, a greater proportion suffer from drug and alcohol abuse, divorces and family problems are more frequent, and their educational level is often found to be lower (Andersson 1981; Croft and Rigby 1994). It is difficult to assess whether such factors are primary or secondary, and the interrelations with smoking and occupation are unclear.

Risk factors

Smoking is a risk factor for prolapsed lumbar intervertebral disc (Kelsey *et al*. 1984; Battie *et al*. 1991), as well as for the first attack (Biering-Sorensen *et al*. 1989) and the severity of low back pain (Frymoyer *et al*. 1983). A dose–effect relation was found between the severity of back pain and both the number of cigarettes smoked and the duration of smoking (Frymoyer *et al*. 1983). Possible explanations include impaired blood supply to the involved spinal segment (Frymoyer *et al*. 1983; Battie *et al*. 1991), carbonmonoxidaemia (Dimberg *et al*. 1989), coughing (Biering-Sorensen *et al*. 1989), and ergonomic factors related to smoking, which may be associated with people's attitudes, life-styles, and behaviour patterns (Frymoyer *et al*. 1983; Dimberg *et al*. 1989; Battie *et al*. 1991).

Vibrations around 5 Hz are the most troublesome in relation to back pain, which may explain the prevalence of back pain in those spending prolonged periods in motor vehicles (Editorial 1989b). A mismatch between physical stress and strength predisposes to back problems (Editorial 1989b).

Walsh *et al*. (1992b) have shown that a person under the age of 60 who is admitted to hospital for a traffic accident or a fall has a 7 per cent chance of developing low back pain as a result of the injury. The link between the injury and subsequent symptoms is often not obvious to the patient.

Health associations

Previous hospital admissions/operations are more frequent in those with a first episode of low back pain (Biering-Sorensen *et al*. 1989). People with back pain appear to have a propensity to use health services. Porter and Oakshot (1994) has shown that cardiovascular and gastrointestinal symptoms are more common in those with narrow spinal canals. Those with wider canals had more post-school qualifications than those with smaller canals. He postulates that the association between a smaller canal and impairment of health and certain intellectual abilities may result from an adverse environment that affects growing systems early in life (Porter and Oakshot 1994).

The associations with smoking (see above) may be linked with atherosclerosis of the abdominal aorta. Kauppila *et al*. (1994) examined 86 males at autopsy, comparing radiological evidence of disc degeneration with the presence of atherosclerosis of the abdominal aorta. They postulated that stenosis of the ostia of the segmental arteries may play a part in lumbar disc degeneration. These findings may explain the increased risk of dying from ischaemic heart disease in a cohort of Finnish farmers aged 30 to 49 years who reported low back pain or sciatica in a postal survey (Penttinen 1994).

The nature of low back pain

Waddell (1982) has helpfully classified low back pain into that caused by mechanical back pain (arising from the lumbar segment without direct nerve-root involvement); direct nerve-root involvement from structures within the lumbar segment; and spinal pathology, which includes inflammatory disorders such as ankylosing spondylitis, tumour, infection, bone disease (e.g. Paget's), and primary neurological disease (Table 6).

Table 6 Some causes of low back pain

Over 90 per cent of all episodes of low back pain may be considered idiopathic when non-spinal conditions (e.g. arthritis of the hip, proximal myopathy) and spinal pathology have been excluded. The important questions relate to whether the pain that arises from the lumbar segment does or does not involve the nerve root.

Many treatments for low back pain are based on the assumption (Frank and Hills 1988) that it is usually associated with excessive mechanical stress at work or leisure on normal structures, normal stresses on degenerate discs/facet joints, or the poorly resolved aftermath of an acute episode (e.g. a lifting injury). The physical nature of the problem has been identified in the epidemiological section above. Thus doctors and therapists have concentrated on physical measures to alleviate the problem. More recently, attention has been paid to psychological approaches to management. Whilst this has been predominately aimed at helping those with prolonged and chronic low-back pain, an intriguing use of the behavioural approach to back-pain management in acute situations is described by Fordyce *et al*. (1986).

Bremner and Simpson (1959) commented that “the patient's ability to 'live with his back' seems to be more important than the form of treatment”. It is likely that the longer an episode continues, the more important it is effectively to manage individuals and their problems. Most programmes for chronic back-pain sufferers and 'back schools' have a major emphasis on psychological management.

Few studies have followed patients over many years. In contrast to the pessimistic outcome described by Caldwell and Glanville (1980), Wilson and Wilson (1964), who studied 117 employees 10 years after an episode of back pain, found that 52 per cent had remained at work over 10 years without time off (including 50 per cent of the heavy manual workers), although over 80 per cent of those at greatest risk had recurrences. Only 13 patients were referred to hospital after failed therapy at work; only two changed job. They concluded that patients should be treated optimistically and not given bed rest or referred to hospital if this could possibly be avoided. This sentiment is now being repeated nearly 30 years later by Waddell (1987), to whom the 'back pain epidemic' is partly contributed to by doctors prescribing rest and partly by social concepts of disability, and who contrasts the traditional medical model of illness with a biopsychosocial model (Fig. 3).

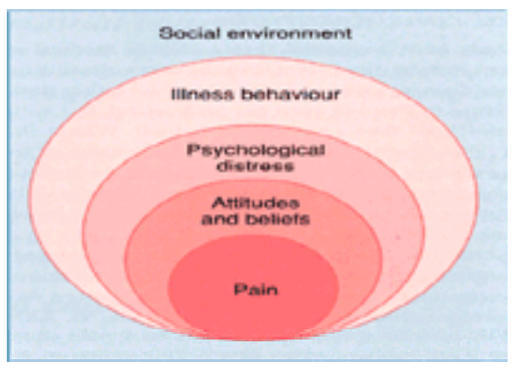


Fig. 3 A biopsychosocial model of low back pain disability — from Report of a CSAG Committee 1994.

Most attacks of acute low back pain are self-limiting and of short duration (Coste *et al*. 1994), although this picture has recently been challenged by Von Korff *et al*. (1993), who followed a cohort of patients in primary care for 1 year. They showed that 69 per cent of patients with pain of recent onset had had pain in the previous month, compared with 82 per cent of those with pain of greater than 6 months' duration. The important fact is that once an individual has had a major episode of back pain, the risk of a recurrent attack is always present (see Occupation above). Management must emphasize the importance of adopting strategies in the patient's life that minimize the risks of such attacks (see Table 7 and Table 8). At this stage a positive attitude is essential, minimizing rest and medication (Von Korff *et al*. 1994).


Things to avoid
Prolonged leaning forward, e.g. by use of long-handled garden tools; bath/change nappy/dress baby on high table with everything handy
Non-stop driving (longer than 30 min)
Soft or unyielding mattress
Low easy chairs
Lifting with a bent spine or with outstretched arms
Jerking or twisting the spine
Overstretching, e.g. reaching into high cupboards
Uneven loading of spine, e.g. shopping should be carried in two evenly weighted bags

Things to consider
Frequent changes in posture during routine activities
Seating modifications at work/hobbies
Use small lumbar supports in seats
Use castors on heavy furniture, e.g. fridges
Use webbings to assist lifting awkward furniture, e.g. wardrobe
Regular exercise, e.g. swimming
Sweep and brush with a sideways action from upright posture
Daily relaxation regime
Use of orthotic heelpads or wearing trainers (sneakers in North America) if standing or walking continuously

Source from Frank and Hills (1988).

Table 7 Some suggestions for those with spinal pain


Task	Method
Washing up	Raise bowl within sink or use high stool to sit on
Ironing	Use adjustable height ironing board and high stool
Loading washing machine/ dishwasher	Kneel to load, use front loaders; for top-loading washing machines use longs
Making beds	Perform kneeling — consider duvets
Cleaning bath	Clean from inside; kneel outside; use short-handled mops and rinse with shower attachment
Brushing teeth	Do sitting*
Washing hair	Stand upright using shower attachment
Picking up	From kneeling or squatting position, use 'helping hand' for light objects if unable to kneel, e.g. for osteoarthritic knees
For older people	
Getting up	High chairs
	High toilet seats
	Appropriate height of bed

Table 8 Hints to avoid bending and stooping

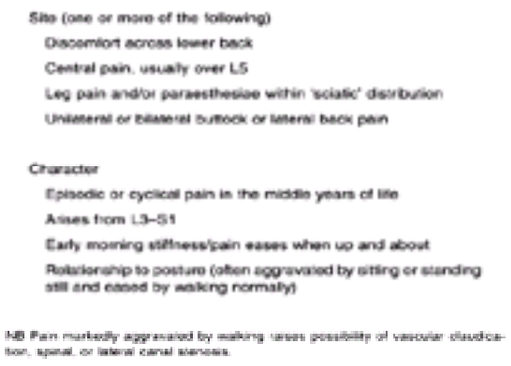
Differential diagnosis

General considerations

It has already been suggested that a strictly medical approach to management may be disadvantageous by concentrating on the exclusion of other diseases presenting as spinal pain rather than dedicating sufficient time to helping patients understand the nature of their problem and how they can best be helped. Whilst the measures outlined (Table 3) will help, much of the control of the situation needs to be grasped by the patient. Ultimately the patients' attitudes to avoiding aggravating factors, to exercise (Philips *et al*. 1991), and to a positive approach may be more important than physical management (Frank and Hills 1988).

It is important to balance safeguarding the requirement not to miss treatable organic pathology with the need to avoid unnecessary and costly investigations. This is best achieved by pattern recognition of common presentations of low back pain (Table 9). Although often diagnosed by exclusion, non-specific low back pain should be diagnosed on positive grounds (Table 9). Sometimes one side predominates, or more rarely the pain is entirely localized to a buttock or leg. Not infrequently the pain may be localized to the midline. Leg pain is not an uncommon presentation, usually felt within the sciatic distribution on the lateral border of the calf, or rarely the

foot. A careful history often gives a pattern of occasionally radiating up to the buttock or back. Sometimes paraesthesias give a clue as to the nature of the underlying pain in the presence of atypical leg pain. There are three important groups of patients to consider: those with non-specific low back pain ('backache'), nerve-root compression, and back pain caused by other conditions (Clinical Standards Advisory Group 1994a). More complex algorithms can be consulted (Waddell 1982; Gavin and Wiesal 1991), and a fuller review of the differential diagnosis is available (Jayson 1992).


Site (one or more of the following)
Discomfort across lower back
Central pain, usually over L5
Leg pain and/or paraesthesiae within 'sciatic' distribution
Unilateral or bilateral buttock or lateral back pain

Character
Episodic or cyclical pain in the middle years of life
Arises from L3–S1
Early morning stiffness/pain eases when up and about
Relationship to posture (often aggravated by sitting or standing still and eased by walking normally)

NB Pain markedly aggravated by walking raises possibility of vascular claudication, spinal, or lateral canal stenosis.

Table 9 Features of non-specific mechanical low back pain

Other factors suggestive of the symptom complex of mechanical low back pain are episodic or cyclical pain in the middle years of life arising from the L3–S1 area, associated with morning stiffness/aggravation of pain that lasts until the individual is up and about (Waddell 1982). A marked relation to body posture is common: the pain is often aggravated by sitting or standing still and eased by walking at a normal pace. Pain markedly aggravated by walking raises the possibility of vascular claudication, or stenosis of the spinal or lateral canals.

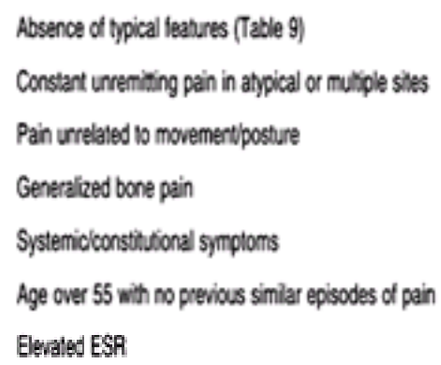
Thoracic pain is much less common than either lumbar or cervical pain (Table 4) and always requires investigation. Often, in older patients, vertebral collapse will be due to osteoporosis, but a high index of suspicion is required to exclude myeloma or secondary malignancy. Tuberculosis classically effects the lower thoracic or upper lumbar vertebrae (see Infections below).

Investigations—exclusion of other pathology

Age is important. Before the age of 25, radiographs are indicated to exclude congenital disorders such as spondylolisthesis (Waddell 1982). Inclusion of the sacroiliac joints will help exclude ankylosing spondylitis. The sudden onset of low back pain after the age of 55 warrants investigation. In a series of 900 patients presenting to an orthopaedic back clinic, 46 per cent of those over 55 had a definite abnormality, including 11 per cent with malignant disease (Waddell 1982). Patients over 55 presenting with episodes of pain similar to those suffered earlier in life do not usually require investigation.

Different clinics will have variable referral patterns. The relative rarity of infections and malignant disease is seen in Table 4: 825 consecutive referrals to a rheumatologist with an interest in back pain are reviewed with 1336 diagnoses; only nine tumours were found, just under 1.1 per cent of the 825 patients, not all of whom had spinal pain. It should be noted how even such a crude diagnostic index demonstrates the frequency of multiple rheumatic diagnoses. Ankylosing spondylitis and metabolic bone disease were the most common diagnoses other than non-specific or degenerative back pain.

Thus most patients with back pain will not require investigation to exclude pathology. In this situation, radiographs will only be needed for reassurance that serious pathology is not present, to explain possible mechanical factors to the patient, or to assist physiotherapists in their management. Usually blood tests are not helpful, but when in doubt an erythrocyte sedimentation rate is the most helpful screening test (Waddell 1982). Table 10 outlines factors that should alert physicians to the possibility that the cause of back pain may be sinister.


Absence of typical features (Table 9)

Constant unremitting pain in atypical or multiple sites

Pain unrelated to movement/posture

Generalized bone pain

Systemic/constitutional symptoms

Age over 55 with no previous similar episodes of pain

Elevated ESR

Table 10 Features suggestive of malignancy

The same principles hold true for those with a previous history of cancer, as they will frequently suffer from mechanical low back pain. As radiographs may not show deposits till late in the disease, a higher threshold for radiological investigation is needed (Table 11). In this situation a bone scan is often the most helpful investigation unless the level of the lesion is clearly established on clinical grounds.

Table 11 Guidelines for radiological investigation

Leg pain

There is a widely held belief that pain in the leg invariably reflects nerve-root pain and usually is caused by disc disease. That this is not necessarily so was demonstrated by Mooney and Robertson (1976) with their pioneering work on the facet joint. They demonstrated that injection into the facet joints caused intense local pain that subsequently radiated posteriorly down the leg and sometimes into the foot. Hamstring spasm subsequent to this was abolished by local anaesthetic. In

some patients, previously reduced tendon reflexes returned to normal. Thus leg pain (even below the knee) can arise from the facet joint (Mooney 1992) and probably most structures within a segment.

Leg pain caused by root irritation is generally more clearly defined, sharper, and often has an element of paraesthesia (Waddell 1982).

Extraspinal causes of pain

Three per cent of apparent back troubles presenting to an orthopaedic clinic are due to extraspinal causes such as retroperitoneal or pelvic pathology, hip disease, peripheral vascular disease, or primary neurological disease (Waddell 1982), and need appropriate investigation.

Pelvic pathology is not easy to exclude on history as many women notice that their pain is more severe in the last few days of their menstrual cycle, easing off on the first or second day of their period. This may also be noted in women with neck pain or headaches and is non-specific. Where there is doubt, abdominal and pelvic examinations are required.

Spinal causes

Children and adolescents often require investigation. Acute pain may be caused by spondylolisthesis or by a developing scoliosis and warrants referral to a specialist orthopaedic clinic. In the absence of spinal disorders, adolescent discitis may cause severe pain and only be diagnosed retrospectively. Physiotherapy is important in this condition.

Spondylolisthesis is not uncommon and is usually unrelated to low back pain. At the L5/S1 level it is usually stable and does not require follow-up. At the L4/L5 level, however, instability may develop (diagnosed by lateral lumbar spine views in flexion and extension) and on occasions requires fusion, particularly if root involvement is present.

Lumbar disc disease

This is considered separately in Chapter 5.18.1, but it is the most frequent major pathology seen in back pain clinics and requires some consideration here. Firstly, physicians must know when to investigate to ensure that they are not missing surgically correctable pathology. It is the presence of nerve-root compression that usually benefits from surgery by decompression. Here, leg pain usually dominates over back pain. The presence of symptoms correlating with objective physical signs (usually of loss of sensation, power and/or reflexes), confirmed by myelography or an equivalent scanning technique, suggests that surgery should be 95 per cent successful (Bell *et al*. 1984). Secondly, it is important to recognize that a central disc prolapse may give bilateral leg pain in the absence of demonstrable signs and with a normal straight-leg raise test. Sphincter disturbances may reflect spinal-cord compression and often require emergency treatment.

Spinal stenosis

Spinal stenosis is a symptom complex of root pain and sensory or motor symptoms that come on during walking and that pass off after a few minutes of sitting down or flexing the spine. Symptoms may be aggravated by extension (Waddell 1982). Although occasionally symptoms are associated with a congenitally narrow spinal canal, more often additional pathology develops in later life. This may reflect disc disease or osteoarthrosis of the facet joints separately or together combining with hypertrophy of the soft tissues (particularly the ligamentum flavum) to compress the cauda equina. If such compression is predominately lateral, usually due to bony outgrowths from an osteoarthritic facet joint, the term 'lateral canal stenosis' is sometimes used. There may be narrowing of the exit foramen, compressing the root, and resulting, when bilateral, in a trifoliate shape to the canal (Porter 1992; see Fig. 4).

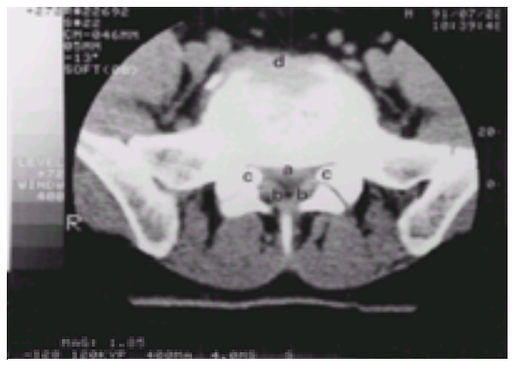


Fig. 4 Spinal stenosis secondary to a combination of degenerate disc protruding posteriorly and laterally (a) and hypertrophy of the ligamenta flava (b) in an individual with osteoarthritis in the facet joints intruding into the lateral recesses (c) to produce symptoms of spinal stenosis. The anterior margin of the disc is degenerate (d).

Infections

Brucellosis is a disease of those working in abattoirs, or with animals, such as vets. Back pain may be very generalized and often associated with evidence of systemic disease. In contrast, most infections of the spine are clearly localized and give rise to constant pain, often in an unusual or atypical site.

Tuberculosis classically affects the lower thoracic or upper lumbar vertebrae and must always be considered for people at risk for this disease, for example, in the United Kingdom, those arrived as immigrants from the Indian subcontinent within the last decade, the frail elderly, or those immunocompromised from therapy or for other reasons. This is in distinction to degenerative lumbar disease that usually affects the L5/S1 or L4/L5 levels (Kelsey *et al*. 1984) or less commonly L3/4. Other organisms may be found, such as staphylococci. Where suspicion exists without localizing changes on plain radiographs and no certain spinal level found, a bone scan may be helpful, although not all tuberculous lesions produce a hot spot. If a level is suspected, CT or MRI will confirm abnormalities. Biopsy under radiological guidance or occasionally an open biopsy will usually provide microbiological and histological evidence of the causal organism. Confirmation of tuberculosis is not always obtained and occasionally a trial of antituberculous chemotherapy is needed.

Tumours

Primary tumours of the spine are rare and outside the scope of this chapter. Benign tumours may be found in the upper lumbar spine and are easily missed if too much reliance is placed on CT scanning as the reliability of CT depends on the correct levels being scanned. Myelography or MRI may be required to exclude such tumours (Bell *et al*. 1984).

Secondary tumours from myeloma, lymphoma or carcinoma present to all back clinics. Suspicion is raised (Table 10) by constant unremitting pain in unusual or multiple sites in an individual with systemic symptoms. Routine haematological and biochemical investigations will usually confirm the presence of disseminated disease. Spinal investigations will follow the pattern outlined above for infection (Table 11).

Ankylosing spondylitis

This is the most common specific condition to present in a back-pain clinic: it is more frequent in men than women and usually presents at a younger age than mechanical or discogenic back pain. Stiffness, particularly in the morning, is usually a cardinal symptom; it is usually relieved by movement and/or exercise. It is important to diagnose ankylosing spondylitis as early as possible as its treatment differs from that of non-specific low back pain and because of its family and other associations. A vigorous exercise programme maintains spinal movement (thus preventing spinal contracture) and may inhibit the process of calcification at the enthesis. There is usually a good response to non-steroidal anti-inflammatory drugs, particularly if a long-acting preparation is taken after food last thing at night as it

inhibits the early-morning pain and stiffness. This facilitates performance of the recommended early-morning exercise programme. In the early stages, plain radiographs of the spine and sacroiliac joints are normal, and CT or MRI scanning of the sacroiliac joints may show early erosive change. Less specific changes may be found by bone scans of these joints. Less commonly the disease presents with localized calcified segments of the spine, even in the neck.

The genetic predisposition to the disease is associated with the presence of the *HLA-B27* genotype. This is not diagnostic of the disease but its detection may be helpful in those patients with a very suspicious clinical picture and negative investigations. Its presence is always compatible with, but never diagnostic of, ankylosing spondylitis. Its absence does not exclude the diagnosis.

Bone disease

Metabolic diseases of bone may be found in the back-pain clinic. Osteoporosis must be considered in women who had an early menopause or have other risk factors. In older patients, as mentioned above, vertebral collapse will be due usually to osteoporosis but myeloma or secondary malignancy may need to be excluded (Table 10). If plain films of the lumbothoracic spine are normal, bone densitometry will determine the bone mass. In rheumatological practice, osteoporosis secondary to steroid use is not infrequent, particularly in older women with polymyalgia rheumatica or arteritis.

Osteomalacia is rarely seen in the United Kingdom, except among female immigrants from the Indian subcontinent who may spend much of their time indoors. An elevated alkaline phophatase in the presence of normal liver function tests confirms the diagnosis. The generalized nature of the pains of osteoporosis and osteomalacia generally makes the diagnosis relatively simple.

Paget's disease is occasionally destructive and then complicates the diagnosis as infection and tumour lie within the differential. Biopsy is sometimes required. More frequently it is found on radiographs, when its significance is uncertain. Activity may be demonstrated on bone scan and by an elevated alkaline phosphatase.

Sickle-cell disease must also be considered in populations at specific risk (Ozoh *et al*. 1990).

Referred pain to head, thorax, and abdomen

This chapter will not detail the clinical features and management of cervical and thoracic pain, but it is important to note the range of presentations of spinal pain. Cervical pain may refer to the scalp, particularly the bifrontal areas, to present as headaches. Facial pain may arise from involvement of the upper cervical segments, or from the extensions of the Vth nerve nuclei into the cervical cord. Dizziness is not infrequent and may not always be due to vertebrobasilar insufficiency. Pain frequently radiates to the shoulders, where it may contribute to symptom complexes associated with capsulitis or painful arc syndromes (perhaps by involvement of the sympathetic nerves), and is more frequent in those with narrow-diameter cervical canals. Similarly, epicondylitis and carpal tunnel syndromes are associated with narrow cervical canals. Less appreciated is the referral of pain, which may be very severe, to the anterior chest wall (Hockaday and Whitty 1967) where it is often superficial, unilateral, and associated with tenderness over the pectoral muscles, ribs or costochondral junctions—a cause of many visits to accident and emergency departments for fear of ischaemic heart disease (Taylor and Dawes 1990).

Acute abdominal pain referred from the spine is usually a diagnosis of exclusion, although a careful history of pain radiating from or to the spine may be obtained. Sometimes the pain can be reproduced by rotation of the spine, or by spinal palpation/mobilization by trained manipulators. Often patients have had unnecessary investigations for gastrointestinal or renal disease.

Acute low back pain

Although acute low-back pain develops insidiously in some, or is noted on waking in others, many episodes have a clearly defined cause that must be noted and explained to the patient. Trauma or inappropriate spinal movements may not give rise to severe pain until the next morning. Where an incident or injury has taken place more than 48 h before the onset of symptoms, experience suggests there is unlikely to be a causal relation, whilst acute leg pain may develop after a period of up to 2 weeks. Provision of educational literature at this stage may prevent referral to hospital, referral for physiotherapy, admission to hospital, and need for laminectomy (Roland and Dixon 1989).

Patient compliance may also be increased if the possibility of progression of pain with occasional need for surgery is discussed at the initial consultation (Hickson 1983).

Patients without leg pain fare better than those with leg pain, which may develop after the onset of pain (Chavennes *et al*. 1986).

The crucial therapeutic interventions in general family practice are to take the patient seriously, and to take a positive view (Roland 1994). In the absence of root signs, or suggestions of a serious problem, advice that activity is helpful, clicks are not harmful, and that pain is not necessarily bad are all important.

Bed rest and getting going

Bed rest

Minimal bed rest only is encouraged, particularly in the absence of nerve-root irritation. A random sample of patients with back pain (mostly acute) from the United States resting for 2 days did as well as those resting for 7 days (Deyo *et al*. 1986). A comparison of bed rest with physiotherapy and education from family practices in Canada showed no difference between the two groups, suggesting bed rest was not advantageous (Morrison *et al*. 1988). Bed rest is usually best in the position of maximum comfort. Often flexing both hips and knees tilts the pelvis, flattening the lordosis with immediate reduction of pain intensity (see Fig. 5). Supporting the hip and knee of the affected leg with cushions or pillows, or even sleeping in a chair, may be appropriate in the presence of femoral root irritation.

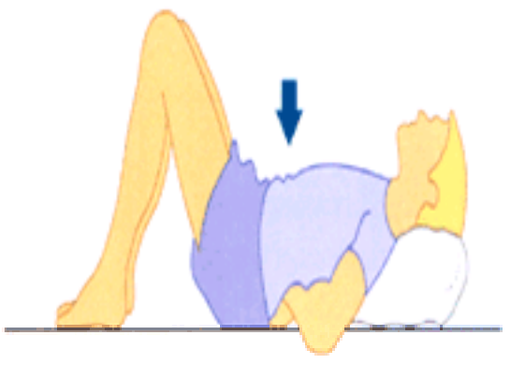


Fig. 5 Flattening of the lumbar lordosis is often helpful in easing the pain following an acute attack; some exercises encourage adoption of this position (reproduced by kind permission of the Northwick Park Hospital Department of Physiotherapy).

For those with very severe attacks of pain, or the elderly, attention must be paid to good-quality rest and adequate support during mobilization. The mattress should not sag, the height of the bed should be appropriate, and domiciliary physiotherapists should help take the mechanical stress off the spine with walking sticks or elbow crutches. The positions of rest are important, sitting often aggravates acute pain.

Mobilization after bed rest

Provision of a corset may greatly facilitate initial mobilization. Social services' or rehabilitation staff may be able to provide blocks to raise chair or bed, a more

appropriate chair, and a toilet raise to facilitate mobility at home. Prolonged bed rest is particularly contraindicated for the very old; joints stiffen up and osteoporosis is enhanced. For this group of back sufferers, domiciliary physiotherapy greatly improves confidence and self-image. A commode by the bed is often essential to prevent long walks to the (often cold) toilet. As with other forms of arthritis, the stairs may become a major barrier to health if the only toilet is upstairs. Elderly people often have coincidental urinary problems (often related to diuretics) and repeated visits upstairs put unnecessary strains on the spine and legs, which may be reduced by banisters (stair rails), having a bed downstairs, and the use of a commode (Frank and Hills 1988; Turner-Stokes and Frank 1992).

When the patient starts to improve, referral for physiotherapy is usually helpful (see below). A contraindication is that of exacerbation of symptoms related to sitting whilst travelling to hospital, particularly if there is involvement of an L5 or S1 root, when domiciliary therapy may still be advised.

Drugs

Analgesia

Ensuring compliance and getting people back to work cannot be done if they are in severe pain. Important judgements are needed as to when, how much, and for how long medication should be prescribed to obtain the best balance of benefit versus risk in taking medication.

The danger is that people will take unnecessary rest to control pain that could be helped by analgesics and that would enhance the ability to regain normal activity patterns quickly. As an inflammatory component is highly likely in an acute episode, a non-steroidal anti-inflammatory drug is often helpful (Deyo 1983), particularly to obtain a reasonable duration of analgesia for a good night's rest. A recent example of a good study is that of Szpalski and Hayez (1994). Non-steroidals are often the only non-opiate derivatives available giving analgesia that lasts through the night. The risks of short courses may be no more than the risks of prescribing the stronger opiate derivatives (see below), which may be the only alternative. Care must be taken in smokers, those with a history of ulcer disease, severe dyspepsia, or elderly people.

Many people suffer side-effects arising from the gastrointestinal tract when taking non-steroidal anti-inflammatory drugs. Whilst omitting drugs giving side-effects is always the best policy, sometimes this is not in the patient's best interests. Here, men or older women may be helped by taking misoprostol separately, or combined with a non-steroidal anti-inflammatory drug such as diclofenac or naproxen. 'Pro-preparations' may have fewer side-effects. Alternatively, side-effects may be minimized by prescribing acid-suppression tablets. Where gastric pathology has been confirmed, for example erosions, omeprazole is the drug of choice. If duodenal pathology is present, ranitidine may be preferred.

Non-steroidal anti-inflammatory drugs can be augmented by analgesia that occasionally may require opiate derivatives (e.g. nefopam, meptazinol or buprenorphine) if the compound paracetamol containing preparations are insufficient. Many require only mild analgesics (e.g. compound paracetamol preparations). Antiemetics are helpful in those inclined to vomit after drugs, as vomiting will almost certainly inhibit healing and aggravate pain. Opiates should rarely if ever be prescribed in the home for acute or chronic back pain. Failure to control pain with these measures requires admission to hospital.

Other drugs

Antispasm medications may be used in the first few days (Deyo 1983), although non-steroidal anti-inflammatory drugs seem more appropriate. Benzodiazepine drugs may be of value for a maximum of 1 week for those having difficulty resting. They may facilitate a good night's sleep, essential to enable individuals to cope with the next day's pain. Those with a history of mood disorders, insomnia or dependency are treated more safely with a tricyclic compound with sedative properties, for example amitriptyline.

Injections

Most interest has been in the effect of local injections of lignocaine and/or corticosteroid around the area of maximal tenderness (Garvey *et al*. 1989; Collée *et al*. 1991), or into the disc (Simmons *et al*. 1992) or facet joint (Lillius *et al*. 1989; Carette *et al*. 1991). There is no good evidence that they offer widespread benefit. The local effect of a mechanical stimulus may be more important than the drug injected (Garvey *et al*. 1989), and injections given by specialists in clinic may be more efficacious than those given in a general practice setting (Collée *et al*. 1991). Injections have been combined with manipulation with good effect (Blomberg *et al*. 1992).

Injections by the epidural route are discussed in the section on intractable pain.

Physical management

Considerable evidence supports the role of physical therapy in acute back pain. Increased compliance with better results from physiotherapy can be obtained by simultaneous use of complementary literature and by planned review of patients after a course of treatment. Benefits of therapy may be lost in those who do not persist with recommended treatment (Manniche *et al*. 1991). Few studies have compared physical with non-physical methods of treating acute low-back pain. In part this probably reflects the difficulties of comparing two differing modes of therapy when the type of therapy cannot be 'blind'. Overman *et al*. (1988) randomly assigned 107 patients in a 'walk-in low back pain clinic' to internists or to physical therapists. The latter referred more patients to physical therapy than did internists and recommended less bed rest and medication. Those managed by the physical therapists expressed greater satisfaction with care and had significantly greater functional improvement.

Meade *et al*. (1990), in a sample of 741 patients for whom manipulation was not contraindicated (derived from hospital or chiropractic clinics), compared hospital outpatient management with that of chiropractors. In the group with symptoms of less than 1 months' duration (acute and subacute back pain), no difference was noted until 2 years, when those treated by chiropractors appeared less disabled. A preliminary study of osteopathic manipulation in acute/subacute low-back pain has suggested benefit (MacDonald and Bell 1990). Reviews of manipulative treatment show that manipulation may speed recovery of a few with acute backache (Koes *et al*. 1991; Shekelle *et al*. 1992; Frank 1993a), views supported by more recent studies (Blomberg *et al*. 1992; Koes *et al*. 1992).

A prospective, randomized study of manipulation, transcutaneous muscle stimulation, massage, and corset showed minimum drop-out in the manipulation group and maximum drop-out in the muscle-stimulation and corset groups (Pope *et al*. 1994). This supports the criteria for use of a corset suggested by Frank and Hills (1988), that corsets should be reserved for those with failed manipulation.

A novel approach by Fordyce *et al*. (1986) compared traditional and behavioural methods of treating acute back pain of 10 days' duration or less and showed no difference at 6 weeks. But at 9 to 12 months those given more supervision and planned withdrawal of treatment (behavioural group) were 'less sick' and claimed 'less impairment' than the group able to control their own schedules of medication, activity, exercise, and follow up (traditional group). This shows the benefits of taking control from the patient and giving it to the therapist.

McKenzie techniques of passive extension and postural correction have been compared with a one-session 'mini back school'. This school consisted of a 45-min session only and perhaps was an unfair comparison to a technique involving considerable therapeutic input. McKenzie techniques were shown to be superior in relieving pain, promoting return to work, reducing sick leave during the initial attack, controlling recurrences during the follow-up year, and promoting movement (Stankovic and Johnell 1990).

A graded-activity programme with a behavioural therapy approach for those with subacute back pain improved mobility, strength, and fitness, resulting in an earlier return to work than in a control group (Lindstrom *et al*. 1992). Mitchell and Carmen (1990) and Mitchell and Carmen (1994) also included elements of 'behavioural support' in their sports injury approach, 'functional restoration', basically hard exercise. This approach reduced absence from work, compensation costs, and disability award costs.

Many studies have not differentiated between acute and chronic back pain. Whilst this can be argued as failing to differentiate between totally nociceptive pain and pain where other factors may be of importance (e.g. pain/sickness behaviour), in practice it has been shown that many people respond to standard conservative physical measures in spite of pain of long duration (Meade *et al*. 1990; Koes *et al*. 1992). Unfortunately, many studies do not define the duration of the current

episode of low back pain or the time of onset of the first episode of pain.

As cervical symptoms often coexist with low-back symptoms, advice about physical and ergonomic measures must be appropriate for all parts of the spine. Thus, as mentioned earlier, lying prone may help some people with low back pain, but aggravate neck problems that have may been mild or asymptomatic.

Exercises may be used in acute pain (see below).

The elderly

Back pain in elderly people is a strangely neglected topic. Its management did not feature in a 330-page review of 'arthritis in the elderly' (Kean 1986). The deleterious effect of musculoskeletal disorders is important (Williams 1985; Hodkinson 1988; Turner-Stokes and Frank 1992).

A recent review of 187 patients referred to my back clinic revealed that 21 per cent were over the age of 65. Some had major degenerative problems (spinal stenosis and root compression). Many also had additional medical problems that complicated their management, for example cardiac and gastrointestinal diseases, peripheral arthritis, and cervical degenerative changes. The differential diagnosis can be difficult in old age when acute pain may be due to osteoporotic collapse or to aggravations of disc or facetal degenerative disease. The two diseases may coexist.

Prolonged rest is usually contraindicated as joints stiffen and osteoporosis is encouraged. For this group, domiciliary physiotherapy greatly improves confidence, mobility, and self-image. If rest is required, attention must be paid to resting posture (usually avoiding sitting), taking pressure off protruding bony points, and adequate support during mobilization. The domiciliary physiotherapy service may assess the patient for their ideal resting posture. Resting sitting is preferable if the femoral roots are involved. A checklist is given in Table 12 to ensure that the everyday stresses on the spine are minimized, advice being given by the domiciliary physiotherapy service, social services department or the hospital occupational therapy service, separately or in combination (depending on local circumstances).

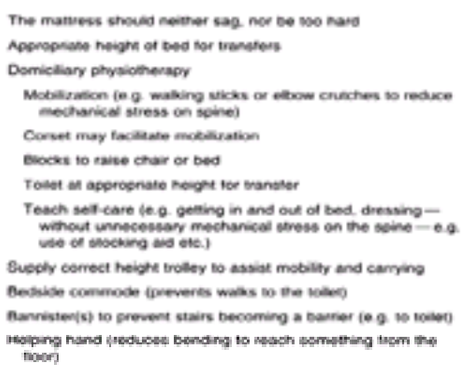


Table 12 Checklist for supporting the elderly person with back pain

Particularly for the old person with acute or chronic back pain, a visit to the home by a physician or a member of the primary-care or community team is an ideal method of teaching self-care. Self-care is not likely to be absorbed in an environment away from the home. A simple check that the tasks of everyday living can be performed without placing unnecessary mechanical stress on the spine (e.g. getting in and out of bed, putting on trousers, shoes or slippers) may encourage a speedier resolution of pain and minimize risks of recurrence if the advice is understood and maintained by the patient and their partner, friend or care assistant.

Elderly patients should be encouraged to acquire a 'helping hand' device (Frank and Hills 1988), which reduces their need to bend the spine to reach something from the floor.

Care needs to be taken with medication. Trials with the simplest analgesics initially may give adequate relief. The risk:benefit ratio in elderly patients shifts towards greater side-effects, requiring increased vigilance on the part of the physician (Hodinka 1991; Girgis and Brooks 1994). Ibuprofen is invariably found safer than any it is compared with in terms of its gastrointestinal toxicity (Bateman 1994; Committee on Safety of Medicines 1994). Ketoprofen has been variously reviewed (Committee on Safety of Medicines 1994) but has been found not to influence control of hypertension (Weiss *et al*. 1991) and to be well tolerated by an elderly population, mean age 72 (Schattenkirchner 1991).

Prevention (Table 13)

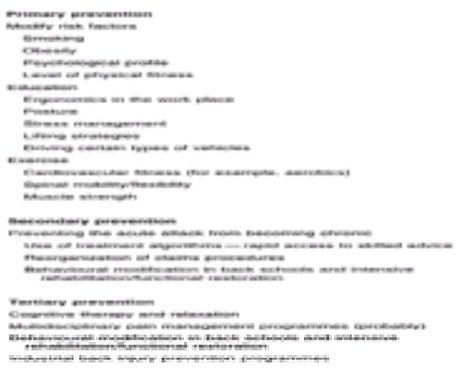


Table 13 Strategies for preventing back pain

Primary prevention

There have been no studies to show the effectiveness of psychological methods of preventing low back pain, although corporate stress-management programmes at three high-risk work sites have reduced the number of accidents and related costs (Weiser and Cedraschi 1992). There have been no studies on the benefits of stopping smoking on low back pain.

A review of ergonomic changes in the United States railroad industry (where the physically demanding nature of the job results in numerous musculoskeletal injuries) showed low back injuries to be the most common. They are also most costly in terms of compensation and productivity. Corrective measures have included storing tools and materials off the ground at between knee and shoulder height, with the heaviest items at knuckle height; devising winches to lift and handle heavy equipment; and designing work-tables, dollies, and carts to handle more easily railroad-car parts and tools. Management took a participatory approach and encouraged workers to design their own tools and material-handling devices to make their tasks easier and safer. This approach resulted in low back injuries and lost work days falling to zero and absenteeism falling from 4 to 1 per cent (Brown *et al*. 1991).

Teaching and social support, by themselves, may not be enough to alter behaviour (Daltroy *et al*. 1993). In a hospital environment, teaching plus ward instruction on lifting techniques may change behaviour of nurses in the short term (Feldstein *et al*. 1993).

Secondary prevention

This topic is addressed in the sections on back schools and intensive rehabilitation, and has been reviewed by Weiser and Cedraschi (1992). It is the combination of intensive physical training, behavioural principles, and attention to the working environment that appears to be effective. Back schools may be more effective in the work environment (Weiser and Cedraschi 1992).

Tertiary prevention

Weiser and Cedraschi (1992) have argued that studies reporting follow-up data of 6 months' duration may be considered preventive. Thus Turner and Jenson (1993) showed at 6 and 12 months that patients given cognitive therapy, relaxation, and cognitive therapy with relaxation all improved significantly compared to a waiting-list control group. There are few suggestions of the efficacy of multidisciplinary pain-control clinics (Weiser and Cedraschi 1992). In the United States, but not necessarily elsewhere, intensive rehabilitation in the form of 'functional restoration programmes' appears effective (Mayer *et al*. 1987).

Once back pain has become chronic or intractable, psychosocial and behavioural interventions are necessary for successful rehabilitation, probably combined with intensive physical programmes that are designed to get people back into employment. Back education from specially trained physiotherapists during pregnancy reduced the amount of back pain and sick leave during pregnancy, and back pain post partum (Ostgaard *et al*. 1994). In California, a 1-year back-injury prevention programme consisting of a combination of education, training, physical fitness activities, and ergonomic improvement showed a net benefit for introducing the programme of $161 108, with a return on investment of 179 per cent (Shi 1993).

Chronic low back pain

Exercises

Exercise regimens may aim to increase range of movement, strengthen muscles, stretch tightened structures, or toughen up physically and mentally (Frank and Hills 1988; Jenkins and Borenstein 1994). Allowing for difficulties in methodology and interpretation, there is evidence that back extension (Manniche *et al*. 1988; Manniche *et al*. 1991; Hansen *et al*. 1993; Manniche *et al*. 1993; Risch *et al* 1993), calisthenics (Donchin *et al*. 1990), and a mixed exercise regimen help people with low back pain (Deyo *et al*. 1990). Isometric flexion exercises (Hume Kendall and Jenkins 1968; Zylbergold and Piper 1981), mobilizing exercises (Hume Kendall and Jenkins 1968), and calisthenics (Donchin *et al*. 1990) are beneficial. Deyo *et al*. (1990) advised a wide variety of exercises, including relaxation, in a pattern similar to some back schools. The McKenzie programme aims to centralize pain through an individualized exercise programme and is twice as effective as traction and back schools alone (DiMaggio and Mooney 1987).

There is little evidence to support the idea that particular theoretical rationales have been justified as patients usually have not been selected to meet criteria to justify particular exercises. However, Manniche *et al*. (1993) has shown that active back-extension exercises appear most effective in women with sedentary jobs, who might be expected to have less physical strength.

Exercises combined with behavioural methods are more effective than exercises alone, both being more effective than inaction (Turner *et al*. 1990). Behavioural principles have also been shown to reduce sickness behaviour (Alaranta *et al*. 1994) and get people back to work quicker (Lindstrom *et al*. 1992).

Both active-flexion and active-extension exercises may increase the range of movement of the spine, but not necessarily reduce the pain and disability (Elnaggar *et al*. 1991).

Aerobic exercise is increasingly recommended and has been shown to augment the benefits of an English back school (Frost *et al*. 1995). Benefits of exercise may not only be physical. There is a suggestion that specific muscle strengthening or aerobic fitness training may improve nurses' job satisfaction as well as decreasing the duration of back-pain episodes (Kaplan and Deyo 1987).

Intensive outpatient physical retraining consisting of pain relief and mobilization, increasing movement and muscle strengthening, and work conditioning has been shown to reduce work absence. The increased costs to health care were more than offset by savings in 'wages loss cost' (Mitchell and Carmen 1990; Mitchell and Carmen 1994). Work conditioning has not been separately evaluated but its use in intensive rehabilitation settings is believed to be of value (Frank *et al*. 1987).

Manual therapy and manipulation

Manipulation (Fig. 6) has been defined as 'abrupt passive movement of a vertebra beyond its physiological range but within its anatomic range' (Spitzer 1987), although some would include movements that are active with the patient moving in one direction whilst the manipulator applies force in another. In all reports, readers should note the methodology as all will use slightly different definitions and varieties of manipulation.

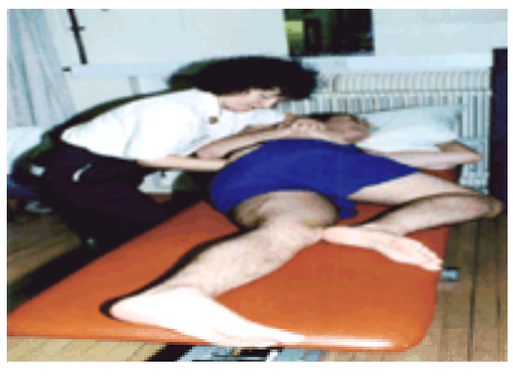


Fig. 6 Mobilization or manipulation of the lumbar spine is helpful in reducing morbidity and duration of symptoms in acute and chronic, but not intractable low back pain (reproduced by kind permission of the Northwick Park Hospital Department of Physiotherapy).

Shekelle *et al*. (1992) and Shekelle (1994) reviewed the studies on manipulation in acute low-back pain. Meade *et al*. (1990), Koes *et al*. (1992), and Blomberg (1992) give support to the view that manual therapy including manipulation is effective in reducing pain of longer duration. Mathews *et al*. (1987) included patients with acute and chronic pain in their study, which demonstrated early benefits from manipulation.

The effect of manipulation in acute pain has been referred to above. Hadler *et al*. (1987) have suggested that manipulation may be more effective in pain of longer duration, and this view is supported by Meade *et al*. (1990) in their comparison of hospital outpatient management with that of chiropractors. That study showed major long-term benefit in a group of people subjected to chiropractic manipulation compared to a group treated by a combination of outpatient hospital techniques, which included hydrotherapy and traction in addition to manipulation/mobilization by Maitland or Cyriax techniques. The benefits were mostly with those having chronic or severe pain. Although it has been argued that their results could be explained by increased numbers of treatment sessions, greater level of skill (more experienced manipulators) or greater communication or interpersonal skills, in my view the type of manipulation was likely to play a part as it is compatible with the theory that loosening stiffened spinal segments may ease current, and reduce subsequent, symptoms by redistributing forces that act on the spine more evenly.

Individually, these studies do not prove that these forms of manual therapy are effective. All are capable of many interpretations. Together, however, they lead me to conclude that these forms of therapy are effective. The beneficial effects on the economy of 'a few days' saved for work by more effective means of treatment are potentially enormous.

Similar arguments might have been used in a controlled comparison of short-wave diathermy with osteopathy (Gibson 1985); here, however, no significant differences were found. Mathews *et al*. (1987) compared manipulation and placebo infrared for those with back pain with or without impaired straight-leg raising and showed the beneficial effects of manipulation between the second and eighth days of therapy.

Other physical treatments

Corsets are frequently prescribed for chronic low back pain, although criteria for their use are not clear. Our criteria (Frank and Hills 1988) assume that they should be reserved for those patients in whom attempts to mobilize the spine have failed. They definitely help some patients in this category (Million *et al*. 1982). Other logical categories include those with pain on physical activities (particularly if it facilitates remaining at work) and those with scoliosis or unstable spondylolisthesis for whom symptoms persist and surgery is not indicated. More recently, Kaplan and Sinaki (1993) have piloted the use of posture training supports, which may be more acceptable than ordinary corsets and help those with a kyphosis. They may be useful adjuncts to other forms of physical management. Corsets are likely to be less effective than active therapy for unselected patients (Hsieh *et al*. 1992).

Traction may also be helpful in a few patients, though demonstration of its benefits in controlled studies is lacking. Patients in hospital may find bed rest in traction upsetting and avoiding this has been recommended (Clinical Standards Advisory Group 1994*a*).

Heterodox management

Whilst most experienced back-pain practitioners agree on pattern recognition of symptoms, many disagree about likely cause and rationale of treatment; this has been summarized elsewhere (Frank 1990). It is crucial that those claiming to help people with low back pain are prepared to allow their treatments to be evaluated. It is likely that osteopathy will be shown to be effective (MacDonald and Bell 1990). A number of studies on chiropractic have now been performed (Assendelft *et al*. 1992); the best on methodological grounds was that of Meade *et al*. (1990), which has shown that orthodox practitioners cannot afford to be complacent. Conversely, the public clearly needs safeguards against those using unproven techniques that may be harmful or unhelpful. Doctors are best advised not to recommend heterodox treatment unless the practitioner is a member of a recognized professional association (Frank 1990).

Medication

The role of non-steroidal anti-inflammatory drugs was covered in the section on acute pain. The risks of serious side-effects from these drugs are probably less than 2 per cent, but as the risks become more clearly understood, so one becomes more cautious. It is worth attempting to wean patients off non-steroidals, giving them the opportunity to have 5- to 10-day courses when necessary for exacerbations of pain (e.g. after lifting or bending inappropriately). The use of non-steroidal anti-inflammatory drugs in the elderly was discussed above.

Similarly, longer-term analgesia may encourage sedation or adverse effects on the gastrointestinal tract and many patients discontinue their use. Having a bottle of tablets that are known to be effective 'on hand' is prudent, and the knowledge that they are available helps counter the fear of aggravation with activity and may inhibit the risks of depression (Frank and Hills 1988).

One of the major advances in the management of chronic pain has been the use of tricyclic antidepressant compounds (Pheasant *et al*. 1983; Ward 1986). Their mode of action is unclear; a suggestion is that they increase serotonergic uptake in the central nervous system (Ward 1986). Benefits are shown for those who are or are not depressed and for those known to have a physical basis for their pain. The effects appear to be independent of any sedative properties of the drugs. However, the sedative side-effects can be used advantageously in those having disturbed sleep and patients often find they facilitate relaxation. Appropriate warnings about sedation and a dry mouth must be given. Patients will need reassurance that tricyclic antidepressants are not addictive and that they are not being prescribed for psychiatric purposes.

Intractable low back pain

A few patients have persistent, disabling pain despite therapy (usually including non-steroidal anti-inflammatory drugs, analgesics, physiotherapy and often advice from heterodox practitioners). Investigations (often including MRI or radiculography) are usually negative. Having excluded spinal and non-spinal pathology and direct nerve-root involvement, patients should be considered as having 'failed conservative treatment'. The term 'intractable' low back pain is used to differentiate failed conservative management from chronic pain, which may never have been treated before, or 'chronic pain syndrome', when changes in behaviour are apparent (Frank 1993a). Of those with failed medical management, up to 10 per cent may have a major underlying medical problem (Wiesel *et al*. 1988).

At this stage, it has to be determined whether physical or psychological factors are most likely to be influencing current symptoms (Frank and Hills 1988; Philips and Grant 1991). Often both will coexist, but on occasions it will be clear that either the physical or the psychological are predominating. Thus, although for the sake of clarity they are described separately, the two aspects of management must go hand in hand as is the situation in most back schools and rehabilitation programmes.

Physical management

Clues that physical aspects are important lie in the pattern of pain. When episodic and related to movement or posture, physical issues are likely to predominate (Table 14). Often faulty posture at work or leisure places abnormal stress on normal tissue or causes recurrent stress on degenerate structures within the lumbar segment. Management consists of education (Fordyce *et al*. 1986; Lehmann *et al*. 1986) into the mechanical and ergonomic factors that are relevant (see 'Education and self-care').

Repetitive lifting
Inappropriate bed or mattress
Sleeping unsupported or in an inappropriate position
Getting in/out of bed incorrectly
Sitting for longer than 30 min (particularly when driving)
Recurrent bending (e.g. when brushing teeth)
Stopping exercise regime when 'better'
Defaulting medication or corset
Lack of physical fitness

Table 14 Physical barriers to successful outcome

Sleep patterns are important. After checking that bedding and mattress are satisfactory (Frank and Hills 1988), it must be established whether sleep loss is due to insomnia, pain, or depression. Constant pain day and night suggests that depression may be present. If pain disturbs sleep a long-acting non-steroidal anti-inflammatory drug may be helpful. On occasions, questioning 'what is thought about when lying awake' gives a clue to social factors that may reflect the causes or consequences of the spinal condition.

Physical measures

Transcutaneous electrical nerve stimulation (**TENS**) is now widely used to try and control pain for which conservative measures were unsuccessful and surgery is not contemplated (Gersh and Wolf 1985; Wynn Parry and Girgis 1988). Proof of efficacy to date is lacking (Lehmann *et al*. 1986; Deyo *et al*. 1990; Herman *et al*. 1994). If successful, TENS is continued at home for months or years with patients having their own machine. These machines cost around £90 each and patients may be

admitted to hospital for a few days' assessment during which the optimum site for placement of electrodes and the duration of treatment is decided. Some centres assess people as outpatients, and some loan stimulators from the pain clinic or the physiotherapy department. In other districts, patients are requested to purchase their own equipment (Anonymous 1989).

Acupuncture is now frequently offered by pain clinics (Coan *et al*. 1980; Lehmann *et al*. 1986) and some physiotherapy departments. Currently it is not possible to predict who will do well with either TENS or acupuncture (Lewith 1984; Anonymous 1989). Both TENS and acupuncture are being increasing used in therapy earlier in the course of a patient's back pain.

Medication; local and epidural injections

Tricyclic antidepressant compounds, as detailed above, are apparently effective for some who may or may not be depressed and for those known to have a physical basis for their pain (Pheasant *et al*. 1983; Ward 1986). Warnings about sedation and dry mouth, and reassurance that tricyclics are not addictive and not being prescribed for psychiatric purposes, are appropriate.

The role of local injections was discussed above in the section on acute pain. For decades, enthusiasts have injected a 'sclerosant' mixture of phenol, dextrose, glycerine, and water into the supraspinous, iliolumbar, and sacroiliac ligaments (Mathews *et al*. 1987; Ongley *et al*. 1987; Klein *et al*. 1993). It seems likely that this is an effective form of treatment, although follow-up beyond 12 months has not been reported.

Epidural injections

Opiates, steroids, and local anaesthetics have all been administered by this route. The use of morphine for the treatment of postoperative pain has not been shown to be helpful, and is hazardous. Epidural steroids are helpful in reducing postoperative pain following laminectomy (Ang *et al*. 1988).

The use of epidural local anaesthetic with or without steroids has been reviewed by Benzon (1986) for the management of acute, chronic, or intractable low back pain. Drugs may be introduced into the epidural space via sacral, lumbar, thoracic or cervical routes, at the level of the presumed lesion, or randomly, with varying doses of steroid or local anaesthetic. These factors complicate any serious evaluation of epidural injections for low back pain.

Controlled trials have given contradictory results in the management of acute back pain with radiculopathy and their use remains controversial (Bush and Hillier 1991; Power *et al*. 1992). I know of no decisive controlled studies demonstrating unequivocal cost-effectiveness in the management of chronic or intractable back pain (Occupational Medicine Forum Committee 1994).

As has been noted with some other means of treatment for low back pain (e.g. traction), epidural injections are widely used by surgeons, physicians and in pain clinics, and has major benefits for patients even though specific indications remain unclear.

Psychological and social management

A significant group of patients presents with pain that is continuous and does not alter over the 24-h period. Often they claim to have suffered non-stop pain for many years. They are best considered as suffering from the 'pain or illness behaviour syndrome' (Chapman and Brena 1982; Waddell 1987), sometimes considered as 'learned helplessness', which is discussed elsewhere (Fordyce 1988). Examination may not reveal signs compatible with the disability described. The pattern of pain may be non-anatomical (Ransford *et al*. 1976) and examination is often accompanied by exaggerated gestures of pain, particularly during the straight-leg raise test (Waddell *et al*. 1980). Insomnia is frequent (see above).

Patients may admit to problems consequent to loss of earnings or to changing marital/sexual relationships. Often unrealistic fears exist (Table 15): that long-standing pain cannot be due to common musculoskeletal causes and must be caused by horrific disorders, for example cancer or multiple sclerosis. The fear of aggravating pain may act as an excessive inhibition, keeping patients inactive and away from work unnecessarily (Waddell *et al*. 1993). The effect on the spouse may be critical. The importance of meeting the spouse or 'significant other' cannot be overstressed (Fordyce 1988), particularly if overprotection or excessive sympathy has encouraged adoption of a 'sick/patient role'.

1. Fear
   - Of Medication — drug dependency, side-effects, effectiveness will mask pain, creating further damage
   - Of physical activity — will increase pain particularly during recovery
   - Of the cause — e.g. cancer, multiple sclerosis
   - Of the consequences — loss of job (status and poverty), disability (e.g. in a wheelchair)
   - Of clicking in the back — frightened of damage to spine
2. Effects on the family — particularly on the partner (potential to develop learned helplessness or pain behaviour)
3. Psychiatric consequences
   - Depression
   - Phobic disorders
   - Substance abuse, e.g. alcohol
   - Post-traumatic stress disorder
4. Previous psychiatric history
5. Other — job dissatisfaction, litigation/compensation

Table 15 Psychological barriers to successful outcome

Other strains may be placed on marital and family relationships (Frank and Hills 1988). Thus a heavy manual worker unable to continue work may find the dependence on the spouse an unbearable strain. Where the marital relationship has been difficult before the onset of back pain, the pain may be used as an excuse to withdraw from the sexual side of the relationship, putting further strain upon it. Not all patients will be able to continue a normal sexual relationship when in severe pain, but many people can be helped by the intelligent use of positioning (Fahrni 1978).

Both hypnosis and relaxation have been evaluated in a small study (McCauley *et al*. 1983), both groups having less pain and depression than control patients.

Sensory deprivation (Shea *et al*. 1991) has been shown to be helpful in chronic (and probably intractable) low back pain. This suggests continued avenues for management of those with failed conservative management. Other psychological factors that may inhibit resolution of symptoms are listed in Table 15. Many are amenable to psychological/psychiatric intervention.

It may take a long time to unravel the many psychological and social problems that have become associated with an individual's 'pain'. Not every problem is soluble. Often the time taken to come to grips with these problems is worthwhile, and back-pain sufferers and their families can learn to cope with life again. Self-help may be facilitated by literature such as 'Living with your pain' (Broome and Jellicoe 1987). Sometimes specialist counselling is required from a social worker or a psychologist.

Psychologists also have an increasingly important role in helping patients to cope with their pain. This may be done singly or in group work. Groups are particularly helpful in boosting patient's confidence through sharing experiences and realizing they are not alone in their suffering. Skinner *et al*. (1990) have shown significant improvements in mood, coping skills, physical disability, and analgesic consumption in a British study of a cognitive behavioural treatment programme, but without any controls.

Psychologists are also helpful in advising the rehabilitation team on possible behavioural approaches to management. Recognition by the patient that psychosocial problems predominate may not occur for some time. Usually both physical and behavioural approaches are combined simultaneously (Fordyce *et al*. 1986; Turner *et al*. 1990; Lindstrom *et al*. 1992), the physiotherapist often being essential to 'allow' the individual to get better even if the programme does not have a major physical component. Incorporation of behavioural models (Fordyce *et al*. 1986) into early physical therapy may offer the best hope of reducing the morbidity of pain behaviour.

Pre-existing psychopathology is frequent in those requiring intensive rehabilitation (Polatin *et al*. 1993) and does not militate against a successful outcome if the treatment programme is structured to manage psychopathology appropriately (Gatchel *et al*. 1994). The results of physically demanding programmes vary between individuals with different psychological profiles (Harkapaa *et al*. 1991, Talo *et al*. 1992).

Depression is a well-accepted consequence of chronic pain. However, other variables such as irritability and increased somatization may be important (Pelz and Merskey 1982; Main *et al*. 1992).

Trauma and litigation

Trauma remains a complicated issue when 'distress' is commonly seen (Waddell *et al*. 1980) and symptoms fail to resolve as expected. The reasons are unclear and do not always revolve around litigation. Negative attitudes may play a part, such as bitterness and an inability to understand 'why me?'. Depression may go undetected (as is often the case with chronic pain and long-standing disability) and phobias may develop (McCarthy *et al*. 1990).

Litigation may be another complicating problem. The adverse effects on prognosis of trauma at work have already been referred to (Troup *et al*. 1981). Teasell and White (1994) argue that many reasons exist to militate against resolution of disability when litigation is involved. They include the need to persuade doctors that their injuries are serious: medical disbelief hardens attitudes; compensation vindicates patients who blame third parties; fears that a return to work would compromise their safety; use of disability to avoid difficult situations; and financial anxieties. There is abundant evidence, however, that settlement of litigation does not relieve symptoms (Editorial 1991).

Once medical reports have been obtained from both parties, patients can be assured that the basis for a settlement is highly likely and that they must now concentrate on positive matters, such as return to work, the development of new interests, or looking after the family whilst the spouse works. Although symptoms do not disappear from such people when litigation is settled, the uncertainty surrounding it is not helpful.

Insight into the possible nature of the causes of unexplained pain after road traffic accidents has been shown by Twomey and Taylor (1993), who performed autopsies after road traffic accidents and showed haemorrhage into the outer annulus of the disc and haemarthroses, and capsular or synovial tears to the facet joints, in addition to traumatic herniation of the disc. The concept of the 'hidden disc' (Simon and Goldstein 1992) is important, emphasizing that trauma may result in a syndrome of '(a) intractable back pain with aggravation of the pain and loss of spinal motion with any physical exercise, (b) leg pain, (c) loss of energy, (d) marked weight loss, (e) profound depression'. This concept is particularly important when expert witnesses for plaintiff and defendants will not always be from musculoskeletal specialties. The psychological consequences of road traffic accidents have been largely ignored in the past. Mayou *et al*. (1993) have shown that 1 in 10 patients had mood disturbances 1 year after a road traffic accident. At 1 year, severe horrific intrusive memories of the accident were present in 6 to 11 per cent of the patients, with travel anxiety being present in 12 to 18 per cent. In a small series of patients involved in litigation, irritability, tearfulness, sleep disturbance, and travel anxiety each occurred in over half of the sample (Frank 1993b). The profound ignorance among orthopaedic surgeons of the assessment, diagnosis, and treatment of emotional problems following trauma creates real difficulties for patients whether they are involved in litigation or not.

Back schools

The term 'back school' may mean any programme with an educational content (Fig. 7) varying from a 1-session programme (Stankovic and Johnell 1990), variable sessions as an outpatient (Berquist-Ullman and Larsson 1977; Klaber-Moffett *et al* 1986; Morrison *et al*. 1988), or an inpatient programme (Harkapaa *et al*. 1990). In this review it will refer only to outpatient education with or without concurrent therapy. Three to four sessions (Berquist-Ullman and Larsson 1977; Klaber-Moffett *et al*. 1986) and 15 sessions (Harkapaa *et al*. 1990) have been shown to decrease back pain and disability.

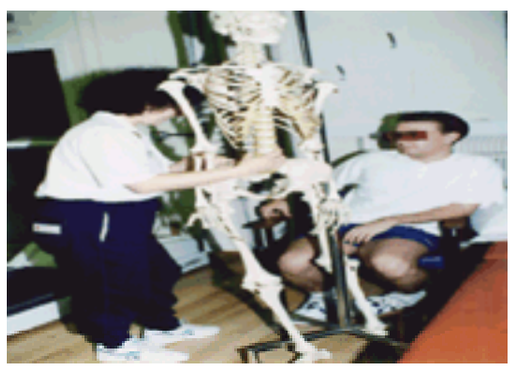


Fig. 7 Patien education has been shown to reduce back pain (reproduced by kind permission of the Northwick Park Hospital Department of Physiotherapy).

This subject has been reviewed by Koes *et al*. (1994). The best studies suggested that back schools are effective, particularly in occupational settings in acute, recurrent, or chronic conditions. The best-rated study (Harkapaa *et al*. 1990) has been discussed in the section on 'Intensive rehabilitation', indicating the wide variety of settings and programmes on to which a back school may be grafted. Results, however, are short term. It appears that attempts to change human behaviour over the long term (greater than 1 year) continue to escape us, with non-compliance and relapse remaining as major problems (Turk and Rudy 1991).

Intensive rehabilitation

Most rehabilitation programmes involve elements of intensity and multiprofessional input covering physical, psychosocial, educational, and vocational components. It also implies a failure of conservative or surgical management. Intensive rehabilitation includes those components of Table 3 under long-term measures. Whilst factors listed in Table 3, Table 7, and Table 8 should be part of any back-pain programme, the main features consist of learning to function in spite of pain, sharing experiences in groups with a skilled leader, and physically 'toughening up'—developing physical tolerance in both fitness (Frost *et al*. 1995) and vocational programmes, for example the Helsinki (Harkapaa *et al*. 1990; Mellin *et al*. 1990; Harkapaa *et al*. 1991), Texas (Mayer *et al*. 1987) and Turku (Alaranta *et al*. 1994) programmes.

It can be postulated that one of the reasons that complex rehabilitation programmes are successful is that opportunities present to sort out some of the social and psychological problems that follow low back pain. A model programme was evaluated in Helsinki by Harkapaa *et al*. (1990), where 459 patients were randomly allocated into control, inpatient, or outpatient management groups. Each group contained, in addition to physical therapies, education and the role of self-care in the prevention and early rehabilitation of low back disability, ergonomics, relaxation techniques, and group work. It is not clear whether the extra expense of the apparently superior inpatient treatment is really cost-effective; in the inpatient group there was some suggestion that inpatient was superior to outpatient rehabilitation at reducing pain, disability, and in enhancing compliance. Intensive rehabilitation would appear to be one of the few indications for admitting patients to hospital (Table 16).

1. Diagnosis and management of atypical pain or severe root compression
2. Pain management
3. Surgery
4. Assessment of intractable pain/pain behaviour
5. Initiation of rehabilitation of intractable pain

Table 16 Indications for hospitalization

The evidence suggests that intensive, multiprofessional programmes that encourage individuals to ignore pain, toughen up, and take away control of pain before giving it back to them are cost-effective. Long-term benefits have yet to be demonstrated, suggesting that the art of changing human behaviour is as yet far from understood (Turk and Rudy 1991).

Education and self-care

The first presentation of back pain is an opportunity to illustrate, for the purposes of patient education, the common predisposing factors. The history is often that of a trivial bending episode, possibly related to twisting, for example, putting a bottle of milk in a low refrigerator or bending over the wash basin to brush one's teeth. A demonstration of correct lifting techniques is important (Fig. 8). Few patients presenting to my back-pain clinic understand the importance of technique in approaching the floor (to pick up an object) or to lift. Whilst squatting is usually recommended (to avoid bending the trunk with straight legs), this is often awkward and inappropriate for older people.

Fig. 8 (a) Incorrect lifting technique: the lumbar lordosis is lost with maximal strain being placed on the spine; (b) correct lifting technique: use of the kneeling position minimizes strain on the back and is more stable than lifting from squatting or bending; painful knees from osteoarthritis or other causes prevent this and increase patients' disability (reproduced by kind permission of the Northwick Park Hospital Department of Physiotherapy).

Approaching the floor from the kneeling position can often overcome this, and older people (who are not using 'helping hands') may find help by using their arms to assist rising. The combination of arthritis of the large joints of the legs and back pain is particularly disabling. 'Helping hands' may help avoid bending.

An educational booklet has been developed and evaluated by general practitioners and given to a random sample of all patients presenting with back pain in five group practices. In the period from 2 weeks to 1 year following receipt of the booklets, there were significantly fewer reconsultations with back pain. Referrals to hospital consultants, for physiotherapy, admissions to hospital, and laminectomies were all less common in the booklet group. It was concluded that the booklets had some effect in altering both the knowledge and the behaviour of patients with back pain (Roland and Dixon 1989).

Bergquist-Ullman (1977) illustrated the potential importance of education into vocational and ergonomic aspects of self-care following episodes of back pain.

The model of the 'back school' has been followed throughout the world, being introduced into programmes at many different stages of the back patient's career. It has now developed into standard practice in many United Kingdom departments of physiotherapy (Frank and Hills 1988). A simple 3-session back school has been shown to improve functional disability and pain levels (Klaber Moffett *et al*. 1986). A similar 6-session outpatient programme has been reported from a community hospital in Canada (Morrison *et al*. 1988). The concept of back schools is discussed in more depth above.

Simple advice for back pain sufferers such as may be provided in simple educational form is given below (living with back pain).

Self-help may be facilitated by literature (Roland and Dixon 1989) or self-help groups (Webb 1982), but group education interventions in patients with low back pain are not conclusively successful (Cohen *et al*. 1994). Sometimes specialist counselling is required from a social worker or a psychologist, who may help overcome some of the problems (Table 15).

Whereas in acute stages patients often prefer professionals to take responsibility for their back care, which is therapeutic (Fordyce *et al*. 1986), this responsibility has to be transferred back to patients as pain becomes chronic and intractable. Thus responsibility to keep fit, maintain their individual exercise programme, remain relaxed so as to avoid physically (and ergonomically) stressing the spine is that of the individual, not of the professionals.

Living with low back pain

In hospital practice it is not infrequent for patients to have persisting low back pain in spite of many forms of therapy. Some patients come to the conclusion that they can live with their pain. Others can be helped to that conclusion, particularly if they are convinced they have received good treatment and that diseases have been excluded.

Such individuals can be helped to minimize mechanical stress on the spine by simple advice, some of which is listed in Table 7 and Table 8. The incorporation of such advice into everyday life can be rewarding and may remove triggering factors that perpetuate the current episode and may precipitate a further recurrence.

Some details of bedding, mattresses, chairs, and cars have been outlined elsewhere (Frank and Hills 1988). The importance of avoiding prolonged sitting must be stressed, particularly for those who recognize that their pain is aggravated by sitting. Approaches to sitting are reviewed elsewhere (Frank and Hills 1988). Continuous passive motion in sitting may be useful when fully developed (Reinecke *et al*. 1994).

Psychological factors may be equally important and are not well understood. Avoiding rushing about, being relaxed, and having time to plan one's activities clearly reduce the risk of doing everyday tasks in an unplanned way, thus decreasing the chance of doing them at a mechanical disadvantage. Simple advice has been produced by the British Psychological Society (Broome and Jellicoe 1987). The development of coping strategies (Rosenthiel and Keefe 1983) is important. Those who may benefit from professional psychological advice include those whose pain was the result of trauma (particularly at work).

Some individuals find sharing their experiences with fellow sufferers helpful and this may be in a group setting with a professional group counsellor (Webb 1982). Others may join branches of the Back Pain Society. The National Ankylosing Spondylitis Society has developed some groups that have a very physically orientated approach to their group meetings. At the request of the members of the group (some of whom are not being treated at the hospital) a physiotherapist helps organize and supervise the physical activities, which sometimes take place in the hydrotherapy pool.

Many appreciate further information about their condition. The Arthritis and Rheumatism Council produces a number of leaflets about rheumatic conditions; those on neck pain, backache, and seating are well worth recommending to patients with spinal pain if their local department does not produce information.

Return to work

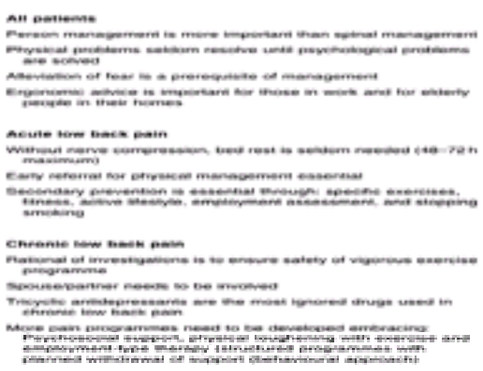
A minority of patients with low back pain may be 'off sick' for prolonged periods and have difficulty returning to work. Problems arise from the fear that work will aggravate the pain or cause it to recur; the prolonged adoption of the 'patient role', which may discourage a positive approach to work; the loss of self-discipline, which may result in poor time-keeping; and the lack of stamina and strength due to prolonged inactivity, which may have been consequent to bed rest. Such people can be helped by a general toughening up in the physiotherapy department and/or supervision of an increasing range of activities by therapists.

A return to normal working may require working through the pain and fears discussed above to develop stamina and the ability to cope with sitting, standing, and lifting. Sessions of therapy (physiotherapy or occupational therapy) should be increased to mimic a full day at work to reassure both patient and staff that a return to work will be viable. Ideally such programmes should include the teaching of coping strategies (Rosenthiel and Keefe 1983) so patients learn how to continue with life in spite of pain. Regrettably the shortage of clinical psychologists in the United Kingdom has inhibited the adequate development of programmes to include such strategies, sometimes with unfortunate results (Leak *et al*. 1989). This type of rehabilitation programme should be available at any large district general hospital and is suitable for outpatients, although, as detailed above, comparisons of inpatient versus outpatient rehabilitation regimens have favoured the more costly inpatient programmes (Harkapaa *et al*. 1990). When rehabilitation fails, or if it is not available locally, patients should be offered a residential assessment in the rehabilitation ward of the district general hospital or at a medical rehabilitation unit. Here, a thorough medical, nursing, psychological, remedial therapeutic, and social assessment can take place as the patient is under observation for the whole of the 24-h period. Such programmes have not been evaluated in the United Kingdom. In the United States, Mayer *et al*. (1987), in an uncontrolled evaluation of a 3-week, outpatient, 'functional restoration programme', combined a physical programme with work simulation/hardening, resulting in an 85 per cent return to work by 1 year. Pre-existing psychopathology does not militate against a successful outcome if the treatment programme is structured to manage psychopathology appropriately (Gatchel *et al*. 1994).

A recent meta-analysis of 37 studies that gave detailed definitions of patients' status, delineated patients' work status before treatment and at follow-up, and documented the proportion of patients employed at follow-up showed that non-surgical treatment for chronic pain does return patients to work, increased rates of return to work are due to treatment, and that benefits of treatment are not temporary (Cutler *et al*. 1994).

Those aged 50 or more are less likely to return to work (Frederickson *et al*. 1988). When a rehabilitation programme clarifies that people will not be able to return to their previous work, referral to the Disability Employment Adviser (United Kingdom) may be helpful. The adviser can discuss with employers alternative work that should be within the individual's capabilities. Where none is available, advice is given about retirement on the grounds of ill health, or other job options. Younger people (under 40–45) may be referred to an Employment Rehabilitation Centre where they are assessed for their retraining potential. The Disability Employment Adviser can register people 'disabled' and can help them to find work in a variety of ways. Although their titles vary, most industrialized nations have schemes that facilitate employment for those disadvantaged through illness or disability.

Service implications (Table 17)



All patients

Person management is more important than spinal management

Physical problems seldom resolve until psychological problems are solved

Alleviation of fear is a prerequisite of management

Ergonomic advice is important for those in work and for elderly people in their homes

Acute low back pain

Without nerve compression, bed rest is seldom needed (48–72 h maximum)

Early referral for physical management essential

Secondary prevention is essential through: specific exercises, fitness, active lifestyle, employment assessment, and stopping smoking

Chronic low back pain

Rational of investigations is to ensure safety of vigorous exercise programme

Spouse/partner needs to be involved

Tricyclic antidepressants are the most ignored drugs used in chronic low back pain

More pain programmes need to be developed embracing: Psychosocial support, physical toughening with exercise and employment-type (family involved programmes with planned withdrawal of support (behavioural approach)

Table 17 Summary of important points to remember

Until recently it was assumed that no particular therapy had demonstrable efficacy. Bed rest, symptomatic relief, and exclusion of other pathology were the prongs of therapy, with surgery for those requiring it. The evidence now shows that back pain and disability can be ameliorated, and often further episodes minimized, by professional intervention at the initiation of the first attack, shortening absence from work via physiotherapeutic and chiropractic (and probably osteopathic) manipulation, physical training, a variety of exercise programmes, and multiprofessional rehabilitation. Minimal rest and a planned withdrawal of support are important developments with demonstrable benefit.

Delays in referral for specialist advice, delays in being seen in hospital, and delays in the provision of therapy are all likely to contribute to non-resolution of acute/subacute back pain, increasing the pool of patients with chronic pain and the risk of pain becoming intractable, with the consequent suffering shared by the patient with the family. These people may progress to pain and illness behaviour, becoming more costly for employers and the state. Research now suggests that early (within 3 months), multiprofessional, integrated programmes can abort this negative process (Lindstrom *et al*. 1992).

The development of proper diagnostic and treatment protocols can decrease accidents and lost work days and be cost-effective (Wiesel *et al*. 1988) if costs to all sections of national life (state and industry) are considered. Services have not developed in line with the research evidence of efficacy, exposing both industry and public sectors to unnecessary disadvantage through avoidable sickness absence.

Chapter References

Abenhaim, L., Suissa, S., and Rossignol, M. (1988). Risk of recurrence of occupational back pain over a three year follow up. *British Journal of Industrial Medicine*, **45**, 829–33.

Alaranta, H. *et al*. (1994). Intensive physical and psychosocial training program for patients with chronic low back pain. *Spine*, **19**, 1339–49.

Aldridge, D. (1989). Europe looks at complementary medicine. *British Medical Journal*, **299**, 1121–2.

Allan, D. B. and Waddell, G. (1989). An historical perspective on low back pain and disability. *Acta Orthopaedica Scandinavica*, **60** (Suppl. 234), 1–23.

Andersson, G. B. J. (1981). Epidemiological aspects on low back pain in industry. *Spine*, **6**, 53–60.

Ang, E. T. *et al*. (1988). Postoperative analgesia: epidural injection of dexamethasone sodium phosphate. *Annales Français Anaesthesie et Reanimation*, **7**, 289–93.

Anonymous (1989). Transcutaneous electrical nerve stimulation to relieve pain. *Drug and Therapeutics Bulletin*, **27**, 25–7.

Armstrong, D. (1991). What do patients want? Someone who will hear their questions. *British Medical Journal*, **303**, 261–2.

Assendelft, W. J. J., Koes, B. W., Van der Heijden, G. J. M. G., and Bouter, L. M. (1992). The efficacy of chiropractic manipulation for back pain: blinded review of clinical trials. *Journal of*

Manipulative and Physiologic Therapeutics, **15**, 487–93.

Bateman, D. N. (1994). NSAIDs: time to re-evaluate gut toxicity. *Lancet*; **343**, 1051–2.

Battie, M. C. *et al*. (1991). Volvo award in clinical sciences. Smoking and lumbar intervertebral disc degeneration. *Spine*, **16**, 1015–21.

Behrens, V., Seligman, P., Cameron, L., Matthias, C. G. T., and Fine, L. (1994). The prevalence of back pain, hand discomfort and dermatitis in the US working population. *American Journal of Public Health*, **84**, 1780–5.

Bell, G. R. *et al*. (1984). A study of computer assisted tomography. Comparison of metrizamide myelography and computed tomography in the diagnosis of herniated lumbar disc and spinal stenosis. *Spine*, **9**, 552–6.

Benzon, H. T. (1986). Epidural steroid injections for low back pain and lumbosacral radiculopathy. *Pain*; **24**, 277–95.

Bergenudd, H. and Nilsson, B. (1994). The prevalence of locomotor complaints in middle age and their relationship to health and socioeconomic factors. *Clinical Orthopedics and Related Research*, **308**, 264–70.

Bergquist-Ullman, M. (1977). Acute low back pain in industry. A controlled prospective study with special reference to therapy and vocational factors. *Acta Orthopaedica Scandinavica* (Suppl. 170).

Biering-Sorensen, F., Thomsen, C. E., and Hilden, J. (1989). Risk indicators for low back trouble. *Scandinavian Journal of Rehabilitation Medicine*, **21**, 151–7.

Bigos, S. J. *et al*. (1991). A prospective study of work perceptions and psychological factors affecting the report of back injury. *Spine*, **16**, 1–6.

Blomberg, S., Svardsudd, K., and Mildenberger, F. (1992). A controlled multicentre trial of manual therapy in low back pain. *Scandinavian Journal of Primary Health Care*, **10**, 170–8.

Blomberg, S., Hallin, G., Grann, K., Berg, E., and Sennerby, U. (1994). Manual therapy with steroid injections—a new approach to treatment of low back pain. A controlled multicentre trial with an evaluation by orthopaedic surgeons. *Spine*, **19**, 569–77.

Bremner, R. A. and Simpson, M. (1959). Management of chronic lumbosacral strain. *Lancet*; **ii**, 949–50.

Broome, A. and Jellicoe, H. (1987). *Living with your pain: a self-help guide to managing pain*. The British Psychological Society in association with Methuen, London.

Brown, R. T., Page, G. B., and McMahan, P. B. (1991). Ergonomics in the US Railroad Industry. *Human Factors Society Bulletin*, **34**, 1–4.

Bush, K. and Hillier, S. (1991). A Controlled study of caudal epidural injections of triamcinolone plus procaine for the management of intractable sciatica. *Spine*, **16**, 572–5.

Butler, D., Trafimow, J. H., Andersson, G. B. J., McNeil, T. W., and Huckman, M. S. (1990). Discs degenerate before facets. *Spine*, **15**, 111–13.

Caldwell, J. and Glanville, H. J. (1980). *The industrial and social implications of low back pain (1963–1980)*. Report to the European Coal and Steel Community, Luxembourg and the DHSS, London.

Carette, S. *et al*. (1991). A controlled trial of corticosteroid injections into facet joints for chronic low back pain. *New England Journal of Medicine*, **325**, 1002–7.

Chapman, S. L. and Brena, S. F. (1982). Learned helplessness and responses to nerve blocks in chronic low back pain patients. *Pain*, **14**, 355–64.

Chavannes, A., Gubbels, J., Post, D., Rutten, G., and Thomas, S. (1986). Acute low back pain: patients' perceptions of pain four weeks after initial diagnosis and treatment in general practice. *Journal of the Royal College of General Practitioners*, **36**, 271–3.

Cherkin, D. C., Deyo, R. A., Loeser, J. D., Bush, T., and Waddell, G. (1994). An international comparison of back surgery rates. *Spine*, **19**, 1201–6.

Clinical Standards Advisory Group (1994*a*). *Back pain* (Report of a CSAG Committee on Back Pain). HMSO, London.

Clinical Standards Advisory Group (1994*b*). *Epidemiology review: the epidemiology and cost of back pain*. HMSO, London.

Coan, R. M. *et al*. (1980). The acupuncture treatment of low back pain: a randomised controlled study. *American Journal of Chinese Medicine*, **8**, 181–9.

Cohen, J. E., Goel, V., Frank, J. W., Bombardier, C., Peloso, P., and Guillemin, F. (1994). Group education interventions for people with low back pain—an overview of the literature. *Spine*, **19**, 1214–22.

Collée, G., Dijkmans, B. A. C., Vandenbroucke, J. P., and Cats, A. (1991). Iliac crest pain syndrome in low back pain. A double-blind, randomized study of local injection therapy. *Journal of Rheumatology*, **18**, 1060–3.

Committee on Safety of Medicines (1994). Relative safety of oral non-aspirin NSAIDs. *Current Problems in Pharmacovigilance*, **20**, 9–11.

Cooper, R. G., Hollis, S., and Jayson, M. I. V. (1992). Gender variation of human spinal and paraspinal structures. *Clinical Biomechanics*, **7**, 120–4.

Coste, J., Delecoeuillerie, G., Cohen de Lara, A., Le Parc, J. M., and Paolaggi, J. B. (1994). Clinical course and prognostic factors in acute low back pain: an inception cohort study in primary care practice. *British Medical Journal*, **308**, 577–80.

Croft, P. R. and Rigby, A. S. (1994). Socioeconomic influences on back problems in the community in Britain. *Journal of Epidemiology and Community Health*, **48**, 166–70.

Cutler, R. B., Fishbain, D. A., Rosomoff, H. L., Elsayed, A-M., Khalil, T. M., and Rosomoff, R. S. (1994). Does non-surgical pain center treatment of chronic pain return patients to work? A review and meta-analysis of the literature. *Spine*, **19**, 643–52.

Daltroy, L. H. *et al*. (1993). Teaching and social support: effects on knowledge, attitudes and behaviors to prevent low back injuries in industry. *Health Education Quarterly*, **20**, 43–62.

Davies, H. T. O., Crombie, I. K., Macrae, W. A., and Rogers, K. M. (1992). Pain clinic patients in northern Britain. *The Pain Clinic*, **5**, 129–35.

Davis, H. (1994). Increasing rates of cervical and lumbar spine surgery in the United States (1979–90). *Spine*, **19**, 1117–24.

Deyo, R. A. (1983). Conservative therapy for low back pain. Distinguishing useful from useless therapy. *Journal of the American Medical Association*, **250**, 1057–62.

Deyo, R. A. (1991). Fads in the treatment of low back pain. *New England Journal of Medicine*, **325**, 1039–40.

Deyo, R. A. and Tsui-Wu, Y-J. (1987). Descriptive epidemiology of low back pain and its related medical care in the United States. *Spine*, **12**, 264–8.

Deyo, R. A., Diehl, A. K., and Rosenthal, M. (1986). How many days of bed rest for acute low back pain? A randomized clinical trial. *New England Journal of Medicine*, **315**, 1064–70.

Deyo, R. A., Walsh, N. E., Martin, D. C., Schoenfeld, L. S., and Ramamurthy, S. (1990). A controlled trial of TENS and exercise for chronic low back pain. *New England Journal of Medicine*, **322**, 1627–34.

DiMaggio, A. and Mooney, V. (1987). The McKenzie programme: exercise effective against back pain. *Journal of Musculoskeletal Medicine*, **4**, 63–74.

Dimberg, L. *et al*. (1989). Sickness absenteeism in an engineering industry—an analysis with special reference to absence for neck and upper extremity symptoms. *Scandinavian Journal of Social Medicine*, **17**, 77–84.

Donchin, M., Woolf, O., Kaplan, L., and Floman, Y. (1990). Secondary prevention of low back pain—a clinical trial. *Spine*, **15**, 1317–20.

Editorial (1989*a*). The back pain epidemic. *Acta Orthopaedica Scandinavica*, **60**, 633–4.

Editorial (1989*b*). Risk factors for back trouble. *Lancet*, **i**, 1305–6.

Editorial. (1991). Neck injury and the mind. *Lancet*, **338**, 728–9.

Elnaggar, I. M., Nordin, M., Sheikhzadeh, A., Parnianpour, M., and Kahanovitz, N. (1991). Effects of spinal flexion and extension exercises on low back pain and spinal mobility in chronic mechanical low back pain patients. *Spine*, **16**, 967–72.

Fahrni, W. H. (1978). *Backache: assessment and treatment*. Musqueam, Vancouver.

Feldstein, A., Valanis, B., Vollmer, W., Stevens, N., and Overton, C. (1993). The back injury prevention project pilot study. *Journal of Occupational Medicine*, **35**, 114–20.

Fordyce, W.E. (1988). Psychological factors in the failed back. *International Disability Studies*, **10**, 29–31.

Fordyce, W. E., Brockway, J. A., Bergman, J. A., and Spengler, D. (1986). Acute back pain: a control-group comparison of behavioural versus traditional management methods. *Journal of Behavioural Medicine*, **9**, 127–40.

Frank, A. O. (1990). Therapies for back pain. *Practitioner*, **234**, 113–15.

Frank, A. O. (1993*a*). Low back pain. *British Medical Journal*, **306**, 901–9.

Frank, A. O. (1993*b*). Psychiatric consequences of road traffic accidents—often disabling and unrecognised. (Letter). *British Medical Journal*, **307**, 1283.

Frank, A. O. and Hills, J. E. (1988). Spinal pain. In *Disabling diseases—physical, environmental and psychosocial management* (ed. A. O. Frank, and G. P. Maguire). Heinemann Medical, Oxford.

Frank, A. O., Henshaw, D. J., and Parkin, D. R. (1987). Industrial therapy for physical disability in a district hospital rehabilitation unit. *Care, Science and Practice*, **5**, 29–33.

Frederickson, B. E., Trief, P. M., VanBeveren, P., Yuan, H. A., and Baum, G. (1988). Rehabilitation of the patient with chronic back pain. A search for outcome indicators. *Spine*, **13**, 351–3.

Frost, H., Klaber-Moffett, J. A., Moser, J. S., and Fairbank, J. C. T. (1995). Evaluation of a fitness programme for patients with chronic low back pain. A randomised controlled trial. *British Medical Journal*, **310**, 151–4.

Frymoyer, J. W. (1993). Introduction. In *Rehabilitation of the spine: science and practice* (ed. S. H. Hochschuler, H. B. Cotler, and R. D. Guyer). Mosby, St Louis.

Frymoyer, J. W. *et al*. (1983). Risk factors in low back pain. *Journal of Bone and Joint Surgery* (A), **65**, 213–18.

Garvey, T. A., Marks, M. R., and Weisel, S. W. (1989). A prospective, randomized double-blind evaluation of trigger-point injection therapy for low back pain. *Spine*, **14**, 962–4.

Gatchel, R. J., Polatin, P. B., Mayer, T. G., and Garcy, P. D. (1994). Psychopathology and the rehabilitation of patients with chronic low back pain disability. *Archives of Physical Medicine and Rehabilitation*, **75**, 666–70.

Gavin, M.B., and Weisel, S.W. (1991). Low back pain. *Current Opinion in Rheumatology*, **3**, 65–70.

Gersh, M. R. and Wolf, S. L. (1985). Applications of transcutaneous electrical nerve stimulation in the management of patients with pain. State of the art update. *Physical Therapy*, **65**, 314–22.

Gibson, T., Grahame, R., Harkness, J., Woo, P., Blagrave, P., and Hills, R. (1985). Controlled comparison of short wave diathermy treatment with osteopathic treatment in non-specific low back pain. *Lancet*, **i**, 1258–61.

Gilbert, J. R., Taylor, D. W., Hildebrand. A., and Evans, C. (1985). Clinical trial of common treatments for low back pain in family practice. *British Medical Journal*, **291**, 789–94.

Girgis, L. and Brooks, P. (1994). Nonsteroidal anti-inflammatory drugs—differential use in older patients. *Drugs and Aging*, **4**, 101–12.

Hadler, N. M., Curtis, P., Gillings, D. B., and Stinnett, S. (1987). A benefit of spinal manipulation as adjunctive therapy for acute low back pain: a stratified controlled trial. *Spine*, **12**, 703–6.

Hanley, E. N. (1992). The cost of surgical intervention for lumbar disc herniation. In *Clinical efficacy and outcome in the diagnosis and treatment of low back pain* (ed. J. N. Weinstein). Raven Press, New York.

Hansen, F. R. *et al*. (1993). Intensive dynamic back-muscle exercises, conventional physiotherapy or placebo-control treatment of low back pain. A randomized, observer-blind trial. *Spine*, **18**, 98–108.

Harber, P., Pena, L., Hsu, P., Billet, E., Greer, D., and Kim, K. (1994). Personal history, training, and worksite as predictors of back pain in nurses. *American Journal of Industrial Medicine*, **25**, 519–26.

Harkapaa, K., Mellin, G., Jarvikoski, A., and Hurri, H. (1990). A controlled study on the outcome of inpatient and outpatient treatment of low back pain. III. Long term follow-up of pain, disability and compliance. *Scandinavian Journal of Rehabilitation Medicine*, **22**, 181–8.

Harkapaa, K., Jarvikoski, A., Mellin, G., Hurri, H., and Luoma, J. (1991). Health locus of control beliefs and psychological distress as predictors for treatment outcome in low back pain patients: results of a 3 month follow-up of a controlled intervention study. *Pain*, **46**, 35–41.

Herman, E., Williams, R., Stratford, P., Fargas-Babjak, A., and Trott, M. (1994). A randomised controlled trial of transcutaneous electrical nerve stimulation (CODETRON) to determine its benefits in a rehabilitation program for acute occupational low back pain. *Spine*, **19**, 561–8.

Hickson, N. (1983). Back problems in general practice. *Australian Family Physician*, **12**, 345–8.

Hockaday, J. M. and Whitty, C. W. M. (1967). Patterns of referred pain in the normal subject. *Brain*, **XC**, 481–96.

Hodinka, L. (1991). Improving the risk-benefit ratio of NSAID therapy in the elderly. *Scandinavian Journal of Rheumatology*, **20** (Suppl. 91), 3–44.

Hodkinson, M. (1988). Multiple disabilities. In *Disabling diseases—physical, environmental and psychosocial management* (ed. A. O. Frank and G. P. Maguire). Heinemann Medical, Oxford.

Hsieh, C. J., Phillips, R. B., Adams, A. H., and Pope, M. H. (1992). Functional outcomes of low back pain: Comparison of four treatment groups in a randomized controlled trial. *Journal of Manipulative and Physiologic Therapy*, **15**, 4–9.

Hume Kendall, P. and Jenkins, J. M. (1968). Exercises for backache: a double-blind controlled trial. *Physiotherapy*, **54**, 154–7.

Jayson, M. I. V. (ed.) (1992). *The lumbar spine and back pain* (4th edn). Churchill Livingstone, Edinburgh.

Jayson, M. I. V. (1994). Mechanisms underlying chronic back pain. Studies are looking at the brain and spinal cord. *British Medical Journal*, **309**, 681–2.

Jenkins, E. M. and Borenstein, D. G. (1994). Exercise for the low back pain patient. *Clinical Rheumatology*, **8**, 191–7.

Kaplan, R. M. and Deyo, R. A. (1987). Back pain in hospital workers. In *Occupational back pain* (ed. R. A. Deyo), *Spine*: *state of the art reviews*, Vol. 2 (1), pp. 61–73.

Kaplan, R. S. and Sinaki, M. (1993). Posture training support: preliminary report on a series of patients with diminished symptomatic complications of osteoporosis. *Mayo Clinic Proceedings*, **68**, 1171–6.

Kauppila, L. I., Penttila, A., Karhunen, P. J., Lalu, K., and Hannikainen, P. (1994). Lumbar disc degeneration and atherosclerosis of the abdominal aorta. *Spine*, **19**, 923–9.

Kean, W. F. (1986). Arthritis in the elderly. *Clinics in the Rheumatic Diseases*, **12** (1).

Keller, R. B. (1992). Tracking physician practice: small area analysis techniques. In *Clinical efficacy and outcome in the diagnosis and treatment of low back pain* (ed. J. N. Weinstein). Raven, New York.

Kelsey, J. L. *et al*. (1984). Acute prolapsed intervertebral disc. An epidemiological study with special reference to driving automobiles and cigarette smoking. *Spine*, **9**, 608–13.

Klaber Moffett, J. A., Chase, S. M., Portek, I., and Ennis, J. R. (1986). A controlled prospective study to evaluate the effectiveness of a Back School in the relief of chronic low back pain. *Spine*, **11**, 120–2.

Klein, R. G., Eek, B. C., DeLong, W. B., and Mooney, V. (1993). A randomised double-blind trial of Dextrose-Glycerine-Phenol injections for chronic low back pain. *Journal of Spinal Diseases*, **1**, 23–33.

Koes, B. W., Assendelft, W. J. J, van der Heijden, G. J. M. G., Bouter, L. M., and Knipschild, P. G. (1991). Spinal manipulation and mobilization for back and neck pain: a blinded review. *British Medical Journal*, **303**, 1298–303.

Koes, B. W. *et al*. (1992). Randomised clinical trial of manipulative therapy and physiotherapy for persistent back and neck complaints: results of one year follow up. *British Medical Journal*, **304**, 601–5.

Koes, B. W., van Tulder, M. W., van der Windt, D. A. W. M., and Bouter, L. M. (1994). The efficacy of back schools: a review of randomised clinical trials. *Journal of Clinical Epidemiology*, **47**, 851–62.

Lahad, A., Malter, A.D., Berg, A. O., and Deyo, R. A. (1994). The effectiveness of four interventions for the prevention of low back pain. *Journal of the American Medial Association*, **272**, 1286–91.

Leak, A., Hull, R. G., Madden, S., Cooper, C., and Frank, A. O. (1989). Agoraphobia, back pain and the rheumatic diseases. (Letter). *British Journal of Rheumatology*, **28**, 554–5.

Lehmann, T. R. *et al*. (1986). Efficacy of electroacupuncture and TENS in the rehabilitation of chronic low back pain patients. *Pain*, **26**, 277–90.

Leigh, J. P. and Sheetz, R. M. (1989). Prevalence of back pain among full time United States workers. *British Journal of Industrial Medicine*, **46**, 651–7.

Lewith, G. T. (1984). Can we assess the effects of acupuncture? *British Medical Journal*, **288**, 1475–6.

Lilius, G. *et al*. (1989). Le syndrome des articulaires. Signification des signes de non-organicite. Une etude clinique randomisee et controlee par placebo. *Revue de Chirurgie Orthopedique*, **75**, 493–500.

Lindstrom, I., Ohlund, C., Eek, C., Wallin, L., Peterson, L-E., and Nachemson, A. (1992). Mobility, strength, and fitness after a graded activity program for patients with subacute low back pain. A randomized prospective clinical study with a behavioural therapy approach. *Spine*, **17**, 641–9.

Lindstrom, I., Ohlund, C., and Nachemson, A. (1994). Validity of patient reporting and predictive value of industrial physical work demands. *Spine*, **19**, 888–93.

MacArthur, C., Knox, G., and Lewis, M. (1993). Epidural anaesthesia during childbirth-association with backache is real. *British Medical Journal*, **307**, 64.

McCarthy, C., Coughlan, R., and O'Connor, P. (1990). Back pain, phobia and female (letter). *British Journal of Rheumatology*, **29**, 239.

McCauley, J. D., Thelen, M.H., Frank, R. G., Willard, R.R. and Callen, K. E. (1983). Hypnosis compared to relaxation in the outpatient management of low back pain. *Archives of Physical Medicine and Rehabilitation*, **64**, 548–52.

MacDonald, R. S. and Bell, C. M. J. (1990). An open controlled assessment of osteopathic manipulation in nonspecific low back pain. *Spine*, **15**, 364–70.

Main, C. J., Wood, P. L. R, Hollis, S., Spanswick, C. C., and Waddell, G. (1992). The distress and risk assessment method. A simple patient classification to identify distress and evaluate the risk of poor outcome. *Spine*, **17**, 42–52.

Manniche, C., Bentzen, L., Hesselsoe, G., Christensen, I., and Lundberg, E. (1988). Clinical trial of intensive muscle training for chronic low back pain. *Lancet*, **ii**, 1473–6.

Manniche, C., Lundberg, E., Christensen, I., Bentzen, L., and Hesselsoe, G. (1991). Intensive dynamic back exercises for chronic low back pain: a clinical trial. *Pain*, **47**, 53–63.

Manniche, C. *et al*. (1993). Clinical trial of postoperative dynamic back exercises after first lumbar discectomy. *Spine*, **18**, 92–7.

Mason, V. (1994). The prevalence of back pain in Great Britain. A report on OPCS Omnibus Survey data produced on behalf of the Department of Health. OPCS, London.

Mathews, J. A. *et al*. (1987). Back pain and sciatica: controlled trials of manipulation, traction, sclerosant and epidural injections. *British Journal of Rheumatology*, **26**, 416–23.

Mayer, T. G., Gatchel, R. J., Mayer, H., Kishino, N. D., Keeley, J., and Mooney, V. (1987). A prospective 2 year study of functional restoration in industrial low back injury. *Journal of the American Medical Association*, **258**, 1763–7.

Mayou, R., Bryant, B., and Duthie, R. (1993). Psychiatric consequences of road traffic accidents. *British Medical Journal*, **307**, 647–51.

Meade, T. W., Dyer, S., Browne, W., Townsend, J., and Frank, A. O. (1990). Low back pain of mechanical origin: randomised comparison of chiropractic and hospital outpatient treatment. *British Medical Journal*, **300**, 1431–7.

Mellin, G., Harkapaa, K., Hurri, H., and Jarvikoski, A. (1990). A controlled study on the outcome of inpatient and outpatient treatment of low back pain. *Scandinavian Journal of Rehabilitation Medicine*, **22**, 189–94.

Million, R., Hall, W., Nilson, K. H., Baker, R. D., and Jayson, M. I. V. (1982). Assessment of the progress of the back pain patient. *Spine*, **7**, 204–12.

Mitchell, R. I. and Carmen, G. M. (1990). Results of a multicentre trial using an intensive active exercise program for the treatment of acute soft tissue and back injuries. *Spine*, **15**, 514–21.

Mitchell, R. I. and Carmen, G. M. (1994). The functional restoration approach to treatment of chronic pain in patients with soft tissue and back injuries. *Spine*, **19**, 633–42.

Mooney, V. (1992). Facet joint syndrome. In *The lumbar spine and back pain* (4th edn) (ed. M. I. V. Jayson). Churchill Livingstone, Edinburgh.

Mooney, V. and Robertson, J. (1976). The facet syndrome. *Clinical Orthopedics and Related Research*, **115**, 149–56.

Moore, J., Phipps, K., Marcer, D., and Lewith, G. (1985). Why do people seek treatment from alternative medicine. *British Medical Journal*, **290**, 28–9.

Morrison, G. E. C, Chase, W., Young, V., and Roberts, W. L. (1988). Back pain: treatment and prevention in a community hospital. *Archives of Physical Medicine and Rehabilitation*, **69**, 605–9.

Occupational Medicine Forum Committee (1994). What is the efficacy of lumbar epidural steroid injections in the treatment of low back pain? If they are given, should they be given under radiographic guidance? *Journal of Occupational Medicine*, **36**, 118–19.

Ongley, M. J., Klein, R. G., Dorman, T. A., Eek, B. C., and Hubert, L. J. (1987). A new approach to the treatment of chronic low back pain. *Lancet*, **ii**, 143–6.

Ostgaard, H. C., Zetherstrom, G., Roos-Hansson, E., and Svanberg, B. (1994). Reduction of back and posterior pelvic pain in pregnancy. *Spine*, **19**, 894–900.

Overman, S.S., Larson, J.W., Dickstein, D.A., and Rockey, P.H. (1988). Physical therapy care for low back pain. Monitored program of first-contact nonphysician care. *Physical Therapy*, **68**, 199–207.

Ozoh, J. O., Onuigbo, M. A. C., Nwankwo, N., Ukabam, S. O., Umerah, B. C., and Emeruwa, C. C. (1990). 'Vanishing' of vertebra in a patient with sickle cell haemoglobinopathy. *British Medical Journal*, **301**, 1368–9.

Pelz, M. and Merskey, H. (1982). A description of the psychological effects of chronic painful lesions. *Pain*, **14**, 293–301.

Penttinen, J. (1994). Back pain and risk of fatal ischaemic heart disease: 13 year follow-up of Finnish farmers. *British Medical Journal*, **309**, 1267–8.

Pheasant, H., Bursk, A., Goldfarb, J., Azen, S. P., Weiss, J. N., and Borelli, L. (1983). Amitriptyline and chronic low back pain. A randomized double-blind crossover study. *Spine*, **8**, 552–7.

Philips, H. C. and Grant, L. (1991). The evolution of chronic back pain problems: a longitudinal study. *Behavior Research and Therapy*, **29**, 435–41.

Philips, H. C., Grant, L., and Berkowitz, J. (1991). The prevention of chronic pain and disability: A preliminary investigation. *Behavior Research and Therapy*, **29**, 443–50.

Polatin, P. B., Kinney, R. G., Gatchel, R. J., LIllo, E., and Mayer, T. G. (1993). Psychiatric illness and low back pain. the mind and the spine—which goes first? *Spine*, **18**, 66–71.

Pope, M. H., Phillips, R. B., Haugh, L. D., Hsieh, C. J., MacDonald, L., and Haldeman, S. (1994). A prospective randomized three-week trial of spinal manipulation, transcutaneous muscle stimulation, massage and corset in the treatment of subacute low back pain. *Spine*, **19**, 2571–7.

Porter, R. W. (1992). Spinal stenosis of the central and root canal. In *The lumbar spine and back pain* (4th edn) (ed. M. I. V. Jayson). Churchill Livingstone, Edinburgh.

Porter, R. W. and Oakshot, G. (1994). Spinal stenosis and health status. *Spine*, **19**, 901–3.

Porter, R. W., Hibbert C. S., and Wicks, M. (1978). The spinal canal in symptomatic lumbar disc lesions. *Journal of Bone and Joint Surgery* (B), **60**, 485–7.

Power, R. A., Taylor, G. J., and Fyfe, I. S. (1992). Lumbar epidural injection of steroid in acute prolapsed intervertebral discs. A prospective study. *Spine*, **17**, 453–5.

Ransford, A. O., Cairns, D., and Mooney, V. (1976). The pain drawing as an aid to the evaluation of patients with low-back pain. *Spine*, **1**, 127–34.

Reinecke, S. M., Hazard, R. G., and Coleman, K. (1994). Continuous passive motion in sitting: a new strategy against low back pain. *Journal of Spinal Diseases*, **7**, 29–35.

Risch, S. V. *et al*. (1993). Lumbar strengthening in chronic low back pain patients. *Spine*, **18**, 232–8.

Roland, M. (1994). Back pain in general practice: taking a positive view. *Musculoskeletal Medicine*, **1**, 2–4.

Roland, M. and Dixon, M. (1989). Randomized controlled trial of an educational booklet for patients presenting with back pain in general practice. *Journal of the Royal College of General Practitioners*, **39**, 244–6.

Rosenthiel, A. K. and Keefe, F. J. (1983). The use of coping strategies in chronic low back pain patients: Relationship to patient characteristics and current adjustment. *Pain*, **17**, 33–44.

Russell, R., Groves, P., Taub, N., O'Dowd, J., and Reynolds, F. (1993). Assessing long term backache after childbirth. *British Medical Journal*, **306**, 1299–303.

Salminen, J. J., Pentti, J., and Terho, P. (1992). Low back pain and disability in 14 year old schoolchildren. *Acta Paediatrica*, **81**, 1035–9.

Sandstrom, J. (1986). Clinical and social factors in rehabilitation of patients with chronic low back pain. *Scandinavian Journal of Rehabilitation Medicine*, **18**, 35–43.

Schattenkirchner, M. (1991). Long-term safety of ketoprofen in an elderly population of arthritic patients. *Scandinavian Journal of Rheumatology*, **20** (Suppl. 91), 27–36.

Shea, D. D. *et al*. (1991). The effect of sensory deprivation in the reduction of pain in patients with chronic low back pain. *Spine*, **16**, 560–1.

Shekelle, P. G., Adams, A. H., Chassin, M. R., Hurwitz, E. L., and Brook, R. H. (1992). Spinal manipulation for low back pain. A review. *Annals of Internal Medicine*, **117**, 590–8.

Shekelle, P. G. (1994). Spine update—spinal manipulation. *Spine*, **19**, 858–61.

Shi, L. (1993). Cost-benefit analysis of a California County's Back Injury Prevention Program. *Public Health Reports*, **108**, 204–11.

Simmons, J. W., McMillin, J. N., Emery, S. F., and Kimmich, S. J. (1992). Intradiscal steroids. A prospective double-blind clinical trial. *Spine*, **17**, S172–5.

Simon, W. H. and Goldstein, G. N. (1992). Hidden disc syndrome. In *Medico-legal consequences of trauma* (ed. W. H. Simon and G. E. Ehrlich), pp. 87–128. Marcel Dekker, New York.

Skinner, J. B., Erskine, A., Pearce, S., Rubenstein, I., Taylor, M., and Foster, C. (1990). The evaluation of a cognitive behavioural treatment programme in outpatients with chronic pain. *Journal of Psychosomatic Research*, **34**, 13–19.

Spitzer, W. O., Le Blanc, F. E., and Dupuis, M. (1987). Scientific approach to the assessment and management of activity-related spinal disorders. A monograph for clinicians. *Spine*, **12**, 7S.

Stankovic, R. and Johnell, O. (1990). Conservative treatment of acute low back pain. A prospective randomized trial: McKenzie method of treatment versus patient education in 'Mini Back School'. *Spine*, **15**, 120–3.

Svenson, H-O., Andersson, G. B. J., Hagstad, A., and Jansson, P-O. (1990). The relationship of low back pain to pregnancy and gynecologic factors. *Spine*, **15**, 371–5.

Szpalski, M. and Hayez, J. P. (1994). Objective functional assessment of the efficacy of Tenoxicam in the treatment of acute low back pain. A double-blind placebo controlled study. *British Journal of Rheumatology*, **33**, 74–8.

Talo, S., Rytokosky, U., and Puuka, P. (1992). Patient classification, a key to evaluate pain treatment: A psychological study in chronic low back pain patients. Volvo award in clinical sciences. *Spine*, **17**, 998–1011.

Taylor, H. G. and Dawes, P. T. (1990). Musculo-skeletal causes of chest pain. *Rheumatology in Practice*, Dec./Jan., pp. 16–20.

Taylor, V. M., Deyo, R. A., Cherkin, D. C., and Kreuter, W. (1994). Low back pain hospitalisation: recent United States trends and regional variations. *Spine*, **19**, 1207–13.

Teasell, R. W. and White, K. (1994). Clinical approaches to low back pain. Part 2. Management, sequelae, and disability and compensation. *Canadian Family Physician*, **40**, 490–5.

Troup, J. D. G., Martin, J. W., and Lloyd, D. C. E. F. (1981). Back pain in industry. *Spine*, **6**, 61–9.

Turk, D. C. and Rudy, T. E. (1991). Neglected topics in the treatment of chronic pain patients—relapse, non compliance and adherence enhancement. *Pain*, **44**, 5–28.

Turner, J. A. and Jenson, M. P. (1993). Efficacy of cognitive therapy for chronic low back pain. *Pain*, **52**, 169–77.

Turner, J.A., Clancy, S., McQuade, K. J., and Cardenas, D. D. (1990). Effectiveness of behavioural therapy for chronic low back pain: a component analysis. *Journal of Consulting and Clinical Psychology*, **58**, 573–9.

Turner-Stokes, L. and Frank, A. O. (1992). Urinary incontinence among patients with arthritis. *Journal of the Royal Society of Medicine*, **85**, 389–93.

Twomey, L.T. and Taylor, J.R. (1993). The whiplash syndrome: pathology and physical treatment. *Journal of Manual and Manipulation Therapy*, **1**, 26–9.

Von Korff, M., Deyo, R. A., Cherkin, D., and Barlow, W. (1993). Back pain in primary care—outcomes at one year. *Spine*, **18**, 855–62.

Von Korff, M., Barlow, W., Cherkin, D., and Deyo, R. A. (1994). Effects of practice style in managing back pain. *Annals of Internal Medicine*, **121**, 187–95.

Waddell, G. (1982). An approach to backache. *British Journal of Hospital Medicine*, (Sept.), 187–219.

Waddell, G. (1987). A new clinical model for the treatment of low back pain. *Spine*, **12**, 632–44.

Waddell, G., McCulloch, J. A., Kummel, E., and Venner, R. M. (1980). Nonorganic physical signs in low back pain. *Spine*, **5**, 117–25.

Waddell, G., Newton, M., Henderson, I., Somerville, D., and Main, C. J. (1993). A fear-avoidance beliefs questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. *Pain*, **52**, 157–68.

Walsh, K., Cruddas, M., and Coggan, D. (1992*a*). Low back pain in eight areas of Britain. *Journal of Epidemiology and Community Health*, **46**, 227–30.

Walsh, K., Cruddas, M., and Coggan, D. (1992*b*). Risk of low back pain in people admitted to hospital for traffic accidents and falls. *Journal of Epidemiology and Community Health*, **46**, 231–3.

Ward, N. G. (1986). Tricyclic antidepressants for chronic low back pain. Mechanisms of action and predictors of response. *Spine*, **7**, 661–5.

Webb, P. (1982). Self help groups for back pain sufferers. *Health Visitor*, **55**, 487–9.

Weiser, S. and Cedraschi, C. (1992). Psychosocial issues in the prevention of chronic low back pain—a literature review. In Common low back pain: prevention of chronicity (ed. M. Nordin and T. L. Vischer). *Clinical Rheumatology*, **6**, 657–84.

Weiss, Y., Chenard, A., De Kermadec, J. M., Tertrain, F., and Abiteboul, M. (1991). Maintenance of blood pressure control in elderly hypertensives on Ketoprofen. *Scandinavian Journal of Rheumatology*, **20** (Suppl. 91), 37–44.

Wells, N. (1985). *Back pain*. Office of Health Economics, London.

Wiesel, S. W., Feffer, H. I., and Borenstein, D. G. (1988). Evaluation and outcome of low back pain of unknown etiology. *Spine*, **6**, 679–80.

Williams, M. E. (1985). The impact of musculoskeletal disease on older persons. In *Arthritis and society* (ed. N. M. Hadler and D. B. Gillings). Butterworths, London.

Wilson, R. N. and Wilson, S. (1964). A ten year follow up of cases of low back pain. *Practitioner*, **192**, 657–60.

Wood, P. H. N. and Badley, E. M. (1980). Epidemiology of back pain. In *The lumbar spine and back pain* (2nd edn) (ed. M. I. V. Jayson), pp. 29–55. Pitman Medical, Tunbridge Wells.

Wynn Parry, C. B. and Girgis, F. (1988). The assessment and management of the failed back, Part 2. *International Disability Studies*, **10**, 25–8.

Zylbergold, R. S. and Piper, M. C. (1981). Lumbar disc disease: comparative analysis of physical therapy treatments. *Archives of Physical Medicine and Rehabilitation*, **62**, 176–9.

1.2.1.2 Spinal problems in children

Charles S. B. Galasko

Children with spinal disorders present with deformity, pain, neurological disturbances, or a combination. Deformity, either scoliosis or kyphosis, is probably the commonest presenting symptom. Frequently the deformity is not noted by the patient, but was detected by a relative, at school, or by a dressmaker.

The incidence of back pain in children is low, as seen by the number of referrals to the orthopaedic clinic (Winter and Lipscomb 1978; Hoffman 1980; Bunnell 1982), but school-based studies suggest that the incidence of backache may be high (Grantham 1977; Fairbank *et al*. 1984; Balague *et al*. 1988; Chan and Ryan 1992), though the vast majority of children did not seek specialist medical advice. Analysis of the referrals to the Royal Manchester Children's Hospital indicated that back pain counted for just under 2 per cent of non-trauma referrals to the orthopaedic clinic (Turner *et al*. 1989), whereas it has been estimated that 80 per cent of adults will visit a medical practitioner at some time because of back pain.

Turner *et al*. (1989) found that the patients fell into two distinct groups. In just over half the patients, a definite underlying pathological cause was found. No specific pathological cause was detected in 29 of the 61 patients. The study has now been extended to 233 patients, 75 of whom had no specific pathological cause (Table 1). There was no significant difference between the two groups in terms of age, sex, or site of the pain and none of the patients in the non-specific group was subsequently found to have developed a specific cause of back pain. In contrast, a survey of 100 children with back pain at the University of Michigan in 1980, revealed that all but 15 per cent had a definite identifiable cause for their complaint (Hensinger 1989). Eighteen per cent had either a tumour or infection, and approximately 33 per cent were found to have a post-traumatic cause such as an occult fracture or spondylolysis/spondylolisthesis.

Non-specific		75
Spondylolysis/spondylolisthesis		26
Infection including discitis		14
Scheuermann's disease		10
Other kyphoses		4
Tumours		10
Psychogenic		9
Disc prolapse		8
Inflammatory		6
Osteoporosis		5
Trauma (excluding sprains)		4
Limb length inequality		2
Secondary to other gait abnormality		7
Renal		2
Rickets		1
Congenital anomalies		1
Scoliosis		46
Neuromuscular		
Too late for surgery	17	
Presurgery	3	
Postsurgery	2	
Sub-total	22	
Minor scoliosis	19	
Fixation	3	
Fusion	5	
Sub-total	25	
Total		233

Table 1 Causes of back pain in childhood in 233 patients

History and examination

A careful history and clinical examination must be made for all children. The history may be very non-specific and, if the child is very young, may be nothing more than a recent refusal to walk, or the development of a postural change. Even older children are often not able to localize their symptoms.

There frequently is a history of trauma, but this does not imply that the pain is traumatic in origin. If the pain occurs during sport spondylosis, spondylolisthesis, or possible injury should be excluded. The presence of night pain is suggestive of a tumour; morning stiffness is suggestive of an inflammatory spondylarthropathy.

The history should not just concentrate on the features of the presenting complaint, but should seek to establish the general health of the patient, the presence or absence of symptoms affecting other parts of the locomotor system, and the psychological background of the patient. Persistent pain that is unrelieved by rest or immobilization, increasing pain, and symptoms of systemic illness suggest infection or tumour. These patients may also present with abdominal pain and guarding.

A neurological history must also be taken, although neurological symptoms are uncommon in children. The presence of radicular pain, paraesthesia, or weakness suggests significant compression of the cord or nerve roots. Change in sphincteric control in a previously continent child requires immediate investigation. A foot deformity may be secondary to a neurological abnormality including spinal dysraphism.

Scoliosis is usually pain free. The combination of pain and scoliosis usually indicates significant underlying pathology, such as tumour, infection, or a prolapsed intervertebral disc.

A family history should be taken and the presence of any other diseases established. Scoliosis, for example, frequently complicates neuromuscular disease (Fig. 1), mesenchymal disorders such as Marfan's syndrome and Ehlers–Danlos syndrome, neurofibromatosis, the osteochondrodystrophies, rheumatoid disease, tumours of the spinal column or spinal cord, and metabolic disorders (Table 2). Scoliosis may also follow irradiation, for example of a Wilms tumour, or a thoracotomy. Kyphosis may also be secondary to an underlying disorder (Table 3; Fig. 2).

Fig. 1 Scoliosis in a patient with Duchenne muscular dystrophy. The patient has a typical collapsing 'C'-shaped scoliosis which involves the entire spine. In neuromuscular scoliosis, progression of the curve is usually associated with increase in pelvic obliquity, which makes sitting uncomfortable and affects seating balance.

Structural	Idiopathic Infantile (0–3 years) (1) Resolving (2) Progressive Juvenile (3–10 years) Adolescent (> 10 years) Neuromuscular Congenital Neurofibromatosis Mesenchymal disorders (e.g. Marfan's; Ehlers–Danlos syndrome) Rheumatoid disease Trauma Fracture Surgical (laminectomy; post thoracoplasty) Irradiation Fibrous/scar contractures (posterior/burns; post burns) Skeletal dysplasias Infection Metabolic disorders Related to lumbosacral joint (spondylolysis and spondylolisthesis; congenital anomalies) Tumours
Non-structural	Postural Hysterical Nerve root irritation (e.g. prolapsed disc; tumour) Inflammatory (e.g. appendicitis) Related to long-length discrepancy Related to contractures about the hip

Table 2 Causes of scoliosis

Postural
Scheuermann's disease
Congenital Failure of formation Failure of segmentation
Neuromuscular
Myelomeningocele
Trauma
Postsurgery (postlaminectomy; following excision of vertebral body)
Postirradiation
Metabolic
Skeletal dysplasia
Collagen disease
Tumour
Infection
Inflammatory
Neurofibromatosis

Table 3 Causes of kyphosis

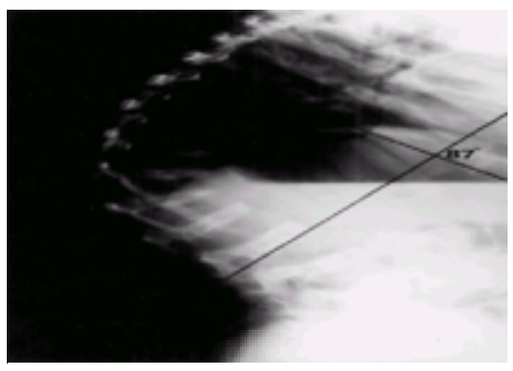


Fig. 2 Kyphosis secondary to an osteochondrodystrophy.

Refusal to walk in a very young child may be the first manifestation of a spinal disorder, such as discitis. A limp may be the presenting symptom. Irritation of the psoas muscle, secondary to a spinal disorder may result in flexion of the hip, a limp, and reduced movements of the hip on clinical examination. Other presenting features of a spinal disorder include postural change and morning stiffness.

The physical examination must include a general examination of the patient, an examination of the spine, and a detailed neurological examination. The spine should be examined standing, sitting, and prone. The presence of any deformity should be noted. If pain is present and if the patient is old enough, the patient should be asked to indicate the site and distribution of the pain. Any tenderness should be elicited and the movements of the spine determined. The immature spine is normally very flexible. Hamstring tightness may result in severe restriction of spinal flexion. Any deformity of the lower limbs should be noted and the child's gait should be observed, as well as the ability of the child to hop. In addition to assessing carefully sensation (including perineal sensation), if a child is old enough, power, reflexes, and any bladder or bowel involvement must be assessed.

The investigations should be based on the clinical findings. A standard 'work-up' is not indicated. The patients whose symptoms are severe enough to warrant referral to an orthopaedic clinic should have plain radiographs taken of their spine. If scoliosis or kyphosis is present, erect radiographs should be taken.

In the series of Turner *et al*. (1989), the plain radiographs provided the diagnosis in 58 per cent of their patients. If infection is considered, a white cell count and erythrocyte sedimentation rate should be carried out. This proved positive in 14 per cent of their patients. A painful scoliosis in a child usually implies a significant lesion producing the deformity, such as a prolapsed disc, tumour, or infection and magnetic resonance imaging (**MRI**) is indicated. Restriction of spinal movement and/or limited straight leg raising might indicate compression of the spinal cord or nerve roots and MRI is indicated. Turner *et al*. (1989) found that radiculography, combined with computed tomography (**CT**), demonstrated compression of the cord or nerve root in 30 per cent of their patients, but this investigation has largely been replaced by MRI where the facilities are available.

Without such investigations the diagnosis may be delayed. Turner *et al*. (1989) recorded a delay of 6 months or more in 16 per cent of patients with serious spinal pathology, all being tertiary referrals, the delay occurring prior to referral to their clinic. These included: two patients originally diagnosed as having Scheuermann's disease, but who were found to have cord tumours; an adolescent girl with spastic diplegia and severe hip flexion contractures, who had undergone soft tissue release of her hip flexors, but in whom a diagnosis of isthmic spondylolysis was made 4 years after the onset of increasing back pain; and a 14-year-old boy whose

prolapsed disc was diagnosed by combined radiculography and CT scanning 3 years after the onset of back pain. Galasko (1990) reported the case of an adolescent male who presented with a 3-year history of severe back pain. Previous investigations, carried out prior to his referral to the orthopaedic clinic, had included radiculography and plain radiographs, which were normal. His pain had not responded to treatment and, before labelling him as 'psychogenic', an orthopaedic opinion was sought. On examination he had a mild scoliosis and movements of his spine were limited. The plain radiographs were normal. Skeletal scintigraphy showed increased uptake in the right L5 pedicle (Fig. 3). Prior to surgery he was given 370 MBq of technetium-99m methyl diphosphonate (MDP) and a sterilizable probe was used to detect the areas of greatest activity. These were excised and histological examination revealed the presence of two osteoid osteomas.

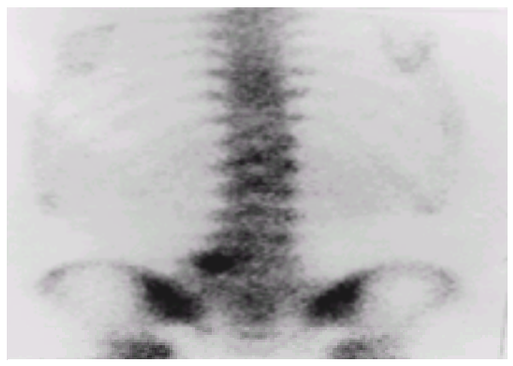


Fig. 3 Patient with osteoid osteoma who had complained of back pain for 3 years prior to referral to the orthopaedic clinic. A technetium-99m MDP scintigram showed increased uptake in the pedicle of L5. At surgery, two areas of increased uptake were found. Both proved to be osteoid osteomas. Following excision of the lesions, his pain settled completely.

Skeletal scintigraphy is a useful investigation in many patients and demonstrated the lesion in 10 per cent of the patients of Turner *et al*. It is useful in the diagnosis of discitis, particularly in a young child who is unable to localize the pain. Disorders of the genitourinary tract may be identified on the scintigram. Sacroiliac joint infection can cause back pain and is best diagnosed by skeletal scintigraphy. These patients complain of back, hip, and buttock pain. The Faber test is usually positive. In this test, the hip is flexed, abducted, and externally rotated, stressing the sacroiliac joint, whilst the pelvis is stabilized on the contralateral side.

Wenger (1993) described four levels of work-up. In level 1, the history usually includes a recent minor injury. Clinical examination is negative and there is no hamstring tightness, local muscle spasm, limitation of straight leg raising or neurological sign. The child requires a careful examination but usually no investigations. Treatment is symptomatic and the patient is seen 1 to 2 weeks later if the symptoms persist. If the symptoms have not resolved, level 2 work-up may be required.

A level 2 assessment is carried out when the history is less clear. Usually, there is no specific injury. The child has often had pain for several weeks. Physical examination is inconclusive without specific neurological signs, but the child may have tight hamstrings or minor spinal asymmetry. There are no systemic signs. The investigations include anteroposterior and lateral radiographs of the part of the spine that seems to be the source of pain. If the lumbosacral area is involved, oblique views are considered to evaluate a possible spondylolysis. If the radiographs are normal, the patient is treated symptomatically and is seen for review after several weeks. If the symptoms have not resolved a level 3 work-up may be required.

A level 3 assessment is required in children who have had significant pain for some weeks, with normal plain radiographs and in whom a diagnosis has not been established. The clinical examination is usually non-specific. There is no neurological deficit, but hamstring tightness, limited straight leg raising, and/or back muscle spasm are commonly present and may be severe. The child may be pyrexial. Investigations include plain radiographs as indicated above. If radiographs have already been taken and they are of poor quality, they should be repeated, since poor quality radiographs are a common reason for initially missing the diagnosis in a child with back pain. The child should have a complete blood count and an erythrocyte sedimentation rate carried out. In most cases, Wenger advised a skeletal scintigram with pinhole collimator views of suspected levels of involvement. Treatment is symptomatic with watchful waiting if the plain radiographs, blood count, erythrocyte sedimentation rate, and skeletal scintigram are normal. If the investigations reveal a specific diagnosis, such as spondylolysis, tumour, or discitis, specific treatment is required.

A level 4 assessment is required in patients who are experiencing severe symptoms. Often these symptoms have been present for weeks or months and may include pyrexia and/or complaints suggesting a neurological deficit. There may be extreme focal back pain and muscle spasm, positive straight leg raising, or neurological deficit on clinical examination. The investigations include a complete blood count, erthrocyte sedimentation rate, and in most cases an urgent CT scan and/or MRI. Often a skeletal scintigram is also included. The underlying spinal condition requires appropriate treatment.

Prevention

Back pain may occur in patients who develop a progressive collapsing scoliosis, secondary to neuromuscular disease such as Duchenne muscular dystrophy. Untreated, well over 90 per cent of patients with Duchenne muscular dystrophy develop a progressive scoliosis, once they are no longer able to walk and stand. The curve is progressive and is associated with increasing pelvic obliquity. The optimum treatment for such curves is spinal stabilization, once the patient has lost his independent ambulation, has developed a progressive scoliosis, or has a scoliosis greater than 20 degrees, and is still fit for surgery in terms of his respiratory and cardiac function (Galasko *et al*. 1992). The use of spinal orthoses may slow the progression of such a curve, but does not stop a curve from progressing once it has reached 30 to 40 degrees (Hsu 1983; Colbert and Craig 1987). Unfortunately, the more effective the orthosis, the more it affects the already limited lung function in these patients. Without treatment the curve continues to progress and is associated with pelvic obliquity which leads to loss of sitting balance, sitting discomfort, and eventually may make sitting impossible because of the associated pain. This pain can be prevented by timely spinal stabilization, which not only stops the progression of scoliosis and maintains sitting balance but also slows down the deterioration in lung function (Galasko *et al*. 1992). Spinal stabilization is also associated with maintenance of sitting balance and comfort in many other neuromuscular conditions including spinal muscle atrophy and Friedreich's ataxia.

Non-specific causes of back pain

The term 'non-specific' is a useful term, but does not adequately describe the lesion. In these patients the symptoms usually settle within a few months, with no specific treatment or physiotherapy. However, this group does include a number of patients with psychogenic elements to their pain. This occurs more commonly in girls. In patients whose pain persisted for more than 1 year, there was a high incidence of other complaints, such as headache or abdominal pain, and frequently there was also a strong family history of back pain (Turner *et al*. 1989). True hysterical pain is rare in childhood or adolescence, but does occur.

In the absence of any neurological symptoms or signs, with a mobile spine and normal straight leg raising, and without the development of a deformity, a 'wait and see' policy is justified, the symptoms being treated by analgesics, physiotherapy, and very occasionally, a simple spinal orthosis. The vast majority of patients with non-specific back pain will settle on this regimen but, if the pain persists, further investigation, such as MRI, CT scanning, or skeletal scintigraphy may be indicated.

Salminen *et al*. (1992) compared spinal mobility and trunk muscle strength in 38 15-year-old children suffering from low back pain with 38 asymptomatic controls. The boys with back pain were over 4 cm taller than those in the control group and in both sexes lumbar extension and straight leg raising were decreased and lumbar flexion was increased. Endurance strength was decreased in the abdominal and back muscles in the children with back pain.

The patients who suffered from sciatica in addition to recurrent low back pain had decreased lumbar flexion and side bending compared with those who suffered from recurrent low back pain without sciatica.

Specific causes of back pain in children and adolescents

Tumours

Tumours may arise from the vertebral column or from within the spinal canal (Galasko 1990). Bone tumours are uncommon and tumours arising from the spine even less common. They may be confused with non-neoplastic disease, and a high index of suspicion is required if an early diagnosis is to be made. An accurate diagnosis is essential for early and appropriate treatment. They may present as accidental findings or with pain, deformity, stiffness, or neurological symptoms.

The commonest presenting symptom is pain, which usually is persistent and is not significantly alleviated by rest. Occasionally, the patient may present with neurological symptoms or the sequelae of neural involvement such as pes cavus, equinus, or a foot droop. The combination of back pain and bilateral sciatica may be due to an intradural tumour. Examination may reveal local tenderness, muscle spasm, scoliosis, and limited movement. Occasionally, a lump may be palpable.

Children with leukaemia occasionally present with back pain, but this can usually be excluded by the blood count. Other serological investigations will depend on the age of the patient and the likely diagnosis. If metastatic neuroblastoma is considered, a urinary vanillyl mandelic acid may be required.

Radiological investigations may include plain radiographs, CT scanning, MRI, and/or angiography. A spinal lytic lesion may be due to infection, tumour-like conditions such as fibrous dysplasia, a benign tumour such as an aneurysmal bone cyst, or a malignant tumour such as a small round cell tumour.

A tissue diagnosis by biopsy is important, the type of biopsy depending on the site of the lesion. The choices include a needle (aspiration), trocar, open, or excision biopsy. Aspiration and trocar biopsy should be carried out under radiological control.

Spinal tumours may be classified as to whether they are primary benign, primary malignant, metastatic, intraspinal, or paraspinal, which includes tumours arising in the paraspinal musculature.

Osteochondroma

This is a benign cartilage-capped exostosis and can arise from any bone that develops from cartilage. It is the commonest primary tumour of bone and about 3 per cent occur in the spine (Novick *et al*. 1982). They usually present as a solitary lesion but may be multiple (diaphyseal aclasis) and rarely cause symptoms. When they do, they present as a painless lump and rarely with pain or neurological dysfunction. There is a definite, albeit small, predisposition to malignant degeneration, particularly in diaphyseal aclasis. This should be suspected in any osteochondroma that becomes painful or increases in size. Treatment of the osteochondroma is by excision, if the symptoms warrant it.

Osteoid osteoma

This lesion accounts for approximately 1 per cent of all tumours. Jackson *et al*. (1977) reviewed 860 osteoid osteomata, of which 85 occurred in the spine. It is commoner in males and usually presents between the ages of 10 and 25 years. Painless lesions have been reported (Pettine and Klassen 1986), although this is rare. The patients classically present with pain which is usually well localized, may be worse at night, may interfere with sleep, and may appear to be exacerbated by rest and activity. The pain is not necessarily relieved by aspirin. It gradually increases in severity and becomes more constant. There may be local tenderness and, frequently, there is an associated scoliosis (Fig. 4). Muscle spasm, limitation of spinal movement, and a limp may be present. Plain radiographs are frequently normal. Skeletal scintigraphy is indicated whenever the lesion is suspected, and has greatly reduced the delay in diagnosis. CT scanning of the area of increased uptake is required to pinpoint the lesion. CT scanning is preferable for localizing bony lesions, whereas MRI gives more information about the soft tissues and bone marrow.

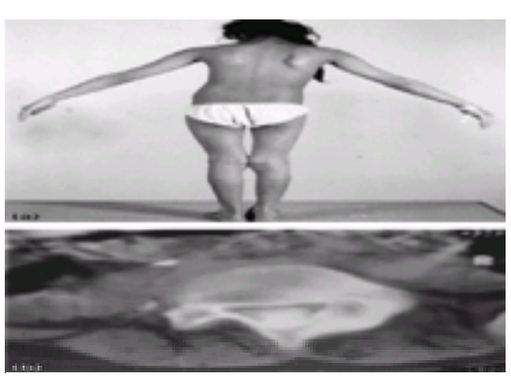


Fig. 4 Osteoid osteoma. This adolescent female presented with a painful scoliosis (a). Her skeletal scintigram showed an area of increased uptake in the right pedicle of L5. Computed tomography confirmed the lesion (b). Following resection of the osteoid osteoma, which was confirmed by histological examination, both her pain and scoliosis disappeared.

It sometimes may be difficult to localize the lesion at surgery, and the increased avidity of these lesions for bone-seeking radionuclides can be used to help pinpoint the lesion and minimize the amount of bone resected. Prior to surgery, the patient is given a bisphosphonate labelled with technetium-99m and is a sterilized probe used to detect the area of greatest activity, which is then excised. If a further area of increased activity is noted, it must also be excised. Failure to relieve pain might indicate that a second lesion was left or the wrong piece of bone resected. If the pain is not relieved, postoperative skeletal scintigraphy, CT scanning, or MRI is indicated to exclude a persistent osteoid osteoma.

An important differential diagnosis is the sclerotic phase of spondylosis. These patients present with a unilateral defect in one pars interarticularis and an area of sclerosis in the contralateral pedicle (Sherman *et al*. 1977). This area of sclerosis is associated with increased uptake of the bone-seeking radionuclide. Single-photon emission computed tomographic (**SPECT**) skeletal scintigraphy is more sensitive than planar skeletal scintigraphy in demonstrating this lesion (Bellah *et al*. 1991). CT scanning of the area of increased radionuclide uptake is required to differentiate osteoid osteoma from this lesion. If it is excised, the result is a bilateral interarticularis defect with resultant instability.

Osteoid osteoma has been reported as the commonest cause of a painful scoliosis (Mehta and Murray 1977). If the scoliosis has been present for less than 12 to 18 months, it usually disappears following excision of the tumour. However, it may persist and become structural. This depends on the age of the patient and the duration of the symptoms (Mehta and Murray 1977). It is my practice to carry out skeletal scintigraphy and/or MRI in all children and adolescents whose back pain has not settled within 2 to 3 months and whose plain radiographs are normal.

Osteoblastoma

This tumour has a predilection for the spine (approximately 40 per cent), particularly the posterior elements (Fig. 5) (Byers 1968; De Souza and Frost 1973; Jackson *et al*. 1977; Akbarnia and Rooholamini 1981). The patients are usually under the age of 30 years, the lesions tend to be larger than osteoid osteomas, and neurological deficit occurs more commonly. As with osteoid osteoma, the patient may complain of pain for 2 or 3 years before the diagnosis is made. The lesion appears lytic and expansile on plain radiographs, with varying degrees of central ossification and a thin rim of periosteal new bone. This tumour is most often benign, but it may recur and even metastasize. Treatment is by excision.

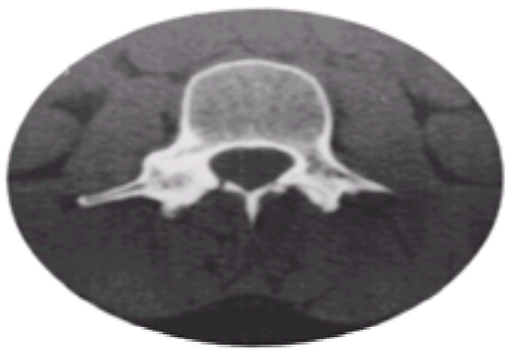


Fig. 5 Osteoblastoma affecting the left pedicle and extending into the transverse process.

Aneurysmal bone cyst

These tumours are rare and account for only 1 per cent of primary bone tumours, but 11 to 22 per cent of the lesions are located in the spine (Hay *et al*. 1978; Capanna *et al*. 1985). The tumours usually occur in the second decade of life and most commonly in the lumbar spine. In 20 to 40 per cent of patients there is involvement of more than one vertebra. The lesions usually present with pain, but may also present with loss of movement and/or scoliosis. Local tenderness may be present and, if a lesion is large, swelling may be observed. Capanna *et al*. (1985) recorded a neurological deficit in 12 of their 22 patients.

The lesions are ill defined and may be missed on the initial radiograph. As the tumour grows a thin cortex is noticed around an expansile lesion with some periosteal new bone. The cavity may contain fine trabeculae seen on tomography or computed tomography. Destruction of the entire vertebral body may cause collapse with vertebra plana. The lesion may extend to the adjacent rib or vertebra and there may be a soft tissue component. Radiculography may show a complete block or an irregular extradural defect. MRI is helpful to delineate the soft tissue spread and cord or nerve root compression. There is usually increased uptake on skeletal scintigraphy, and angiography shows arteriovenous shunting.

Needle biopsy may be associated with profound bleeding. Treatment is by surgical excision, radiotherapy, or their combination. The lesion is usually curetted out, but Hay *et al*. (1978) reported a 25 per cent recurrence rate following incomplete excision. If feasible, selective embolization, prior to resection of the tumour, is advisable to minimize the extent of intraoperative bleeding; cryosurgery has been employed for a similar reason. Capanna *et al*. (1985) reported poor results with radiotherapy. Furthermore, the latter may result in growth disturbance or radiation myelopathy. Sarcomatous change has been reported following irradiation of a spinal aneurysmal bone cyst (Tillman *et al*. 1968).

Langerhans cell histiocytosis (eosinophilic granuloma)

This may occur as a solitary lesion, may be polyostotic, or may be associated with systemic involvement. Lichtenstein (1953) coined the phrase histiocytosis X (now called Langerhans cell histiocytosis) to include all forms of eosinophilic granuloma. Letterer–Siwe disease and Hand–Schüller–Christian disease represent the acute and chronic form, respectively, of the systemic illness. The disease may affect infants, children, or young adults. The condition is probably not a true neoplasm and accounts for less than 1 per cent of all tumour and tumour-like conditions of bones.

It most frequently presents in an adolescent with back pain of some weeks' duration. Neurological deficit is rare. It is characterized radiographically by areas of well-circumscribed osteolysis and is the most common cause of vertebra plana in the skeletally immature spine. Considerable collapse of the vertebral body may occur without neurological compromise (Fig. 6), but spinal cord involvement may occur from acute collapse or occasionally from extradural spread of the lesion. In the spine, the lesion occurs most commonly in the thoracic region. Occasionally, several adjacent vertebral bodies are involved; this occurs more commonly in the systemic form. The diagnosis can often be made from the radiographs, but biopsy may be required.

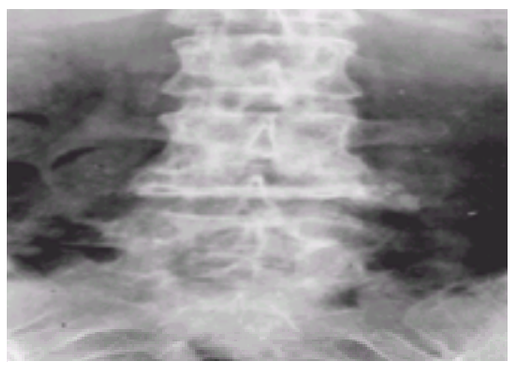


Fig. 6 Vertebra plana secondary to eosinophilic granuloma. There are many causes of severe collapse of a vertebral body, but in an adolescent the commonest cause is eosinophilic granuloma.

Treatment of eosinophilic granuloma depends upon the presence or absence of neurological compromise and the degree of vertebral collapse.

The solitary form of eosinophilic granuloma is self-limiting and restoration is variable, but the younger the patient the more likely a vertebra plana will reconstitute to an almost normal configuration. Treatment usually consists of an hyperextension orthosis, which maintains alignment of the spine whilst reconstitution is occurring. Following complete collapse of the vertebral body, immobilization of the spine is probably all that is required.

Neurological deficit, secondary to soft-tissue tumour extension, is treated with radiotherapy; but, if it is secondary to bony impingement, decompression and fusion are required. Chemotherapy is required for the systemic forms of the condition. Radiation is no longer considered appropriate or necessary for a single lesion unless there is cord compression.

Other benign tumours

Other benign tumours may affect the spine but are extremely rare in childhood. Haemangiomas usually occur after the age of 40 years. Giant-cell tumours rarely occur in the spine and, when they do, they usually involve the sacrum.

Malignant tumours

These are rare. The commonest primary malignant tumour of bone is a myeloma, which usually occurs after the age of 50 years.

It may be difficult to make the diagnosis of a malignant tumour from plain radiographs. The tumours may cause neural compression. Total excision may be impossible. Biopsy is essential. Many of these tumours are radioinsensitive. Irradiation is also limited by the dose to the spinal cord.

Ewing's sarcoma usually occurs in the 5 to 20 years age group. It rarely arises in the spine but when it does, usually presents with localized pain. Tumour growth may result in compression of the neural elements. The tumour may metastasize to the spine from a primary Ewing's sarcoma in the pelvis or long bone. It occasionally is

accompanied by a low-grade fever, elevated erythrocyte sedimentation rate, and leucocytosis. The latter findings are frequently associated with a fulminating course. The radiographic appearance is usually one of moth-eaten destruction and may be difficult to differentiate from infection. There usually is a large soft tissue component. Accurate tissue diagnosis is essential. The prognosis is particularly poor, but radiotherapy may result in marked amelioration of symptoms and good local control in many patients. The addition of chemotherapy has improved the survival. Surgical decompression and spinal stabilization is occasionally indicated.

Malignant lymphoma (reticulum cell sarcoma) of bone is a rare condition. It occurs most commonly in the young adult and may eventually metastasize. The radiographic appearance is of a sclerotic response, and an 'ivory' vertebra may be seen. Treatment is by radiation therapy.

Osteosarcoma is the second commonest primary malignant tumour of bone, but primary involvement of the spine is rare, and spinal involvement is usually due to metastases from osteosarcoma at other sites. The tumour occurs most commonly in the second decade of life. Radiographs show a combination of lytic bone destruction with bone formation. A tissue diagnosis is essential.

Treatment is by resection and spinal reconstruction (whenever possible), chemotherapy, and local irradiation. Unfortunately, radical excision of a primary malignant tumour of the spine is often impossible without compromising the integrity of the spinal cord. Spinal cord or cauda equina compression is treated by debulking the tumour, decompressing the neural elements, and stabilizing the spine.

Other malignant tumours, such as chondrosarcoma, fibrosarcoma, and chordoma usually occur in the adult.

The commonest malignant tumour of bone is metastatic cancer, but this usually affects the adult. In the child, the malignancies most likely to metastasize to the spine are neuroblastoma and leukaemia. These patients usually present with pain which may be localized or may be neuritic in nature. In the majority of patients, the symptoms are controlled by localized radiotherapy and/or chemotherapy. Surgical decompression and stabilization may be required for spinal cord or cauda equina compression, and surgical stabilization for instability secondary to metastatic bone destruction.

Intra- and extramedullary spinal tumours

Primary intramedullary spinal cord tumours may be present for months, or even years, before the diagnosis is made (Reimer and Onofrio 1985). Fearnside and Adams (1978) reported a median delay of 2 years before the diagnosis of a cauda equina tumour was made. The symptoms and physical signs in 57 patients with intradural tumours are shown in Table 4 and Table 5. The mean delay in diagnosis was 4.2 years (Pena *et al*. 1992) and the initial diagnosis was incorrect in 62 per cent.

Symptom	No.	%
Low back pain	28	49.12
Lower limb weakness	14	24.56
Sciatic pain	12	21.05
Thoracic back pain	7	12.28
Altered sensation in legs	4	7.02
Urinary incontinence	4	7.02
Night pain	3	5.26
Upper limb weakness	2	3.51
Thigh pain	1	1.75
Neck pain	1	1.75

(After Pena *et al*. 1992).

Table 4 Presenting symptoms in 57 patients with intradural spinal tumours

Physical sign	No.	%
Muscle weakness	40	70.18
Spastic paraparesis	22	38.60
Spastic paraplegia	2	3.51
Extensor plantar reflexes	25	43.86
Clonus	19	33.33
Absent/reduced reflexes	26	45.61
Sensory level	19	33.33
Localized sensory loss	18	31.58
Absent/impaired vibration/position sense	20	35.09
Positive Romberg test	3	5.26
Urinary retention	8	14.04
Scoliosis	4	7.02
Pes cavus	4	7.02
Claw toes	1	1.75
Drop foot	1	1.75
Tight Achilles tendon	3	5.26
Visual disturbance	2	3.51
Café-au-lait spots	2	3.51

(After Pena *et al*. 1992).

Table 5 Abnormal physical sign at the time of diagnosis in 57 patients with intradural spinal tumours

Woltman *et al*. (1951) found the mean delay in diagnosis of intramedullary tumours was 4.8 years and for extramedullary tumours 2.9 years. Some run a fluctuant neurological course and, in others, symptoms can be precipitated by trivial spinal injury.

Between 40 and 50 per cent of patients complain of a dull aching back pain at the level of the tumour, whereas 20 per cent complain of radicular pain. The presence of nocturnal back pain should raise the suspicion of a spinal tumour (Stern 1978). The tumour may present with spinal rigidity, muscle spasm, scoliosis, or disturbance of gait. On examination there may be neurological signs including weakness, absent or reduced reflexes, an extensor plantar response, spastic paraparesis, abnormalities of sensation, and urinary retention. Foot deformities may be present such as claw toes, pes cavus, or tight Achilles tendon. Up to 50 per cent of patients have abnormal plain radiographs, with increasing width of the spinal canal, medial erosion of the pedicles and scalloping of the vertebral body. Radiculography, CT scanning, and/or MRI show widening of the spinal cord and there may be a block. MRI is the investigation of choice. Treatment is by neurosurgical removal of the tumour. If an extensive laminectomy is required, the spine should be stabilized at the time of surgery.

Approximately two-thirds of intraspinal tumours are extramedullary. The clinical presentation partly depends on the pathological type, but progressive loss of cord function predominates. Pain occurs more frequently than with intramedullary tumours. The extramedullary lesions occur more commonly in adults.

Tachdjian and Matson (1965) reported that the initial diagnosis was incorrect in 70 per cent of 115 children with spinal tumours of all types.

Infection

Spinal osteomyelitis

This occurs much less commonly than osteomyelitis affecting the appendicular skeleton. It occurs most commonly in the adult but in young children there is often more of a systemic reaction. Frequently, the child is ill with a high pyrexia and radiographically significant bone destruction is noted. The organism most commonly responsible for the infection is *Staphylococcus aureus*. It may be difficult to differentiate pyogenic infection from tuberculosis (Fig. 7) and a needle biopsy may be required to isolate the organism before instituting treatment. Treatment normally consists of appropriate antibiotics and immobilization of the spine with an orthosis. Surgery is occasionally required.

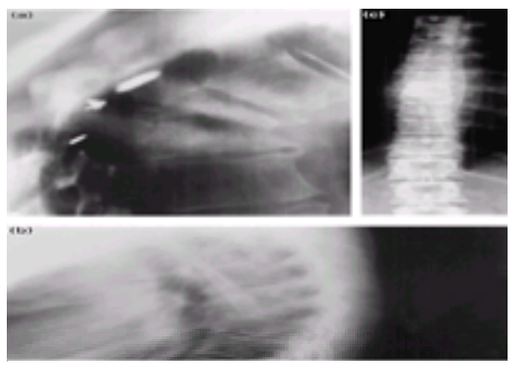


Fig. 7 Spinal infection. It may be difficult to differentiate a pyogenic infection (a) from tuberculosis (b). Needle biopsy may occasionally be required. *Staphylococcus aureus* was cultured from the needle biopsy taken from the patient shown in (a). Tuberculosis is usually associated with a paravertebral abscess (c), which usually does not occur with pyogenic spondylitis.

Tuberculosis

Tuberculosis may affect any bone or joint, but is most commonly found in the spine. Typically, a disc and adjacent vertebrae are involved (Fig. 7(b and c)), although several adjacent vertebrae may be affected. Collapse of the disc with damage to adjacent vertebrae is the characteristic radiographic appearance and produces a localized kyphosis, which may be very severe depending on the amount of destruction (Fig. 8). The tuberculous pus spreads along the soft tissue planes and may extend posteriorly to involve the spinal canal. Usually a paravertebral abscess is evident on the plain radiographs (Fig. 7(c)).

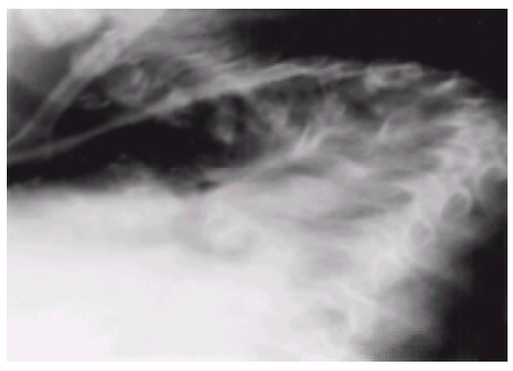


Fig. 8 Tuberculosis of the spine associated with significant destruction and a severe kyphosis.

The patient presents with pain frequently associated with a localized kyphosis. There may be a cold abscess, depending upon the spread of pus. Neurological symptoms may arise from pressure on the cord or cauda equina secondary to deformation of the spine, or compression of the neural tissues by fragmented bone and/or pus.

The erythrocyte sedimentation rate is raised and the Mantoux (tuberculin) test is positive. The typical radiographic appearance is of localized destruction of a disc and the adjacent two vertebrae, although the destruction may be more extensive. Usually a soft tissue abscess is seen in the paravertebral tissues.

The diagnosis can only be proved by identifying the organism or demonstrating the typical histological features, but this is not always required. In most cases, infection can be treated by antituberculous chemotherapy and immobilization of the spine in a plaster jacket or orthosis, although the latter may not be essential (Medical Research Council 1985). Large abscesses should be drained after chemotherapy has been instituted and, if there is any suggestion of spinal cord or cauda equina compression, the neural tissues must be decompressed and the spine stabilized, usually with an anterior bone graft.

Plain radiographs provide most of the information necessary for diagnosis and treatment. The main value of CT and MRI is in the preoperative evaluation of the small proportion of patients who require surgical treatment (Hoffman *et al*. 1993). Skeletal scintigraphy or MRI may be useful in diagnosing a recurrent infection (Fig. 9).

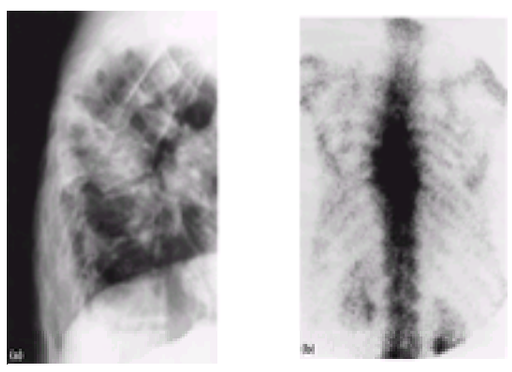


Fig. 9 (a) Recurrent tuberculosis. This patient had previously been treated for tuberculosis but returned some years later complaining of recurrent back pain. From the radiographs it was impossible to determine whether there was any active infection. (b) The skeletal scintigram showed increased uptake at the site of the previous infection. This appearance, together with the symptoms, suggested that she had reactivated tuberculosis and required a further course of therapy.

Discitis

This implies an infection of the discs. There are two varieties (Galasko 1989): postoperative discitis, which is a rare complication of discectomy and occurs in adults; and discitis of childhood.

Discitis of childhood

This most frequently occurs in children between the ages of 2 and 6 years, who present with irritability and a painful, tender, and stiff back (Fig. 10). The degree of stiffness is usually striking. Often the child cannot bend down and pick up an object from the floor without squatting. Sometimes the child limps. The tenderness may be localized, but in a younger child it may be diffuse across the back. There is no neurological involvement. The erythrocyte sedimentation rate may be elevated but the white cell count is frequently normal. There may be a history of recent febrile illness, such as a sore throat or earache (Wenger *et al*. 1978). Widespread paravertebral muscle spasm and limited back movement are characteristic signs. Pain may radiate to the abdomen, anterior thigh, or lower limbs. The presentation may be similar to that of an irritable hip or other hip disorders because of a limp in an older child, refusal to walk in a child under the age of 3 years, and radiation of the pain to the anterior thigh.

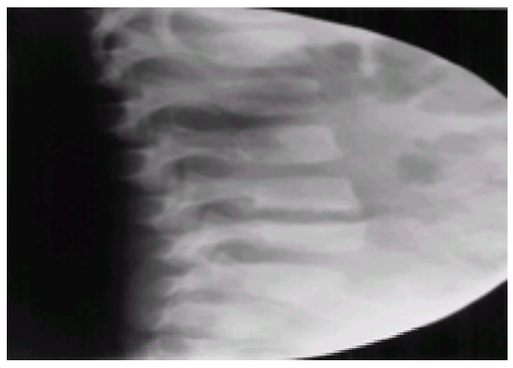


Fig. 10 Discitis. There is narrowing of the affected disc with involvement of the adjacent vertebrae.

In the early case, the radiographs may be normal, but subsequent radiographs show narrowing of the intervertebral disc (Fig. 10) followed by erosion of the adjacent vertebrae. A kyphosis may develop. Occasionally, fusion of the adjacent vertebrae may occur after the infection has settled. Increased uptake of radionuclide may be seen in skeletal, and particularly gallium-67, scintigrams (Norris *et al*. 1978; Bruschwein *et al*. 1980) prior to the development of radiographic signs. MRI is more sensitive but may have to be carried out under a general anaesthetic. Organisms are found in only half of the patients (Wenger *et al*. 1978) and, in many centres, no attempt is made to find the organism. The patient is usually treated with antibiotics. Where an organism has been found, the commonest is *Staphylococcus aureus* and the antibiotics most frequently used are flucloxacillin or a cephalosporin. Blood cultures should be taken prior to starting antibiotic therapy, which may be given intravenously for the first 3 to 4 days. The spine is also immobilized in a plaster jacket for approximately 6 weeks. If the patient does not respond to empiric antibiotic treatment, needle aspiration or biopsy may be required. A CT-guided needle biopsy is usually carried out by a radiologist.

Not all patients make a full recovery. Jansen *et al*. (1993) followed up 35 children for 12 to 35 years, with an average follow-up of 17 years. Fifteen patients still complained of backache. Flexion of the spine was normal in 32 patients whereas extension was restricted in 30. Twenty-six patients had a blocked vertebra and 28 patients had narrowing of the vertebral canal.

Wenger (1993) suggested that discitis and vertebral osteomyelitis were variations of a similar disorder, the infection beginning as a microabscess in the vertebral body adjacent to the vertebral endplate. This is a rapidly growing portion of the vertebral body, anatomically similar to the metaphysis of a long bone where osteomyelitis usually begins. The disc is involved early by the infection because of the vascular channels. Wenger suggested that vertebral osteomyelitis occurred when the infection progressed to the point of causing destructive changes in the vertebral body with collapse.

Discitis occurs much more frequently than vertebral osteomyelitis. However, patients with vertebral osteomyelitis usually have a very short history, whereas some patients with discitis may have been symptomatic for a number of weeks before the diagnosis is made. Wenger concluded that both conditions are due to infection. In both disorders, the commonest organism is *Staphylococcus aureus* and Wenger suggested that the difference may be due to the response of the host to the infection.

In contrast, Ryöppy *et al*. (1993) carried out an operative biopsy in 16 of 18 patients who were treated for non-specific discitis. The mean age at admission was 3 years and 3 months. The characteristic clinical findings were restriction of spinal mobility and an elevated erythrocyte sedimentation rate. Radiographic narrowing of the affected disc space was seen 4 to 5 weeks after the initial symptoms. A histological diagnosis of chronic or subacute non-specific inflammation was made in 10 patients. There were non-specific changes in 2 and normal tissue was found in 5. The patients were either treated with bed rest and antibiotics or bed rest alone. They found that all patients made a full clinical recovery, their follow-up ranging from 5 months to 9 years. Bacterial staining and cultures were negative in all the patients. Ryöppy and colleagues queried whether discitis was infective and suggested that antibiotics should only be given to patients with systemic signs, with positive bacterial cultures from other foci, or to patients who did not respond to adequate immobilization.

Intramedullary spinal cord abscess

This is a rare condition (Bartels *et al*. 1995). The presentation can be confusing, mimicking thoracic or abdominal disease. It is associated with neurological signs. MRI is the investigation of choice. An early diagnosis and emergency decompression of the intramedullary abscess is required to offer the patient the best opportunity of regaining neural function. Appropriate antibiotics are given. Dexamethasone is required to reduce the oedema. A high initial dose is given and the dose is gradually reduced.

Prolapsed intervertebral disc

There are three types of underlying pathology:

1. a bulging unripe disc;
2. a sequestrated disc;
3. occasionally, a slipped vertebral epiphysis.

Although it is not as common as in the adult, disc herniation can occur in children and adolescents. The symptoms are frequently less severe in children and there is less commonly a precipitating event. De Orio and Bianco (1982) suggested that disc herniation in children was more likely to be due to cumulative trauma than to a single event, i.e. the result of a fatigue failure of the annular fibres. MRI has shown that several of the discs are abnormal in an adolescent with a prolapsed intervertebral disc, although only a single disc prolapses. The commonest levels are at L4/5 and L5/S1.

Back pain is the commonest presenting symptom. Typically, the pain is aggravated by activity and relieved by rest. It may be increased by coughing and sneezing. Local tenderness may be present. There usually is significant spinal spasm and the patient frequently has a shuffling gait. The straight leg raising is usually restricted and other sciatic nerve tension signs are usually positive, but objective neurological findings are uncommon. The plain radiographs are usually normal and the diagnosis is usually based on CT scanning, radiculography, or MRI.

Unlike the adult, the adolescent disc prolapse does not frequently respond to conservative treatment and surgical excision is required more often. In the majority of patients, excision relieves the lower back pain, stiffness, and hamstring spasm (Silvers *et al*. 1994) but it may take months, following removal of the disc, for the stiffness to settle. De Luca *et al*. (1994) found that patients treated by discectomy did better than patients treated conservatively.

Slipping of the vertebral epiphysis occurs only in adolescence, is more common in males, and is usually associated with heavy lifting. The posterior, inferior epiphysis of L4 is most frequently affected. The slipped epiphysis with its adjacent disc is displaced into the vertebral canal (Lowrey 1973; Keller 1974; Lippitt 1976; Handel *et al*. 1979: Banerian *et al*. 1990). It is analogous to a slipped upper femoral epiphysis and occurs in the same age group. The symptoms and signs are similar to those of a prolapsed disc and there are often signs of neurological impairment. A small bony fragment (the edge of the vertebral endplate) may be seen within the spinal canal on plain radiographs. The rim of the epiphysis is usually seen within the canal on CT scans and MRI. The radiculographic appearance is of a large anterior defect or a complete block. Surgical excision of both the extruding disc and the bony ridge is required for relief of symptoms.

Tertti *et al*. (1991) compared 39 15-year-old children with low back pain with 39 asymptomatic control children. When examined by MRI, disc degeneration was found in 15 (38 per cent) of the patients with pain and 10 (26 per cent) of the controls. The difference was not significant. There were no differences between the two groups in Scheuermann-type changes in the lumbar spine, transitional vertebrae, or disc space narrowing. Only disc protrusion was more common in patients in the low back pain group (8 as opposed to 1 in the control group, $p < 0.05$). The authors concluded that disc degeneration was frequently found in this age group and is often asymptomatic.

Scheuermann's osteochondritis

Scheuermann's osteochondritis is a common cause of a thoracic kyphosis in the adolescent and can be painful, although the patient usually presents with a painless kyphosis. The diagnosis is made on plain lateral radiographs of the spine.

Scheuermann's osteochondritis has been classified into two types. Type 1 affects the thoracic spine. It is the commonest form (Fig. 11) and is associated with wedging of the thoracic vertebrae. The typical radiographic appearance (Sorensen 1964) is of wedging of more than 5 degrees of three or more adjacent vertebrae with endplate irregularity. It is usually associated with excessive lumbar and cervical lordosis, tight hamstrings, tight pectoralis muscles, tight hip flexors and anterior pelvic tilt. Progressive curves of more than 45 to 50 degrees are often treated by bracing (Fig. 12) until growth stops (Bradford *et al*. 1974; Sachs *et al*. 1987), but orthotic treatment is only likely to be successful in co-operative, skeletally immature patients. Surgery is sometimes indicated in curves greater than 70 degrees. The majority of patients require physiotherapy, which should include postural awareness, hamstring and pectoral stretching, and trunk extensor strengthening.

Fig. 11 Patient with painful thoracic Scheuermann's osteochondritis. Note the involvement of the anterior ring epiphyses, with slight wedging of the affected vertebrae.

Fig. 12 Scheuermann's osteochondritis. The kyphosis may be associated with a secondary lumbar lordosis.

Type 2 Scheuermann's osteochondritis affects the thoracolumbar and lumbar spine. It is less severe in magnitude but is more commonly painful (Greene *et al*. 1985). It is treated with extension exercises and bracing may sometimes be required.

There is a relationship between lumbar Scheuermann's osteochondritis and hard physical labour in immature teenagers (Greene *et al*. 1985) and it occurs more commonly in young athletes (Micheli 1979), suggesting that it might be due to trauma to the vertebral growth plates. When present, the pain is usually aggravated by activity and responds to simple rest and, if necessary, immobilization in an orthosis.

Other forms of kyphosis (Table 3)

There are two types of congenital kyphosis. Failure of formation is associated with a posterior hemivertebra and can cause paraplegia. Ideally, a posterior fusion *in situ* is required under the age of 2 years. The fusion may have to be re-explored and augmented 6 months after the original operation. If the diagnosis is made later, anterior and posterior fusion are required with anterior resection of the short ligaments, discs, and all the cartilage remnants. A patient with a paraparesis usually requires anterior cord decompression, and anterior and posterior fusion.

Failure of segmentation is associated with an anterior bar. The kyphosis progresses more slowly and the risk of paraplegia is much less.

Neurosurgeons often have to carry out a laminectomy to remove a tumour from a spinal cord or canal. In growing children this will nearly always result in a progressive kyphosis. The longer the laminectomy, the worse the deformity, particularly if the facet joints have had to be removed. Sometimes the laminectomy is combined with posterior fusion. If the child develops a progressive kyphosis, anterior and posterior fusion is required. Bracing can be used as a temporizing measure, but the only definitive treatment for a progressive post-laminectomy kyphosis is fusion.

Spinal deformity is the commonest skeletal problem in patients with neurofibromatosis. It can take the form of a scoliosis or a kyphosis. A scoliosis with significant rotation may present as a kyphoscoliosis. The treatment of a neurofibromatous kyphosis is by anterior and posterior fusion.

Progressive kyphosis secondary to skeletal dysplasia (Fig. 2) may also cause progressive paraparesis and may require anterior and posterior fusion.

In patients with a progressive kyphosis, an early fusion *in situ* is preferable to a fusion with correction of deformity once the latter has progressed to an advanced stage. Surgery for kyphosis carries a definite, albeit small, risk of paraplegia, usually consequent upon vascular damage to the spinal cord, particularly in the thoracic region. In general terms, fusion *in situ* is less likely to be associated with this complication than a fusion combined with correction of the deformity but, nevertheless, neurological sequelae can still occur. Severe kyphosis may continue to worsen after growth has ceased.

Postural kyphosis is by definition a flexible deformity. It is often associated with rounded shoulders and an increased lumbar lordosis. It is treated with postural awareness and an exercise programme. Occasionally, bracing is required.

Spondylolysis and spondylolisthesis

Spondylolysis probably represents a stress or fatigue fracture of the pars interarticularis (Wiltse *et al*. 1975). Although the condition probably occurs in childhood, the symptoms may not occur for many years, and are frequently precipitated by minor trauma. Many individuals remain asymptomatic. Fredrickson *et al*. (1984) found an incidence of spondylolysis with or without spondylolisthesis in 4.4 per cent of schoolchildren at the age of 6 years, increasing to 6 per cent by adolescence. Very few had symptoms. Symptoms commonly occur during the adolescent growth spurt, but may occur in older patients. Jackson *et al*. (1976) found that the incidence was four times higher than expected in female gymnasts, some of whom initially had normal radiographs, and Wiltse *et al*. (1975) reported acute spondylolysis in soldiers who carried heavy back packs and performed exercises to which they were unaccustomed.

The pain is generally localized to the low back and there may be some tenderness and muscle spasm. The origin of the pain is not certain but it has recently been shown that there are nerve endings within the pars defect tissue (Schneiderman *et al*. 1995). Hamstring tightness is commonly found in the symptomatic patient. A

careful neurological examination is essential, but there is rarely any neurological impairment in children with a spondylolysis or grade I spondylolisthesis.

If large, the defect in the pars interarticularis may be seen on routine anteroposterior and lateral radiographs of the spine but, if the defect is unilateral, is small, or is not associated with a spondylolisthesis, it may only be seen on oblique radiographs of the lumbar spine (Fig. 13).

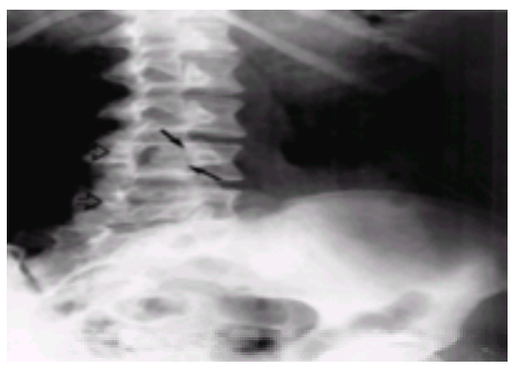


Fig. 13 Spondylolysis. The defect in the pars interarticularis (arrowed) may only be noted on the oblique view. This patient has been treated by a spinal fusion (open arrows).

Occasionally the lesion is due to an acute fracture of the pars interarticularis, when the radiographic appearance is of a narrow gap with irregular edges, whereas in the chronic lesion the edges are smooth. A stress fracture may be detected on a skeletal scintigram, CT scan, or MRI before it becomes evident on radiographs. There may be reactive sclerosis and hypertrophy of one pedicle with a contralateral spondylolysis in the same vertebra (see above). This probably is due to repeated trauma in the presence of an unstable neural arch. Radiographically, and scintigraphically, the lesion can be confused with an osteoid osteoma, but CT scanning will indicate whether a nidus is present (see above).

Asymptomatic spondylolysis does not require treatment. The spondylolytic defect may heal with immobilization in a cast or orthosis in children or adolescents with an acute onset of symptoms and a clearly documented injury (Micheli 1979).

If the spondylolysis has been present for a long time, restriction of vigorous activities and isometric spinal and abdominal muscle strengthening exercises are frequently successful in controlling the pain. If the symptoms are more severe or more persistent, immobilization in a plaster cast or orthosis may be required. The majority of adolescents respond to simple measures. Persistence of hamstring tightness usually indicates that treatment is not successful. If the symptoms do not settle, or if the patient is not prepared to curtail his/her activities, surgical fusion may be required.

Spondylolysis is frequently asymptomatic and the presence of a spondylolytic defect on the radiographs does not necessarily mean that the patient's symptoms are due to the lesion. Other causes of back pain must be excluded.

The vast majority of spondylolistheses (forward slipping of one vertebra upon another) in childhood and adolescence are secondary to spondylolysis (isthmic spondylolisthesis). Clinically, a step may be palpable in the spine and a transverse crease may be noted across the back (Fig. 14). There is a lumbosacral kyphosis.

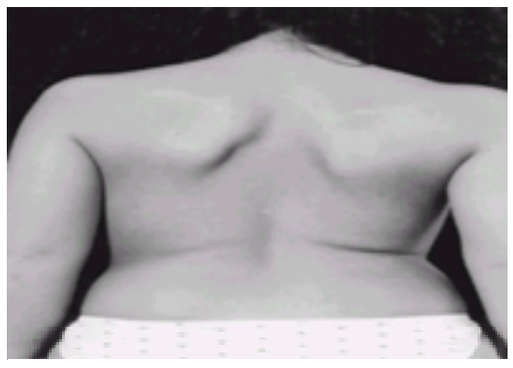


Fig. 14 Spondylolisthesis. Occasionally an abnormal crease may be seen across the back in a patient with a significant spondylolisthesis. Frequently a step is palpable in the spine at the level of the slip.

Occasionally dysplastic spondylolisthesis is noted. Here, the posterior structures are poorly developed, with elongation and attenuation of the pars interarticularis. There is usually marked limitation of forward flexion and significant hamstring tightness. If required, treatment is by fusion.

Degenerative spondylolisthesis does not occur in children and pathological spondylolisthesis is rare. It either results from a generalized skeletal dysplasia, such as osteopetrosis or a mucopolysaccharidosis, or from localized bone destruction, for example following infection or a tumour.

Isthmic spondylolisthesis may progress, particularly during the rapid adolescent growth spurt. Progression seldom occurs after the age of 25 years. In general terms, the spondylolisthesis is more likely to progress if the slip on the initial radiographs is greater than 30 per cent. Severe slip occurs more frequently in females and patients with dysplasia at the lumbosacral junction, including spina bifida. However, these two variables have no statistical value for predicting progression.

The more severe a spondylolisthesis, the less likely are the symptoms to respond to conservative measures. Immobilization in a lightweight plaster cast or orthosis is a useful adjunct in the preoperative assessment. In general, if the symptoms are related to activity and are greatly improved by bed rest or immobilization in a plaster cast or orthosis, a localized fusion is all that is required (Fig. 15). If the symptoms are not relieved by bed rest and/or immobilization, or if there is significant neurological impairment, MRI is indicated and, if necessary, spinal fusion should be combined with decompression. There may be an associated prolapsed disc which requires discectomy at the time of fusion.

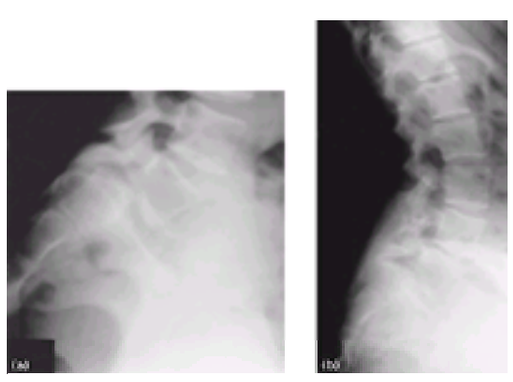


Fig. 15 (a) Patient with spondylolisthesis of long standing and without neurological signs. L5 has slipped forward by about 50 per cent. Note the bone that has developed on the anterior surface of the first sacral segment as a response to the chronic slip. A lumbar kyphosis is clinically evident. (b)This was treated by a

posterior fusion with total relief of her symptoms.

The indications for surgery are slip progression, persistent deformity and abnormal gait, refractory pain, neurological symptoms, and a high slip angle of more than 50 per cent, in a skeletally immature individual. Decompression alone has no place in the management of spondylolisthesis. It leads to increased instability, slip progression, and increased lumbosacral kyphosis.

The rare individual with severe deformity may benefit from reduction of the slippage along with fusion. These are patients with a severe lumbosacral kyphosis for whom *in situ* fusion may be inadequate for postural improvement and symptom relief. However, Poussa *et al*. (1993) compared the results of *in situ* fusion in 11 patients with slips of more than 50 per cent with reduction, internal fixation, and fusion in a similar group of 11 patients and found that the results were better for *in situ* fusion because there were fewer complications and fewer reoperations. The clinical results were the same in both groups. McGuire and Amundson (1993) found no difference in the fusion rate between in situ fusion compared with internal fixation without reduction in a prospective randomized study.

Pedicle sclerosis

Back pain in young athletes may be due to stress-related microfractures in the pars. Radiographs may show a defect or sclerosis. However, the plain radiographs, including oblique views, may appear normal. Under these circumstances, skeletal scintigraphy including SPECT scintigraphy is indicated. This may show an area of increased activity in the pars. The treatment is immobilization in a spinal orthosis for 3 to 6 months. Failure to detect and treat the lesion in the sclerotic phase may lead to a stress fracture and spondylolysis. Biopsy is contraindicated (see above).

Spinal dysraphism

The condition is associated with some evidence of nerve tissue anomaly plus a bony anomaly. The bony anomaly is usually a spina bifida occulta. The neural anomaly is cord tethering which is due to or associated with a variety of disorders including a thickened filum terminale, lipomyelomeningocele, fibrous bands, diastematomyelia, meningocele, neuroenteric cysts, and dermoid cysts. It may be associated with cutaneous manifestations such as a hairy patch, dimple, haemangioma, lipoma, or an area of thin atrophic skin. It may present with back pain; neurological symptoms including leg weakness, fatigue, and urinary incontinence; deformity such as limb length inequality, foot deformity including clawed toes, pes cavus, and foot size inequality; trophic ulceration; and/or sciatica. Scoliosis may be the first manifestation and patients with a congenital scoliosis should undergo MRI to exclude the presence of spinal cord tethering prior to corrective surgery.

The onset may be insidious or may occur acutely after minor trauma. It is thought that the neurological symptoms are due to tethering of the spinal cord with stretching of the cord or nerve roots. However, deterioration may occur at any time during childhood or even in adult life and does not just occur during the growth spurts.

MRI is the investigation of choice. If this is not available, myelography should be carried out with computed tomography. Surgery is not without its hazards and should be carried out by an experienced neurosurgeon. There is some debate as to the indications. In general terms, most neurosurgeons agree that surgery is indicated in newborn children with a lumbosacral lipoma and tethered cord but no neurological abnormalities, patients with neurological symptoms including bladder involvement, and patients with an associated scoliosis who will be undergoing corrective surgery. There is some debate as to whether surgical release is indicated in the asymptomatic adolescent or adult.

Foot size inequality (Fig. 16) or limb length inequality may be secondary to spinal dysraphism and examination of the spine should be undertaken in a child who presents with either of these conditions. Widening of the interpedicular distance may be seen on plain radiographs of the spine and is an indication for MRI.

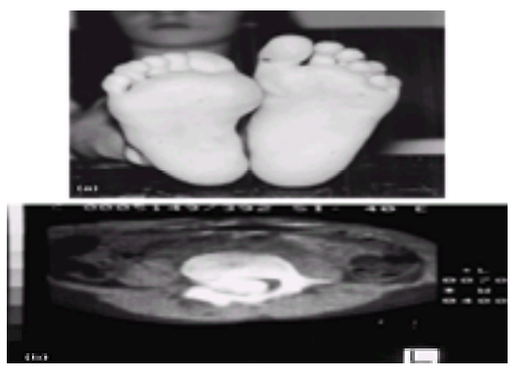


Fig. 16 (a) Patient who presented with inequality of foot size. (b) This was found to be secondary to a thickened and adherent filum terminale.

Spina bifida occulta occurs commonly and in the majority of patients is benign, is not associated with any symptoms, and there is no associated cord tethering.

Diastematomyelia may be seen on plain radiographs but radiculography, CT scanning (Fig. 17), or MRI may be required to make the diagnosis.

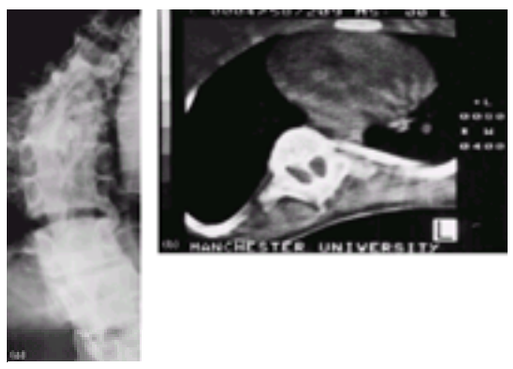


Fig. 17 (a) Patient with congenital scoliosis. A bony spur can just be made out in the centre of the scoliotic segment. (b) CT scan showing the diastematomyelia. This patient had a complete bony spur dividing the spinal canal into two. The diastematomyelia may produce neurological symptoms. Traction on the spinal cord during elective surgery for scoliosis can produce paraplegia and, if a patient has a diastematomyelia, the spur must be resected prior to embarking upon scoliotic corrective surgery.

Congenital deformities of the spine may be associated with congenital deformities elsewhere, particularly in the urinary tract.

Inflammatory disorders

Juvenile ankylosing spondylitis usually occurs in young adults, but can occur in a younger age group. In about 8 to 9 per cent of patients with seronegative

spondylarthropathy, the symptoms start between the ages of 10 and 15 years (Hart 1955; Schaller 1977). However, there is usually a considerable delay before the diagnosis is established. In addition to back pain, these patients may develop a transient arthropathy of the larger peripheral joints. The condition occurs most commonly in males and examination usually reveals diminished chest expansion and restriction of spinal movement. The vast majority of patients have the HLA B27 antigen. Radiographs are usually normal early on in the course of the disease, but there is increased sacroiliac uptake of radionuclide, best revealed by measuring the sacroiliac uptake ratio. Treatment is by non-steroidal anti-inflammatory agents and mobilizing exercises.

Seronegative spondylarthropathies in childhood are often misdiagnosed as juvenile rheumatoid arthritis (Kredich and Patrone 1990). Reiter's syndrome, psoriatic arthritis, and the arthritis associated with inflammatory bowel disease may occur in young adults, but may also be present in children. Extra-articular manifestations include inflammation of the eyes, skin, gastrointestinal tract, and genitourinary tract, and enthesopathies should be sought. The treatment is different to that required for juvenile chronic arthritis (juvenile rheumatoid arthritis) (see Chapter 5.5.2).

Juvenile osteoporosis

Juvenile osteoporosis is a rare disease, particularly when compared with postmenopausal or senile osteoporosis. It may lead to vertebral collapse (Fig. 18) and kyphosis. In general terms, osteoporosis in children may be primary or secondary. The latter complicates disorders such as Cushing's syndrome, the use of steroids, inflammatory spondylarthropathies, juvenile chronic arthritis, and has been described in association with diabetes. Soejima and Landing (1986) showed that osteoporosis was a regular feature of juvenile-onset diabetes mellitus and suggested that the degree of bone matrix and mineral deficiency in such patients was greater than usual. Osteoporosis also occurs with disuse. For example, severe disuse osteoporosis may be seen in patients with severe neuromuscular disorders who are confined to wheelchairs and have very little residual function. Other known causes are coeliac disease, Turner's syndrome, and homocystinuria. Serious disorders such as leukaemia must also be considered in the differential diagnosis.

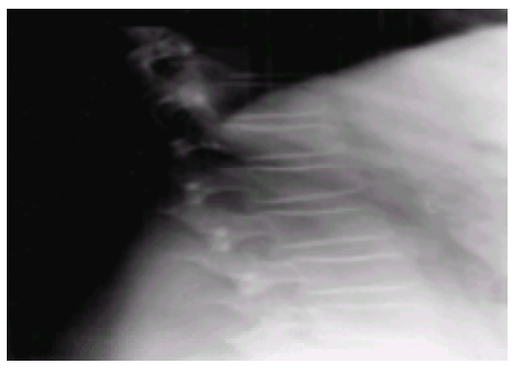


Fig. 18 Osteoporosis. This patient was treated with high doses of steroids for focal multinodular myositis. Within 3 months he developed significant collapse of several vertebrae.

Juvenile idiopathic osteoporosis is a rare condition. The onset is usually between 6 and 13 years of age. It is commoner in males. The patient usually presents with pain, resultant upon a compression fracture of the vertebral body. Height loss of the vertebrae is common. Radiographs show osteopenia.

The patients may also present with a fracture of the lower extremities following a minor fall. Juvenile osteoporosis should be suspected in any child between 6 and 18 years of age who sustains a fracture without a reasonable cause, and these patients should have a spinal radiograph taken. The main differential diagnosis is osteogenesis imperfecta.

The cause of idiopathic juvenile osteoporosis is unknown, although some have postulated a deficiency of 1,25-dihydroxy vitamin D (Marder *et al*. 1982; Saggese *et al*. 1991), and treatment with 1,25-dihydroxycholecalciferol has been suggested.

Routine biochemical measurements are usually within the normal range when due account is taken of the normal increase in alkaline phosphatase and urinary hydroxyproline in the preadolescent growth phase, but Dent and Friedman (1965) and Jowsey and Johnson (1972) reported a tendency for high values of urine calcium in some patients. Dent and Friedman (1965) also reported malabsorption of calcium in some patients. In his experience of 21 patients, Smith (1995) found that routine biochemical measurements were within the normal range as were vitamin D metabolites. In his series the main presenting symptoms were long bone fracture (7), back pain (7), difficulty in walking (5), progressive kyphosis (1), and failure to thrive (1). Some patients had more than one of these symptoms simultaneously. Although pain in the back was not a constant first symptom, the vertebrae were always radiologically abnormal. In the absence of overt fracture, the most striking sign was pain and difficulty in walking. Apart from bone fragility, no patient had features of osteogenesis imperfecta.

Patients with juvenile osteoporosis are likely to develop crush fractures of their vertebral bodies, irrespective of the cause of their osteoporosis. The vertebral trabecular bone density is abnormally low (Fredericks *et al*. 1990), the lowest values being found in the spine and hip soon after the onset of symptoms (Smith 1995). Bone biopsy shows either the features of osteoporosis alone or of an excess of osteocytes sometimes associated with woven bone or changes similar to those found in type IV osteogenesis imperfecta (Smith 1995). Smith (1995) suggested that the primary defect was a failure of osteoblastic activity, whereas Hoekman *et al*. (1985) suggested that it was uncontrolled activity of the metaphyseal osteoclasts and that treatment with bisphosphonates may be beneficial. Collagen studies are usually normal although minor abnormalities were found in occasional patients (Smith 1995).

Radiographs show metaphyseal compression fracture in the long bone and progressive loss of height in the vertebrae. When the condition improves, improvement is first noted in the vertebral bodies. Unlike osteogenesis imperfecta, patients with idiopathic juvenile osteoporosis do not have Wormian bones or the other features of osteogenesis imperfecta. Radiographs of the vertebral column show vertebral collapse, biconcave vertebrae, wedged vertebrae, and loss of vertebral height.

In most patients, significant improvement occurs after the adolescent growth spurt is complete, although mild kyphosis and probably a low normal bone density may persist (Smith 1995). Because improvement may occur spontaneously, it is difficult to assess the effectiveness of the many treatments that have been given. These include appropriate sex hormones (Wright *et al*. 1995), additional calcium, physiological amounts of vitamin D, bisphosphonates (Hoekman *et al*. 1985; Levis *et al*. 1993), and calcitonin (Jackson *et al*. 1988), given separately or together. However, improvement does not always occur. Smith (1995) found that 11 of 14 patients, who had been followed up until growth had ceased, improved substantially or completely, but the remaining 3 were disabled and were eventually confined to a wheelchair.

If osteoporosis is suspected, it is essential to exclude possible causes. For example, patients with leukaemia may present with back pain. The pain may be associated with osteoporosis and vertebral compression fractures. Chemotherapy often includes steroids such as prednisolone or dexamethazone, which may cause osteoporosis. This is reversible and settles after treatment is discontinued. Once these causes of osteoporosis have been excluded, a skin biopsy for fibroblast culture and collagen studies is indicated to exclude the possibility of osteogenesis imperfecta. If there is a reduction in 1,25-dihydrovitamin D, treatment with 1,25-dihydroxycholecalciferol may be worthwhile. Bone biopsy may be helpful in excluding haematological disorders. If the diagnosis of idiopathic juvenile osteoporosis is established and if the investigations are normal, there is probably no indication for the use of sex hormones or bisphosphonates but symptomatic treatment is indicated.

Trauma

Trauma of the spine probably occurs quite often in children and adolescents. Sardelic and Ryan (1989) surveyed 2489 students, aged 11 to 18 years. Only 4.5 per cent (111) had had disabling low back pain and 0.4 per cent (11) had been admitted to hospital. However, 53.1 per cent had complained of pain, the commonest age of onset being 13 years. Of those who experienced pain, 58.4 per cent could recall a preceding back injury. Grantham (1977) found that 11.5 per cent of the male adolescent population of the school he studied complained of back pain, and Fairbank *et al*. (1984) found that 26 per cent of the students they studied, who ranged in age from 14 to 17 years, suffered from back pain. It is likely that trauma is responsible for most cases of non-specific back pain, the pain usually resolving rapidly

without treatment. Trauma may precipitate pain from underlying conditions, such as spondylolysis.

In children, trauma is very rarely associated with a significant injury, but fractures of the spine do occur in childhood (Fig. 19). If stable (e.g. minor compression fractures), the treatment is symptomatic but, if unstable, the fracture must be treated on its merits.

Fig. 19 Compression fractures of L1, L2, and L3 following a fall from a height. The patient was treated with bed rest until the acute pain had settled. He was then mobilized with a spinal orthosis, which he wore for some weeks, and physiotherapy.

Deformity

Spinal deformity in children is usually pain free. If it is associated with pain, it usually implies a significant underlying pathological process which requires investigation and treatment (see above). However, some patients with idiopathic scoliosis complain of mild back pain that responds to physiotherapy.

Scoliosis

Non-structural scoliosis

There is usually a list to one side without associated rotation. It is frequently secondary to an underlying disorder and provided the underlying disorder is corrected, the list will disappear. However, if the underlying disorder is allowed to persist, the scoliosis may become structural.

The commonest causes for a non-structural scoliosis include postural scoliosis, nerve root irritation (for example caused by a prolapsed intervertebral disc or tumour), inflammation (for example appendicitis), lower limb length discrepancy, or contractures around the hip with pelvic obliquity. It may rarely be psychogenic.

Structural scoliosis

This occurs most commonly in adolescents. School screening studies suggest that approximately 3 per cent of adolescent girls have a curve greater than 10 degrees, but in only a tenth of such patients is treatment indicated. The causes of scoliosis are shown in Table 2.

The aetiology of idiopathic scoliosis is not known. However, it occurs most commonly in adolescent girls and there is a genetic predisposition. It has been suggested that it is a balance disorder involving central vestibular centres. There are many other suggestions, none of which have been confirmed. The deformity consists of a lateral curvature of the spine which has also rotated so that the ribs on the convex side become more prominent and, in the lumbar spine, the paravertebral muscles are more prominent on the convex side. This rotation forms the basis of the bending test for the early diagnosis. The child bends forward and the examiner looks for a rotatory prominence of the ribs/paravertebral region. The degree of rotation can be measured.

The severity of the curve is measured on a posteroanterior view of the spine. This minimizes breast irradiation compared with an anteroposterior film. Patients with congenital scoliosis have a high incidence of associated congenital renal abnormalities and renal ultrasound may be necessary. Left-sided curves, particularly in boys, may be associated with intraspinal anomalies.

Modern methods of instrumentation may help to derotate the spine at the time of surgery (if indicated) but, in patients with a severe rib prominence, a costoplasty is often carried out at the time of posterior fusion and instrumentation. Division of the most prominent ribs, usually near the costovertebral joint, allows the rib cage to be remoulded and reduces the severe rib prominence.

Treatment depends on the severity of the curve, the age of the patient, and the underlying cause. In general terms, idiopathic and congenital scoliosis tend to affect one part of the spine, whereas neuromuscular scoliosis tends to affect the thoracic and lumbar spine, the patient developing an extensive 'C'-shaped collapsing spine. This usually involves the sacrum and pelvis, and progression of the curve is usually associated with progressive pelvic obliquity.

Idiopathic or congenital scoliosis may be associated with respiratory involvement if the curve is very severe and develops in the first few years of life; whereas some forms of neuromuscular scoliosis are always associated with respiratory changes. For example, the vital capacity is reduced by 4 per cent for every 10-degree increase in curve in patients with Duchenne muscular dystrophy (Kurz *et al*. 1983).

Unless severe, idiopathic or congenital scoliosis tend not to progress once growth has ceased, but neuromuscular scoliosis maintains its tendency to progress throughout life.

The treatment of each type of scoliosis is beyond the scope of this chapter. In general terms, adolescent idiopathic scoliosis is treated by regular observation until skeletal maturity if the curve is less than 25 to 30 degrees. Moderate curves of 30 to 45 degrees are frequently treated by bracing until skeletal maturity, whereas curves greater than 40 to 45 degrees in a growing adolescent are often treated by surgical correction and spinal fusion. Progressive congenital scoliosis is usually treated by spinal fusion. Progressive neuromuscular scoliosis is usually treated by spinal stabilization and fusion. In some disorders, such as Duchenne muscular dystrophy, spinal stabilization and fusion is indicated in much smaller curves whilst the patient is still fit for surgery (Galasko *et al*. 1992). In neuromuscular scoliosis, spinal stabilization and fusion may be associated with improved maintenance of lung function. The type of fusion depends on the site and severity of the curve.

Box 1 Important points to remember

1. Children with spinal disorders may present with deformity, pain, neurological disturbance, or a combination.

2. A young child may present with refusal to walk. An older child may present with a limp.

3. Symptoms and signs may be minor despite the child having a serious and even possibly a life-threatening disorder.

4. All patients require a careful history and clinical examination.

5. The history may be non-specific.

6. The examination must include a general examination of the patient, the spine, and the lower limbs and a detailed neurological examination.

7. The investigations should be based on the history and clinical findings. A standard 'work up' is not indicated.

8. A painful scoliosis usually implies a significant lesion producing the deformity.

9. Spinal tumours are rare. They may present as accidental findings or with pain, deformity, stiffness, or neurological symptoms. Because the initial symptoms and signs may be 'minor', the diagnosis may be significantly delayed particularly with intraspinal tumours.

10. Spinal infection occurs rarely in children. Discitis of childhood usually occurs between the ages of 2 and 5. The degree of stiffness is usually striking.

11. Disc herniation can occur in childhood and adolescence. The symptoms are usually less severe than in the adult and significant spasm is usually present.

12. There are several causes for kyphosis. Those associated with a posterior hemivertebra may cause paraplegia. Ideally, a posterior fusion *in situ* is required under the age of 2 years. Post-laminectomy kyphosis may require anterior and posterior fusion.

13. Asymptomatic spondylolysis does not require treatment. The presence of a spondylolytic defect on a radiograph does not necessarily mean that the patient's symptoms are due to the lesion. Other causes of back pain must be excluded.

14. Foot deformities such as pes cavus, clawed toes, foot size inequality and limb length inequality may be due to spinal dysraphism. Cutaneous manifestations in the low back may also be present. Cord tethering may lead to significant neurological loss. The onset may be insidious.

15. The spine may be involved in inflammatory disorders.

16. Juvenile osteoporosis is rare. It may be primary or secondary.

17. Trauma of the spine probably occurs commonly but rarely causes disabling symptoms or significant problems.

18. Spinal deformity in children and adolescents is usually pain free. If it is painful it frequently is secondary to a significant underlying pathological process.

19. Back pain may be referred from other sites, e.g. renal disease.

20. Back pain in children often indicates severe underlying pathology. In the absence of any neurological symptoms or signs and in the absence of a secondary deformity, conservative measures are indicated in patients with a mobile spine and normal straight leg raising. If symptoms persist for more than 2 months, further investigation is required.

21. In the majority of children and adolescents spinal deformity requires no active treatment. Underlying neuromuscular and other conditions must be excluded. Severe or progressive deformities, particularly if they are likely to be associated with progressive impairment of pulmonary or neurological function, require correction and surgical stabilization.

Kyphosis (see above)

Lordosis

This is usually postural, but may be secondary to congenital deformities, dorsal kyphosis (Fig. 12), neuromuscular disease associated with an hyperextension spinal contracture, or fixed flexion contractures of the hips.

Referred pain

Back pain may be the presenting feature of renal disease, gastrointestinal disorders, or gynaecological disorders even in premenarchal adolescents (Letts and Haasbeek 1990), including haematocolpus and haematometra. A child with pneumonia may present with back pain.

Conclusions

Back pain that requires referral to an orthopaedic surgeon occurs infrequently in children and adolescents, but it often indicates serious underlying pathology. A careful history and clinical examination is essential (see Box 1). If there is any associated neurological impairment, immediate investigation is required. However, in the absence of any neurological symptoms or signs and in the absence of a secondary deformity, conservative measures are indicated in patients with a mobile spine and a normal straight leg raising but, if the symptoms persist for more than 2 months, further investigation is required.

Spinal deformity is also relatively uncommon and in the vast majority of patients requires no active treatment. Underlying neuromuscular and other conditions must be excluded. Severe or progressive deformities, particularly if they are likely to be associated with progressive impairment of pulmonary or neurological function, require correction and surgical stabilization.

Chapter References

Akbarnia, B.A. and Rooholamini, S.A. (1981). Scoliosis caused by benign osteoblastoma of the thoracic or lumbar spine. *Journal of Bone and Joint Surgery*, **63A**, 1146–55.

Balague, F., Dutoit, G., and Waldburger, M. (1988). Low back pain in schoolchildren. *Scandinavian Journal of Rehabilitation Medicine*, **20**, 175–9.

Banerian, K.G., Wang, A.M., Samberg, L.C., Kerr, H.H., and Wesolowski, D.P. (1990). Association of vertebral end plate fracture with pediatric lumbar intervertebral disk herniation: value of CT and MR imaging. *Radiology*, **177**, 763–5.

Bartels, R.H.M.A., Gonera, E.G., van der Spek, J.A.N., Thijssen, H.O.M., Mullaart, R.A., and Gabreëls, F.J.M. (1995). Intramedullary spinal cord abscess. A case report. *Spine*, **20**, 1199–204.

Bellah, R.D., Summerville, D.A., Treves, S.T., and Micheli, L.J. (1991). Low-back pain in adolescent athletes: detection of stress injury to the pars interarticularis with SPECT. *Radiology*, **180**, 509–12.

Bradford, D.S., Moe, J., Montalvo, F.J., and Winter, R.B. (1974). Scheuermann's kyphosis and roundback deformity. Results of Milwaukee brace treatment. *Journal of Bone and Joint Surgery*, **56A**, 740–58.

Bruschwein, D.A., Brown, M.L., and McLeod, R.A. (1980). Gallium scintigraphy in the evaluation of disc-space infections. *Journal of Nuclear Medicine*, **21**, 925–7.

Bunnell, W.P. (1982). Back pain in children. *Orthopedic Clinics of North America*, **13**, 587–604.

Byers, P.D. (1968). Solitary benign osteoblastic lesions of bone: osteoid osteoma and benign osteoblastoma. *Cancer*, **22**, 43–57.

Capanna, R., Albisinni, U., Picci, P., Calderoni, P., Campanacci, M., and Springfield, D.S. (1985). Aneurysmal bone cysts of the spine. *Journal of Bone and Joint Surgery*, **67A**, 527–31.

Chan, S. and Ryan, M.D. (1992). Low-back pain in school children in the fifth and sixth grade. *Journal of Orthopaedic Rheumatology*, **5**, 43–47.

Colbert, A.P. and Craig, C. (1987). Scoliosis management in Duchenne muscular dystrophy: prospective study of modified Jewett hyperextension brace. *Archives of Physical Medicine and Rehabilitation*, **68**, 302–4.

De Luca, P.F., Mason, D.E., Weiand, R., Howard, R., and Bassett, G.S. (1994). Excision of herniated nucleus pulposus in children and adolescents. *Journal of Pediatric Orthopedics*, **14**, 318–22.

De Orio, J.K. and Bianco, A.J., Jr. (1982). Lumbar disc excision in children and adolescents. *Journal of Bone and Joint Surgery*, **64A**, 991–6.

De Souza, Dias, L. and Frost, H.M. (1973). Osteoblastoma of the spine. A review and report of eight new cases. *Clinical Orthopaedics and Related Research*, **91**, 141–51.

Dent, C.E. and Friedman, M. (1965). Idiopathic juvenile osteoporosis. *Quarterly Journal of Medicine*, **34**, 177–210.

Fairbank, J.C.T., Pynsent, P.B., Van Poortvliet, J.A., and Phillips, H. (1984). Influence of anthropometric factors and joint laxity in the incidence of adolescent back pain. *Spine*, **9**, 461–4.

Fearnside, M.R. and Adams, C.B.T. (1978). Tumours of the cauda equina. *Journal of Neurology, Neurosurgery and Psychiatry*, **41**, 24–31.

Fredericks, B.J., de Campo, J.F., Sephton, R., and McCredie, D.A. (1990). Computed tomographic assessment of vertebral bone mineral in childhood. *Skeletal Radiology*, **19**, 99–102.

Fredrickson, B.E., Baker, D., McHolick, W.J., Yuan, H.A., and Lubicky, J.P.L. (1984). The natural history of spondylolysis and spondylolisthesis. *Journal of Bone and Joint Surgery*, **66A**, 699–707.

Galasko, C.S.B. (1989). The management of bone and joint infection. *British Journal of Hospital Medicine*, **42**, 32–44.

Galasko, C.S.B. (1990). Tumours of the spine. In *Spinal disorders in the child and adolescent* (ed. J.D. Hsu), pp. 101–19. Hanley and Belfus, Philadelphia.

Galasko, C.S.B., Delaney, C., and Morris, P. (1992). Spinal stabilisation in Duchenne muscular dystrophy. *Journal of Bone and Joint Surgery*, **74B**, 210–14.

Grantham, V.A. (1977). Backache in boys—a new problem? *Practitioner*, **218**, 226–9.

Greene, T.L., Hensinger, R.N., and Hunter, L.Y. (1985). Back pain and vertebral changes simulating Scheuermann's disease. *Journal of Pediatric Orthopedics*, **5**, 1–7.

Handel, S.F., Twiford, T.W., Jr, Reigel, D.H., and Kaufmann, H.H. (1979). Posterior lumbar apophyseal fractures. *Radiology*, **130**, 629–33.

Hart, F.D. (1955). Ankylosing spondylitis. A review of 184 cases. *Annals of the Rheumatic Diseases*, **14**, 77–83.

Hay, M.C., Paterson, D., and Taylor, T.K.F. (1978). Aneurysmal bone cysts of the spine. *Journal of Bone and Joint Surgery*, **60B**, 406–11.

Hensinger, R.N. (1989). Back pain in children. In *Lumbar spine syndromes*: *evaluation and treatment* (ed. G.B.T. Andersson and T.W. McNeill), pp. 125–52. Springer-Verlag, New York.

Hoekman, K., Papapoulos, S.E., Peters, A.C., and Bijvoet, O.L. (1985). Characteristics and bisphosphonate treatment of a patient with juvenile osteoporosis. *Journal of Clinical Endocrinology and Metabolism*, **61**, 952–6.

Hoffman, E.B., Crosier, J.H., and Cremin, B.J. (1993). Imaging in children with spinal tuberculosis. *Journal of Bone and Joint Surgery*, **75B**, 233–9.

Hoffman, H.J. (1980). Childhood and adolescent lumbar pain. *Clinical Neurosurgery*, **27**, 553–76.

Hsu, J.D. (1983). The natural history of spine curvature progression in the non-ambulatory Duchenne muscular dystrophy patient. *Spine*, **8**, 771–5.

Jackson, D.W., Wiltse, L.L., and Cirincoine, R.J. (1976). Spondylolysis in the female gymnast. *Clinical Orthopaedics and Related Research*, **117**, 68–73.

Jackson, E.C., Strife, C.F., Tsang, R.C., and Marder, H.K. (1988). Effect of calcitonin replacement therapy in idiopathic juvenile osteoporosis. *American Journal of Diseases of Children*, **142**, 1237–9.

Jackson, R.P., Reckling, R.W., and Mantz, F.A. (1977). Osteoid osteoma and osteoblastoma: similar histologic lesions with different natural histories. *Clinical Orthopaedics and Related Research*, **128**, 303–13.

Jansen, B.R.H., Hart, W., and Schreuder, O. (1993). Discitis in childhood. *Acta Orthopaedica Scandinavica*, **64**, 33–6.

Jowsey, J. and Johnson, K.A. (1972). Juvenile osteoporosis: bone findings in seven patients. *Journal of Pediatrics*, **81**, 511–17.

Keller, R.H. (1974). Traumatic displacement of the cartilaginous vertebral rim: a sign of intervertebral disc prolapse. *Radiology*, **110**, 21–4.

Kredich, D. and Patrone, N.A. (1990). Pediatric spondyloarthropathies. *Clinical Orthopaedics and Related Research*, **259**, 18–22.

Kurz, L.T., Mubarak, S.J., Schultz, P., Park, S.M., and Leach, J. (1983). Correlation of scoliosis and pulmonary function in Duchenne muscular dystrophy. *Journal of Pediatric Orthopedics*, **3**, 347–53.

Letts, M. and Haasbeek, J. (1990). Hematocolpos as a cause of back pain in premenarchal adolescents. *Journal of Pediatric Orthopedics*, **10**, 731–2.

Levis, S., Gruber, H.E., Cohn, D., Howard, G.A., and Roos, B.A. (1993). Juvenile osteoporosis treated with pamidronate. *Calcified Tissue International*, **52** (Suppl.), S41.

Lichtenstein, L. (1953). Histiocytosis X. Integration of eosinophilic granuloma of bone, 'Letterer–Siwe disease' and 'Schüller–Christian disease' as related manifestations of a single nosologic entity. *Archives of Pathology*, **56**, 84–102.

Lippitt, A.B. (1976). Fracture of a vertebral body end plate and disc protrusion causing subarachnoid block in an adolescent. *Clinical Orthopaedics and Related Research*, **116**, 112–15.

Lowrey, J.J. (1973). Dislocated lumbar vertebral epiphysis in adolescent children. Report of three cases. *Journal of Neurosurgery*, **38**, 232–4.

Marder, H.K., Tsang, R.C., Hug, G., and Crawford, A.C. (1982). Calcitriol deficiency in idiopathic juvenile osteoporosis. *American Journal of Diseases in Children*, **136**, 914–17.

McGuire, R.A. and Amundson, G.M. (1993). The use of primary internal fixation in spondylolisthesis. *Spine*, **18**, 1662–72.

Medical Research Council (1985). A 10-year assessment of controlled trials of inpatient and outpatient treatment and of plaster-of-Paris jackets for tuberculosis of the spine in children on standard chemotherapy. *Journal of Bone and Joint Surgery*, **67B**, 103–10.

Mehta, M.H. and Murray, R.O. (1977). Scoliosis provoked by painful vertebral lesions. *Skeletal Radiology*, **1**, 223–30.

Micheli, L.J. (1979). Low back pain in the adolescent: differential diagnosis. *American Journal of Sports Medicine*, **7**, 362–4.

Norris, S., Ehrlich, M.G., Keim, D.E., Guiterman, H., and McKusick, K.A. (1978). Early diagnosis of disc-space infection using gallium-67. *Journal of Nuclear Medicine*, **19**, 384–6.

Novick, G.S., Pavlov, H., and Bullough, P.G. (1982). Osteochondroma of the cervical spine: Report of 2 cases in preadolescent males. *Skeletal Radiology*, **8**, 13–15.

Pena, M., Galasko, C.S.B., and Barrie, J.L. (1992). Delay in diagnosis of intradural spinal tumours. *Spine*, **17**, 1110–16.

Pettine, K.A. and Klassen, R.A. (1986). Osteoid osteoma and osteoblastoma of the spine. *Journal of Bone and Joint Surgery*, **68A**, 354–61.

Poussa, M., Schlenzka, D., Seitsalo, S., Ylikoski, M., Hurri, H., and Österman, K. (1993). Surgical treatment of severe isthmic spondylolisthesis in adolescents. Reduction or fusion *in situ. Spine*, **18**, 894–901.

Reimer, R. and Onofrio, B.M. (1985). Astrocytomas of the spinal cord in children and adolescents. *Journal of Neurosurgery*, **63**, 669–75.

Ryöppy, S., Jääskeläinen, J., Rapola, J., and Alberty, A. (1993). Nonspecific diskitis in children. A nonmicrobial disease? *Clinical Orthopaedics and Related Research*, **297**, 95–9.

Sachs, B., Bradford, D., Winter, R., Lonstein, J., Moe, J., and Willson, S. (1987). Scheuermann's kyphosis: follow-up of Milwaukee brace treatment. *Journal of Bone and Joint Surgery*, **69A**, 50–7.

Saggesse, G., Bertelloni, S., Baroncelli, G.I., Perri, G., and Calderazzi, A. (1991). Mineral metabolism and calcitriol therapy in idiopathic juvenile osteoporosis. *American Journal of Diseases of Children*, **145**, 457–62.

Salminen, J.J., Maki, P., Oksanen, A., and Pentti, J. (1992). Spinal mobility and trunk muscle strength in 15-year-old schoolchildren with and without low-back pain. *Spine*, **17**, 405–11.

Sardelic, F. and Ryan, M.D. (1989). Low back pain in adolescents. *Journal of Orthopaedic Rheumatology*, **2**, 23–30.

Schaller, J.G. (1977). Ankylosing spondylitis of childhood onset. *Arthritis and Rheumatism*, **20**, 398–401.

Schneiderman, G.A., McLain, R.F., Hambly, M.F., and Nielsen, S.L. (1995). The pars defect as a pain source. *Spine*, **20**, 1761–4.

Sherman, F.C., Wilkinson, R.H., and Hall, J.E. (1977). Reactive sclerosis of a pedicle and spondylolysis in the lumbar spine. *Journal of Bone and Joint Surgery*, **59A**, 49–54.

Silvers, H.R., Lewis, P.J., Clabeaux, D.E., and Asch, H.L. (1994). Lumbar disc excisions in patients under the age of 21 years. *Spine*, **19**, 2387–92.

Smith, R. (1995). Idiopathic juvenile osteoporosis: experience of twenty-one patients. *British Journal of Rheumatology*, **34**, 68–77.

Soejima, K. and Landing, B.H. (1986). Osteoporosis in juvenile-onset diabetes mellitus: morphometric and comparative studies. *Pediatric Pathology*, **6**, 289–29.

Sorensen, H.K. (1964*). Scheuermann's kyphosis*: *clinical appearances, radiography, aetiology and prognosis*. Munksgaard, Copenhagen.

Stern, W.E. (1978). Localisation and diagnosis of spinal cord tumours. *Clinical Neurosurgery*, **25**, 480–94.

Tachdjian, M.O. and Matson, D.D. (1965). Orthopaedic aspects of intraspinal tumours in infants and children. *Journal of Bone and Joint Surgery*, **47A**, 223–48.

Tertti, M.O., Salminen, J.J., Paajanen, H.E.K., Terho, P.H., and Kormano, M.J. (1991). Low-back pain and disk degeneration in children: a case–control MR imaging study. *Radiology*, **180**, 503–7.

Tillman, B.P., Dahlin, D.C., Lipscomb, P.R., and Stewart, J.R. (1968). Aneurysmal bone cyst. An analysis of ninety-five cases. *Mayo Clinic Proceedings*, **43**, 478–95.

Turner, P.G., Green, J.H., and Galasko, C.S.B. (1989). Back pain in childhood. *Spine*, **14**, 812–14.

Wenger, D.R. (1993). Back pain in children. In *The art and practice of children's orthopaedics* (ed. D.R. Wenger and M. Rang), pp. 455–86. Raven Press, New York.

Wenger, D.R., Bobechko, W.P., and Gilday, D.L. (1978). The spectrum of intervertebral disc-space infection in children. *Journal of Bone and Joint Surgery*, **60A**, 100–8.

Wiltse, L.L., Widell, E.H., and Jackson, D.W. (1975). Fatigue fracture: the basic lesion in isthmic spondylolisthesis. *Journal of Bone and Joint Surgery*, **57A**, 17–22.

Winter, R.B. and Lipscomb, P.R. (1978). Back pain in children. *Minnesota Medicine*, **61**, 141–7.

Woltman, H.W., Kernohan, J.W., Adson, A.W., and Craig, W. McK. (1951). Intramedullary tumours of spinal cord and gliomas of intradural portion of filum terminale. *Archives of Neurology and Psychiatry*, **65**, 378–95.

Wright, N.M., Metzger, D.L., and Key, L.L. (1995). Estrogen and diclofenac sodium therapy in a prepubertal female with idiopathic juvenile osteoporosis. *Journal of Pediatric Endocrinology and Metabolism*, **8**, 135–9.

1.2.2.1 The upper limbs in adults

Adel G. Fam

Applied anatomy

Shoulder

Shoulder movements are a synthesis of motion at four articulations: glenohumeral, acromioclavicular, sternoclavicular, and scapulothoracic (Lucas 1973; Rothman *et al*. 1975; Poppen and Walker 1976; Bateman 1978; Polley and Hunder 1978; Bechtol 1980; Sarrafian 1983; Williams *et al*. 1989).

The sternoclavicular joint is a spheroidal joint between the medial end of the clavicle and the manubrium sterni and first costal cartilage. An intra-articular fibrocartilaginous disc divides the joint into two cavities. The joint capsule is reinforced by the anterior and posterior sternoclavicular ligaments.

The acromioclavicular joint is a spheroidal joint between the lateral end of the clavicle and the acromion. An intra-articular fibrocartilaginous disc divides the joint into two compartments. The joint capsule is strengthened by the superior and inferior acromioclavicular ligaments. The coracoclavicular ligament (conoid and trapezoid parts) extends between the distal clavicle and the coracoid process of the scapula. It stabilizes both the clavicle and scapula, and maintains a close relation between the two bones during shoulder movements, thus limiting scapular rotation around the acromioclavicular joint. A subcutaneous, often non-communicating, bursa may be present over the joint.

Movements at the sternoclavicular and acromioclavicular joints enable slight rotation of the clavicle along its long axis, and allow elevation or depression (as in shoulder shrugging), and flexion or extension (as in forward or backward thrusting) of the shoulder girdle.

The glenohumeral joint is a ball-and-socket articulation between the glenoid fossa of the scapula and the humeral head (Fig. 1). The lax articular capsule, and the small area of contact between the shallow glenoid cavity and the spheroidal humeral head, permit a wide range of movements. The stability of the joint depends upon a number of static and dynamic stabilizers. Static stabilizers include the capsule, glenoid labrum, and ligaments (glenohumeral and coracohumeral). The capsule has two apertures: one for the long biceps tendon and the second for the subscapularis bursa. Dynamic stabilizers play a greater part in the stability of the shoulder, and include two musculotendinous layers: an inner stratum, made of the rotator cuff muscles (supraspinatus, infraspinatus, teres minor and subscapularis), and the tendon of the long head of the biceps, and an outer stratum made of the deltoid, teres major, pectoralis major, latissimus dorsi, and trapezius muscles.

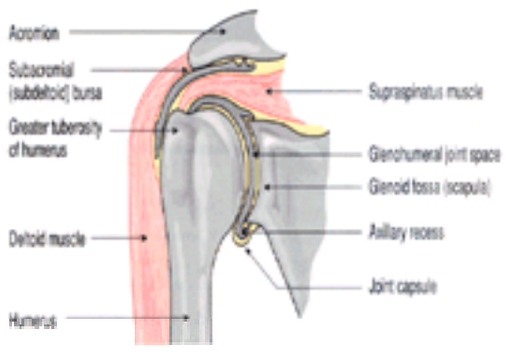


Fig. 1 The shoulder.

The muscles of the inner stratum stabilize and retain the humeral head in the glenoid cavity during shoulder movements, while simultaneously providing abduction (through the supraspinatus, which inserts into the superior part of the greater tuberosity), external rotation (through the infraspinatus and teres minor, which insert into the posterior aspect of the greater tuberosity), and internal rotation (through the subscapularis, which inserts into the lesser tuberosity). At the initiation of shoulder abduction, the rotator cuff muscles and the long biceps tendon depress and fix the humeral head against the glenoid cavity to counteract the upward pull of the more powerful deltoid muscle. The mechanism whereby these two groups of muscles combine to produce abduction, the one elevating (deltoid muscle) and the other stabilizing the humeral head (rotator cuff and biceps tendons), is termed 'force-couple' (Lucas 1973; Rothman *et al*. 1975; Bechtol 1980).

The muscles of the outer stratum are the prime movers of the shoulder, although the trapezius acts through movements of the scapula and clavicle. These muscles provide abduction, flexion, extension, adduction and some degree of rotation, and together with the rotator cuff muscles (which provide more rotation of the humeral head) permit a wide range of movement at the shoulder.

The coracoacromial arch (coracoid, coracoacromial ligament, and acromion) acts as a secondary socket for the humeral head under which the rotator cuff tendons and long head of the biceps tendon glide, with the subacromial bursa lying in between. The arch protects the humeral head and the rotator cuff from direct trauma. The subacromial bursa facilitates movements of the greater tuberosity beneath the rigid coracoacromial arch during shoulder abduction, and acts as a cushion minimizing the impingement of the acromion on the supraspinatus, infraspinatus, teres minor, and long biceps tendons during abduction (Fig. 1) (Lucas 1973). The bursa, which extends beneath the deltoid (subdeltoid part), communicates with the cavity of the glenohumeral joint in about one-third of adults. The synovium of the shoulder has two extracapsular outpouchings: the synovial tendon sheath of the long biceps tendon, and the subscapularis bursa lying beneath the subscapularis tendon (Fig. 2). The subcoracoid bursa lies between the shoulder capsule and coracoid process, but it rarely communicates with the subacromial bursa (Lucas 1973; Rothman *et al*. 1975; Poppen and Walker 1976; Bateman 1978; Polley and Hunder 1978; Bechtol 1980; Sarrafian 1983; Williams *et al*. 1989).

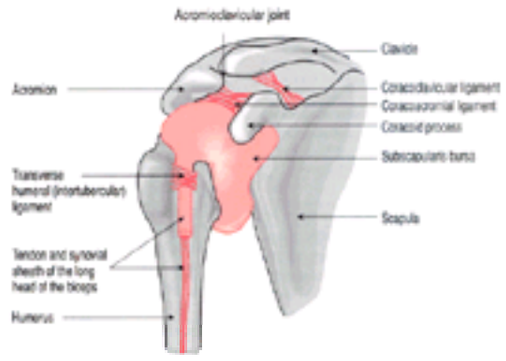


Fig. 2 The shoulder (synovial membrane and its outpouchings).

Scapulothoracic articulation

The scapula is connected to the posterior aspect of the chest wall by the axioappendicular muscles. The scapula provides the origin for the rotator cuff muscles and deltoid, and the trapezius inserts into its superior aspect. Scapulothoracic movements, including rotation, elevation, depression, protrusion, retraction and circumduction, are important for the normal functioning of the shoulder (Lucas 1973; Rothman *et al*. 1975; Poppen and Walker 1976 ; Bateman 1978; Polley and Hunder 1978; Bechtol 1980; Sarrafian 1983; Williams *et al*. 1989).

Shoulder abduction involves synchronous movements of the glenohumeral, sternoclavicular, and acromioclavicular joints and rotation of the scapula on the chest wall. The initial 30° of abduction, achieved by contraction of the supraspinatus, takes place at the glenohumeral joint with little movement of the scapula. Beyond 30°, an approximate 2:1 ratio exists between movements at the glenohumeral joint and the scapula. The combined movement is referred to as the 'scapulohumeral rhythm' (Lucas 1973). Abduction of the shoulder (normal range, 180°) is associated with external rotation of the humerus. The prime movers are the supraspinatus and deltoid muscles. The flexors of the shoulders (normal range, 180°) are the deltoid and coracobrachialis muscles. The normal range of shoulder adduction across the front of the chest is about 50°. The pectoralis major muscle is the main adductor. The normal range of extension (posterior flexion) is about 50°. The latissimus dorsi, teres major, and deltoid muscles are the principal extensors. With the shoulder abducted to 90° and the elbow flexed at a right angle, the normal range of internal and external rotation is 90° each. The normal range of external rotation with the elbow placed by the side at the waist is about 45°, and internal rotation is to 55° (before its motion is stopped by the body), or to 120° if the patient can reach behind the back to touch the inferior angle of the opposite scapula. The prime movers of internal rotation are subscapularis, and pectoralis major. The infraspinatus and teres minor are the main external rotators.

Elbow

The elbow is a relatively stable hinge joint formed by the humeroulnar (trochleo-ulnar), humeroradial (capitelloradial) and proximal radio-ulnar articulations (Fig. 3), with the trochleo-ulnar being the principal joint (Polley and Hunder 1978; Williams *et al*. 1989). The stability of the joint depends on its congruity, anterior capsule, and ulnar and radial collateral ligaments (Morrey and An 1983). The common flexor tendon (pronator teres, flexor carpi radialis, palmaris longus, flexor carpi ulnaris, and flexor digitorum superficialis) takes origin from the medial epicondyle of the humerus. The common extensor tendon (brachioradialis, extensor carpi radialis longus and brevis, extensor digitorum communis and anconeus) originates from the lateral epicondyle. The annular ligament is a strong, cup-shaped band that encircles the radial head at the proximal radio-ulnar joint. The paraolecranon grooves are the depressions between the ulnar olecranon process and the medial and lateral epicondyles. The ulnar nerve runs in a groove (cubital tunnel) behind the medial epicondyle.

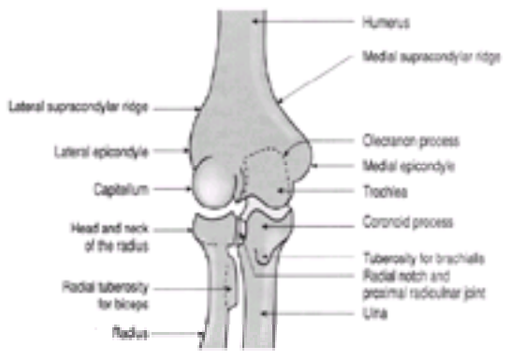


Fig. 3 The elbow.

The trochleo-ulnar, capitelloradial, and proximal radio-ulnar joints share a common synovial cavity. The subcutaneous olecranon bursa overlies the olecranon process but does not communicate with the joint (Polley and Hunder 1978; Williams *et al*. 1989).

The neutral position or position of complete extension of the elbow is designated 0°. Some normal individuals (particularly muscular athletes) lack 5° to 10° of full extension, while others (particularly women) may demonstrate 5° to 10° of hyperextension. The normal range of flexion at the cubital (humeroulnar and humeroradial) joint is 150° to 160°. The brachialis, biceps, and brachioradialis muscles are the main elbow flexors. Extensors of the elbow include the triceps and anconeus muscles.

Due to the oblique shape of the trochlea, extension of the elbow is associated with slight lateral (valgus) angulation of 5° to 15°. This is known as the 'carrying angle' of the elbow. The normal angle is about 5° in men and 10 to 15° in women. During flexion of the elbow, the ulna becomes more parallel with the humerus.

Pronation of forearm and hand (palm of the hand turned backward) to 90°, and supination (palm turned forward) to 90° occur at the proximal and distal radio-ulnar joints. During these movements the radial head pivots on the capitellum while the distal radius rotates around the ulna. The biceps and supinator muscles are the primary supinators of the radio-ulnar joints, and the pronator teres and pronator quadratus are the principal pronators.

Wrist (radiocarpal) joint

This is an ellipsoid joint between the distal radius and articular disc proximally and the scaphoid, lunate, and triquetrum distally (Fig. 4) (Polley and Hunder 1978; Markison and Kilgore 1987; Williams *et al*. 1989; Fam 1994a). The articular capsule is strengthened by the radiocarpal (dorsal and palmar), and collateral (radial and ulnar) ligaments. The articular disc, or triangular fibrocartilage of the wrist, joins the radius to the ulna. Its base is attached to the ulnar border of the distal radius and its apex to the base of the ulnar styloid process. The synovial cavity of the distal radio-ulnar joint is L-shaped and extends distally beneath the triangular fibrocartilage but is usually separated from the radiocarpal joint (Polley and Hunder 1978; Markison and Kilgore 1987; Williams *et al*. 1989; Fam 1994a).

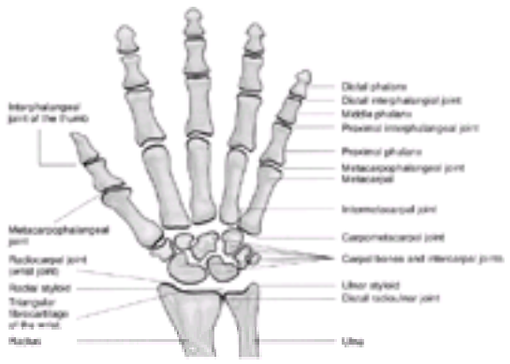


Fig. 4 Bones and joints of the wrist and hand.

The radiocarpal, intercarpal, midcarpal (located between the proximal and distal rows of the carpal bones), carpometacarpal, and intermetacarpal joints often intercommunicate through a common synovial cavity (Fig. 4).

The carpal bones form a volar concave arch or carpal tunnel: pisiform and hook of the hamate on the ulnar side, and scaphoid tubercle and crest of the trapezium on the radial side. The four bony prominences are joined by the flexor retinaculum (transverse carpal ligament), which forms the roof of the carpal tunnel. The palmaris longus (absent in 10 to 15 per cent of the population) partly inserts into the flexor retinaculum and partly fans out into the palm forming the palmar aponeurosis (fascia). The aponeurosis divides distally into four digital slips that attach to the finger flexor tendon sheaths, metacarpophalangeal joint capsules, and proximal phalanges (Polley and Hunder 1978; Markison and Kilgore 1987; Williams *et al*. 1989; Fam 1994a).

Tendons crossing the wrist are enclosed for part of their course in tenosynovial sheaths. The common flexor tendon sheath encloses the long flexor tendons of the fingers (flexor digitorum superficialis and flexor digitorum profundus) and extends from approx. 2.5 cm proximal to the wrist crease to the mid-palm. It runs with the flexor pollicis longus tendon sheath and the median nerve through the carpal tunnel (Fig. 5). The tendon sheath of the little finger is usually continuous with the common flexor sheath. The flexor pollicis longus tendon to the thumb runs through a separate tenosynovial sheath, but may join the common flexor sheath. The flexor carpi radialis is invested in a short tendon sheath as it crosses the volar aspect of the wrist between the split radial attachment of the flexor retinaculum. The flexor retinaculum straps down the flexor tendons as they cross at the wrist. The ulnar nerve, artery, and vein cross over the retinaculum but are sometimes covered by a fibrous band—the superficial part of the transverse carpal ligament—to form the ulnar tunnel or Guyon's canal (Polley and Hunder 1978; Markison and Kilgore 1987; Williams *et al*. 1989; Fam 1994a).

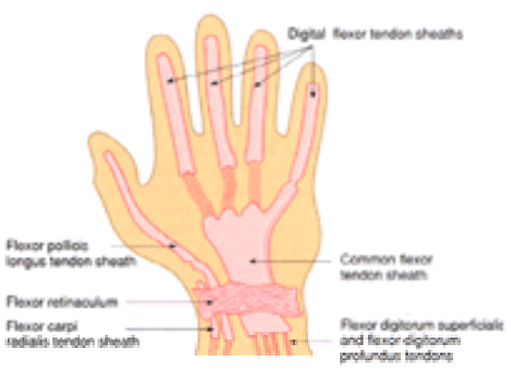


Fig. 5 Flexor tendon sheaths of the wrist and hand.

On the dorsum of the wrist, the extensor tendons pass through six tenosynovial, fibro-osseous tunnels beneath the extensor retinaculum (dorsal carpal ligament): abductor pollicis longus and extensor pollicis brevis, usually in a single sheath (first extensor compartment, most radial); extensor carpi radialis longus and brevis; extensor pollicis longus; extensor digitorum communis and extensor indicis proprius; extensor digiti minimi; and extensor carpi ulnaris (sixth extensor compartment, most ulnar) (Fig. 6). Each tenosynovial sheath extends about 2.5 cm proximally and distally from the retinaculum. The extensor retinaculum, by its deep attachments to the distal radius and ulnar, binds down and prevents bowstringing of the extensor tendons as they cross the wrist.

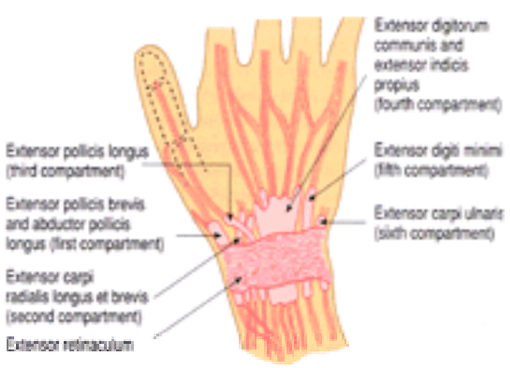


Fig. 6 Extensor tendon sheaths of the wrist.

Movements of the wrist include palmar flexion (flexion), dorsiflexion (extension), ulnar deviation, radial deviation, and circumduction. The intercarpal joints contribute to wrist movements, particularly palmar flexion. Prime wrist palmar flexors are flexor carpi radialis, flexor carpi ulnaris, and palmaris longus. The prime dorsiflexors are the extensor carpi radialis longus and brevis, and the extensor carpi ulnaris (Polley and Hunder 1978; Markison and Kilgore 1987; Williams *et al*. 1989; Fam 1994a).

The first carpometacarpal joint is a saddle-shaped, very mobile joint between the trapezium and the base of the first metacarpal. It allows 40 to 50° of thumb flexion–extension (parallel to the plane of the palm) and 40 to 70° of adduction–abduction (perpendicular to the plane of the palm). These movements are important in bringing the thumb into opposition with the fingers.

The metacarpophalangeal joints are modified hinge joints that lie about 1 cm distal to the knuckles (metacarpal heads) (Fig. 4). Their capsule is strengthened by the radial and ulnar collateral ligaments on the sides and by the palmar (volar) plate on the volar surface. The collateral ligaments are loose in the neutral position, allowing radial and ulnar deviations, but become tight in the fully flexed position preventing side-to-side motion (referred to as 'sagittal cam effect'). The deep transverse metacarpal ligament joins the volar plates of the second to fifth metacarpophalangeal joints. The metacarpophalangeal joint of the thumb is large and has two sesamoid bones overlying its volar surface (Polley and Hunder 1978; Markison and Kilgore 1987; Williams *et al*. 1989; Fam 1994a).

Where the long extensor tendon of the digit reaches the metacarpal head it is joined by the tendons of the interossei and lumbricales, and it expands over the dorsum of the metacarpophalangeal joint and digit to form the extensor hood or expansion (Fig. 7). The expansion divides over the dorsum of the proximal phalanx into an

intermediate slip, which is inserted principally into the base of the middle phalanx, and two collateral slips, which are inserted into the base of the distal phalanx.

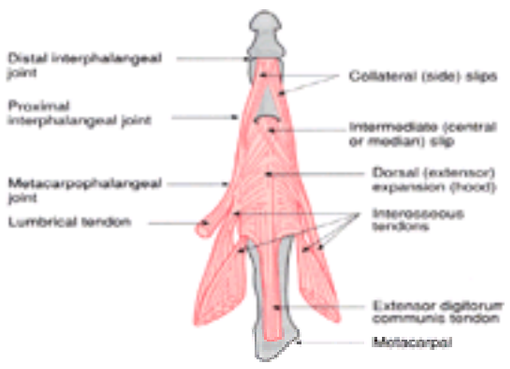


Fig. 7 Extensor expansion of the finger.

The first metacarpophalangeal joint permits 50 to 70° palmar flexion and 10 to 30° dorsiflexion. Radial and ulnar deviations are limited to less than 10 to 20°. The other metacarpophalangeal joints allow 90° palmar flexion, 30° dorsiflexion, and 35° of radial and ulnar movements. The extensor pollicis brevis, extensor indicis proprius, extensor digitorum communis, and extensor digiti minimi dorsiflex the metacarpophalangeal joints. The palmar flexors are the flexor pollicis brevis, lumbricales, interossei, and flexor digiti minimi brevis assisted by the long flexors. Radial and ulnar movements of the second to fifth metacarpophalangeal joints are a function of the intrinsic muscles (Polley and Hunder 1978; Markison and Kilgore 1987; Williams *et al*. 1989; Fam 1994a).

The interphalangeal joints

The proximal and distal interphalangeal joints of the fingers and the interphalangeal joints of the thumbs are hinge joints (Fig. 4). Their capsules are strengthened by the collateral ligaments on the sides and by the volar plate on the palmar surface. The volar plates serve to limit hyperextension, particularly at the proximal interphalangeal joints. Unlike the metacarpophalangeal joints, the radial and ulnar collateral ligaments remain taut in all positions, providing side-to-side stability throughout the range of movement.

The flexor tendon sheaths for the fingers enclose the flexor digitorum superficialis and profundus tendons to their insertions on the middle and distal phalanges, respectively. The sheaths extend from just proximal to the metacarpophalangeal joints to the bases of the distal phalanges (Fig. 5). The flexor pollicis longus tendon sheath of the thumb extends proximally to the carpal tunnel. The flexor sheath of the little finger is often continuous with the common flexor tendon sheath of the wrist. Segmental condensations, or annular pulleys, in the digital flexor sheaths prevent bowstringing of the tendons, and are mechanically critical for full digital flexion (Polley and Hunder 1978; Markison and Kilgore 1987; Williams *et al*. 1989; Fam 1994a).

The proximal interphalangeal joints do not normally hyperextend. They allow 100 to 120° palmar flexion. The distal interphalangeal joints permit 50 to 80° palmar flexion and 5 to 10° dorsiflexion. The interphalangeal joint of the thumb allows 80 to 90° palmar flexion and 20 to 35° dorsiflexion. The flexor digitorum superficialis flexes the proximal interphalangeal joints, and the flexor digitorum profundus flexes the distal interphalangeal joints of the fingers. The prime dorsiflexors are the interossei and lumbrical muscles. The flexor pollicis longus flexes the interphalangeal joint of the thumb and the extensor pollicis longus dorsiflexes the joint (Polley and Hunder 1978; Markison and Kilgore 1987; Williams *et al*. 1989; Fam 1994a).

General considerations and differential diagnosis of regional rheumatic pain syndromes of the upper extremities in adults (also see Chapter 5.14)

General considerations

Upper extremity regional pain syndromes are a common presentation in general, rheumatological, and orthopaedic practices. Table 1 provides a classification of painful disorders of the shoulder, elbows, wrists, and hand based upon the site of origin and the predominant location of pain.

Table 1 Classification of painful upper-limb disorders

Precise diagnosis of regional rheumatic pain syndromes in the upper limb depends upon greater awareness of these disorders, knowledge of basic regional anatomy, a detailed history, a thorough physical examination of the joints, periarticular structures, cervical spine nerve and blood supplies to the upper limb, and a few carefully selected diagnostic studies.

The onset of many musculoskeletal pain syndromes in the upper limb is often insidious. An antecedent history of repetitive, excessive, or unaccustomed physical activity is common. Less frequently, the pain results from a specific activity or injury. With ageing, tendons become less flexible and less elastic, making them more susceptible to injury. A shortened musculotendinous unit, from lack of regular stretching exercises, is also more prone to injury.

In addition to plain radiographs and arthrography, the newer methods of imaging, such as ultrasonography, computed tomography (**CT**), and magnetic resonance imaging (**MRI**), have improved diagnostic accuracy in shoulder disorders. Arthroscopy of the shoulder is particularly useful in the diagnosis and treatment of subacromial and intra-articular conditions.

The following is a brief summary of the most common causes of shoulder, elbow, wrist, and hand pain (excluding arthritis). Not included in this chapter are nerve entrapment syndromes, pain arising from sternoclavicular joints, pain originating from the cervical spine, spinal cord and spinal nerves, pain referred from the thoracic outlet, intrathoracic and intra-abdominal structures, and pain due to reflex sympathetic dystrophy syndrome, fibromyalgia, and vascular disorders.

Disorders of the shoulder region

Shoulder pain is a common musculoskeletal symptom of diverse causes (Table 1). The pain may have its origin in the glenohumeral joint, acromioclavicular joint or periarticular structures, or be referred from the cervical spine (felt at the supraspinatus region), thoracic outlet, subdiaphragmatic conditions (felt at the tip of the

shoulder), or brachial plexus (Bland *et al*. 1977; Kapandji 1982; Neviaser 1983; Chard and Hazleman 1987; Vecchio *et al*. 1995).

Periarticular disorders are by far the most common cause of shoulder pain in adults (Bland *et al*. 1977; Kapandji 1982; Neviaser 1983; Chard and Hazleman 1987; Vecchio *et al*. 1995). The anatomical configuration of the shoulder joint is such that during abduction the rotator cuff and long biceps tendon are subjected to impingement between the greater tuberosity and coracoacromial arch (Neer 1983; Miniaci and Fowler 1993; Frieman *et al*. 1994). This leads to tendinitis and tears of the rotator cuff, which compromises its function as a stabilizer and depressor of the humeral head, resulting in chronic glenohumeral instability, superior migration of the humeral head (due to unopposed upward pull of the deltoid), and further impingement on both the rotator cuff and long biceps tendons. Osteophytes on the inferior surface of the acromioclavicular joint can also aggravate this impingement. Thus, the spectrum of chronic impingement syndrome ranges from mild tendinitis of the rotator cuff, with or without inflammation of the adjacent subacromial bursa, to bicipital tendinitis, tears of the rotator cuff and chronic glenohumeral instability with superior migration of the humeral head, and secondary glenohumeral osteoarthritis ('cuff tear arthropathy'). The causes of shoulder impingement include repetitive low-grade trauma or unaccustomed activities, excessive overhead use in sport or work, lack of conditioning, ageing factors, and structural abnormalities (e.g. acromial spur) that compromise the cuff space. It is likely that both chronic impingement and degenerative tendinopathy (presumably the result of ischaemia) contribute to the development of rotator cuff disease. Vascular compromise of the rotator cuff tendons in diabetic patients is an important predisposing factor (Frieman *et al*. 1994).

The pathology of shoulder impingement syndrome can be divided into three overlapping stages (Neer 1983; Miniaci and Fowler 1993; Frieman *et al*. 1994). Stage I consists of oedema, swelling, and haemorrhage of the rotator cuff and subacromial bursa. Stage II is characterized by tendon tears, inflammation, fibrosis, and thickening of the rotator cuff, bicipital tendon, and subacromial bursa. Stage III is marked by fraying or rupture of the rotator cuff and/or long biceps tendon, instability, and secondary glenohumeral osteoarthritis. Stage I often begins in young individuals under 25 years of age. Stage II usually occurs in patients aged 25 to 40 years and stage III in those over 40 years of age. The term 'impingement syndrome' replaces the old expression 'painful arc syndrome' (Kessel and Wastson 1977), and includes such disorders as rotator cuff tendinitis, subacromial bursitis, and bicipital tendinitis.

Rotator cuff tendinitis

Rotator cuff tendinitis is the most common cause of shoulder pain. The disorder may be acute or chronic (Bland *et al*. 1977; Kapandji 1982; Neer 1983; Neviaser 1983; Chard and Hazleman 1987; Birrer *et al*. 1989; Miniaci and Fowler 1993; Snyder 1993; Frieman *et al*. 1994; Vecchio *et al*. 1995). In young persons an acute sport injury due to use of the arms in an overhead position (e.g. baseball, racquet ball, tennis, swimming) is common (Miniaci and Fowler 1993). In middle-aged and elderly individuals the onset is more gradual, and an antecedent history of repetitive movements above the shoulder level or of strenuous or unaccustomed arm activity is common. Symptoms include aching pain in the shoulder, pain with movements, particularly abduction and internal rotation, night pain when rolling on to the affected side, restriction of shoulder movements, and, sometimes, weakness due to a traumatic or degenerative tear of the rotator cuff. The patient typically experiences pain on active abduction, especially between 60 and 120°, and sometimes when lowering the arm. Patients complain of difficulty with overhead work or lifting and with reaching behind their back when dressing. The pain is often felt over the lateral aspect of upper arm and deltoid insertion rather than in the point of the shoulder. Less commonly, it radiates to the scapula, root of the neck or elbow. Clinical findings include a painful arc between 60 to 120° of abduction, limitation of active movement by pain, and tenderness localized to the rotator cuff and greater tuberosity (Kessel and Wastson 1977; Neer 1983; Birrer *et al*. 1989; Miniaci and Fowler 1993; Snyder 1993; Frieman *et al*. 1994). There is often pain on resisted movement of the affected tendon. The 'supraspinatus' or 'empty can' sign (pain on resisted elevation of the arm to 90° midway between abduction and forward flexion with the thumb pointing downwards in internal rotation) is often positive (Birrer *et al*. 1989). Neer's 'impingement sign' (one hand of the examiner forward flexing the patient's arm while the other hand restricting scapular rotation) is positive: the patient develops pain in the overhead position near the end of full flexion (Neer 1983). In Hawkin's 'impingement sign', the humerus is both forward flexed to 90° and internally rotated while on the other hand restricting scapular movements, causing impingement of the greater tuberosity against the anterior acromion with reproduction of patient's symptoms (Miniaci and Fowler 1993; Frieman *et al*. 1994). The 'impingement test' is also positive: pain relief and greater range of abduction following injection of 2 to 5 ml of 2 per cent of lidocaine into the subacromial bursa (Neer 1983; Miniaci and Fowler 1993).

Calcification of the cuff tendons may be asymptomatic, but can present either as an acute condition in the younger patient, or as a chronic calcific tendinitis associated with aching pain and symptoms of impingement (Re and Karzel 1993). In the acute phase (acute calcific rotator cuff tendinitis) the onset is often explosive, with severe pain, exquisite local tenderness, muscle spasm, and marked painful restriction of all shoulder movements. An acute calcific subacromial bursitis may also be present when calcific material ruptures into the adjacent bursa.

Shoulder radiographs in acute cuff tendinitis may appear either normal, or show oval or round calcific deposits in the region of the rotator cuff tendons, particularly that of the supraspinatus (Fig. 8). These deposits often disappear following resolution of an attack of acute calcific tendinitis. In chronic tendinitis the presence of sclerosis, cystic changes, and osteophytes of the greater tuberosity is suggestive of chronic impingement and insertional damage (Snyder 1993) (Fig. 9). Tendinitis, rotator cuff tears, and thickening or fluid distension of the subacromial bursa may be demonstrated by MRI and by grey-scale, dynamic or power Doppler sonography (Ryu *et al*. 1993; Snyder 1993).

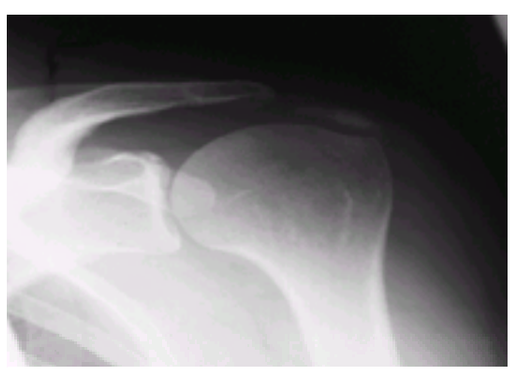


Fig. 8 Acute calcific tendinitis of left shoulder.

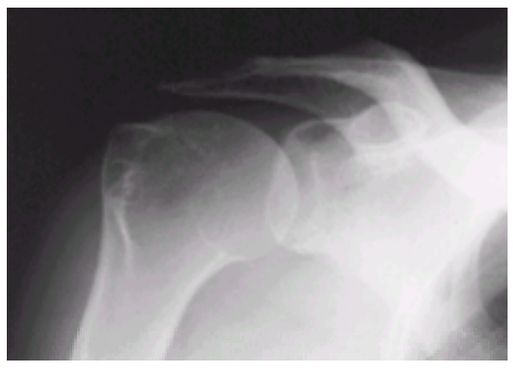


Fig. 9 Sclerosis and cystic changes of the greater tuberosity in a patient with chronic shoulder impingement.

Initial management of rotator cuff tendinitis consists of rest, non-steroidal anti-inflammatory drugs, and physical treatments including heat application, and range of movement exercises (Miniaci and Fowler 1993; Snyder 1993; Dalton 1994; Frieman *et al*. 1994). Early mobilization of the shoulder is important in minimizing disability, and reducing the risk of adhesive capsulitis or reflex sympathetic dystrophy. If symptoms persist for more than 2 to 4 weeks, injection of a depot corticosteroid into the nearby subacromial bursa (posterior subacromial approach) is often beneficial. Once the pain is controlled, strengthening exercises are important to restore the function of the rotator cuff muscles and reduce the likelihood of further injury. Diagnostic imaging with MRI, ultrasound, or arthrography to determine the presence and size of a possible cuff tear is indicated in those not responding to 6 to 12 weeks of non-operative treatment. Surgical (or arthroscopic) subacromial decompression (anterior acromioplasty with reshaping the anterior acromion and section of the coracoacromial ligament) is reserved for patients with

chronic symptoms refractory to medical treatment (Miniaci and Fowler 1993; Snyder 1993; Dalton 1994; Frieman *et al*. 1994).

Treatment of acute calcific rotator cuff tendinitis consists of cold application and non-steroidal anti-inflammatory drugs (Re and Karzel 1993). If symptoms persist, a local injection into the subacromial bursa of lidocaine followed by a depot corticosteroid is helpful. Surgical or arthroscopic excision of the calcific deposits is rarely necessary in chronic cases (Re and Karzel 1993).

Rotator cuff tears

Rotator cuff tears can be partial or complete, acute or chronic. They are further classified as small (£1 cm), medium (1–3 cm), large (3–5 cm) and massive (³5 cm) (Snyder 1993; Post *et al*. 1983). In young adults acute tears can result from direct trauma, unexpected falls or a sport injury. In older patients, minor less severe trauma, superimposed on an already degenerated cuff tendon from chronic impingement and age-related attritional changes, can lead to partial tears (Snyder 1993; Post *et al*. 1983).

Clinical features include shoulder pain on abduction, night pain, varying degrees of weakness of abduction and external rotation, local tenderness, and loss of range of movement. The supraspinatus and impingement signs are usually positive. In complete tears, the 'drop-arm' sign is positive: inability to actively maintain 90° of passive shoulder abduction. Subacromial crepitus and rupture of the long biceps tendon may be present (Snyder 1993; Post *et al*. 1983).

Radiographic findings include sclerosis and cystic degeneration of the greater tuberosity, superior migration of the humeral head with narrowing of the subacromial (acromiohumeral) space (less than 6 mm indicates a tear), and secondary glenohumeral osteoarthritis (Snyder 1993; Post *et al*. 1983). The diagnosis of a complete tear can be confirmed by single- or double-contrast arthrography showing a communication between the glenohumeral joint and subacromial bursa (Fig. 10) (Snyder 1993; Post *et al*. 1983). In partial tears there is no communication but the arthrogram may show a crater-like defect in the rotator cuff. Ultrasonography (Newman *et al*. 1994) and MRI (Recht and Resnick 1993) are non-invasive methods that are useful in identifying rotator cuff tears. Arthroscopy is particularly valuable in assessing the size of the tear, and the status of the labrum, long biceps tendon, and glenohumeral joint (Snyder 1993).

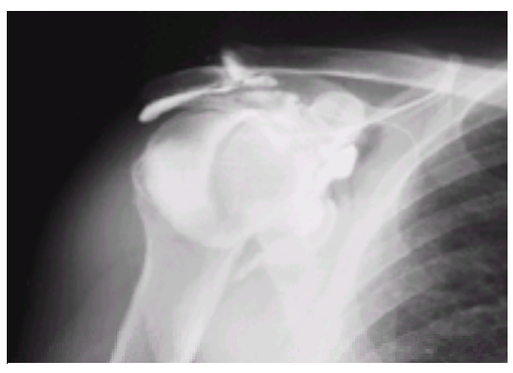


Fig. 10 Right shoulder arthrogram in a patient with a rotator cuff tear showing a communication between glenohumeral joint and subacromial bursa. Note abnormal communication with the acromioclavicular joint.

Partial tears are treated conservatively with rest, physical therapy, and non-steroidal anti-inflammatory drugs (Snyder 1993; Dalton 1994). The role of local corticosteroid injections has not been established, but they are not recommended within 4 to 6 weeks of an acute cuff tear. In patients with persistent symptoms a subacromial injection of a depot corticosteroid is usually beneficial.

Early surgical repair is recommended for acute, complete ruptures of the cuff in the young active patient (Snyder 1993). In contrast, chronic complete tears are treated non-operatively in the first instance. Functional disability and failure to achieve pain relief are the main indications for surgical (Snyder 1993; Frieman *et al*. 1994) or arthroscopic (Burkhart 1993) treatment; acromioplasty, and debridement to relieve impingement, and whenever possible, repair of tears of the rotator cuff and long biceps tendon. However, complete pain relief or return to full function are uncommon after surgery.

Bicipital tendinitis

Bicipital tendinitis is often related to chronic impingement of the bicipital tendon by the acromion. The long biceps tendon becomes fibrotic, frayed, and may rupture. Bicipital tendinitis usually occurs in association with rotator cuff tendinitis and glenohumeral instability (Bland *et al*. 1977; Neviaser 1980; Kapandji 1982; Neer 1983; Neviaser 1983; Chard and Hazleman 1987; Miniaci and Fowler 1993; Frieman *et al*. 1994; Vecchio *et al*. 1995). Primary, isolated, bicipital tendinitis is rare, and develops as an overuse injury due to repetitive stresses applied to the tendon in certain sports such as weight lifting and ball throwing.

The pain is felt over the anterior aspect of the shoulder, but may radiate into the biceps muscle. It is increased by overhead activities, shoulder extension, and elbow flexion. Localized tenderness is present over the tendon in the bicipital groove. Signs of chronic impingement and glenohumeral instability are often present. Bicipital tendon pain can be reproduced by resisted supination of the pronated forearm with the elbow 90° flexed (Yergason's sign), shoulder flexion against resistance (Speed's test), passive extension of the shoulder, or by resisted flexion of the elbow (Bland *et al*. 1977; Neviaser 1980; Kapandji 1982; Neer 1983; Neviaser 1983; Chard and Hazleman 1987). Rupture of the long biceps tendon typically occurs at the superior edge of the bicipital groove. It produces a characteristic bunching up of the lateral half of the muscle belly, best seen with resisted elbow flexion and supination ('Popeye sign').

Subluxation of the bicipital tendon is also related to chronic impingement. It results from traumatic rupture of the intertubercular ligament, which normally straps the tendon in the bicipital groove. Pain in the anterior aspect of the shoulder, and a clicking sensation of the shoulder 'going out and popping back in', are the principal symptoms (Neviaser 1980). Patients are usually able to reduce the tendon themselves. Tenderness in the bicipital groove and a snap over the tendon when the arm is passively abducted to 90° then moved through internal and external rotation, are usually present (Neviaser 1980).

Treatment of bicipital tendinitis consists of physiotherapy and non-steroidal anti-inflammatory drugs. Corticosteroid injections into the bicipital sheath, with care being taken not to inject the tendon itself, are often beneficial (Bland *et al*. 1977; Neviaser 1980; Kapandji 1982; Neer 1983; Neviaser 1983; Chard and Hazleman 1987). Surgical subacromial decompression is required in patients with chronic impingement. Repair of subluxation and/or rupture of the bicipital tendon is indicated in the younger, more active patient.

Subacromial bursitis

Chronic subacromial bursitis, with thickening and fibrosis of the bursal wall, often occurs in association with rotator cuff tendinitis and chronic impingement. Primary subacromial bursitis due to trauma, infection, rheumatoid arthritis, or other arthritic disorders is rare. Abduction is painful, and there is local tenderness and sometimes fluid distension of the bursa, which is visible just anterior to the acromion with the shoulder extended (Bland *et al*. 1977; Kapandji 1982; Neer 1983; Neviaser 1983; Chard and Hazleman 1987).

Treatment of acute traumatic subacromial bursitis is symptomatic, with ice packs, rest, non-steroidal anti-inflammatory drugs, range-of-movement exercises, and if indicated, a subacromial corticosteroid injection. Surgical excision and debridement of the bursa may be indicated in those requiring operative treatment for chronic impingement.

Adhesive capsulitis

Adhesive capsulitis, also known as frozen shoulder, pericapsulitis and obliterative bursitis, is characterized by progressive, global restriction of shoulder movements associated with pain and functional disability (Bland *et al*. 1977; Kapandji 1982; Neer 1983; Neviaser 1983; Bulgen *et al*. 1984; Chard and Hazleman 1987; Birrer *et al*. 1989; Hulstyn and Weiss 1993). The condition is rare before the age of 40 years, and is bilateral in about 15 per cent of patients. A period of immobility of the

shoulder is the most commonly identified predisposing factor. The onset is insidious, and the capsulitis is often secondary to rotator cuff tendinitis, rotator cuff tears, bicipital tendinitis, or glenohumeral arthritis. The condition may also coexist with diabetes mellitus, hypothyroidism, lung carcinoma, myocardial infarction, cardiac surgery, cerebrovascular events, and shoulder trauma. Primary or idiopathic shoulder capsulitis is rare. No autoimmune disturbance or association with histocompatibility antigen has been demonstrated in this disorder.

Limited histological studies demonstrate an initial phase with increased vascularity, sparse chronic inflammatory cellular infiltrate, and fibroblastic proliferation. This is followed by fibrous thickening and contraction of the capsular folds, recesses, and surrounding ligamentous structures. The capsule often adheres to the humeral neck, and the axillary pouch becomes obliterated causing restricted movements (Bland *et al*. 1977; Kapandji 1982; Neviaser 1983; Chard and Hazleman 1987; Hulstyn and Weiss 1993; Vecchio *et al*. 1995).

Clinical features

The natural history of adhesive capsulitis can be divided into three overlapping phases, each lasting a few months: a painful ('freezing') phase, an adhesive ('frozen') phase, and a resolution ('thawing') phase (Bland *et al*. 1977; Kapandji 1982; Neviaser 1983; Chard and Hazleman 1987; Hulstyn and Weiss 1993; Vecchio *et al*. 1995). The painful or 'freezing' phase is characterized by the insidious onset of shoulder pain, diffuse tenderness, night pain, and progressive restriction of shoulder movements. The classic finding is limitation of external rotation in the absence of glenohumeral arthritis. During the adhesive or 'frozen' phase, pain and local tenderness decrease but shoulder movements become globally more restricted. Disuse atrophy of the deltoid and scapular muscles is common. In the resolution or 'thawing phase', pain becomes less evident and there is slow increase in the range of shoulder movements. Although adhesive capsulitis is often a self-limiting disorder, the extent and rate of recovery are variable. About 90 per cent of patients recover use of the extremity within 12 to 18 months, but about 10 per cent develop more prolonged pain and functional disability.

The diagnosis of adhesive capsulitis depends primarily on the clinical findings. The erythrocyte sedimentation rate and acute-phase reactants are usually normal. Radiographs may show changes of chronic rotator-cuff tendinitis with or without calcification. Double-contrast arthrography is diagnostic, showing a marked decrease in the joint volume to under 10 ml (normal values: 20–35 ml), loss of the normal axillary recess, and reduced filling of the bicipital tendon sheath (Hulstyn and Weiss 1993). Dynamic sonography, demonstrating limitation of sliding movements of the supraspinatus tendon during abduction of the arm, is a reliable, non-invasive, rapid technique for the diagnosis of this disorder (Ryu *et al*. 1993). Arthroscopy allows evaluation of any underlying shoulder pathology. Increased uptake on bone scanning is a non-specific finding and does not predict outcome or response to therapy (Hulstyn and Weiss 1993).

Treatment of adhesive capsulitis consists of non-steroidal anti-inflammatory drugs, physical therapy and corticosteroid injections into both the glenohumeral joint and subacromial bursa (Bland *et al*. 1977; Kapandji 1982; Neviaser 1983; Bulgen *et al*. 1984; Chard and Hazleman 1987; Petri *et al*. 1987; Hulstyn and Weiss 1993; Vecchio *et al*. 1995). Physiotherapy to relieve pain and improve the range of movement is important. Gentle range-of-movement exercises, beginning with pendulum exercises and wall climbing with the fingers, and finally active strengthening exercises, are required to restore joint mobility. Other physical methods have been advocated but proof of their efficacy is lacking. Included among these are transcutaneous electrical nerve stimulation, ultrasound, laser, magnetotherapy, and gentle manipulation.

Oral corticosteroids (beginning with 30 mg prednisone per day with gradual tapering off within 1–2 months), by virtue of their anti-inflammatory effects, appear to have a beneficial effect in the early rehabilitation and pain relief of shoulder capsulitis (Petri *et al*. 1987; Baslund *et al*. 1990; Hulstyn and Weiss 1993).

In the chronic adhesive phase, three treatments have been used with inconsistent results: infiltration brisement, manipulation, and open surgical release. Brisement, or hydraulic saline distension of the shoulder joint under local anaesthesia, by releasing and dissecting the capsule from the underlying glenoid and humerus, has been shown to reduce pain and increase shoulder mobility in some patients (Hulstyn and Weiss 1993). Favourable results have also been obtained with arthrographic distension and rupture of the capsule through injection under pressure, of a 30-ml mixture containing 8 ml of 1 per cent lidocaine, 80 mg methylprednisolone, and 20 ml of water-soluble contrast material (Rizk *et al*. 1994). Manipulation of the shoulder under anaesthesia, to rupture the inferior capsule, is a controversial form of therapy. Restoration of shoulder movements following the procedure is variable. The treatment can cause soft tissue damage and shoulder dislocation, and long-term studies suggest that manipulation does not significantly alter the rate of recovery (Hulstyn and Weiss 1993). Operative capsulotomy, with surgical release of the coracohumeral ligament and of subacromial adhesions, is occasionally beneficial in patients with chronic symptoms refractory to conservative treatment (including injections, physiotherapy, oral corticosteroids, and manipulations) (Hulstyn and Weiss 1993).

Glenohumeral instability

Subtle, traumatic tears of the rotator cuff and capsule leading to chronic, multidirectional, glenohumeral instability are a frequent cause of shoulder pain, particularly in the young, active athlete (Bland *et al*. 1977; Kapandji 1982; Neviaser 1983; Chard and Hazleman 1987; Vecchio *et al*. 1995). Although the findings may resemble those of rotator cuff tendinitis (with which it often coexists), laxity of the glenohumeral joint can be demonstrated by a number of tests. These include anterior and posterior 'drawer' signs (the humeral head and proximal humerus are moved forward and backward in the glenoid fossa), the 'sulcus' sign (applying distal traction on the upper arm while palpating the gap between the humeral head and the acromion), and the 'containment' sign (with the patient lying supine and the shoulder abducted 90°, the examiner slowly extends and externally rotates the arm). A patient with anterior glenohumeral instability experiences pain that is relieved when downward pressure is applied to the subluxed humeral head (Bland *et al*. 1977; Kapandji 1982; Neviaser 1983; Bulgen *et al*. 1984; Chard and Hazleman 1987; Petri *et al*. 1987; Hulstyn and Weiss 1993; Vecchio *et al*. 1995).

Treatment of glenohumeral instability occurring in association with rotator cuff tendinitis consists of non-steroidal anti-inflammatory drugs, and an exercise programme to strengthen the dynamic stabilizers: rotator cuff and biceps muscle. A stabilizing operative procedure is indicated in resistant cases.

Traumatic lesions of the acromioclavicular joint

Acute and chronic trauma to the acromioclavicular joint can lead to disruption of the joint capsule, dislocation, or damage to the fibrocartilaginous disc. Late sequelae include chronic acromioclavicular instability, acromioclavicular separation, osteolysis of the lateral end of the clavicle, and secondary osteoarthritis (Bland *et al*. 1977; Kapandji 1982; Neviaser 1983; Bulgen *et al*. 1984; Chard and Hazleman 1987; Petri *et al*. 1987; Hulstyn and Weiss 1993; Vecchio *et al*. 1995).

Local pain, tenderness, and sometimes swelling of the acromioclavicular joint are the main clinical findings. Pain at the acromioclavicular joint can be localized by performing various 'stress' tests: the 'adduction stress' test (passive adduction of the extended shoulder behind the back), and the 'cross-arm acromioclavicular loading' test (shoulder abducted 90° and then adducted across the chest at shoulder height). A painful arc from 90° of abduction upwards is characteristic.

Treatment consists of physiotherapy, non-steroidal anti-inflammatory drugs, and sometimes, an intra-articular injection of corticosteroid. A surgical stabilizing procedure is indicated in chronic cases.

Disorders of the elbow region

Lateral epicondylitis, medial epicondylitis, and olecranon bursitis are the three most common soft-tissue lesions at the elbow.

Lateral epicondylitis (tennis elbow)

This is one of the most common painful lesions of the upper extremity. It affects 1 to 3 per cent of the adult population, mostly those between 40 and 60 years of age (Leach and Miller 1987; Birrer *et al*. 1989; Gellman 1992; Foley 1993; Warhold *et al*. 1993; Geoffroy *et al*. 1994). The dominant arm is most frequently involved. Although about 40 per cent of tennis players suffer with lateral epicondylitis, less than 10 per cent of patients in clinical practice acquire the disorder through playing tennis. It occurs most often in non-athletes who overuse their arms in recreational or occupational activities (Leach and Miller 1987; Gellman 1992; Foley 1993; Warhold *et al*. 1993; Geoffroy *et al*. 1994).

The exact mechanism of lateral epicondylitis is not entirely clear. It is generally regarded as a cumulative trauma overuse disorder due to repetitive mechanical overloading of the origin of the common forearm extensor muscles (Leach and Miller 1987; Gellman 1992; Foley 1993; Warhold *et al*. 1993; Geoffroy *et al*. 1994). An abnormality of local microvascular control, presumably due to sympathetic dysfunction, has also been postulated (Smith *et al*. 1994a). Limited pathological studies have demonstrated small tears, granulation tissue, sparse round-cell infiltration, and fibrous degeneration of the attachment of the common extensor tendon into the

lateral epicondyle, particularly that portion derived from the extensor carpi radialis brevis (Leach and Miller 1987; Gellman 1992; Foley 1993; Warhold *et al*. 1993; Geoffroy *et al*. 1994). Fibrocartilaginous and bony transformation, and an increase in glycosaminoglycan content within the tendon close to its insertion, have also been observed (Chard *et al*. 1994).

The typical patient with tennis elbow is a middle-aged or elderly individual who is active in gardening, hobbies, or sports. With ageing, muscles and tendons become less flexible and less elastic, making them more susceptible to overuse or work-related injuries, particularly near their bony attachments. The onset is often insidious, and bilateral occurrence is not uncommon. A rapid onset following a direct blow to the lateral epicondyle or a sports-related injury is more frequent in the younger, more active patient (Leach and Miller 1987; Gellman 1992; Foley 1993; Warhold *et al*. 1993; Geoffroy *et al*. 1994). Pain is localized to the lateral epicondyle but may extend both distally and proximally. It may occur during handshakes, turning doorknobs, carrying a briefcase, lifting or gripping, resulting in restricted hand activities. Localized tenderness just distal and slightly anterior to the lateral epicondyle is the hallmark of tennis elbow (Leach and Miller 1987; Gellman 1992; Foley 1993; Warhold *et al*. 1993; Geoffroy *et al*. 1994). The elbow pain can be increased by a number of manoeuvres: resisting dorsiflexion of the wrist with the elbow in extension, extending the elbow with the wrist flexed and pronated, and by resisted forearm pronation. The 'chair-lift' test, in which the patient lifts a chair or any similarly heavy object with the forearm pronated, will cause sharp pain at the lateral epicondyle. The range of movement of the elbow joint is normal, but in chronic cases a slight flexion deformity may be present. The grip strength, measured by a dynamometer, is usually diminished.

Elbow radiographs are often normal, but periarticular calcification or exostosis may be present in those with chronic symptoms. Infrared thermography may show localized increased heat in the region of the lateral epicondyle. MRI may show an extensor tendon tear in young athletes with acute onset of symptoms (Warhold *et al*. 1993). MRI in patients with chronic lateral epicondylitis has demonstrated unexplained increased signal intensity of the anconeus, a muscle not known to be involved in this disorder (Coel *et al*. 1993).

Treatment of lateral epicondylitis consists of an initial period of rest from the workplace or sporting activity, the identification and modification of activities that aggravate symptoms, applications of ice packs or heat, and non-steroidal anti-inflammatory drugs (Leach and Miller 1987; Gellman 1992; Foley 1993; Warhold *et al*. 1993; Geoffroy *et al*. 1994). A forearm brace, a wrist cock-up splint to provide 20° extension, or strapping of forearm muscles 4 to 5 cm distal to the elbow may be of benefit in some patients. Tennis-elbow bands worn during activity change the direction of muscle pull and short-circuit the force of contraction by reducing the effective length of the muscle. There are two types of band, static and counterforce. The static band wraps around the proximal forearm and applies equal pressure to all areas of the forearm. The counterforce band applies most of the pressure directly over the extensor muscle mechanism constraining full muscle contraction, thereby reducing injury-producing tension on the attachment of the extensor tendon. A local corticosteroid injection (methylprednisolone acetate 10 mg or triamcinolone hexacetonide 4 mg) into the site of maximal tenderness at the anterior part of the lateral epicondyle near the origin of extensor carpi radialis brevis produces relief in about 90 per cent of patients. Complications are rare: they include postinjection flare and infection. Repeated corticosteroid injections can cause loss of subcutaneous fat, depigmentation, and calcification at the site of injection. Steroid injections may be repeated after 4 to 8 weeks in the event of either failure to respond or relapse, but should not be repeated more than three to four times as they are unlikely then to improve the outcome. After the pain has subsided, graded exercises to stretch and strengthen the extensor forearm muscles are important for the prevention of recurrences and long-term rehabilitation. A number of other treatments have been used but their efficacy remains unproven. Included among these are pulsed therapeutic ultrasound, use of a whirlpool, medical laser-light therapy, acupuncture (Molsberger and Hille 1994), and manipulation. Ultrasound, by virtue of its ability to cross myofascial planes and concentrate near bone, has theoretical but unproven advantages.

Lateral epicondylitis is a self-limiting disorder; most patients improve within a year, but some may continue to have chronic symptoms. Surgical treatment is indicated in those with refractory symptoms for more than 6 to 12 months (about 10 per cent of patients) (Leach and Miller 1987; Gellman 1992; Foley 1993; Warhold *et al*. 1993; Geoffroy *et al*. 1994). Release of the origin of the common extensor tendon is the surgical treatment of choice (Verhaar *et al*. 1993). Other operative procedures include excision of granulation tissue and repair of extensor tendon tears, distal tenotomy to lengthen the common extensor tendon, division of the tendinous origin of extensor carpi radialis brevis, and combined release of the common extensor tendon and resection of the proximal third of the annular ligament (Verhaar *et al*. 1993).

Medial epicondylitis (golfer's elbow)

Medial epicondylitis affects the common flexor tendon at the medial epicondyle and is considerably less common and less disabling than lateral epicondylitis (Leach and Miller 1987; Birrer *et al*. 1989; Glazebrook *et al*. 1994). It commonly occurs in individuals who overuse their arms, and its pathology is essentially similar to that of tennis elbow. Cumulative repetitive strains, and increased activity of the common flexor muscles of the forearm in golfing, baseball pitching or work-related activities, can lead to disruption and tears of the common flexor muscle–tendon unit (particularly that portion derived from the pronator teres), near its origin from the medial epicondyle. Electromyographic activity recorded from forearm flexor muscles is significantly greater in golfers with medial epicondylitis than in those without (Glazebrook *et al*. 1994).

The onset is often insidious, with pain localized to the medial epicondyle. The pain may spread widely down the forearm, and may occur during lifting, grasping, and other activities involving the hands. There is tenderness at and just distal to the medial epicondyle. Resisted palmar flexion of the wrist with the elbow extended exacerbates the pain. Flexion of the fingers rather than of the wrists may sometimes elicit pain at the medial epicondyle.

Modification of forearm activities, ice packs or heat, non-steroidal anti-inflammatory drugs, and occasionally a local corticosteroid injection are sufficient therapy for most patients with medial epicondylitis. Enlarged golf-club handles and medial forearm counterforce braces (Glazebrook *et al*. 1994) may provide subjective relief in some patients. Exercises to strengthen the wrist flexors and forearm pronators are important. Surgical release of the common extensor tendon is indicated in those with chronic symptoms (Leach and Miller 1987; Birrer *et al*. 1989).

Olecranon bursitis

The subcutaneous olecranon bursa is a frequent site for traumatic, inflammatory, and septic bursitis (Fam 1992; Zimmermann *et al*. 1995).

Non-septic olecranon bursitis can result from either a discrete injury, for example an acute blow (traumatic olecranon bursitis), or from repetitive, minor, or occult trauma to the elbow, for example excessive leaning (idiopathic olecranon bursitis). 'Dialysis elbow' refers to olecranon bursitis that may occur in patients on haemodialysis due to sustained pressure on the resting elbow with the venous access during dialysis. 'Miner's elbow' and 'student's elbow' also refer to traumatic olecranon bursitis due to excessive friction on the elbow (Fam 1992).

Pain is usually minimal except when pressure is exerted on the swollen bursa. Bursal fluid distension and local tenderness are present but elbow joint movements are usually unimpaired and painless. Bursal fluid is either clear or blood-tinged, with reduced viscosity and a low, predominantly mononuclear, leucocyte count (£2000×10^6/l). Immunophenotypic characterization of bursal fluid mononuclear cells in idiopathic olecranon bursitis showed a preponderance of activated T-cell subpopulations and monocyte/macrophages; this suggested an immunological role for these cells in the development and perpetuation of non-septic bursitis (Smith *et al*. 1994b).

Treatment of traumatic and idiopathic olecranon bursitis consists of aspiration, a compressive elastic bandage, protection of the elbow from further friction, non-steroidal anti-inflammatory drugs, and one or more intrabursal injections of corticosteroid as required (Fam 1992; Zimmermann *et al*. 1995). Surgical bursectomy is rarely indicated in those with recurrent bursitis.

Inflammatory olecranon bursitis may be due to rheumatoid arthritis, psoriatic arthritis, or gout. It is associated with local tenderness, bursal distension, and an inflammatory synovial effusion with a high, predominantly polymorphonuclear, leucocyte count (> 2000 × 10^6/l). Treatment includes aspiration, non-steroidal anti-inflammatory drugs, one or more intrabursal injections of steroid, and therapy for the underlying disorder. Local sclerotherapy with intrabursal tetracycline has recently been shown to be effective in the recurrent treatment of rheumatoid bursitis resistant to intrabursal steroid injections (Hassell *et al*. 1994).

Septic olecranon bursitis usually results from the direct introduction of bacteria through a skin abrasion (Fam 1992; Fam 1994b; Zimmermann *et al*. 1995). It typically occurs in otherwise healthy adult men engaged in physical work involving frequent trauma to the elbows (e.g. plumbers, miners, gardeners, construction workers, etc.). *Staphylococcus aureus* is the most frequent pathogen.

Clinical findings include local pain, swelling, tender bursal effusion, warmth, erythema, peribursal cellulitis, and diffuse oedema. Bursal fluid is opaque, yellow, with a high, predominantly polymorphonuclear, leucocyte count (>10 000–50 000 × 10^6/l). The diagnosis is confirmed by culture of the causative organism from the bursal

fluid and, rarely, from blood.

Treatment of septic olecranon bursitis consists of appropriate antibiotic therapy and repeated bursal needle aspirations (Fam 1994b; Zimmermann *et al*. 1995). Surgical bursotomy is indicated in patients with resistant or loculated septic bursitis.

Other less common elbow tendinitis

Tendinitis of the musculotendinous insertion of the biceps tendon is an uncommon cause of antecubital pain. The pain is increased by resisted flexion and supination of the forearm. It is associated with local tenderness at the distal insertion of the biceps tendon into the radial tuberosity, restriction of full elbow extension, and sometimes a local swelling (Sheon *et al*. 1982). The condition is traumatic in origin. Heat, non-steroidal anti-inflammatory drugs, rest, modification of activity, and occasionally a local corticosteroid injection are often beneficial.

Triceps tendinitis is a less common traumatic lesion, characterized by posterior elbow pain, tenderness, and sometimes swelling of the triceps tendon insertion into the olecranon process. Treatment consists of modification of precipitating activities, non-steroidal anti-inflammatory drugs, and physiotherapy (Sheon *et al*. 1982).

Disorders of the wrist and hand

Pain in the wrist and hand is of diverse causes (Table 1). Only three relatively common painful disorders will be discussed in this section: de Quervain's tenosynovitis, trigger finger or thumb, and Dupuytren's contracture.

De Quervain's tenosynovitis

De Quervain's stenosing tenosynovitis of the abductor pollicis longus and extensor pollicis brevis is commonly due to occupational or vocational repetitive activity, involving pinching with the thumb while moving the wrist in radial and ulnar directions (Thompson and Phelps 1990; Thorson and Szabo 1992). This results in frictional inflammation, thickening, and stenosis of the fibrous tendon sheath as it passes over the distal radius beneath the extensor retinaculum. It may also occur in association with rheumatoid arthritis, psoriatic arthritis, direct trauma, and pregnancy, and during the postpartum period (Nygaard *et al*. 1989).

Most patients report several weeks or months of pain on the radial aspect of the wrist and at thumb base during pinch grip, grasping, and other hand activities. The affected tendon sheath is tender and often swollen. Finkelstein's test is positive (passive ulnar deviation of the wrist with the fingers flexed over the thumb placed in the palm stretches the tendons and reproduces the pain over the radial side), and a tendon crepitus may be palpable (Nygaard *et al*. 1989; Thompson and Phelps 1990; Thorson and Szabo 1992).

Treatment consists of heat application, non-steroidal anti-inflammatory drugs, and splinting (Nygaard *et al*. 1989; Thompson and Phelps 1990; Witt *et al*. 1991; Thorson and Szabo 1992; Sampson *et al*. 1994; Weiss *et al*. 1994). A radial-gutter light support splint immobilizes the wrist in slight extension and radial deviation and the first metacarpophalangeal joint in slight extension. The interphalangeal joint of the thumb is left unrestricted. Modification of hand activities with avoidance of tasks that require repetitive thumb movements or pinch grasping is helpful. In those with persistent pain, one or more corticosteroid injections into the affected tenosynovial sheath are often beneficial (Witt *et al*. 1991). Surgical decompression, with or without tenosynovectomy, is indicated in those with persistent or recurrent symptoms for more than 6 months (Witt *et al*. 1991; Sampson *et al*. 1994; Weiss *et al*. 1994).

Trigger finger or thumb (stenosing digital tenosynovitis)

Trigger finger or thumb, also known as stenosing digital tenosynovitis, or snapping finger or thumb, is the most common repetitive strain injury of the hand (Thompson and Phelps 1990; Thorson and Szabo 1992). The pathological lesion is a tenosynovitis of the flexor tendons of the finger or thumb, resulting in fibrosis and constriction localized to the first annular pulley that overlies the metacarpophalangeal joint (Freiberg *et al*. 1989; Canoso 1990; Thompson and Phelps 1990; Thorson and Szabo 1992). A nodular thickening of the tendon often develops at the site of stenosis; this interferes mechanically with the normal gliding of the tendon, resulting in pain over the area of the pulley and 'snapping', 'triggering', or 'catching' movement of the finger or thumb. Pain over the sheath with resisted flexion, and pain on stretching the tendon passively in extension, are common. Intermittent locking of the digit in flexion may also develop, particularly upon arising in the morning. Passive extension of the proximal interphalangeal joint of the finger or interphalangeal joint of the thumb may produce a crepitus and a popping sensation as the digit is straightened. Examination reveals tenderness over the area of the proximal pulley, linear tenderness and swelling of the flexor tendon sheath, tendon crepitus, and, often, limitation of digital flexion and extension. A nodular swelling of the tendon can usually be palpated in the palm just proximal to the metacarpophalangeal joint as it moves during finger or thumb flexion and extension (Freiberg *et al*. 1989; Canoso 1990; Thompson and Phelps 1990; Thorson and Szabo 1992). The most common cause of trigger finger or thumb is overuse trauma of the hands from repetitive gripping activities, with increased pull and friction on the flexor tendons. One digit is often affected, usually the thumb, middle, or ring fingers in that order. Other causes of flexor digital tenosynovitis include rheumatoid arthritis, psoriatic arthritis, diabetes mellitus, and infections, including tuberculosis, *Mycobacterium marinum*, and sporotrichosis (Canoso 1990).

Management consists of modification of hand activity, local heat application, gentle exercises and non-steroidal anti-inflammatory drugs as required (Freiberg *et al*. 1989; Canoso 1990; Thompson and Phelps 1990; Thorson and Szabo 1992; Sampson *et al*. 1994). Extension splinting of the affected digit at night prevents painful flexion during sleep. Although trigger thumb may resolve spontaneously (Schofield and Citron 1993), one or more injections of corticosteroid into the affected flexor tendon sheath are effective in the majority of patients (Canoso 1990; Thompson and Phelps 1990; Anderson and Kaye 1991; Thorson and Szabo 1992; Schofield and Citron 1993; Sampson *et al*. 1994). Surgical release with transection of the fibrous annular pulley of the finger or thumb flexor sheath is required for those with chronic symptoms not responding to medical treatment (Freiberg *et al*. 1989; Canoso 1990; Thompson and Phelps 1990; Thorson and Szabo 1992; Sampson *et al*. 1994).

Dupuytren's contracture

This is a relatively common condition characterized by nodular thickening and contraction of the palmar fascia, often leading to disabling flexion contractures of the fingers (Wooldridge 1988; McFarlane 1990). In most patients, Dupuytren's contracture affects the ulnar side of both hands. The fourth finger is usually affected earliest, followed by the fifth, third, and second fingers in decreasing order of frequency. Fibrous nodules, composed of whorls of proliferating fibroblasts and myofibroblasts in the superficial layers of the palmar fascia, are the earliest abnormality. The dermis is invaded by fibroblastic cells, resulting in puckering, dimpling, and tethering of the overlying skin. There is usually little pain initially. However, after a variable period of months or years, the aponeurotic thickening may extend distally to involve the digits. The fingers become flexed at the metacarpophalangeal joints by taut fibrous bands radiating from the palmar fascia (Fig. 11), and the hand cannot be placed flat on a table (positive 'table top' test). Although there is no direct involvement of the joints or tendons, progressive flexion deformity of the fingers can lead to severe functional impairment.

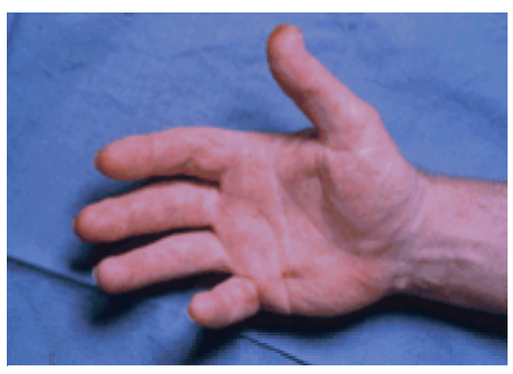


Fig. 11 Dupuytren's contracture.

The aetiology of Dupuytren's contracture is poorly understood (Wooldridge 1988; McFarlane 1990). The disorder is rare in non-Caucasian individuals. Its incidence

rises with increasing age and the sex ratio is predominantly male (5:1). Familial predisposition is frequent, suggesting an autosomal-dominant pattern with variable penetrance. A pathogenetic role for local repetitive injury and occupational trauma remains unproven. An association with cigarette smoking, producing microvascular occlusion, has been suggested (An *et al*. 1988). Dupuytren's contracture has been observed in association with idiopathic epilepsy, alcohol abuse, diabetes mellitus, chronic pulmonary disease, and reflex sympathetic dystrophy syndrome. The association of Dupuytren's disease with other localized fibroses, such as nodular plantar fibromatosis, nodular fasciitis of the popliteal fascia, Peyronie's disease, and dorsal knuckle pads, has led to the concept of a 'Dupuytren's diathesis' (Wheeler and Meals 1981; Wooldridge 1988; McFarlane 1990). The cause of these fibrosing diseases is unknown but seems to be determined by a dominant gene with high penetrance in males, usually middle-aged white men of Celtic ancestry (Wooldridge 1988; McFarlane 1990).

The pathological lesion in Dupuytren's disease is characterized by marked fibroblastic proliferation and vascular hyperplasia. This is followed by dense, disorderly, collagen deposition with thickening of the palmar fascia, and nodule formation (Gelberman *et al*. 1980; Wooldridge 1988; McFarlane 1990). The abnormal fascia contains elevated total amounts of collagen with increased content of reducible cross-links and hydroxylysine; about 25 per cent of the collagen is type III, which is not normally present in the palmar fascia (Gelberman *et al*. 1980; McFarlane 1990). Ultrastructurally, contractile, smooth muscle-like fibroblasts or myofibroblasts, surrounded by bundles of disarrayed collagenous fibrils and completely or partially occluded capillaries, are present in the fibrotic nodules and cords (Gelberman *et al*. 1980; McFarlane 1990; McCann *et al*. 1993). Although myofibroblasts are not specific to Dupuytren's contracture, they are believed to be responsible for contraction of the palmar fascia and finger deformities (Gelberman *et al*. 1980; McCann *et al*. 1993). The finding of isolated foci of smooth-muscle -actin (an antibody marker for myofibroblasts) -positive fibroblasts dispersed in the dermis remote from the main Dupuytren's tissue may explain the high recurrence rate of Dupuytren's disease after fasciectomy (McCann *et al*. 1993). The vasoactive prostaglandins E_2 and F_{2a} are present in increased concentrations in Dupuytren's nodules, and are thought to influence myofibroblast contractility and contribute to the formation of the contracture (Badalamente *et al*. 1988). Increased expression of growth factors, including platelet-derived growth factor-b, transforming growth factor-b, and basic fibroblast growth factor in Dupuytren's fascial lesions, suggests that these mediators may play a part in fibroblastic proliferation in this disorder (Baird *et al*. 1993; Terek *et al*. 1995). Production of oxygen-derived free radicals may also be an important feature of the pathogenesis of Dupuytren's and other fibrotic conditions (Duthie and Francis 1988). Excessive formation of superoxide, hydrogen peroxide, and hydroxyl radicals resulting from microvascular occlusion and relative ischaemia of the palmar fascia can lead to tissue damage and enhanced fibroblastic proliferation (Duthie and Francis 1988). Flow cytometry of inflammatory cells from Dupuytren's lesions has demonstrated a predominance of CD3+ T lymphocytes and increased expression of major histocompatibility complex class II proteins (Baird *et al*. 1993a; Baird *et al*. 1993b). An association between Dupuytren's contracture, HLA-DR3, and autoantibodies to collagen types I to IV has also be demonstrated (Neumuller *et al*. 1994). Factor XIIIa-positive dermal dendrocytes of macrophage lineage are present in and around Dupuytren's nodules; these cells may be an important local source of fibrogenic cytokines (Sugden *et al*. 1993). These observations suggest that Dupuytren's disease may represent a T-cell-mediated autoimmune disorder (Baird *et al*. 1993; Sugden *et al*. 1993; Neumuller *et al*. 1994).

Dupuytren's contracture runs a variable course: some patients show little change or incapacity over a period of many years, while in others fascial contraction progresses rapidly with severe deformity and impairment of hand function within a short period of time (McFarlane 1990).

The treatment depends on the rate of progression and severity of the lesions. Recent observations suggest that MRI may be useful in assessing the extent and degree of cellularity, and hence 'activity' of the lesion (Yacoe *et al*. 1993). Local heat, stretching exercises, and the use of protective padded gloves during heavy manual grasping tasks are often helpful in patients with mild contracture (Wooldridge 1988; McFarlane 1990). Many patients learn the benign nature of the condition and adapt to it. In more severe lesions, with pain and inability to straighten the fingers, intralesional injections of corticosteroid may be beneficial (Wooldridge 1988; McFarlane 1990). Repeated intralesional infiltrations of interferon-g, a cytokine produced by T-helper lymphocytes that inhibits fibroblastic proliferation and collagen formation, is a promising but untested new treatment (Pittet *et al*. 1994). A limited or total palmar fasciectomy, with or without skin-graft replacement, is indicated in those with advanced disease with progressive digital contracture of more than 30°, a positive table-top test, and functional impairment (Ketchum and Hixson 1987; Wooldridge 1988; McFarlane 1990). The risk of recurrence is increased in young patients with active bilateral disease and cellular nodules, and in those with a strong family history and/or other ectopic fibrotic lesions (Ketchum and Hixson 1987; Wooldridge 1988; Yacoe *et al*. 1993).

Chapter References

An, H. S., Southworth, S. R., Jackson, W. T., and Russ, B. (1988). Cigarette smoking and Dupuytren's contracture of the hand. *Journal of Hand Surgery*(A), **13**, 872–4.

Anderson, B. and Kaye, S. (1991). Treatment of flexor tenosynovitis of the hand 'trigger finger' with corticosteroids. A prospective study of the response to local injection. *Archives of Internal Medicine*, **151**, 153–6.

Badalamente, M. A., Hurst, L. C., and Sampson, S. P. (1988). Prostaglandins influence myofibroblast contractility in Dupuytren's disease. *Journal of Hand Surgery*(A), **13**, 867–71.

Baird, K. S., Alwan, W. H., Crossan, J. F., and Wojciak, B. (1993 *a*). T-cell-medicated response in Dupuytren's disease. *Lancet*, **341**, 1622–3.

Baird, K. S., Crossan, J. F., and Ralston, S. H. (1993 *b*). Abnormal growth factor and cytokine expression in Dupuytren's contracture. *Journal of Clinical Pathology*, **46**, 425–8.

Baslund, B., Thomsen, B. S., and Jensen, E. M. (1990). Frozen shoulder: current concepts. *Scandinavian Journal of Rheumatology*, **19**, 321–5.

Bateman, J. E. (1978). *The shoulder and neck* (2nd edn). Saunders, Philadelphia.

Bechtol, C. O. (1980). Biomechanics of the shoulder. *Clinical Orthopedics and Related Research*, **146**, 37–41.

Birrer, R. B., Garrick, J. G., and Hench, P. K. (1989). Syndromes of chronic use and overuse. *Patient Care*, **23**, 215–41.

Bland, J. H., Merrit, J. A., and Boushey, D. R. (1977). The painful shoulder. *Seminars in Arthritis and Rheumatism*, **7**, 21–47.

Bulgen, D. Y., Binder, A. I., Hazleman, B. L., Dutton, J., and Roberts, S. (1984). Frozen shoulder: prospective clinical study with an evaluation of three treatment regimens. *Annals of the Rheumatic Diseases*, **43**, 353–60.

Burkhart, S. S. (1993). Arthroscopic debridement and decompression for selected rotator cuff tears. Clinical results, pathomechanics and patient selection based on biomechanical parameters. *Orthopedic Clinics of North America*, **24**, 111–23.

Canoso, J. J. (1990). Bursitis, tenosynovitis, ganglions and painful lesions of the wrist, elbow and hand. *Current Opinions in Rheumatology*, **2**, 276–81.

Chard, M. D. and Hazleman, B. L. (1987). Shoulder disorders in the elderly (a hospital study). *Annals of the Rheumatic Diseases*, **46**, 684–7.

Chard, M. D., Cawston, T. E., Riley, G. P., Gresham, G. A., and Hazleman, B. L. (1994). Rotator cuff degeneration and lateral epicondylitis: a comparative histological study. *Annals of the Rheumatic Diseases*, **53**, 30–4.

Coel, M., Yamada, C. Y., and Ko, J. (1993). MR imaging of patients with lateral epicondylitis of the elbow (tennis elbow): importance of increased signal of the anconeus muscle. *American Journal of Roentgenology*, **161**, 1019–21.

Dalton, S. E. (1994). The conservative management of rotator cuff disorders. *British Journal of Rheumatology*, **33**, 663–7.

Duthie, R. B. and Francis, M. J. O. (1988). Free radicals and Dupuytren's contracture. *Journal of Bone and Joint Surgery*(B), **70**, 689–91.

Fam, A. G. (1992). Bursitis and tendinitis: a practical approach to diagnosis. *Geriatrics*, **8**, 35–42.

Fam, A. G. (1994 *a*). Wrist and hand. In *Rheumatology*(1st edn) (ed. J. H. Klippel and P. A. Dieppe), pp. (5)10.1–10.10. Mosby, St. Louis.

Fam, A. G. (1994 *b*). Strategies in the management of septic bursitis. *Medicine North America*, **17**, 299–304.

Foley, A. E. (1993). Tennis elbow. *American Family Physician*, **48**, 281–8.

Freiberg, A., Mulholland, R. S., and Levine, R. (1989). Nonoperative treatment of trigger fingers and thumbs. *Journal of Hand Surgery*(A), **14**, 553–8.

Frieman, B. G., Albert, T. J., and Fenlin, J. M., Jr (1994). Rotator cuff disease: a review of diagnosis, pathophysiology, and current trends in treatment. *Archives of Physical Medicine and Rehabilitation*, **75**, 604–9.

Gelberman, R. H., Amiel, D., Rudolph, R. M., and Vance, R. M. (1980). Dupuytren's contracture. An electron microscopic, biochemical and clinical correlative study. *Journal of Bone and Joint*

Surgery(A), **62**, 425–32.

Gellman, H. (1992). Tennis elbow (lateral epicondylitis). *Orthopedic Clinics of North America*, **23**, 75–82.

Geoffroy, P., Yaffe, M. J., and Rohan, I. (1994). Diagnosing and treating lateral epicondylitis. *Canadian Family Physician*, **40**, 73–8.

Glazebrook, M. A., Curwin, S., Islam, M. N., Kozey, J., and Stanish, W. D. (1994). Medial epicondylitis, an electromyographic analysis and an investigation of intervention strategies. *American Journal of Sports Medicine*, **22**, 674–79.

Hassell, A. B., Fowler, P. D., and Dawes, P. T. (1994). Intra-bursal tetracycline in the treatment of olecranon bursitis in patients with rheumatoid arthritis. *British Journal of Rheumatology*, **33**, 859–60.

Hulstyn, M. J. and Weiss, A-P. C. (1993). Adhesive capsulitis of the shoulder. *Orthopedic Review*, **22**, 425–33.

Kapandji, I. A. (1982). The shoulder. *Clinics in Rheumatic Diseases*, **8**, 595–616.

Kessel, L. and Wastson, M. (1977). The painful arc syndrome, clinical classification as a guide to management. *Journal of Bone and Joint Surgery*(B), **59**, 166–72.

Ketchum, L. D. and Hixson, F. P. (1987). Dermofasciectomy and full thickness grafts in the treatment of Dupuytren's contracture. *Journal of Hand Surgery* (A), **12**, 659–63.

Leach, R. E. and Miller, J. K. (1987). Lateral and medial epicondylitis of the elbow. *Clinics in Sports Medicine*, **6**, 259–72.

Lucas, D. B. (1973). Biomechanics of the shoulder joint. *Archives of Surgery*, **107**, 425–32.

McCann, B. G., Logan, A., Belcher, H., Warn, A., and Warn, R. M. (1993). The presence of myofibroblasts in the dermis of patients with Dupuytren's contracture. A possible source for recurrence. *Journal of Hand Surgery* (B), **18**, 656–61.

McFarlane, R. M. (1990). Dupuytren's disease. In *Plastic surgery*(ed. J. G. McCarthy, J. W. May Jr, and J. W. Littler), pp. 5053–86. Saunders, Philadelphia.

Markison, R. E. and Kilgore, E. (1987). Hand. In *Clinical surgery*(ed. J. H. Davis *et al*.), pp. 2293–353. Mosby, St. Louis.

Miniaci, A. and Fowler, P. J. (1993). Impingement in the athlete. *Clinics in Sports Medicine*, **12**, 91–109.

Molsberger, A. and Hille, E. (1994). The analgesic effect of acupuncture in chronic tennis elbow pain. *British Journal of Rheumatology*, **33**, 1162–5.

Morrey, B. F. and An, K-N. (1983). Articular and ligamentous contributions to the stability of the elbow joint. *American Journal of Sports Medicine*, **11**, 315–19.

Neer, C. S., II (1983). Impingement lesions. *Clinical Orthopedics and Related Research*, **173**, 70–7.

Neumuller, J., Menzel, J., and Millesi, H. (1994). Prevalence of HLA-DR3 and auto-antibodies to connective tissue components in Dupuytren's contracture. *Clinical Immunology and Immunopathology*, **71**, 142–8.

Neviaser, R. J. (1980). Lesions of the biceps and tendinitis of the shoulder. *Orthopedic Clinics of North America*, **11**, 343–8.

Neviaser, R. J. (1983). Painful conditions affecting the shoulder. *Clinical Orthopedics and Related Research*, **173**, 63–9.

Newman, J. S., Adler, R. S., Bude, R. O., and Rubin, J. M. (1994). Detection of soft-tissue hyperemia: value of power Doppler sonography. *American Journal of Roentgenology*, **163**, 385–9.

Nygaard, I. E., Saltzman, C. L., Whitehouse, M. B., and Hankin, F. M. (1989). Hand problems in pregnancy. *American Family Physician*, **39**, 123–6.

Petri, M., Dobrow, R., Neiman, R., Whiting-O'Keefe, Q., and Seaman, W. E. (1987). Randomized, double-blind, placebo-controlled study of the treatment of the painful shoulder. *Arthritis and Rheumatism*, **30**, 1040–5.

Pittet, B. *et al*. (1994). Effect of gamma-interferon on the clinical and biologic evolution of hypertrophic scars and Dupuytren's disease: an open pilot study. *Plastic and Reconstructive Surgery*, **93**, 1224–35.

Polley, H. and Hunder, G. (1978). The shoulder, elbow, wrist, carpal, metacarpophalangeal, proximal and distal interphalangeal joints. In *Rheumatologic interview and physical examination of the joints* (2nd edn), pp. 55–148. Saunders, Philadelphia.

Poppen, N. K. and Walker, P. S. (1976). Normal and abnormal motion of the shoulder. *Journal of Bone and Joint Surgery*(A), **58**, 195–201.

Post, M., Silver, R., and Singh, M. (1983). Rotator cuff tear, diagnosis and treatment. *Clinical Orthopedics and Related Research*, **173**, 78–91.

Re, L. P., Jr and Karzel, R. P. (1993). Management of rotator cuff calcifications. *Orthopedic Clinics of North America*, **24**, 125–32.

Recht, M. P. and Resnick, D. (1993). Magnetic resonance imaging studies of the shoulder. Diagnosis of lesions of the rotator cuff. *Journal of Bone and Joint Surgery*(A), **75**, 1244–53.

Rizk, T. E., Gavant, M. L., and Pinals, R. S. (1994). Treatment of adhesive capsulitis (frozen shoulder) with arthrographic capsular distension and rupture. *Archives of Physical Medicine and Rehabilitation*, **75**, 803–7.

Rothman, R. H., Marvel, J. P., Jr, and Heppenstall, R. B. (1975). Anatomic considerations in the glenohumeral joint. *Orthopedic Clinics of North America*, **6**, 341–52.

Ryu, K. N., Lee, S. W., Rhee, Y. G., and Lim, J. H. (1993). Adhesive capsulitis of the shoulder joint: usefulness of dynamic sonography. *Journal of Ultrasound Medicine*, **12**, 445–9.

Sampson, S. P., Wisch, D., and Badalamente, M. A. (1994). Complications of conservative and surgical treatment of de Quervain's disease and trigger fingers. *Hand Clinics*, **10**, 73–82.

Sarrafian, S. (1983). Gross and functional anatomy of the shoulder. *Clinical Orthopedics and Related Research*, **173**, 11–19.

Schofield, C. B. and Citron, N. D. (1993). The natural history of adult trigger thumb. *Journal of Hand Surgery* (B), **18**, 247–8.

Sheon, R. P., Moskowitz, R. W., and Goldberg, V. M. (1982). The elbow. In *Soft tissue rheumatic pain: recognition, management, prevention*, pp. 81–92. Lea and Febiger, Philadelphia.

Smith, R. W., Papadopolous, E., Mani, R., and Cawley, M. I. D. (1994 *a*). Abnormal microvascular responses in lateral epicondylitis. *British Journal of Rheumatology*, **33**, 1166–8.

Smith, D. L., Bakke, A. C., Campbell, S. M., and Beckstead, J. H. (1994 *b*). Immunocytologic characteristics of mononuclear cell populations found in nonspecific olecranon bursitis. *Journal of Rheumatology*, **21**, 209–14.

Snyder, S. J. (1993). Evaluation and treatment of the rotator cuff. *Orthopedic Clinics of North America*, **24**, 173–92.

Sugden, P., Andrew, J. G., Andrew, S. M., and Freemont, A. J. (1993). Dermal dendrocytes in Dupuytren's disease: a link between the skin and pathogenesis? *Journal of Hand Surgery* (B), **18**, 662–6.

Terek, R. M., Jiranek, W. A., Goldberg, M. J., Wolfe, H. J., and Alman, B. A. (1995). The expression of platelet-derived growth-factor gene in Dupuytren's contracture. *Journal of Bone and Joint Surgery*(A), **77**, 1–9.

Thompson, J. S. and Phelps, T. H. (1990). Repetitive strain injuries. How to deal with 'the epidemic of the 1990's'. *Postgraduate Medicine*, **88**, 143–9.

Thorson, E. and Szabo, R. M. (1992). Common tendinitis problems in the hand and forearm. *Orthopedic Clinics of North America*, **23**, 65–74.

Vecchio, P., Kavanagh, R., Hazleman, B. L., and King, R. H. (1995). Shoulder pain in a community-based rheumatology clinic. *British Journal of Rheumatology*, **34**, 440–2.

Verhaar, J., Walenkamp, G., Kester, A., Mameren, H. V., and Der Linden, T. V. (1993). Lateral extensor release for tennis elbow. *Journal of Bone and Joint Surgery*(A), **75**, 1034–43.

Warhold, L. G., Osterman, A. L., and Skirven, T. (1993). Lateral epicondylitis: how to treat it and prevent recurrence. *Journal of Musculoskeletal Medicine*, **10**, 55–73.

Weiss, A-PC., Akelman, E., and Tabatabai, M. (1994). Treatment of De Quervain's disease. *Journal of Hand Surgery* (A), **19**, 595–8.

Wheeler ES, Meals RA (1981). Dupuytren's diathesis: a broad spectrum disease. *Plastic and Reconstructive Surgery*, **68**(5), 781–783.

Williams, P. L., Warwick, R., Dyson, M., and Bannister, L. H. (ed.) (1989). Joints of the upper limb. *Gray's anatomy* (37th edn), pp. 499–516. Churchill Livingstone, Edinburgh.

Witt, J., Pess, G., and Gelberman, R. H. (1991). Treatment of de Quervain tenosynovitis: a prospective study of the results of injection of steroids and immobilization in a splint. *Journal of Bone and Joint Surgery*(A), **73**, 219–22.

Wooldridge, W. E. (1988). Four related fibrosing diseases. When you find one, look for another. *Postgraduate Medicine*, **84**, 269–74.

Yacoe, M. E., Bergman, A. G., Ladd, A. L., and Hellman, B. H. (1993). Dupuytren's contracture: MR imaging findings and correlation between MR signal intensity and cellularity of lesions. *American Journal of Roentgenology*, **160**, 813–17.

Zimmermann, B., III, Mikolich, D. J., and Ho, G., Jr (1995). Septic bursitis. *Seminars in Arthritis and Rheumatism*, **24**, 391–410.

1.2.2.2 The lower limbs in adults

Robert W. Simms

Applied anatomy

Applied anatomy of the hip

The hip joint is a diarthrodial ball-and-socket joint and comprises the head of the femur and the acetabulum of the pelvis (Fig. 1). The acetabular cavity comprises the bones of ilium, ischium, and pubis, which present a horse-shoe shaped cavity to accommodate the head of the femur. This arrangement facilitates one of the strongest ball-and-socket-type joints in the human body. The glenoid labrum or lip, a circular fibrocatilagenous rim, forms an additional reinforcement to the hip joint. At the centre and lower portion of the acetabular cavity is a fat pad lined by synovium. The transverse ligament bridges the lower portion of a gap in the labrum and gives rise to the ligamentum teres, an intracapsular ligament which attaches to a small depression in the head of the femur and carries blood vessels to the femur. These vessels provide nutrients to a small area of the head of the femur adjacent to the ligament attachment. The head of the femur itself is lined with hyaline cartilage which varies somewhat in thickness over its surface; it is thickest over the superior and posterior aspects (3 mm) and thinnest at the lateral margins (2–2.5 mm). The capsule of the hip is a dense, fibrous structure which attaches to the acetabulum, labrum, and the transverse ligament. The capsule is reinforced by three ligaments—anteriorly by the iliofemoral ligament, inferiorly by the pubofemoral ligament, and posteriorly by the ischiofemoral ligament. The capsule is lined by synovial membrane on its inner surface.

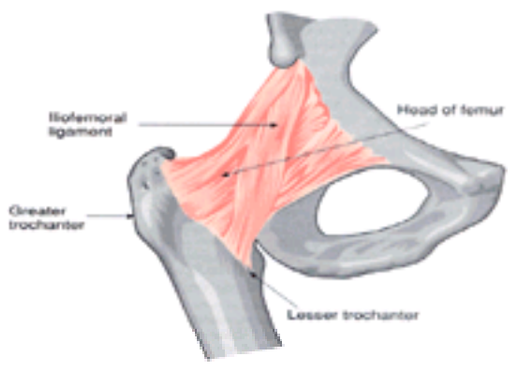


Fig. 1 The anterior view of the hip joint.

A number of bursae have been identified on or about the hip joint (Fig. 2 and Fig. 3). The most important clinically are the trochanteric bursae, the iliopectineal or iliopsoas bursae, and the ischiogluteal bursae (Larsson and Baum 1986). The trochanteric bursae actually comprises three bursae—the gluteus maximus bursa at the site of the insertion of the gluteus maximus muscle and the gluteus medius and gluteus minimus bursae at the sites of the respective muscle attachments (Fig. 2). Clinically the most important bursa is the gluteus maximus bursa. The iliopectineal or iliopsoas bursa lies anteriorly between the psoas muscle and the joint capsule. The ischiogluteal bursa lies over the ischeal tuberosity and facilitates gliding of the gluteus maximus over the tuberosity.

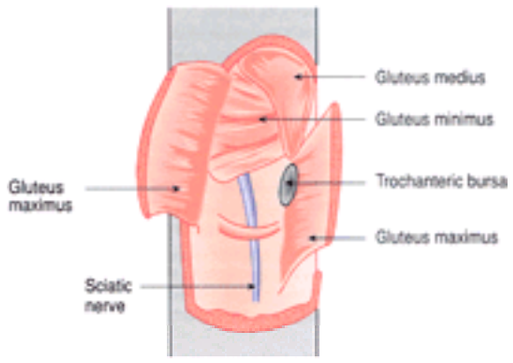


Fig. 2 The posterior view of the hip joint.

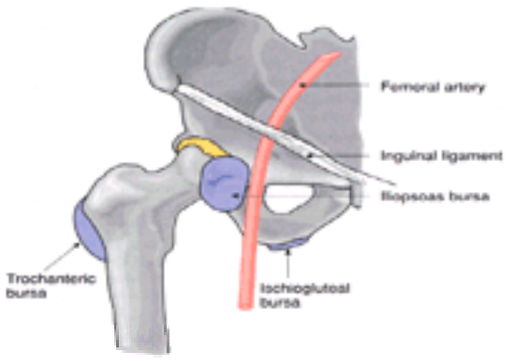


Fig. 3 The bursae of the hip joint.

Hip joint movement is facilitated by several powerful muscle groups. Flexion is accomplished chiefly by the iliopsoas with accessory function of the rectus femoris, pectineus, sartorius, and adductor longus. Extension is accomplished principally by the gluteus maximus and the hamstrings with accessory function by the ischial head of the adductor magnus. Abduction is primarily accomplished by the gluteus medius with accessory function by the gluteus minimus. Adduction is accomplished primarily by the adductor magnus with accessory function by the adductor longus, adductor brevis, pectineus, and gracilis. External rotation is accomplished by the gluteus maximus, quadratus femoris, and piriformis with accessory function by the sartorius and gracilis. Internal rotation is accomplished by the gluteus minimus with accessory function by the gluteus maximus, gluteus medius, adductor longus, adductor brevis, adductor magnus, pectineus, iliacus, and psoas. Normal hip range of motion varies with the manner in which it is examined. With the knee in flexion the hip can be flexed to 120°. With the knee extended, the hip can be flexed to only approximately 90°, being limited by the pull of the hamstrings. Internal rotation is normally to 40° with external rotation to 45°. Abduction is normally to 45° and adduction to 20 to 30°.

The normal gait cycle involves two basic phases—the stance phase and the swing phase (Fig. 4) (Morris *et al*. 1994). The stance phase begins with heal contact. Full weight bearing begins with forefoot contact and ends with heel lift. The stance phase then ends with lift off and starts the swing phase. The swing phase then continues until heel contact and the cycle repeats. Under normal conditions the stance phase comprises approximately 65 per cent of the gait cycle whereas the swing phase comprises approximately 35 per cent (Fig. 4). Hip disorders may cause the patient to walk with a shortened stance phase or with an antalgic gait since he will try to minimize the time spent on full weight bearing. Chronic hip disorders frequently produce weakness of the gluteus medius or a Trendelenberg gait. Here, hip abduction is impaired and the pelvis on the opposite side drops during the stance phase. The body leans toward the diseased side when weight bearing is on the diseased hip.

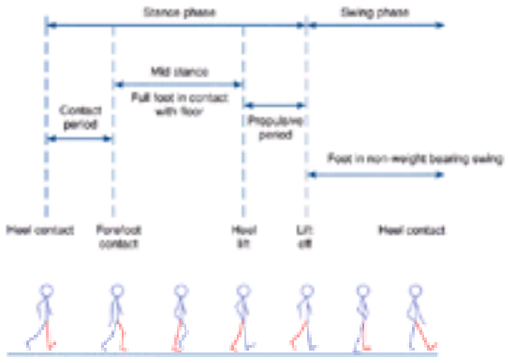


Fig. 4 The gait cycle.

Applied anatomy of the knee

The knee is the most complex and the largest joint in the human body (Fig. 5). It is a modified hinge joint with three compartments; the medial and lateral tibiofemoral compartments and the patellofemoral compartment. The knee is not a simple hinge joint since it is capable of flexion, extension, and rotation. The latter movement occurs as the knee approaches full extension and is due to the shape of the joint surfaces and ligament tension, which causes internal rotation of the femur—the so-called 'screwed-home position'. Stability is provided by the cruciate ligaments and menisci and the capsule with its associated capsular ligaments.

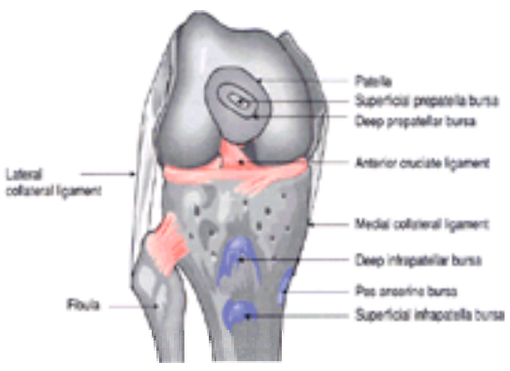


Fig. 5 The anterior view of the knee joint with associated bursae.

The menisci or semilunar cartilages are crescent-shaped structures consisting of fibrocartilage, which transmit up to 70 per cent of the force through the tibiotalar joint and also have a lubricating role (Fig. 6). They are predominately avascular, although the peripheral 10 to 30 per cent and the anterior and posterior horns receive blood supply from the geniculate vessels and therefore they have the capability for repair. The remainder of the menisci receive nutrients passively from the synovial fluid and have little capacity for repair.

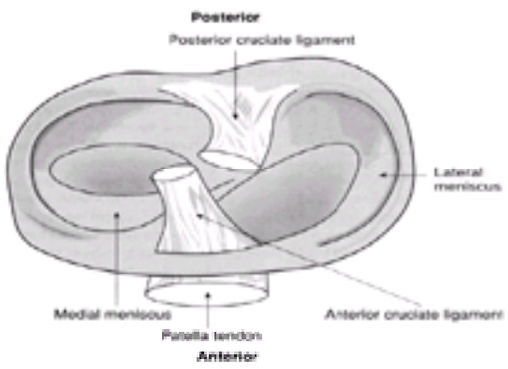


Fig. 6 The meniscii of the knee.

The anterior and posterior cruciate ligaments together with the medial and lateral collateral ligaments are accessory ligaments of the capsule of the knee (Fig. 6). The anterior and posterior cruciate ligaments cross each other within the joint and are formed by twisting rope-like fibres of collagen of different lengths. The cruciate ligaments are the principal knee stabilizers in the anteroposterior plane with the anterior cruciate preventing anterior slippage of the femur relative to the tibia, and the posterior cruciate preventing posterior slippage. The cruciates also provide rotational stability. The collateral ligaments provide medial and lateral stability. The iliotibial band is a fascial band connecting the ilium with the lateral tibia and contributes to the lateral stability of the knee (Fig. 7).

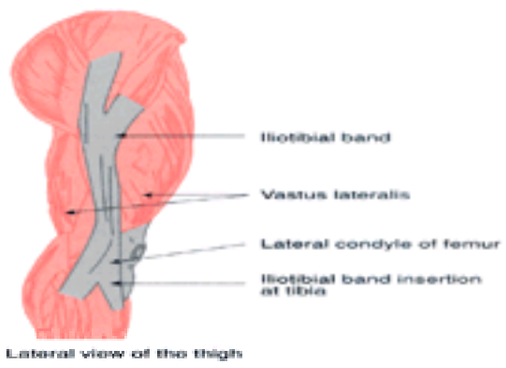


Fig. 7 The iliotibial band.

The patella is the largest sesamoid bone in the human body and its undersurface possesses the thickest articular cartilage. The primary function of the patella is to increase the lever arm of the quadriceps. The stability of the patella is provided statically by the bony interface between the trochlea of the femur and the undersurface of the patella. Dynamic stability is provided predominately by the extensor mechanism. There is considerable variation in the articular anatomy of the patella. The most common shape is a relatively larger lateral facet when compared to the medial facet. The shape of the femoral trochlea also varies considerably—it may be shallow and broad or deeply V-shaped or in between these extremes. The quadriceps mechanism and its attachment to the knee is complex. Most anterior is the rectus femoris which inserts anteriorly and most superficially to the patella. Medially the vastus medialis inserts obliquely at an angle of 60° to 70° and provides dynamic medial stability for the patella. This counteracts the lateral dynamic force on the patella imposed by the majority of the quadriceps mechanism which is a consequence of the normal slight varus angulation of the lower leg in relation to the femur. Patellar tracking is influenced by the direction of pull of the quadriceps and the position of attachment of the patellar tendon into the tibial tubercle (Tria *et al*. 1992). This relationship is expressed by the patella Q angle, which is the angle formed by a line between the centre of the patella extending proximally to the anterosuperior iliac spine and extending distally to the centre of the tibal tubercle (Fig. 8). The normal Q angle is 10° in men and 15° in women. Femoral neck anteversion and external tibial torsion increases the Q angle, whereas femoral neck retroversion and internal tibial torsion decreases the Q angle (Tria *et al*. 1992). An excessive Q angle may predispose to patella subluxation or dislocation, since as the Q angle increases the patella tends to track more laterally.

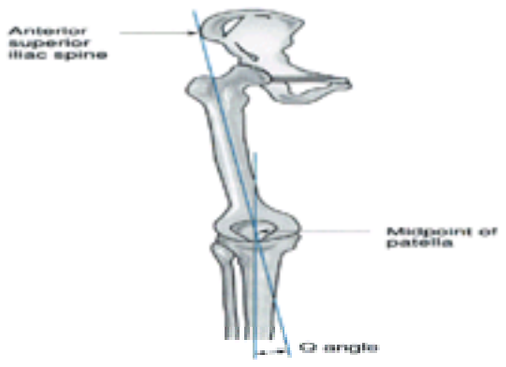


Fig. 8 The patella Q angle.

There are a number of bursae surrounding the knee joint (Fig. 5). Anteriorly is the subcutaneous, prepatellar bursa which lies over the lower pole of the patella. Inferiorly are the infrapatellar bursae; the superficial infrapatellar bursa which lies between the patella tendon and the tibial tubercle and the deep infrapatellar bursa which lie between the patellar tendon and the proximal tibia. Medial and inferior to the joint is the pes anserine bursa which lies between the medial collateral ligament and the insertion of the adductor muscles of the thigh; the sartorius, gracilis, and semitendinosus. Posteriorly is the gastrocnemius–semimembranous bursa which is a complex structure and comprises three components; a base which is located under the capsular insertion of the gastrocnemius muscle, the medial extent between the heads of the gastrocnemius and the semimembranous muscle, and a small subfascial extension. In over 50 per cent of individuals over the age of 50 years there is communication between the gastrocnemius–semimembranous bursa and the knee joint (Canoso 1981). When a communication exists synovial fluid from any aetiology may enter the bursa. If a large volume of synovial fluid expands into the gastrocnemius–semimembranous bursa suddenly it may bulge into the popliteal fossa or it may rupture into the calf producing the so-called ruptured Baker's cyst or pseudothrombophlebitis (Rauschning 1980).

Applied anatomy of the foot and ankle

The foot comprises 26 bones and 38 muscles along with a variable number of sesamoids which are held together by 125 interconnecting ligaments and must absorb up to six times the body weight with each step (Fig. 9). The foot is usually divided into three sections; the forefoot, the midfoot, and the hindfoot. The forefoot consists of the five metatarsals and is separated from the midfoot by the tarsometatarsal joint of Lisfranc. The midfoot contains the three cuneiforms, the navicular, and the cuboid and is separated from the hindfoot by the transverse midtarsal joint. The hindfoot consists of the calcaneus and the talus.

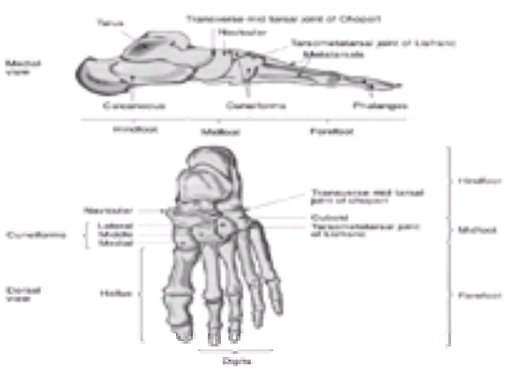


Fig. 9 The bones of the foot.

The ankle joint comprises three articulations; the tibiotalar joint, the subtalar joint, and the talonavicular joint. The tibiotalar joint or 'true' ankle joint is the result of the articulation of the dome of the talus, the roof or plafond of the distal tibia, and the distal tibiofibular joint. The joint is stabilized medially by the medial bony protrusion of the distal tibia, the medial malleolus, and the distal lateral fibula forming the lateral malleolus. Motion at the tibiotalar joint consists of 20° of dorsiflexion and 50° of plantar flexion. A complex set of ligaments additionally stabilizes the ankle. The deltoid ligament (comprising four ligaments; the posterior tibiotalar ligament, the tibiocalcaneal ligament, the tibionavicular ligament, and the anterior tibiotalar ligament) provides medial stability. The lateral collateral ligament (comprising the posterior talofibular ligament, the calcaneofibular ligament, and the anterior talofibular ligament) provides lateral stability.

The subtalar joint is composed of the articulation of the calcaneous and the talus, is surrounded by its own distinct capsule, and does not articulate with other joints. Normal range of motion of the subtalar joint is 5° of inversion and 5° of eversion. The subtalar joint provides inversion and eversion of the heel and facilitates walking on uneven terrain. The transverse midtarsal joint of Chopart is composed of the talonavicular joint and the calcaneocuboid joint and permits multiaxial motion—inversion and eversion of the midfoot and forefoot and, to a lesser degree, dorsiflexion and plantar flexion and abduction and adduction. It is stabilized by the bifurcate ligament which has two components; the calcaneonavicular segment and the calcaneocuboid component. The tarsometatarsal joint of Lisfranc comprises the interconnected second to fifth tarsometatarsal joints. A complex set of ligaments on their dorsal and plantar aspects provide stability and consists of two principal components; the dorsal tarsometatarsal ligaments and the plantar tarsometatarsal ligaments.

The metatarsophalangeal joints are analogous to the metacarpophalangeal joints of the hands and are stabilized by the deep transverse metatarsal ligament and the plantar ligaments. The first metatarsophalangeal joint plays a critical role in normal gait. Under most conditions, 65° to 75° of dorsiflexion of the hallux on the first metatarsal is required for normal gait and for the hallux to function in propulsion (Mahan 1994). Disorders such as hallux valgus and hallux rigidus affect gait adversely by limiting dorsiflexion of the first metatarsophalangeal joint.

The interphalangeal joints of the toes are analogous to those of the hand and consist of hinge joints stabilized by collateral ligaments and a plantar capsular ligament with a fibrous plate. The plantar aponeurosis or plantar fascias runs from the plantar aspect of the calcaneous to the region of the metatarsal heads. It comprises two distinct layers; the superficial and deep. The superficial layer blends with subcutaneous tissue, whereas the deep layer joins the deep transverse metatarsal ligament and the flexor tendons.

Attached to the foot and ankle are the tendons of the extrinsic muscles which form three compartments in the lower leg; the anterior compartment containing the dorsiflexors or extensors of the foot and ankle, the lateral compartment containing the peroneal muscles which act as plantar flexors and abductors, and the posterior compartment which contains the plantar flexors (Fig. 10). Posterior to the ankle is the Achilles tendon, formed by the conjoined tendons of the gastrocnemius and the soleus muscles which functions to plantar flex the foot. Medially at the ankle is the tibialis posterior tendon complex, comprising the tendons of the flexor digitorum longus, the tibialis posterior, and the flexor hallucis longus which pass underneath the flexor retinaculum posterior to the medial malleolus. Between the sheaths of the flexor digitorum longus and the flexor hallucis longus tendons lies the posterior tibial artery and nerve within the tarsal tunnel. Tenosynovitis of either the tibialis posterior, flexor hallucis longus, or flexor digitiorum longus tendons may produce functional compression of the posterior tibial nerve producing tarsal tunnel syndrome (Grumbine *et al*. 1990). The tibialis posterior is the principal inverter of the foot and maintains the longitudinal arch by its insertion to the navicular and medial and intermediate cuneiforms and bases of the second, third, and forth metatarsals (Sarrafian *et al*. 1983). Laterally at the ankle, passing under the superior and inferior peroneal retinacula, is the common sheath of the peroneus longus and brevis tendons, which are the principal everters and abductors of the foot. Anteriorly, the extrinsic extensor tendons of the foot (the extensor digitorum longus, the peroneus tertius, and the extensor hallucis longus tendons) within their tendon sheaths pass under the fibrous superior and inferior extensor retinacula.

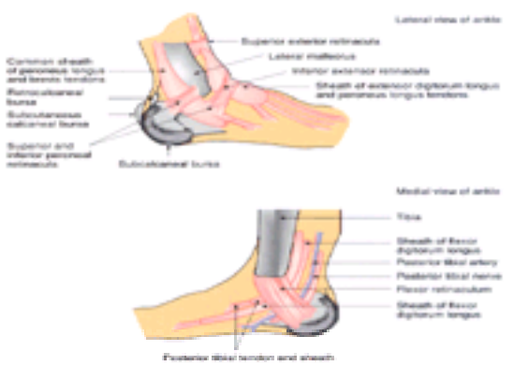


Fig. 10 The ankle: lateral and medial views.

There are several pertinent bursae of the foot and ankle (Fig. 10 and Fig. 11). The retrocalcaneal bursa lies between the calcaneus anteriorly and the Achilles tendon posteriorly (Fig. 11). This bursa has a synovial lining which abuts the Achilles fat pad. Anteriorly, the bursal wall is composed of fibrocartilage and posteriorly is contiguous with the Achilles tendon (Frey *et al*. 1992; Canoso *et al*. 1984). The bursa itself is horseshoe shaped with the average length of the legs measuring 22 mm and the width 4 mm, with the width of the body measuring 8 mm (Frey *et al*. 1992). Between the skin and the Achilles tendon posteriorly is the subcutaneous calcaneal bursa. At the plantar aspect of the midcalcaneous is the subcalcaneal bursa.

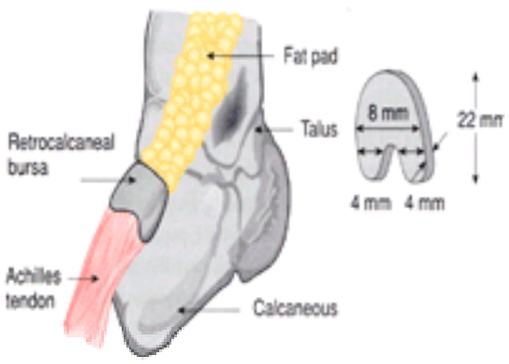


Fig. 11 The retrocalcaneal bursa.

General approach to regional disorders of the lower extremities

Lower extremity regional pain syndromes are common in rheumatology and orthopaedic practices although precise population-based incidence data are not available. A detailed survey and examination of soft tissue abnormalities in 123 unselected medical students revealed that approximately 10 per cent had either bursitis, tendinitis, or chondromalacia patellae (Raskin and Lawless 1982). Lower extremity regional pain conditions often present a challenge to the clinician because of their complexity, frequent confusion with other conditions, and lack of precise diagnostic tests. However, careful history and physical examination combined with knowledge of the relevant anatomy will generally permit identification of these conditions. Table 1 provides a classification of the regional disorders of the lower extremity based on location.

Table 1 Classification of painful lower limb disorders

Many of the conditions producing lower extremity regional pain are bursal disorders. It is worthwhile to review some general concepts relevant to bursae. Bursae are synovial-lined sacs which are located to facilitate tissue gliding. Bursae can be generally divided into deep and subcutaneous varieties. Subcutaneous bursae generally overlie bony prominences and permit skin movement without undue stretching and tearing and they develop after birth (Canoso 1981; Canoso and Yood 1979). Deep bursae are located between divergently pulling tendons or between tendons and bone and generally form in early fetal development. Normal bursal fluid has characteristics similar to synovial fluid—a high content of hyaluronic acid, low leucocyte count (300/mm ^{3}) with a predominance of mononuclear cells (Canoso 1981; Canoso *et al*. 1983). A variety of disorders can affect bursae, particularly the subcutaneous variety. This includes trauma, infection, and systemic or metabolic conditions such as rheumatoid arthritis or gout. Septic subcutaneous bursitis is particularly important to exclude. In the lower extremity, subcutaneous bursae around the knee are the most common sites of infection. Prepatellar bursitis is the second most common form of septic bursitis behind olecranon bursitis, accounting for 27 per cent of 49 cases of septic bursitis in one series (Raddatz *et al*. 1987). Identifiable risk factors for septic bursitis include skin breakdown, trauma, immunosuppression (diabetes, concurrent immunosuppressive medication), and occupations involving frequent or sustained pressure on bursae, such as roofing or carpentry (Raddatz *et al*. 1987). By far the most common organism responsible for septic bursitis is *Staphylococcus aureus*, accounting for approximately 80 per cent of all cases (Raddatz *et al*. 1987; Ho *et al*. 1978; Ho and Ticer 1979). Less frequently other organisms may be cultured (Meys *et al*. 1992; McDermott *et al*. 1994). Aspiration of symptomatic subcutaneous bursae is generally required to exclude infection. In the case of the prepatellar bursa, aspiration should be performed on the lateral site of the bursa to prevent a chronic draining sinus which can occur if aspiration is performed over the apex of the bursa. Bursal fluid leucocyte counts which exceed 5000/mm^{3} should be suspected of being septic and empirical antibiotic coverage directed against *S. aureus* should be considered in this setting. Optimal treatment consists of frequent aspiration until no fluid reaccumulates and either oral or intravenous antibiotics (Ho and Su 1981). The length of time required to achieve culture sterility with antibiotic therapy of septic bursitis (mean of 4 days) was correlated with the duration of symptoms prior to diagnosis in one series (Ho and Su 1981). In patients treated within 2 weeks from the onset of symptoms, culture sterility occurred within 1 week of therapy, while longer duration of symptoms was associated with delayed response (Ho and Su 1981).

Hip region pain

Trochanteric bursitis

There are three bursae in the greater trochanteric region of the hip. The most important clinically, and the largest, is the gluteus maximus bursa which separates the fibres of the gluteus maximus from the greater trochanter (Larsson and Baum 1986). The two other bursae lie between the gluteus medius and greater trochanter and gluteus minimus and greater trochanter respectively. Patients with trochanteric bursitis present with a deep, aching pain over the lateral upper thigh which is often intensified by walking but it may be worse at night, especially when the patient is lying on the affected side. The course is variable with an acute phase lasting several days followed by gradual improvement and resolution over a period of days or weeks. In some cases the course may be quite protracted with symptoms which persist for months. The diagnosis is confirmed with palpation of the posterior aspect of the greater trochanter which elicits local tenderness in this area. Resisted abduction of hip on the affected side may also be painful.

The aetiology of trochanteric bursitis is unknown; however, potential risk factors include local trauma, overuse activities such as jogging, and leg length discrepancies (primarily on the side with the longer leg). These factors are thought to lead to increased tension of the gluteus maximus on the iliotibial band and produce bursal inflammation. The differential diagnosis of trochanteric bursitis includes lumbar radiculopathy (particularly the L1 and L2 nerve roots), meralgia paraesthetica (entrapment of the lateral cutaneous nerve of the thigh as it passes under the inguinal ligament), true hip joint disease, and intra-abdominal pathology. These potentially confounding conditions can generally be excluded by careful history and physical examination.

Treatment of trochanteric bursitis is best accomplished by avoiding aggravating activity and local injection of a long-acting corticosteroid preparation. Referral to physical therapy in resistant cases may be helpful, particularly for exercises to stretch the iliotibial band.

Ischiogluteal bursitis

The ischiogluteal bursa lies over the ischeal tuberosity and facilitates gliding of the gluteus maximus muscle over the tuberosity (Fig. 3). In the sitting position, the gluteus maximus slides superiorly exposing the ischeal tuberosity to potential friction.

Prolonged sitting on hard surfaces or repeated leg flexion and extension in the sitting position may predispose to repeated friction on the ischiogluteal bursa and lead to bursitis, also known as 'weaver's bottom.' Rarely the ischiogluteal bursa may become infected (Lambie *et al*. 1989). The diagnosis is made by the characteristic history and by eliciting tenderness on palpation of the ischeal tuberosity with the patient lying supine and with the hip and knee flexed. Treatment consists of using a cushioning 'doughnut' seat and performing trunk and knee-to-chest stretching exercises while lying on the cushion. A local injection of corticosteroid is useful in refractory cases, although care should be taken to avoid the sciatic nerve which lies just lateral to the ischiogluteal bursa.

Iliopsoas bursitis

The iliopsoas bursa (also called the iliopectineal bursa) lies over the anterior surface of the hip joint under the iliopsoas muscle and between the hip joint capsule (Fig. 3). In most cases of reported iliopsoas bursitis, communication between the hip joint and the bursa occurs, either due to pre-existing communication (which occurs in approximately 15–20 per cent of normal cases) and associated excessive synovial fluid accumulation or due to rupture of the hip joint capsule (Canoso 1981). The condition has been described in association with a number of hip disorders including osteoarthritis, rheumatoid arthritis, synovial chrondromatosis, pigmented villonodular synovitis, osteonecrosis, and septic arthritis (Broadhurst 1995; Flanagan *et al*. 1995; Generini *et al*. 1993; Letourneau *et al*. 1991; Manueddu *et al*. 1991). Typically, patients present with an inguinal mass which may be painful and may be associated with secondary femoral vein obstruction or femoral nerve compression (Savarese *et al*. 1991; Lavyne *et al*. 1982). Computed tomography is considered the optimal modality to diagnose iliopsoas bursitis. Treatment of the underlying hip disorder is generally sufficient, although surgical excision is sometimes required.

Knee pain

Knee pain is among the most common articular complaints. Non-arthritic conditions are a large proportion of knee complaints, particularly in younger age groups.

Knee pain in adolescence (see also Chapter 1.2.2.3)

A variety of knee pain syndromes occur in the adolescent (age 10–18 years) which are unique to this age group. Osgood–Schlatter disease is one such condition and generally presents with activity-related pain around the tibial tubercle. This condition is thought to result from submaximal, repetitive tensile stresses acting on the immature junction of the patellar ligament, tibial tubercle, and tibia, causing mild avulsion injury followed by attempts at osseous repair with proliferation of cartilage and bone (Ogden *et al*. 1976; Kujala *et al*. 1985). Pain is generally intermittent and increased with activities, especially those which require kneeling, squatting, or repeated jumping. Athletically active boys were found to have a fivefold risk of Osgood–Schlatter disease compared to inactive controls in one study (Kujala *et al*. 1985). An increased Q angle has also been associated with Osgood–Schlatter disease (Kujala *et al*. 1985). Physical examination is notable for swelling and

prominence of the tibial tubercle with local tenderness. In approximately 50 per cent of cases a discrete ossicle is noted at the tibial tubercle. Radiographs of the knee may be useful to exclude tumours and infection. Treatment with ice, anti-inflammatory agents, an appropriately contoured knee pad, and maintaining hamstring and quadriceps flexibility are usually sufficient to control symptoms which may take up to 12 months to resolve. Progressive or disabling symptoms may require a course of immobilization for 7 to 10 days.

Sinding–Larsen–Johansson disease is another cause of anterior knee pain in the adolescent and is similar to Osgood–Schlatter disease, except that tenderness is localized to the inferior pole of the patella, or occasionally at the junction of the quadriceps tendon and the patella (Medlar and Lyne 1978). This condition is thought to be the result of persistent traction at the cartilagenous junction of the patella and the patella ligament. Radiographs may show ossification at the junction of the patella and the ligament. The treatment programme outlined for Osgood–Schlatter disease is generally successful in the management of Sinding–Larsen–Johansson disease.

Prepatellar bursitis

Prepatellar bursitis presents as a painful, red, swelling anterior to the kneecap and is seen most often in people who spend a lot of time kneeling, for example roofers or carpet fitters (Raddatz *et al*. 1987). Active knee extension is usually quite painful although passive knee flexion is usually only minimally limited. Infection and gout should be excluded by aspirating any bursal fluid. In most cases rest and avoiding sustained kneeling results in resolution. Septic bursitis is usually due to *S. aureus* and requires antibiotic therapy as well as repeated drainage if fluid reaccumulates (Raddatz *et al*. 1987; Ho *et al*. 1978; Ho and Su 1981; Kerr 1993). Rarely, recurrent episodes of inflammation require surgical excision of the bursa (Kerr 1993).

Infrapatellar bursitis

The deep infrapatellar bursa lies between the upper portion of the tibial tuberosity and patellar ligament and it is separated from the knee joint synovium by a fat pad (Fig. 5). Infrapatellar bursitis presents in a similar fashion to prepatellar bursitis, although the location of swelling and tenderness is in the soft tissues on either side of the patellar ligament just proximal to its insertion on the tibial tubercle. As a result, inflammation in this bursa may be more difficult to detect than in a subcutaneous bursa (Taylor 1989). The risk factors for the development of infrapatellar bursitis are the same as for prepatellar bursitis (Meys *et al*. 1992; Taylor 1989).

Anserine bursitis

The anserine bursa lies under and about the pes anserinus—the insertion of the thigh adductor complex consisting of the sartorius, gracilis, and semitendinosus muscles (Fig. 5). It is located about 5 cm below the medial aspect of the joint space. Rarely, fluid distension of the bursa has been reported. Another bursa, the subtendinea musculi sartorii bursa, lies in this region which is located approximately 0.5 to 1 cm cranial to the anserine bursa. Anserine bursitis is a term loosely applied to pain and associated local tenderness in this region, although 'medial ligament' syndrome and pes anserinus tendinitis may be impossible to separate. Nocturnal pain, often leading to the use of a pillow between the knees, is characteristic. Rarely, a palpable mass may be detected (Zeiss *et al*. 1993; Voorneveld *et al*. 1989). Women seem more likely to develop anserine bursitis, perhaps because of a broader pelvic area and greater angulation of the adductors at the knee joint producing more tension on these attachments (Larsson and Baum 1985). Obesity and osteoarthritis of the knees appear to be additional predisposing factors (Larsson and Baum 1985). Optimal treatment includes the use of local corticosteroid injection mixed with local anaesthetic.

Gastrocnemius–semimembranous bursitis (Baker's cyst)

Any cause of synovitis in the knee can lead to leakage into the popliteal space when a communication exists between the gastrocnemius–semimembranous bursa. A communication between the gastrocnemius–semimembranous bursa and the anterior knee joint occurs in 40 per cent of normal adults and increases in frequency with age (Morris *et al*. 1994). Symptoms consist of fullness and tightness in the popliteal space which increases with walking. A Baker's cyst can be detected clinically by palpation of fullness in the medial third of the popliteal fossa, although there is wide variation in the size of Baker's cysts, ranging from politeal cysts to large calf cysts (Smith *et al*. 1988). Baker's cysts become softer with semiflexion and harder with extension of the knee (Foucher's sign) (Wigley 1982). Knee effusions are generally present although they may be small and are often overlooked. Baker's cysts may dissect into the muscles of the calf, simulating thrombophlebitis (Hench *et al*. 1966). The presence of crescent bruising beneath the medial malleolus, the 'haemorrhagic crescent sign', is felt to be a useful clinical clue to the presence of popliteal cyst rupture or calf haematoma (Kraag *et al*. 1976). Rupture of a Baker's cyst should be confirmed by Doppler ultrasound which can also exclude thrombophlebitis of the popliteal vein (Fam *et al*. 1982). To confirm a ruptured Baker's cyst the ultrasonographer's attention must be directed to the soft tissue plains of the calf where dissecting fluid is generally seen. Rarely, rupture of a Baker's cyst causes associated thrombophlebitis or compressive neuropathy due to extensive venous or neural compression (Nakano 1978).

Treatment of an intact or ruptured Baker's cyst requires treating the cause of the excess or abnormal synovial fluid production in the knee. Optimal management includes aspiration of the knee joint with synovial fluid analysis and culture. Direct posterior aspiration of the Baker's cyst should not be performed due to the close proximity of neurovascular structures. In non-infectious synovitis of the knee, intra-articular instillation of corticosteroid is an effective treatment.

Iliotibial band syndrome (Runner's knee)

The iliotibial band is a fascial band connecting the ilium with the lateral tibia, connecting to the tensor fasciae latae and the gluteus maximus (Fig. 7). Repetitive flexion and extension, particularly with high intensity running, can lead to inflammation of the iliotibial band or at its associated bursa overlying the lateral femoral condyle (Sutker 1981; Martens *et al*. 1989). Excessive foot pronation resulting in tightening of the iliotibial band and increased friction across the femoral epicondyle has also been proposed as a cause of iliotibial band syndrome (Bouche *et al*. 1994). Examination reveals tenderness localized to the lateral femoral condyle approximately 2 cm above the joint line, pain with weight bearing on the knee flexed 30° to 40°, and a positive Ober test (the patient lies on his or her side with the lower leg flexed to eliminate lordosis of the lumbar spine, the knee of the upper leg is flexed to 90°, and the thigh is abducted and extended—a positive test occurs when the hip remains abducted when the examiner's supporting hand is removed). Treatment consists of rest, locally applied heat, and orthotics when excessive foot pronation is present. Occasionally, a local injection of corticosteroid into the painful area is required. Partial resection of the iliotibial band over the lateral femoral condyle is reserved for the most resistant cases (Martens *et al*. 1989).

Pelligrini–Stieda disease

This condition results from calcification of a haematoma at the femoral insertion of the medial collateral ligament following injury. Examination reveals local tenderness and pain with application of valgus stress to the knee. Radiographs show characteristic calcification of the insertion of the medial collateral ligament. The condition is usually self limited. Treatment consists of rest, non-steroidal anti-inflammatory agents, and, occasionally, local corticosteroid injection.

Plica syndrome

Synovial plica are normal folds of the synovium that are remnants of the embryonic development of the synovial sac and are found in approximately 50 per cent of cadaveric knees (Galloway and Jokl 1990). With the advent of arthroscopy and direct visualization of these structures, they have been implicated in previously undiagnosed knee symptoms (Patel 1978). Pathological plica appear to result from two mechanisms:

1. direct trauma with resultant haemorrhage, oedema, and progressive fibrosis;
2. overuse often associated with minor irregularities in knee mechanics causing progressive inflammation with recurrent synovitis, oedema, and fibrosis with resultant thickening of the plica which then irritates surrounding tissue (O'Dwyer and Peace 1988; Tindel and Nisonson 1992).

Patients with symptomatic plica in the knee complain of snapping or popping of the knee at particular degrees of flexion. Medial plica are most often implicated as a cause of knee symptoms, presumably occurring when the plica rubs across the femoral condyle (Galloway and Jokl 1990). Examination often reveals tenderness at the joint line and the palpation of the plica, especially with the knee in 20 of flexion. Most patients respond to a combination of quadriceps strengthening and flexibility exercises and anti-inflammatory medication. Surgical resection is sometimes advocated but remains a controversial therapy (Tindel and Nisonson 1992).

Internal derangements

Internal derangements of the knee refers to disruption of the normal functioning of the ligaments and menisci. Disruption of the ligaments is most often the result of

significant, often athletic, trauma. Acute meniscal injury in young adults is usually the result of a twisting force applied to the weight-bearing knee and results in a longitudinal tear which, if it is large enough, may cause the knee to lock (a 'bucket-handle' tear). Degenerative tears are more frequent in older adults and typically consist of radial tears. Chronic tears of this nature may occur without obvious trauma but are generally accompanied by symptoms of knee locking or giving way. Examination generally reveals pain with forced extension and generally normal or only mildly reduced flexion. McMurray's test (flexing and extending the knee while the tibia is internally and then externally rotated with a positive test indicated by pain and an associated click) may he helpful if positive but is relatively insensitive (Simonsen *et al*. 1984; Gillies and Seligson 1979). In the United States, magnetic resonance imaging (**MRI**) is widely used to diagnose suspected meniscal tears and to help plan the therapeutic approach. Early studies with MRI suggested a relatively high rate of false-positive findings, although currently, with improved technology and experience, this appears to be less of a problem but there still may be substantial variation in accuracy among centres (Watt 1991; Fischer *et al*. 1991; Reicher *et al*. 1986). Symptomatic peripheral tears (i.e. the outer one-third) of the meniscus in the young adult are generally repaired since there is potential for healing. Partial menisectomy is recommended if tears involve the avascular inner two-thirds of the meniscus. Non-operative treatment can be considered in the older patient with a degenerative meniscal tear, although arthroscopic resection may be required for persistent symptoms. Total menisectomy is generally avoided since long-term follow-up studies have demonstrated high rates of osteoarthritis following this procedure (Cooper 1995).

Patellofemoral pain syndrome

Chondromalacia patellae is a term which has been used synonymously with patellofemoral pain. Most authorities now urge that the term be abandoned in this setting given the almost universal presence of asymptomatic cartilage changes on the medial facet of the patella and the observation via arthroscopy that most, if not many, individuals with anterior knee symptoms have normal articular surfaces (Tria *et al*. 1992; Griffiths and Pinder 1981; Kelly and Insall 1992). Chrondromalacia patellae is now recommended to connote gross pathological observation made operatively or at autopsy (Tria *et al*. 1992). Patellofemoral pain is now the recommended term applied to ill-defined, anterior knee pain.

Patellofemoral pain syndrome is most commonly use to describe poorly localized, anterior knee pain, frequently accompanied by the 'grab' sign (patients cover the entire front of the knee with the hand when asked to identify the location of the discomfort). Pain frequently occurs after prolonged sitting with flexed knees, the so-called 'theatre' sign. Physical examination may reveal tenderness of the medial or lateral facets of the patella, pain with compression of the patella on the femoral condyle, or a patellar 'shrug' sign (pain when pressure is applied to the patella while the patient contracts the quadriceps). An abnormally increased Q angle (greater than 15° in males and greater than 20° in females) appears to predispose to patellofemoral pain (Tria *et al*. 1992). The Q angle is easily measured by determining the angle formed between a line drawn from the anterior superior iliac spine to the midpoint of the patella and a line drawn from the tibial tubercle to the same point on the patella (Fig. 8). Patellofemoral pain syndrome is thought to result from either anatomical abnormalities (such as anatomical misalignment of the patella tracking mechanism, quadriceps dysplasia, or patellofemoral ligament imbalance) or from repetitive microtrauma to the patella surface (Tria *et al*. 1992). Standard radiographs are usually not helpful in the evaluation of anterior knee pain—particularly for patellar tracking since most tracking abnormalities occur during dynamic use. Non-operative treatment of idiopathic anterior knee pain is successful in 75 to 90 per cent of patients (Tria *et al*. 1992). Treatment modalities include avoidance of overuse, exercises, orthoses, and anti-inflammatory medications.

Foot pain

Heel pain

The most common causes of heel pain are Achilles tendinitis, retrocalcaneal bursitis, and plantar fasciitis. Achilles tendinitis is generally caused by repetitive trauma and microscopic tears of the tendon at its insertion on the calcaneous (Puddu *et al*. 1976). Occasionally Achilles tendinitis may be seen in patients with seronegative spondylarthropathies without a history of overuse. The usual presentation is a gradual onset of pain with foot push off. Physical examination reveals tenderness and occasionally thickening of the tendon. Rest, anti-inflammatory medication, heel lift gentle stretching exercises, and local heat application are effective treatment measures. The Achilles tendon is vulnerable to rupture in the elderly.

The retrocalcaneal bursa resides between the Achilles tendon and a fat pad posterior to the talus (Fig. 11). Retrocalcaneal bursitis is associated with posterior heel pain which is made worse with passive dorsiflexion of the ankle and may be associated with tender swelling on both sides of the insertion of the tendon (Canoso *et al*. 1984; Frey *et al*. 1992). Causes include repetitive trauma due to athletic activity, rheumatoid arthritis, and all of the seronegative spondylarthropathies (Goldenstein-Schainberg *et al*. 1992; Hernandez *et al*. 1991). Treatment is the same as for Achilles tendinitis, although cautious use of local corticosteroid injection may be necessary in resistant cases.

Plantar fasciitis is attributed to repetitive microtrauma to the attachment of the plantar fascia at the calcaneous producing periostitis and degenerative changes in the origin of the plantar fascia (Karr 1994; Kwong *et al*. 1988). Localized pain with weight bearing on the undersurface of the heel is the typical presentation. Examination shows local tenderness over the anteromedial portion of the plantar surface of the calcaneous, with the symptom worsening on passive dorsiflexion of the toes. Radiographs may show a plantar calcaneal spur. Associated conditions include enthesopathy due to any of the seronegative spondylarthropathies (Gerster 1980). Subcalcaneal or infracalcaneal bursitis may be difficult to distinguish from plantar fasciitis, although passive dorsiflexion of the toes dose not increase symptoms. Treatment consists of a heel pad or cushion, rest, application of local heat, and anti-inflammatory agents. Local injection of corticosteroid at the insertion of the plantar fascia may be useful. The vast majority of patients respond to these conservative measures. Surgical approaches, such as fasciotomy and spur excision, are infrequently required in resistant cases (Karr 1994).

Metatarsalgia

Metatarsalgia is the symptom of pain across the plantar surface of one or more metatarsophalangeal joints and may be due to diverse causes including muscle imbalance, fat pad atrophy, Morton's neuroma, hallux valgus or rigidus, callosities, flat (pes planus) or cavus foot, arthritis of the metatarsophalangeal joints, intermetatarsophalangeal bursitis, tarsal tunnel syndrome, and arterial insufficiency. Symptomatic treatment (for patients with muscle imbalance and fat pad atrophy) includes the use of a metatarsal pad and flexion exercises. Arch supports are recommended for patients with flat or pronated feet.

Hallux valgus and rigidus are common deformities of the great toe which produce pain in the first metatarsophalangeal joint. Hallux valgus may result from excessively narrow foot wear or high heels or from osteoarthritis of the first metatarsophalangeal joint (Caughlin 1984). It is often associated with bursitis over the medial aspect of the metatarsophalangeal joint and, as a result of altered weight-bearing, may result in secondary callosities under the second metatarsophalangeal joint and a hammer toe deformity of the second toe (Schoenhaus and Cohen 1992). Treatment consists of the use of accommodating footwear and a bunion pad. Surgical correction may be required for resistant cases or marked deformities.

Hallux rigidus produces progressive pain and loss of motion in the first metatarsophalangeal joint. The causes are potentially multiple and include osteoarthritis of the first metatarsophalangeal and congenital deformities of the hallux (Mahan 1994). Conservative treatment is designed to reduce the need for first metatarsophalangeal dorsiflexion and consists of wide, stiff soled shoes. Surgical procedures such as the bunionectomy, implant arthroplasties, and first metatarsophalangeal fusion are advocated for severely affected patients (Mahan 1994).

Morton's neuroma is a relatively common cause of foot pain and is caused by neurofibroma or perineural fibrosis of the large superficial branch of the external plantar nerve between the metatarsal heads (Miller 1987). The most common cause is probably repetitive microtrauma of the common digital nerve. Women seem to be affected more often than men, probably because of increased use of tight-fitting footwear. The typical presentation is paraesthesias or dysaesthesias in interdigital web spaces, particularly between the third and fourth interspaces. Pain is increased by weight bearing or with the use of tight-fitting footwear. Examination reveals interspace tenderness with palpation. The diagnosis may be confirmed with an injection of 1 per cent lidocaine into the interspace, which results in almost immediate relief in symptoms. Treatment consists of using low heel and wider width footwear. Local injection of corticosteroids or surgical resection may be required in resistant cases (Miller 1987).

Tarsal tunnel syndrome most often refers to compression of the posterior tibial nerve as it courses around the medial malleolus with the posterior tibial artery and the tibialis posterior, flexor digitorum longus, and flexor hallucis tendons under the flexor retinaculum (Fig. 10). Tarsal tunnel syndrome has diverse potential aetiologies including mechanical factors, such as excessive subtalar joint pronation, tenosynovitis of the tendons accompanying the posterior tibial nerve, synovitis of the ankle joint, and trauma (Grumbine *et al*. 1990; Grabois *et al*. 1981). The typical presentation is with dysaesthesias involving the plantar aspect of the foot that is often more prominent at night. Examination reveals a Tinel's sign with percussion of the flexor retinaculum. Reduced vibratory sensation and decreased two point discrimination may be present on the plantar aspect of the foot and toes. The diagnosis of tarsal tunnel syndrome may be confirmed by nerve conduction studies documenting delay in conduction of the posterior tibial nerve across the ankle (Goodgold *et al*. 1965). MRI may also be useful in the evaluation of the relevant anatomical structures

(Erickson *et al*. 1990). Conservative treatment consists of local corticosteroid injection, use of non-steroidal anti-inflammatory drugs, and orthotic devices. Surgical decompression of the posterior tibial nerve is reserved for resistant cases.

Chapter References

Bouche, R., Sullivan, K., and Ichikawa, D. (1994). Athletic injuries. In *Disorders of the lower extremities. Musculoskeletal* (ed. L. Oloff), pp. 234–59. WB Saunders, Philadelphia.

Broadhurst, N. (1995). Iliopsoas tendinitis and bursitis. *Australian Family Physician*, **24**, 1303.

Canoso, J. (1981). Bursae, tendons, and ligaments. *Clinics in Rheumatic Diseases*, **7**, 189–221.

Canoso, J. and Yood, R. (1979). Reaction of superficial bursae in response to specific disease stimuli. *Arthritis and Rheumatism*, **22**, 1361–4.

Canoso, J., Stack, M., and Brandt, K. (1983). Hyaluronic acid content of deep and subcutaneous bursae of man. *Annals of the Rheumatic Diseases*, **43**, 308–12.

Canoso, J., Wohlgethan, J., Newberg, A., and Goldsmith, M. (1984). Aspiration of the retrocalcaneal bursae. *Annals of the Rheumatic Diseases*, **43**, 308–12.

Caughlin, M. (1984). Hallux valgus: causes, evaluation and treatment. *Postgraduate Medicine*, **75**, 174–87.

Cooper, C. (1995). Epidemiology of osteoarthritis. In *Rheumatology* (eds J. Klippel and P. Dieppe). pp.7:3.1–7:3.4. Mosby, St. Louis.

Erickson, S., Quinn, S., and Kneeland, J. (1990). MR imaging of the tarsal tunnel and related spaces: normal and abnormal findings with anatomic correlation. *American Journal of Radiology*, **155**, 323–8.

Fam, A., Wilson, S., and Holmberg, S. (1982). Ultrasound evaluation of popliteal cysts in osteoarthritis of the knee. *Journal of Rheumatology*, **9**, 428–34.

Fischer, S., Fox, J., Pizzo, W.D., Friedman, M., Snyder, S., and Ferkel, R. (1991). Accuracy of diagnoses from magnetic resonance imaging of the knee. *Journal of Bone and Joint Surgery*, **73A**, 1–10.

Flanagan, F.L., Sant, S., Coughlan, R.J., and O'Connell, D. (1995). Symptomatic enlarged iliopsoas bursae in the presence of a normal plain hip radiograph. *British Journal of Rheumatology*, **34**, 365–9.

Frey, C., Rosenberg, Z., Shereff, M., and Kim, H. (1992). The retrocalcaneal bursa: anatomy and bursography. *Foot and Ankle*, **13**, 203–7.

Galloway, M. and Jokl, P. (1990). Patella plica syndrome. *Annals of Sports Medicine*, **5**: 38–41.

Generini, S. and Matucci-Cerinic, M. (1993). Iliopsoas bursitis in rheumatoid arthritis. *Clinical and Experimantal Rheumatology*, **11**, 549–51.

Gerster, J. (1980). Plantar fasciitis and Achilles tendinitis among 150 cases of seronegative spondarthritis. *Rheumatology and Rehabilitation*, **19**, 218–22.

Gillies, H. and Seligson, D. (1979). Precision in the diagnosis of meniscal lesions: a comparison of clinical evaluation, arthrography and arthroscopy. *Journal of Bone and Joint Surgery*, **61A**, 343–6.

Goldenstein-Schainberg, C., Homsi, C., Rodrigues Pereira, R.M., and Cossermelli, W. (1992). Retrocalcaneal bursitis in juvenile chronic arthritis. *Annals of the Rheumatic Diseases*, **51**, 1162–3.

Goodgold, J., Kopell, H., and Spielholz, N. (1965). Tarsal tunnel syndrome: objective diagnostic criteria. *New England Journal of Medicine*, **273**, 742–5.

Grabois, M., Puentres, J., and Lidsky, M. (1981). Tarsal tunnel syndrome in rheumatoid arthritis. *Archives of Physical and Medical Rehabilitation*, **62**, 401–3.

Griffiths, I. and Pinder, I. (1981). Chondromalacia patella: a clinical and arthroscopic study. *Annals of the Rheumatic Diseases*, **40**, 617–25.

Grumbine, N., Radovic, P., Parsons, R., and Scheinin, B. (1990). Tarsal tunnel syndrome: comprehensive review of 87 cases. *Journal of the American Podiatry Association*, **80**, 457–61.

Hench, P., Reid, R., and Reames, P. (1966). Dissecting popliteal cyst simulating thrombophlebitis. *Annals of Internal Medicine*, **64**, 1259–64.

Hernandez, P.A,, Hernandez, W.A., and Hernandez, A. (1991). Clinical aspects of bursae and tendon sheaths of the foot. *Journal of the American Podiatry Association*, **81**, 366–72.

Ho, G. and Su, E. (1981). Antibiotic therapy of septic bursitis. *Arthritis and Rheumatism*, **24**, 905–11.

Ho, G. and Ticer, A. (1979). Comparison of nonseptic and septic bursitis. *Archives of Internal Medicine*, **139**, 1269–73.

Ho, G., Tice, A., and Kaplan, S. (1978). Septic bursitis in the prepatellar and olecranon bursae. *Annals of Internal Medicine*, **89**, 21–7.

Karr, S. (1994). Subcalcaneal heel pain. *Orthopedic Clinics of North America*, **25**, 161–75.

Kelly, M. and Insall, J. (1992). Historical perspectives of chondromalacia patellae. *Orthopedic Clinics of North America*, **23**, 517–21.

Kerr, D.R. (1993). Prepatellar and olecranon arthroscopic bursectomy. *Clinical Sports Medicine*, **12**, 137–42.

Kraag, G., Thevathasan, E., Gordon, D., and Walker, I. (1976). The hemorrhagic crescent sign of acute synovial rupture. *Annals of Internal Medicine*, **85**, 477–8.

Kujala, U., Kvist, M., and Heinonen, O. (1985). Osgood-Schlatter's disease in adolescent atheletes: retrospective study of incidence and duration. *American Journal of Sports Medicine*, **13**, 236–41.

Kwong, P., Kay, D., and Voner, R. (1988). Plantar fasciitis: mechanics and pathomechanics of treatment. *Clinical Sports Medicine*, **7**, 119–26.

Lambie, P., Kaufman, R., and Beardmore, T. (1989). Septic ischial bursitis in systemic lupus erythematosus presenting as a perirectal mass. *Journal of Rheumatology*, **16**, 1497–9.

Larsson, L. and Baum, J. (1985). The syndrome of anserina bursitis: an overlooked diagnosis. *Arthritis and Rheumatism*, **28**, 1062–5.

Larsson, L. and Baum J. (1986). The syndromes of bursitis. *Bulletin of Rheumatic Diseases*, **36**, 1–8.

Lavyne, M., Voorhies, R., and Coll, R. (1982). Femoral neuropathy caused by an iliopsoas bursal cyst. *Journal of Neurosurgery*, **56**, 584–6.

Letourneau, L., Dessureault, M., and Carette, S. (1991). Rheumatoid iliopsoas bursitis presenting as unilateral femoral nerve palsy. *Journal of Rheumatology*, **18**, 462–3.

Mahan, K. (1994). Joint preservation techniques in hallux limitus/rigidus repair. In *Musculoskeletal disorders of the lower extremities* (ed. L. Oloff). WB Saunders, Philadelphia.

Manueddu, C.A., Hoogewoud, H.M., Balague, F., and Waldeburger, M. (1991). Infective iliopsoas bursitis. A case report. *International Orthopaedics*, **15**, 135–7.

Martens, M., Librecht, P., and Burssens, A. (1989). Surgical treatment of the iliotibial band friction syndrome. *American Journal of Sports Medicine*, **17**, 651–4.

McDermott, M., O'Connell, B., Mulvihill, T.E., and Sweeney, E.C. (1994). Chronic brucella infection of the supra-patellar bursa with sinus formation. *Journal of Clinical Pathology*, **47**, 764–6.

Medlar, R. and Lyne, E. (1978). Sinding-Larsen-Johansson disease. *Journal of Bone and Joint Surgery*, **60**, 1113–9.

Meys, E., Michaux, L., Lambert, M., and Triki, R., Nagant de Deuxchaisnes, C. (1992). Septic Streptococcus milleri prepatellar bursitis. *Clinical Rheumatology*, **11**, 109–11.

Miller, S. (1987). Morton's neuroma: a syndrome. In *Comprehensive textbook of foot surgery* (ed. E. McGlamry), pp. 38–57. Williams and Wilkins, Baltimore.

Morris, J., Berenter, R., and Kosai, D. (1994). Biomechanics of musculoskeletal diseases. In *Musculoskeletal disorders of the lower extremities* (ed. L. Oloff), pp. 65–82. WB Saunders, Philadelphia.

Nakano, K. (1978). Entrapment neuropathy from Baker's cyst. *Journal of the American Medical Association*, **239**, 135.

O'Dwyer, K. and Peace, P. (1988). The plica syndrome. *Injury*, **19**, 350–61.

Ogden, J. and Southwick, W. (1976). Osgood-Schlatter's disease and tibial tuberosity development. *Clinical Orthopedics*, **116**, 180–6.

Patel, D. (1978). Arthroscopy of the plicae: synovia folds and their significance. *American Journal of Sports Medicine*, **6**, 217–28.

Puddu, G., Ippolito, E., and Postacchini, F. (1976). A classification of Achilles tendon disease. *American Journal of Sports Medicine*, **4**, 145–51.

Raddatz, D., Hoffman, G., and Franck, W. (1987). Septic bursitis: presentation, treatment and prognosis. *Journal of Rheumatology*, **14**, 1160–3.

Raskin, R. and Lawless, O. (1982). Articular and soft tissue abnormalities in a 'normal' population. *Journal of Rheumatology*, **9**, 284–8.

Rauschning, W. (1980). Anatomy and function of the communication between knee joint and popliteal bursae. *Annals of the Rheumatic Diseases*, **39**, 354–8.

Reicher, M., Hartzman, S., Duckwiler, G., Bassett, L., Anderson, L., and Gold, R. (1986). Meniscal injuries: detection using MR imaging. *Radiology*, **159**, 753–7.

Sarrafian, S. (1983). *Anatomy of the foot and ankle: descriptive, topographic, functional*. Lipincott, Phildelphia.

Savarese, R.P., Kaplan, S.M., Calligaro, K.D., and DeLaurentis, D.A. (1991). Iliopectineal bursitis: an unusual cause of iliofemoral vein compression. *Journal of Vascular Surgery*, **13**, 725–7.

Schoenhaus, H. and Cohen, R. (1992). Etiology of the bunion. *Journal of Foot Surgery*, **31**, 25–9.

Simonsen, O., Jensen, J., Mouritsen, P., and Lauritzen, J. (1984). The accuracy of clinical examination of injury of the knee joint. *Injury*; **16**, 96–101.

Smith, T., Shawe, D., Crawley, J., and Gumpel, J. (1988). Use of single photon emission computed tomography (SPECT) to study the distribution of Yttrium-90 in patients with Baker's cysts and persisitent synovitis of the knee. *Annals of the Rheumatic Diseases*, **47**, 553–8.

Sutker, A. (1981). Iliotibial band syndrome in distance runners. *Physical Sports Medicine*, **9**, 69–73.

Taylor, P. (1989). Inflammation of the deep infrapatellar bursa of the knee. *Arthritis and Rheumatism*, **32**, 1312–14.

Tindel, N. and Nisonson, B. (1992). The plica syndrome. *Orthopedic Clinics of North America*, **23**, 613–18.

Tria, A., Palumbo, R., and Alicea, J. (1992). Conservative care for patellofemoral pain. *Orthopedic Clinics of North America*, **23**, 545–54.

Voorneveld, C., Arenson, A., and Fam, A. (1989). Anserine bursal distention: diagnosis by ultrasonography and computed tomography. *Arthritis and Rheumatism*, **32**, 1335–8.

Watt, I. (1991). Magnetic resonance imaging in orthopaedics. *Journal of Bone and Joint Surgery*, **73-B**, 539–50.

Wigley, R. (1982). Popliteal cysts: variations on a theme of Baker. *Seminars in Arthritis and Rheumatism*, **12**, 1–10.

Zeiss, J., Coombs, R.J., Booth, R.L. Jr, and Saddemi, S.R. (1993). Chronic bursitis presenting as a mass in the pes anserine bursa: MR diagnosis. *Journal of Computer Assisted Tomography*, **17**, 137–40.

1.2.2.3 The arm and leg in children

Patience H. White

Musculoskeletal problems are frequent in childhood but incidence and prevalence data are incomplete. For example musculoskeletal complaints are the second most common complaint after headaches recorded in a questionnaire survey of school-aged children (Oster and Nielson 1972). In the National Center for Health Statistics questionnaire, the prevalence of arthritis and rheumatic complaints was second only to epilepsy at 2.2/1000 children under the age of 16 years (National Center for Health Statistics 1976). Also, in a Finnish prospective study of the incidence of arthralgias and arthritis documented by paediatric rheumatologists, the overall incidence of musculoskeletal conditions was 109/100 000, with transient synovitis of the hip being the most common at 51.9/100 000 children less than 16 years of age (Kunnamo *et al*. 1986).

In considering a child with a musculoskeletal complaint, the examiner must clarify first if the problem is local or part of a systemic process and, if local, whether it involves the periarticular structures (tendons, ligaments, bursas), intra-articular structures, or is due to referred pain from another source (Table 1). This chapter will focus on the diagnosis of regional problems and touch on only a few of the entities which cause generalized aches and pains that can present as a localized process seen in the paediatric age group. More generalized processes, such as infections of bones and joints, connective tissue diseases (e.g. juvenile arthritis), postinfectious processes (e.g. rheumatic fever), and tumours, are covered in detail in other chapters but must always be considered in the differential diagnosis of regional pain syndromes (Table 2).

Local versus systemic aetiology
Location of pain
Periarticular
Intra-articular
Referred

Table 1 Approach to the child with a regional musculoskeletal complaint

Trauma
Local mechanical disorders
Infection
Connective tissue diseases
Tumours
Metabolic disorders
Genetic abnormalities
Vascular disorders
Generalized aching conditions

Table 2 Major causes of regional pain syndromes peculiar to children

There are multiple causes of regional musculoskeletal pain which include both serious conditions, such as infections and malignancy, as well as self-limited problems such as bursitis and tendonitis. Thus, in any one area of the body, the differential diagnosis of a pain syndrome can be extensive. To guide the examiner to the most likely possibility, it is important to clarify the following considerations; the age of the child at the onset of the problem, history of trauma, constitutional symptoms, the severity, location and quality of pain, and if there is any family history of similar problems.

The age of the child at the time the painful condition started will help focus on the possible causes. For instance, certain conditions are seen most commonly in particular age groups; for example hip pain presenting in a 5 to 10 year old is likely to be Legg–Calve–Perthes' disease, whereas similar pain in an adolescent is most likely to be due to a slipped capital femoral epiphysis. A common cause of regional musculoskeletal pain is trauma. Other physical signs of injury such as bruising and swelling often direct the assessment towards a search for a fracture, tendon injury, or tear. Constitutional complaints and signs, such as fever, anorexia, malaise, and weight loss, in association with a regional pain syndrome draw attention towards the possibility of infection, malignancy, or other systemic disease. Recurrent, predictable symptoms following a specific activity suggest local ligamentous or bursal inflammation or internal derangement of the joint, and may be seen more often in active children with hypermobile joints. Equally important is the intensity, timing, and location of the pain because this information may help define its cause. Severe, constant pain with constitutional symptoms brings infection and malignancy to the top of the differential diagnostic list. Milder pain or no pain associated with upper or lower extremity disease (e.g. limp) is often seen in chronic arthritis due to juvenile arthritis or aseptic necrosis. Morning pain associated with morning stiffness is highly suggestive of a chronic inflammatory condition such as juvenile arthritis, whereas pain occurring at night without constitutional symptoms may be due to an osteoid osteoma or growing pains. As mentioned earlier, the physician should attempt to discern the anatomical location of the pain syndrome. The examiner often needs to know the local anatomy to determine if the painful process is periarticular, intra-articular, or referred from another area. Periarticular pain is often worse at the end of the day and occasionally keeps the child awake at night. Last, but not least, inquiry into family history should be made. Known inherited diseases in families very often give a clue to the child's current, local musculoskeletal problem, such as sickle-cell disease or haemophilia.

An accurate physical examination to localize the problem is a key to finding the cause of the pain. Examine the areas for local tenderness, swelling, range of motion, and muscle size and symmetry. Also look for growth abnormalities that can occur in chronic conditions. On physical examination, periarticular disorders or referred pain syndromes, in contrast to intra-articular disorders, will often result in localized tenderness and the active range of motion of the joint will be less than the passive range of motion (Klaiman and Gerber 1996).

The arm (Table 3)

Shoulder
Subdeltoid bursitis
Bicipital tendinitis
Elbow
Nursemaid's elbow
'Little-league' elbow
Panner's disease
Olecranon bursitis
Wrist and hand
Trigger thumbs
De Quervain's tenosynovitis
Carpal tunnel syndrome
Diabetic cheiroarthropathy
Reflex sympathetic dystrophy (algodystrophy)

Table 3 Arm regional pain syndromes seen in children

Shoulder complaints can resemble those seen in adults. In particular, subdeltoid bursitis and bicipital tendinitis are also seen in children. Anatomically the subdeltoid bursa lies next to the supraspinatus tendon making it difficult clinically to separate out which structure is inflamed. Both result in pain in the lateral aspect of the shoulder worsened by raising the child's hand above his or her head or resisting adduction of the shoulder. Bicipital tendonitis causes pain in the anterior aspect of the shoulder on resisted forward flexion of the shoulder. Pain can also be brought on by having the child supinate the flexed forearm against the examiner's resisting hand. This manoeuvre is called Yergason's sign.

In the elbow children present with particular problems not seen in adults. A common injury in the preschool-aged child is the nursemaid's elbow, or pulled elbow, a result of a strong pull on the forearm or wrist which tears the annular ligament surrounding the radial neck. The child presents with a painful flexed and pronated elbow and the radiographic examination is normal. 'Little-league' elbow is most commonly seen in baseball pitchers between 9 and 13 years of age and is most likely due to an overuse syndrome. Pain and swelling may be localized to the medial, lateral, or posterior aspect of the elbow region. The most common site is in the medial epicondyle; fragmentation and/or calcification in the soft tissues is sometimes visible on the plain radiograph. Tennis elbow can also occur in middle to late childhood.

Avascular necrosis can occur in childhood in multiple locations. Panner's disease is avascular necrosis of the capitellum of the humerus; it occurs most often in middle or late childhood. The child may complain of pain and have swelling around the elbow with slight loss of motion, but the diagnosis is made by the sclerosis and irregularity of the capitellum on radiography. Comparison views of the opposite elbow are helpful in diagnosing mild instances (Woodward and Bianco 1975). Olecranon bursitis is rarer than prepatellar bursitis of the knee because the olecranon bursas develops between the ages of 7 and 10 years and the prepatella bursa develops at a younger age. Pain and swelling of the olecranon bursa can be secondary to infection or idiopathic inflammation found in students or sports enthusiasts such as gymnasts, weight lifters, or wrestlers (Zimmerman *et al*. 1995).

Regional pain disorders of the wrist and hand are often secondary to fractures in childhood. Fractures of the distal forearm are very common in children and fractures of the carpal bones are rare. The most common carpal fracture is a fracture of the navicular and it is seen in adolescents. The trigger thumb seen in children mimics trigger finger in the adult in its presenting signs and symptoms as do de Quervain's stenosing tenosynovitis and carpal tunnel syndrome. It should be remembered that tendonitis of the wrist in an adolescent may be the first sign of disseminated gonoccocaemia. Diabetic cheiroarthropathy, though rare in diabetic adults, is commonly seen in adolescents with onset of diabetes in early childhood. The adolescent presents with painless thickening and tightening of the soft tissues around the fingers resulting in progressive stiffness and flexion contractures, first of the distal interphalangeal joints and the proximal interphalangeal joints (Sergic *et al*. 1976). Another disorder, reflex sympathetic dystrophy (algodystrophy), can occur in the upper or lower extremity. It is discussed below as it may occur more commonly in the lower extremity in children.

The leg (Table 4)

Generalized leg pains
Growing pains
Shin splints
Stress fractures
Leg length inequality
Hip disorders
Transient synovitis
Legg–Calvé–Perthes' disease
Slipped capital femoral epiphysis
Congenital dislocation
Developmental hip dysplasia
Knee disorders
Patella femoral alignment disorders
Plica
Fat pad syndrome
Osgood–Schlatter disease
Internal derangements
Osteochondritis dissecans
Blount's disease
Osgood–Schlatter disease
Baker's cyst
Prepatellar bursitis
Foot and ankle disorders
Pes planus
Stress fracture
Aseptic necrosis
Accessory navicular
Talus dysplasia
Avascular necrosis
Tarsal coalition

Table 4 Leg regional pain syndromes

Generalized leg pains are very frequent in childhood and the most common causes are growing pains, shin splints, and stress fractures. Growing pain occurs in approximately 10 to 20 per cent of school-aged children and young adolescents. It has an aching or sometimes crampy quality and is usually localized to both lower extremities, most often in the thigh, shin, or calf. The pain is never associated with a limp and often occurs in the evening and can interrupt sleep. The child is completely normal by the morning and all laboratory and radiographic investigations are normal. The term 'growing pains' is a misnomer; the aetiology of this disorder is unknown and it occurs after the time of most rapid growth (Peterson 1986).

Shin splints result in pain after activity that is relieved by rest. The pain occurs along the medial or lateral tibial shaft and is thought to be due to microtears of muscle, tendons, and ligaments, muscle inflammation, periostitis, and muscle compartment syndromes. A similar pain syndrome is found with stress fractures. These occur anywhere in the weight-bearing skeleton but are most common in the diaphysial region of the tibia. They present with localized pain that worsens with activity. Plain radiographs will often be negative until new bone formation occurs, which may take at least 2 weeks. A technetium bone scan can be positive much earlier and will differentiate this condition from shin splints which should have negative radiographic studies (Devas 1963).

Unequal leg length can be an important sign of a regional musculoskeletal disorder. Leg length inequality is defined as greater than 2-cm difference between the length of the legs and can result in a limp, scoliosis, and leg or back pain. The leg length discrepancy can be due to overgrowth early in childhood or early epiphyseal closure in adolescence. Some causes are hemihypertrophy, fractures, infections, tumours, intra-articular abnormalities in the hip, knee, or feet (such as juvenile arthritis or avascular necrosis), and neurological disorders such as poliomyelitis (Siffert 1987; Rothenberg 1988). Accurate measurements of the true or apparent discrepancy in the leg length can be difficult and often the best determination is done by plain radiographs of the pelvis centred at the level of the femoral heads when the child is standing.

Regional pain syndromes are much more common in the lower than in the upper extremities. The most common presentation is that of a limp (see Chapter 1.1.2). Abnormalities in the pelvis or back such as discitis or spinal cord lesions can refer pain to the lower extremities. Similarly, hip inflammation may refer pain to the knee. If, on examining the painful area, no focal abnormalities are found, a referral source for the pain should be suspected (Rose and Doughty 1991).

Hip pain

Pain from the hip region is usually localized to the anterior groin, around the greater trochanter, and down the anterolateral thigh to the knee. The differential diagnosis of hip pain includes tumours of the pelvis, spine, or proximal femur, infections of the hip joint, discitis, congenital dislocation, Legg–Calvé–Perthes' disease,

slipped capital femoral epiphysis, and toxic or transient synovitis of the hip.

The most common disorder causing hip pain in children is transient synovitis of the hip. It occurs between the ages of 2 and 12 years, is a self-limited, unilateral condition, and is frequently preceded by a mild upper respiratory infection. Children can have mild hip or knee pain or can present with a painless limp. The hip motion is painful and is limited in internal rotation and abduction. The circumference of the thigh in the affected limb may be decreased. Occasionally a fever is present. The complete blood count and sedimentation rate are usually normal but occasionally the latter can be elevated. A technetium bone scan shows a mild diffuse uptake over the affected hip joint (Wingstrand 1986). The mainstay of treatment is restriction of activities which ranges from axillary crutches to bed rest with traction if the pain is severe. The majority of children with toxic synovitis are better in 2 to 3 weeks. Toxic synovitis is a diagnosis of exclusion and other more serious diagnoses, such as aseptic necrosis, Lyme disease, or poststreptococcal arthritis, must be ruled out.

Legg-Calve-Perthes' disease is avascular necrosis of the femoral head and is seen in children between 3 and 12 years of age with the ratio of boys to girls between 4 and 5:1. The incidence is 1:20 000 and 20 per cent of cases are familial. Both hips are affected in 10 to 18 per cent of children. Almost 90 per cent of children have retarded bone age and hip pain with a loss of hip motion, especially in rotation and abduction. Radiographic findings depend on the stage of the subchondral fracture (Fig. 1). Before the appearance of bony changes on plain radiography, a bone scan will be positive with increased uptake in the femoral head (McCarthy 1988). There is considerable debate, not only about what is the best method of treatment, but also what is the best classification system to identify the stages and severity of the disease. In general, if the child is less than 5 years old at onset and had minimal involvement of the femoral head by any classification system, they are more likely to have a good outcome. Treatment aims at improving the range of motion of the affected hip and containing the hip within the acetabulum by either non-operative or operative means (Wenger *et al*. 1991).

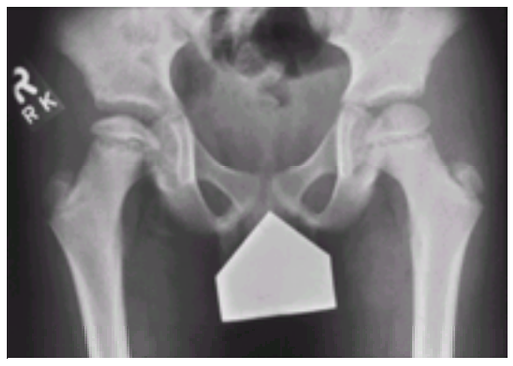


Fig. 1 Legg–Calvé–Perthes' disease—8-year-old child with flattened dense right femoral head characteristic of femoral capital collapse from ischemic necrosis or Legg–Calvé–Perthes' disease.

Slipped capital femoral epiphysis is the gradual or abrupt slippage of the femoral neck anteriorly and superiorly from the femoral head through the growth plate. The most common age of onset is between 8 and 16 years and boys are more likely to develop it than girls in a ratio of 2.2:1. Most cases are unilateral but up to one-third will be bilateral. Approximately 50 to 80 per cent of the affected individuals are obese. Most children have painful weight bearing and have limited range of motion. A pathognomonic sign for this condition is seen when flexing of the involved hip will produce a concomitant external rotation of the hip. Slipped capital femoral epiphysis is associated with endocrine disorders. The diagnosis can be missed unless both anteroposterior and lateral radiographic views are taken as a mild slip can be missed on an anteroposterior view (Busch and Morrisey 1987; Crawford 1988) (Fig. 2). Treatment of a slipped capital femoral epiphysis in the early stages is aimed at preventing further slip and correction of the deformity while minimizing the risk of avascular necrosis and chondrolysis. If the slip of the femoral head is minimal (less than 1 cm in the lateral projection), surgical stabilization is the best treatment. It is more difficult to find agreement in the literature as to the best surgical approach if the slip is chronic and greater than 1 cm.

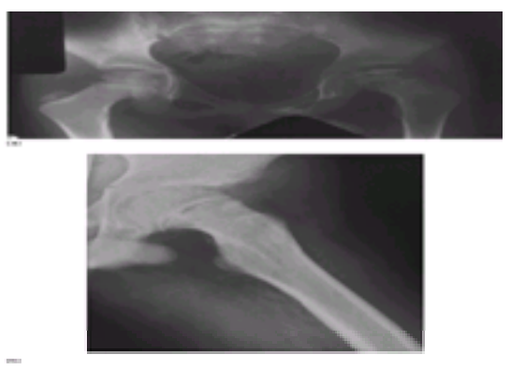


Fig. 2 Slipped capital femoral epiphysis—12-year-old, overweight boy. In the anteroposterior view (a) note the widening of the growth plate and blurring of the metaphyseal side of the growth plate of the left femoral head. The abduction/externally rotated view (b) or frog leg view shows that the femoral capital epiphysis appears to have slipped medially and posteriorly.

Congenital dislocation and subluxation of the hip are quite common with a prevalence of 1.5 per 1000 live births. Girls are affected eight times more commonly than boys and the dislocation is bilateral in 50 per cent. It is important to make the diagnosis early in the new-born period so as to lessen the long-term sequelae, which often result in a total hip replacement. Barlow's manoeuvre of flexion, adduction, and axial pressure in a posterior direction can demonstrate dislocation occurring in the hip. Another physical sign which is significant at any age is shortening of the affected femur. This can be detected when both hips are flexed to 90° with the child lying down on his or her back; the knee on the side of the affected hip is lower than its normal counterpart. If children are not diagnosed until a later age, they present with painful weight bearing and limited range of motion in the affected hip. Radiographic studies may not be diagnostic and the examiner must rely on the physical examination to make the diagnosis (Hensinger 1979).

As in adults, trochanteric bursitis or inflammation of the fascia lata overlying the greater trochanter occurs in children. The pain is localized over the greater trochanter and is worse with walking or lying on the affected side. The pain can be reproduced if pressure is applied to the lateral aspect of the greater trochanter or the fascia lata is put on stretch.

Knee disorders

Patellofemoral disorders of the knee are most commonly seen in adolescence to young adulthood. The disorder is thought to result from malalignment and/or instability of the patellofemoral joint. Symptoms include dull or aching knee pain most commonly with activity such as climbing stairs or squatting and, when the symptoms are severe, there may be intermittent giving way or buckling of the knee due to quadriceps weakness. Locking of the knee is more consistent with an internal derangement or meniscal disorder of the knee. Pain can be precipitated if the patella is compressed or restrained when the quadriceps is contracting. Atrophy of the vastus medialis, patella alta, increased genua valgus, femoral anteversion, and hypermobility are also seen (Lonner and Harwin 1996). Laboratory tests are normal. Radiographs are normal unless there is significant malalignment. Arthroscopic examination of the knee reveals fissuring, fibrillation, and fragmentation of the retropatellar anterior cartilage (Radin 1984; Fulkerson and Shea 1990). Mild patellofemoral pain can be improved by a programme of isometric exercises, designed both to stretch the tight iliotibial band and to strengthen the hamstrings and quadriceps muscle. If exercises fail, surgery must be considered.

A plica is an abnormal synovial band that originates in the suprapatellar area and extends down to the medial side of the knee. It can become trapped between the

tibia and femur resulting in pain, giving way, and locking of the knee. The condition is often diagnosed at arthrography (Reid *et al*. 1980).

The fat-pad syndrome or Hoffa's disease is characterized by intermittent pain and swelling in the anterior knee joint underneath the patellar tendon. Singling–Larsen–Johanssen syndrome occurs in adolescence; it results in pain at the insertion of the patellar tendon where it attaches to the patella and is worsened by activity. Radiographs show calcification at the insertion of the patellar tendon.

Internal derangement of the knee can be secondary to a meniscal tear, torn cruciate ligament, or loose body within the knee joint such as synovial chondromatosis. A child with a meniscal tear presents with a significant past history of trauma. No matter what the aetiology, locking, giving way, pain, and intermittent swelling are found in association with the presence of foreign bodies within the joint space (Renshaw 1986).

Osteochondritis of the knee in children occurs most commonly in the medial femoral condyle. Its prevalence peaks in late childhood and adolescence and it is more common in boys. In about 10 per cent of cases the disease is bilateral. Early in the condition, when the articular cartilage is intact, symptoms consist of vague aching after activity and occasional 'giving-out' of the knee. Localized tenderness of the condyle palpated when the knee is flexed can be the only physical finding. At late stages (Fig. 3), when the fragment can partially separate from the condyle, symptoms such as limping, swelling, and locking occur. Radiographs demonstrate that the subchondral bone of the fragment is separated from the condyle by a thin radiolucent line or the fragment may be displaced (Renshaw 1986). If the articular cartilage is intact, the treatment consists of restricting activity to the child's tolerance level. Healing can take up to several months. If it is possible that the fragment is loose, arthroscopy to remove small pieces or surgery to replace significant bony components should be undertaken. Genu varum (bowed legs) is normal in children under age 2 years. The most common pathological condition that can cause genu varum is Blount's disease. Blount's disease is a disorder of the posterior or the medial tibial physis and those most commonly affected are black males. Radiographs show diminished height of the medial tibial plateau (Fig. 4).

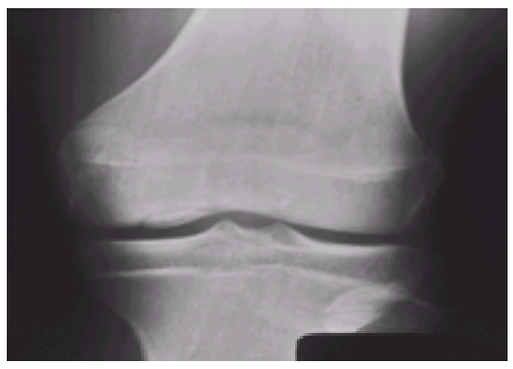


Fig. 3 Osteochondritis dissecans—12-year-old adolescent with a dense ovoid subchondral body on the medial femoral chondyle next to the interchondylar notch.

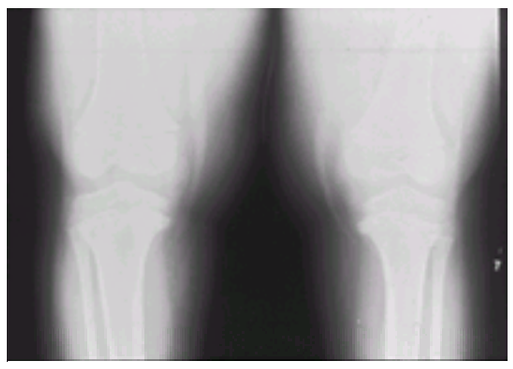


Fig. 4 Blount's disease—4-year-old child with a varus configuration of the right knee. The height of the medial tibial plateau appears to be diminished with sclerosis of the medial tibial metaphysis.

Osgood–Schlatter disease, osteochondritis of the tibial tubercle, is thought to be an overuse syndrome involving the attachment of the patellar tendon to the tibial tubercle and is often caused by recurrent microtrauma. This condition is seen in preadolescence and adolescence and is more common in males. There is pain and swelling around the tibial tubercle which is worsened by strenuous physical activity. Radiographs may be normal in the early stages but can show slight avulsion or fragmentation of the tubercle or both (Renshaw 1986, p. 110). This condition is usually mild and often asymptomatic. If the pain is significant or recurrent, restriction of activity or immobilization of the knee may be necessary.

Another regional condition is a Baker's or popliteal cyst. This is a cystic mass filled with synovial fluid or gelatinous fluid resembling that in ganglion cysts of the wrist. This cystic mass can arise from the joint itself or from the semimembranosus bursas. A cyst in a child is usually a self-limited condition, resolving spontaneously over several months. In contrast to Baker's cysts seen in adults, Baker's cysts in children are not usually associated with intra-articular pathology, and thus investigation of the knee of an asymptomatic child with a Baker's cyst is often unhelpful (Renshaw 1986, p. 111).

There are two benign tumours that must be considered in the differential of regional musculoskeletal pain of the knee; pigmented villonodular synovitis and synovial chrondromatosis. Both occur in teenagers and affect the knee most commonly. Pigmented villonodular synovitis results in a bloody arthrocentesis and has a characteristic histopathology including a cellular infiltrate of the synovium with haemosiderin-laden stromal cells. Synovial chrondromatosis is a condition in which intrasynovial cartilaginous nodules develop and may project into the joint space, resulting in symptoms of pain and locking. Unless the nodules calcify, this diagnosis must be made by arthroscopy.

Prepatellar bursitis is a common problem of the periarticular region of the knee in adults. It can occur in children, but is much less common than in adults. Because of its superficial location, it is often inflamed secondary to a bacterial infection from a nearby cellulitis.

Ankle and foot problems

Foot pain in young children is rare; it becomes more common as the child grows older. Foot discomfort in children can be divided into the following categories; hypermobile flat feet, trauma, infections, apophysitis, neurological disorders, tumours, and boney abnormalities such as avascular necrosis and tarsal or subtalar coalition.

Flat feet are usual in infants and common in children. It can be described as a loss of the longitudinal arch of the weight-bearing foot joint. Pain is unusual and, if it occurs, it is located in the arch, on the medial side of the ankle, or at the insertion of the plantar fascia at the metatarsal heads or calcaneus. If pain is present it is important to distinguish between a flexible planovalgus (flat) foot and a peroneal spastic flat foot. A flexible flat foot is supple with full ankle and subtalar joint motion and the arch is restored when one stands on the toes. In the peroneal spastic flat foot, ankle motion is normal but there is limited inversion and eversion of the foot and/or subtalar joint. Also the arch is not restored when not weight bearing. Peroneal spastic flat foot may be caused by a tarsal coalition or by inflammatory arthritis (Craig and Goldberg 1993). Tarsal coalition is the leading cause of peroneal spastic flat foot. Calcaneonavicular coalition is most common and the next most common is the talocalcaneal coalition which is seen most often between the ages of 12 and 16 years and is best demonstrated by computed tomography (**CT**) scan (Fig. 5).

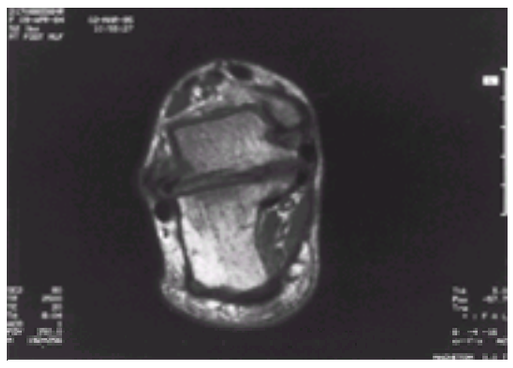


Fig. 5 Subtalar fusion or coalition—a CT scan of the hindfoot of a 12-year-old child demonstrating the lack of definition of the cortex of the subtalar joint on the medial side of the heel.

Flexible flat foot or hypermobile pes planus is a normal part of the development of the foot. At the age of 1 year when the child starts to walk the midfoot is flat to the floor due to normal ligamentous laxity. The foot usually develops the normal longitudinal arch by 5 years. Children with flexible flat feet due to ligamentous laxity form a good arch when standing on tiptoe. No treatment with corrective shoes is thought necessary for the asymptomatic flat foot. Occasionally the hypermobile flat foot may result in pain at the medial side of the arch on weight bearing. Mild genovalgus, slight flexion of the knees and hips, as well as lordosis can occur due to this pes planus and result in lower extremity pain. Often mild symptoms can be improved with a moulded insert.

Another cause of painful flat feet is the hypermobile foot with a short Achille's tendon. The diagnosis is made when dorsiflexion of the ankle with an inverted foot does not reach 90°. Heel-cord stretching exercises are usually helpful. Rarely, an accessory navicular will result in foot pain necessitating its excision. An accessory navicular is found in 50 per cent of all feet; most are asymptomatic.

The most common non-neuromuscular cause of the rigid flat foot in the 8- to 16-year-old age group is tarsal coalition which occurs when two or more of the tarsal bones are joined, either by a bony bar or a fibrocartilaginous bridge. Clinically the adolescent has a rigid flat foot associated occasionally with peroneal spasm. The coalitions are diagnosed by radiography but special views are often needed to demonstrate the coalition, for example an oblique view of the hindfoot is necessary to demonstrate a calcaneonavicular bar.

In adolescents a common cause of foot pain is stress fractures of the second and third metatarsal shafts. Like other stress fractures, they may be difficult to see early in the radiographs. They can either be documented by a technetium bone scan or callus formation on a plain radiograph can be awaited. This takes approximately 3 weeks to form.

There are three tendon–bone junctions in the foot that present problems in preadolescents. Calcaneal apophysitis (Sever's disease) occurs in both girls and boys between the ages of 8 and 10 years, usually in those who are just beginning to play competitive sports. The pain localizes to the insertion of the Achille's tendon where it inserts on to the calcaneus, and adolescents with this condition often have difficulty walking on their heels. Radiographs are normal, often showing a dense fragmented calcaneal apophysis, which is a normal finding at this age. A second condition seen in children aged 10 to 12 years is swelling and pain at the insertion of the peroneus brevis into the apophysis of the base of the fifth metatarsal. There can be a secondary ossification centre in the apophysis and this can cause pain, particularly if narrow shoes are worn. Finally, an accessory navicular bone can result in foot pain. In 75 per cent of cases an accessory navicular will fuse with the main navicular bone. Pain can occur over the accessory navicular and is usually relieved with rest (Wilkings 1988).

Reflex sympathetic dystrophy can occur in the hand but is much more common in the foot. This condition is often initiated by trauma and presents with a diffusely swollen, painful foot. The pain is worse at night and is not relieved by rest. Often the symptoms appear out of proportion to the physical findings. As the condition progresses, the foot can become hypersensitive and demonstrate vascular changes such as being cool, sweaty, and swollen; finally, muscle wasting and contractures can develop if no treatment with physiotherapy is begun. Radiographs can show osteopenia and diffuse increased uptake is found in a technetium bone scan (Forster 1985).

Two painful conditions of the foot are due to avascular necrosis; Kohler's disease and Freiberg's infraction. Kohler's disease occurs in girls and boys between the ages of 4 and 6 years. Pain with exercise slowly develops over the tarsal navicula and radiographs demonstrate a 'pancake' condition of the navicula. Freiberg's infarction is aseptic necrosis of the metatarsal head and most commonly affects the head of the second metatarsal. Very active adolescent girls are most affected and pain is localized to the metatarsal head involved.

There are many tendinous attachments and bursas around the foot and inflammation, such as plantar fasciitis, 'pump bumps' underneath the bursa near the insertion of the Achille's tendon on the calcaneus, and Achille's tendinitis due to type IIa or IV hyperproteinaemia, are often seen in many of these structures. Similarly, bunions or hallux valgus are most common in adolescence but occur in all age groups. Many of these conditions are described elsewhere as they are more common in adults (Wilkings 1988).

Generalized conditions that can present as regional musculoskeletal pain syndromesTable 5)

Hypermobility
Hypomobility
Genetic disorders
Fibromyalgia
Hyperostosis

Table 5 Generalized conditions often presenting as regional pain in children

There are several systemic conditions that can present to the rheumatologist with regional musculoskeletal complaints. It is only after a careful, general examination that the systemic condition may be revealed. One example is hypermobility which occurs in 12 to 19.5 per cent of the population. Joint laxity predisposes periarticular structures to injury resulting in musculoskeletal pain around one or several joints. This painful syndrome is more common in girls with a family history of hypermobility (Biro *et al.* 1983). Pain due to joint laxity is also a feature of other inherited disorders of joint laxity, such as Marfan's syndrome, Ehlers–Danlos syndrome, and Larsen's syndrome.

Disorders that result in acquired hypomobility result in joint contractures and little signs of inflammation are found. The mucopolysaccharidoses and mucolipidoses can present with limited motion of the fingers and, occasionally, other joints. Fabry's disease requires the attention of a rheumatologist because patients with this disease develop degenerative changes and flexion contractures of the fingers. Severe, burning pain in the fingers and toes occurs in 80 per cent of children or young adults with Fabry's disease (Sheth and Bernhard 1979). Gaucher's disease also has hip and knee pain as a component of the syndrome. The Stickler syndrome is the most common connective disease dysplasia in the American midwest. Ankle, knee, and wrists become enlarged in childhood and cause intermittent arthralgias (Liberfarb *et al*. 1981). The epiphyseal dysplasia syndromes can result in pain around or in a joint and a plain radiograph assists in the diagnosis. Another genetic

syndrome that often presents with regional musculoskeletal pain is sickle-cell disease. Synovial, capsular, and tendon infarction can occur in sickle-cell disease resulting in joint pain (Diggs 1967).

Fibromyalgia occurs in the paediatric age group. This condition usually presents as diffuse musculoskeletal aching in a 9 to 15 year old complaining of fatigue. There is a preponderance of girls. The diagnosis is one of exclusion and is based on the demonstration of tender points often found in periarticular locations associated with bursas, epicondyles, and insertion of tendons. Occasionally the child complains of pain in only one or two of these periarticular areas and the cause of the regional pain syndrome is revealed only after looking for multiple tender points (Yanus and Masi 1985; Sherry *et al*. 1991). Myopathies can also present as leg aching or tenderness and should be considered in the differential diagnosis of regional pain syndromes.

Conditions associated with hyperostosis often result in regional pain syndromes. Primary hypertrophic osteoarthropathy is an autosomal dominant disorder seen in adolescent boys with 'spade-like' enlargement of the hands and feet and pain in the distal long bones (Calabro *et al*. 1966). Secondary hypertrophic osteoarthropathy is associated most commonly with malignancy but can occur in pulmonary disease (cystic fibrosis, infections, fibrosis), cardiovascular disease (subacute bacterial endocarditis, cyanotic congenital heart disease), gastrointestinal disease (inflammatory bowel disease, bacterial or parasitic colitis, subphrenic abscess, cirrhosis), and endocrine (thyroid) conditions. Most children present with clubbing of the fingers, periostitis and pain in the long bones, and arthritis of the large joints (Petty *et al*. 1976). Infantile cortical hyperostosis of Caffey's disease is an uncommon disease with hyperostosis in which a child has a symmetrical painful enlargement of the mandible or shoulder girdles or long bones. It is self-limiting and usually occurs before 6 months of age (Prieur 1990).

In conclusion, focal pain syndromes in children and adolescents are common and the physician must decide if the problem is due to a local condition or one that is part of a systemic illness.

Chapter References

Biro, R., Gewanter, H.C., and Baum, J. (1983). The hypermobility syndrome. *Pediatrics*, **72**, 701.

Busch, M.T. and Morrisey, R.T. (1987). Slipped capital femoral epiphysis. *Orthopedic Clinics of North America*, **18**, 637–49.

Calabro, J.E., Marchesano, J.M., and Abruzzo, J.L. (1966). Idiopathic hypertrophic osteoarthropathy (pachydermoperostitis). Onset before puberty. *Arthritis and Rheumatism*, **9**, 496.

Craig, L. and Goldberg, M.J. (1993). Foot and leg problems. *Pediatrics in Review*, **14**, 395–400.

Crawford, A.H. (1988). Current concepts review. Slipped capital femoral epiphysis. *Journal of Bone and Joint Surgery*, **70A**, 1422–7.

Devas, M.B. (1963). Stress fractures in children. *Journal of Bone and Joint Surgery*, **45B**, 528.

Diggs, L.W. (1967). Bone and joint lesion in sickle disease. *Clinical Orthopedics*, **52**, 119–43.

Forster, R.S. (1985). Reflex sympathetic dystrophy in children. *Orthopedics*, **8**, 474–7.

Fulkerson, J.P. and Shea, K.P. (1990). Current concepts review. Disorders of patellofemoral alignment. *Journal of Bone and Joint Surgery*, **72A**, 1424–9.

Hensinger, R.N. (1979). Congenital dislocation of the hip. *Ciba Foundation Symposium*, **31**, No. 1.

Klaiman, M.D. and Gerber, L.H. (1996). General considerations for managing tendon injuries. *Bulletin of Rheumatic Diseases*, **45**, 1–6.

Kunnamo, I., Kallio, P., and Pellconen, P. (1986). Incidence of arthritis in urban Finnish children. *Arthritis and Rheumatism*, **29**, 1232.

Liberfarb, R.M., Hirose, T., and Holmes, L.B. (1981). The Wagner–Stickler syndrome: a study of 22 families. *Journal of Pediatrics*, **99**, 394–9.

Lonner, B.S. and Harwin, S.F. (1996). Chondromalacia patellae: review paper—Part I. *Contemporary Orthopedics*, **32**, 81–6.

McCarthy, R.E. (1988). Avascular necrosis of the femoral head in children. *Instructional Course Lectures*, **37**, 59–65.

National Center for Health Statistics (1976). Prevalence of chronic skin and musculoskeletal conditions. *United States Health Interview Study Series*, 10, No. 124, DHSW (PHA) Pub. No. 79–1552.

Oster, J. and Nielson, A. (1972). Growing pains: a clinical investigation of a school population. *Acta Paediatrica Scandinavica*, **61**, 328–34.

Peterson, H. (1986). Growing pains. *Pediatric Clinics of North America*, **33**, 1365.

Petty, R.E. *et al*. (1976). Secondary hypertrophic osteoarthropathy, an unusual case of arthritis in childhood. *Arthritis and Rheumatism*, **19**, 902.

Prieur, A.M. (1990). Miscellaneous conditions associated with musculoskeletal complaints in children. *Current Opinion in Rheumatism*, **2**, 824–31.

Radin, E.L. (1984). Chondromalacia of the patellae. *Bulletin of Rheumatic Diseases*, **34**, 1.

Reid, G.D. *et al*. (1980). Pathological plica of the knee mistaken for arthritis. *Journal of Rheumatology*, **7**, 573.

Renshaw, T.S. (1986). *Pediatric orthopedics.* Philadelphia, W.B. Saunders.

Rose, C.D. and Doughty, R.A. (1991). Zeroing in on the clause of a child's limp. *Journal of Musculoskeletal Medicine*, **8**, 22–32.

Rothenberg, R.J. (1988). Rheumatic disease aspects of leg length inequality. *Seminars in Arthritis and Rheumatism*, **17**, 196–205.

Sergic, A. *et al*. (1976). Joint contracture—common manifestation of childhood diabetes mellitus. *Journal of Pediatrics*, **88**, 584–8.

Sherry, D.D. *et al*. (1991) Psychosomatic musculoskeletal pain in childhood: clinical and psychological analyses of 100 children. *Pediatrics*, **88**, 1093–8.

Sheth, K.J. and Bernhard, G.C. (1979). The arthropathy of Fabry disease. *Arthritis and Rheumatism*, **22**, 781–3.

Siffert, R.S. (1987). Current concepts review. Lower limb-length discrepancy. *Journal of Bone and Joint Surgery*, **69A**, 1100–6.

Wenger, D.R., Ward, W.T., and Hearing, J.A. (1991). Current concepts review. Legg-Calvé-Perthes disease. *Journal of Bone and Joint Surgery*, **73A**, 778–88.

Wilkings, K.E. (1988). The painful foot in the child. *Instruction Course Lectures*, **37**, 77–85.

Wingstrand, H. (1986). Transient synovitis of the hip in the child. *Acta Orthopaedica Scandinavica Supplementum*, **57**, 219.

Woodward, A.H. and Bainco, A.J. (1975). Osteochondritis diseases of the elbow. *Clinical Orthopedics*, **110**, 35.

Yanus, M.B. and Masi, A.T. (1985). Juvenile primary fibromyalgia syndrome: a clinical study of thirty-three patients and matched controls. *Arthritis and Rheumatism*, **28**, 138.

Zimmerman, B., *et al*. (1995). Septic bursitis. *Seminars in Arthritis and Rheumatism*, **24**, 391–410.

1.2.3 Extra-articular features of rheumatic diseases

Ian D. Griffiths

Introduction

The so-called rheumatic or musculoskeletal disorders encompass a wide variety of disorders, of which many are multisystemic. In some of these disorders the articular features are frequently mild and often inconsequential, for example in primary Sjögren's syndrome or the antiphospholipid syndrome.

This chapter does not attempt to catalogue all the extra-articular features encountered in the rheumatic diseases but addresses the problem of underlying rheumatic diseases that may present in a non-rheumatological fashion. In practice this may occur before the rheumatic features have manifested themselves, when the non-articular features clinically dwarf the rheumatic symptoms, or when the underlying rheumatic disease produces only mild symptoms or subtle clinical joint manifestations.

The most critical factor in establishing the diagnosis at presentation is the competence of the attending doctor in obtaining a history and performing an examination which will elicit features of the underlying musculoskeletal disorder (Doherty *et al*. 1990). Failure to identify a musculoskeletal disorder, or to recognize its relevance in the context of the overall clinical picture, is more than an 'academic' nicety. Most clinicians have experience of patients who have been invasively (and expensively) investigated for symptoms, signs, or laboratory abnormalities which could have been readily accounted for if the underlying diagnosis had been recognized. More alarming are the patients who have been inappropriately treated (e.g. lymph node biopsies which have been misclassified and subsequently treated as lymphoma by the pathologist who was not made aware of the coexisting diagnosis of rheumatoid arthritis).

The development of medical specialization means that undifferentiated diseases may be directed to a wide variety of disciplines depending upon the predominant symptoms at the time. There remains a tendency in standard medical textbooks to under recognize the frequency with which non-specific systemic features, such as fever or weight loss, may be the presenting feature of a rheumatic disorder.

This chapter considers 'extra-articular' presentation of musculoskeletal disorders under three broad headings: general systemic features, specific systems, and clinical patterns.

General systemic features

These are very common in inflammatory joint and connective tissue disorders (Table 1). It is often the combination of features which suggests an underlying cause, for example swinging fever and rash in adult onset Still's disease. Similarly, in the non-inflammatory disorders, the coexistence of problems may point towards a diagnosis, for example lethargy and poor sleep patterns in fibromyalgia. This section considers some of the more common systemic problems.

Table 1 Generalized systemic features which may be the presenting complaint of an underlying rheumatological disorder

Weight loss

Musculoskeletal disorders are hampered by the lack of objective, numerical clinical measurements for management. Weight is a simple measure and a history of weight loss at the onset or during the early stages of several diseases is not uncommon (Fig. 1).

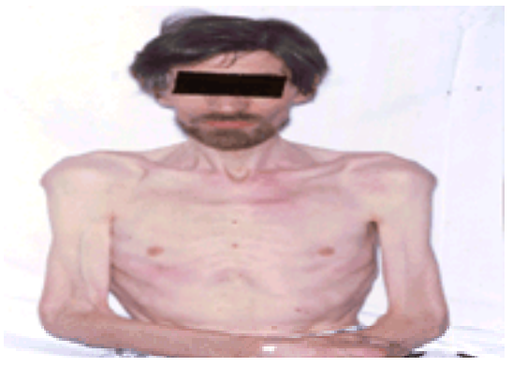


Fig. 1 Profound weight loss in a man with inflammatory polyarthritis, subcutaneous nodules, and low grade persistent pyrexia. The final diagnosis, on histology, was multicentric reticulohistiocytosis.

There is no agreed definition of significant weight loss. In practice, working definitions have included weight loss of 2 kg or more, or a decrease of 10 per cent in body weight. It is rare to have weight loss greater than 10 kg due to arthritis. The weight loss tends to plateau after a few months, unlike that associated with malignancy or endocrine disease. Calculation of Body Mass Index (weight (kg)/ height (m) ^{2}) may offer some standardization but is not routinely undertaken.

In clinical practice, weight loss is most commonly encountered as an early feature of rheumatoid arthritis. It occurs particularly in elderly onset rheumatoid arthritis, being present in 50 per cent of subjects in one series (Terkeltaub *et al*. 1983). Often the articular features are mild and atypical. The weight loss is often associated with other systemic features, such as lymphadenopathy and low grade pyrexia, so suggesting a wide range of diagnostic possibilities, particularly lymphoma. There is the suggestion that reduction of Lean Body Mass in rheumatoid arthritis may be related to tumour necrosis factor-a levels (Roubenoff *et al*. 1992).

However, as may be expected, weight loss is not confined to rheumatoid arthritis and can occur in association with any inflammatory joint or connective tissue disorder. It is particularly in the elderly (if associated with rather non-specific symptoms such as in polymyalgia rheumatica, polyarteritis nodosa, or a related inflammatory vasculitis) that the major diagnostic dilemmas emerge, especially in relationship to underlying neoplasms. The evaluation depends upon detailed clinical, radiological, and laboratory assessment, which should include routine urine testing, full blood count, erythrocyte sedimentation rate (or other measure of acute phase response), biochemical profile (including thyroid function and protein electrophoresis), autoantibodies (including antineutrophil antiplasmic antibodies), and a chest radiograph. Neoplasms which in their turn may mimic musculoskeletal disorders include carcinoma of the bronchus, renal carcinoma, lymphomas, and primary and secondary neoplasms of bone.

Other disorders of the gastrointestinal tract that may present with musculoskeletal features and weight loss include inflammatory bowel disease, Whipple's disease, and adult coeliac disease. It is often the presence of an unusual feature, for example the failure of an iron deficiency anaemia to respond to oral iron, that suggests the diagnosis.

True weight loss in children due to juvenile chronic arthritis or related disorders is extremely uncommon, but a downward shift to lower percentile lines on both weight and height growth curves is frequent in active disease.

Fever

Pyrexia is a common feature of many rheumatic disorders that have an inflammatory component (Fig. 2 and Fig. 3), including septic arthritis, crystal induced arthritis, chronic inflammatory joint disease, and connective tissue disorders (Pinals 1994). In most of these cases the associated rheumatic disorder will be apparent; the major exception is septic arthritis in a young child, where the presentation may be as non-specific as the 'irritable, febrile child'.

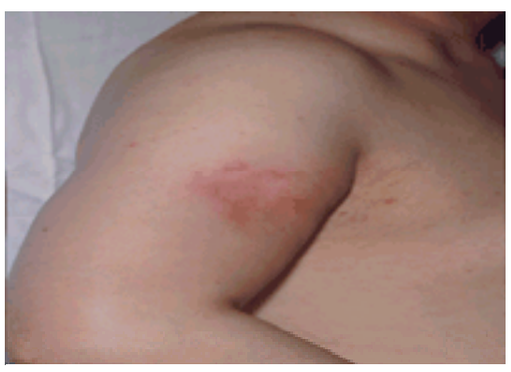


Fig. 2 Man who presented with fever, marked raised erythrocyte sedimentation rate (greater than 100) and right upper arm pain due to a septic arthritis of the right shoulder tracking down the long head of the biceps to present as swelling in the right upper arm.

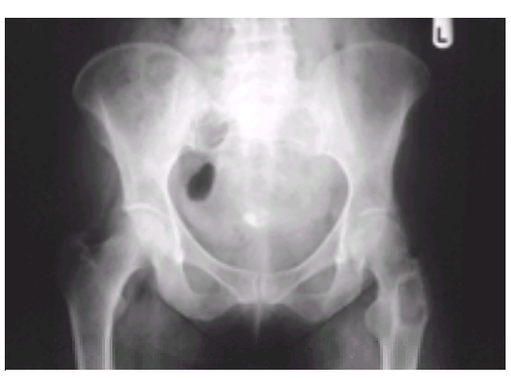


Fig. 3 A 60-year-old woman who presented with left hip pain, weight loss, and low grade pyrexia. Radiographs show erosion in the left greater trochanter and calcified pyometrium. Aspiration and culture of the left greater trochanter bursa revealed *Mycobacterium tuberculosis*.

Fever as an early or presenting feature of a rheumatic disease is more difficult to evaluate. Fever patterns can be considered as:

1. intermittent, when the temperature is swinging but returns to normal between episodes;
2. remittent, when the temperature varies but does not return to normal;
3. sustained, when no variation is occurring.

If no cause for the fever is found after 1 week of investigation, it is usually labelled as fever of unknown origin. A review of patients presenting with persistent fever of unknown origin in the 1980s showed that rheumatological disorders in the wide sense equalled infections as the most common underlying cause, accounting for a third of the cases where a cause was found (Knockaert *et al*. 1992). The most frequent rheumatological disorder was temporal arteritis. A similar finding emerged

when patients with intermittent fever of unknown origin were studied but in that group an underlying disorder was found less frequently (Knockaert *et al*. 1993).

If sepsis has been excluded with reasonable confidence, the intermittent fever pattern with a diurnal variation, with temperatures of 39°C or above occurring in the later afternoon to evening and returning to normal by morning (quotidian), is very suggestive of systemic onset juvenile chronic arthritis in the child or adult Still's disease. The finding of the typical evanescent macular rash at the height of the fever would further support the diagnosis. The fever may precede the arthritis in both juvenile and adult forms of the disease, sometimes by many months, causing diagnostic problems (Martin *et al*. 1994). Approximately 10 per cent of fevers of unknown origin in children are eventually found to be associated with juvenile chronic arthritis.

In adults, there is often a considerable delay before the diagnosis of adult onset Still's is established (Pouchot *et al*. 1991; Wouter and Van de Putte 1986). Several reasons exist for this; the patient often presents as a fever of unknown origin to the general medical department, other 'non-articular' features may dominate the clinical picture, for example deranged liver function tests, or pleuropericardial disease. While no diagnostic tests exist, the absence of a marked leucocytosis, normal acute phase reactants, and the absence of marked elevation of the serum ferritin (>1000 mg/l) all weigh heavily against the diagnosis.

Continuous or remittent fever is an integral feature of some forms of chronic infectious disorders associated with arthritis, for example brucellosis and Lyme disease (O'Connell 1995), and certain forms of vasculitis, for example Kawasaki's syndrome. The more complex situation arises in the patient with a low grade pyrexia, often with a raised erythrocyte sedimentation rate, and in whom infection and neoplasia has been excluded. Rheumatological disorders that can be easily overlooked include Reiter's syndrome, particularly in the younger adult as there tends to be a lack of recognition of the sometimes severe systemic nature of Reiter's. Also, Reiter's syndrome increasingly presents as an 'incomplete' form in which the precipitating event may be overlooked, particularly if it was a transient gastrointestinal disturbance. The clinical finding of an enthesitis, sacroiliitis, or asymptomatic oral ulcers would support the diagnosis.

In the middle-aged to elderly population temporal arteritis, primary Sjögren's syndrome and polymyalgia rheumatica, may present as low grade pyrexia with elevated erythrocyte sedimentation rate, with little evidence of the underlying disorder. Specific enquiry about dry eyes and mouth coupled with a simple bedside screening test (e.g. Schirmer's test) may help and the striking diurnal symptoms of polymyalgia may suggest the diagnosis.

Malaise, lethargy, and fatigue

These symptoms give rise to difficulties in medical definition but are frequent clinical complaints. While these are common and well recognized accompaniments of chronic inflammatory joint and connective tissue disorders, they may be the presenting or sole clinical feature of some disorders which fall within the rheumatological spectrum. Arbitrarily they have been divided on the basis of a normal erythrocyte sedimentation rate and raised erythrocyte sedimentation rate.

Normal erythrocyte sedimentation rate

Fibromyalgia is probably the most common associated problem. The subject is typically a middle-aged woman with a non-restorative sleep pattern and diffuse pains in a girdle distribution with multiple symmetrical tender points. Clinically, overlap seems to exist with other syndromes, where profound tiredness and lethargy may exist in the absence of clear physical or laboratory abnormalities, for example post viral fatigue syndromes, myalgic encephalomyelitis and chronic fatigue syndromes (Goldenberg 1993). Metabolic and endocrine disorders such as periodic hypokalaemic paralysis, Addison's disease, and hypothyroidism need be excluded.

Raised erythrocyte sedimentation rate

Primary Sjögren's syndrome frequently presents under this guise, and lethargy is well recognized as one of the most distressing features of primary Sjögren's. There are no specific features in the history to point to the diagnosis and the patient just complains of intense fatigue. Sleep disturbance may be found but trigger points are not present; otherwise the sex and age distribution is very similar to fibromyalgia. As well as the raised erythrocyte sedimentation rate, polyclonal hypergammaglobulinaemia and serological abnormalities such as positive rheumatoid factor, antinuclear factor, and anti-Ro/La antibodies are often present.

Although less common disorders, profound lethargy with a raised erythrocyte sedimentation rate may be the presenting features of bacterial endocarditis, polyarteritis (and the other vasculitides), cryoglobulinaemia, Waldenström's macroglobulinaemia, polymyalgia, and a wide variety of malignancies of which lymphoma, bronchogenic neoplasm, and renal cell carcinoma are probably the most commonly encountered as presenting under a rheumatological guise.

Lymphadenopathy

This alarming clinical feature may occur as a presenting one in both children and adults.

In children, generalized lymphadenopathy is usual in systemic juvenile chronic arthritis and Kawasaki's disease (Bissenden and Hall 1990). It has to be remembered, however, that leukaemia in children may also present as polyarthralgia or a true polyarthritis.

In the adult, lymphadenopathy is common in rheumatoid arthritis and the connective tissue disorder, and may occur early in the disease. One study reviewing patients who had undergone lymph node biopsies for unexplained lymphadenopathy and in which the histology had shown reactive hyperplasia, found that one-third of them developed a connective tissue disorder within a year of biopsy. A more common clinical problem arises when marked lymphadenopathy develops in a patient with connective tissue disease or rheumatoid arthritis. If, on clinical grounds, it is felt desirable to have a lymph node biopsy to exclude neoplastic disease, then it is important that the histopathologist is informed about the associated connective tissue disorder, as the histological differential diagnosis of marked reactive hyperplasia and giant follicular cell lymphoma may be difficult without immunohistological analysis (Kelly *et al*. 1987).

Persistent lymphadenopathy above the clavicles usually causes greatest clinical concern. A review of 13 patients with a variety of rheumatological disorders, who had lymph node biopsies of either supraclavicular or cervical lymph nodes because of persisting lymphadenopathy, failed to detect changes other than 'reactive' hyperplasia in any of the biopsies (Kelly *et al*. 1987).

Specific systems

This section considers some individual systems and their disorders that may be the presenting feature of an underlying rheumatological disorder. It cannot be completely comprehensive as anecdotal reports exist about virtually all known medical disorders being associated with some rheumatological conditions (e.g. systemic lupus erythematosus). It attempts to identify those non-articular 'rheumatological' disorders which are commonly seen either prior to or in a very early stage of the rheumatic disease process (Table 2).

Table 2 Specific system involvement that may be the presenting feature of an underlying rheumatological disorder

Mucocutaneous systems

Ocular system

In practice, the commonest disorder is acute anterior uveitis occurring in association with spondylarthropathies or reactive arthritis. Other eye problems which should prompt a clinical/serological evaluation for a rheumatological disorder include chronic anterior uveitis (pauciarticular juvenile chronic arthritis), uveitis (Behçet's syndrome), episcleritis (relapsing polychondritis), keratoconjunctivitis (primary Sjögren's syndrome), visual loss (temporal arteritis), and cortical blindness (systemic lupus erythematosus).

Oropharyngeal system

Presenting disorders include xerostomia (primary Sjögren's syndrome), painful oral ulceration (Behçet's syndrome and systemic lupus erythematosus), painless oral ulceration (reactive arthritis), and apparent 'toothache' (temporomandibular joint involvement). Occasionally, apparently unrelated events may suggest a rheumatological disorder (e.g. severe dental caries and primary Sjögren's syndrome).

Cutaneous system

The vast range of cutaneous features associated with rheumatological disorders is dealt with elsewhere. Table 2 lists some cutaneous disorders that may be the presenting feature of rheumatological disorders. In clinical practice Raynaud's, psoriasis, various rashes, and vasculitis are the most commonly encountered.

Cardiovascular system

Table 2 lists cardiovascular problems which may be the presenting features. Again, in clinical practice, the most commonly seen are pericarditis (with rheumatoid arthritis, systemic lupus erythematosus, and systemic juvenile chronic arthritis) (Kelly *et al*. 1990), valvular heart disease (with systemic lupus erythematosus and spondylarthropathies), peripheral ischaemia (with polyarthritis and other forms of medium/large vasculitis), congenital complete heart block (with maternal anti-Ro-positive systemic lupus erythematosus /Sjögren's syndrome) (McCredie *et al*. 1990).

Cardiomyopathies (e.g. with amyloid, polymyositis, and scleroderma) tend to occur later in the natural history of the underlying disease.

Respiratory system

Lung and respiratory tract involvement is common to many rheumatological disorders. Breathlessness is a common presenting feature of many rheumatological disorders. In clinical practice the most frequently encountered are pleural disease (rheumatoid arthritis and systemic lupus erythematosus), pulmonary fibrosis (rheumatoid arthritis and connective tissue disorders), asthma (Churg-Strauss vasculitis), pulmonary opacities on routine chest radiography (rheumatoid arthritis, Wegener's granulomatosis), and bronchial 'irritability' (Sjögren's syndrome). Sinus disorders with Wegener's granulomatosis are common.

Gastrointestinal system

Although the associations of rheumatic diseases and gastrointestinal disorders, particularly infectious and chronic inflammatory diseases, is well established, it is more common to see these problems causing diagnostic confusion because of their rheumatological presentation—for example Crohn's disease, where the gastrointestinal features may be less apparent in the early stages.

Areas of overlap encountered in clinical practice tend to relate to abdominal pain arising from spinal disease, mesenteric vasculitis, or polyserositis (e.g. familial Mediterranean fever), abnormal biochemical tests (particularly liver function tests) detected on 'routine' tests and secondary to an underlying chronic inflammatory rheumatological disease, and apparent bowel disturbance in the irritable bowel/fibromyalgia overlapping group of patients (Wolfe 1989).

Dysphagia in systemic sclerosis is very rarely a presenting complaint.

Urogenital and renal systems

Renal system

Kidney involvement is frequently the presenting feature of the connective tissue disorders (e.g. systemic lupus erythematosus, polyarteritis). Renal tubular acidosis (detected either clinically as secondary osteomalacia or biochemically) may be the presenting feature of primary Sjögren's syndrome (Maher 1989). Renal calculi may develop prior to clinical gout. Acute hypertensive renal failure can be the presenting feature of systemic sclerosis.

Male urogenital tract

Urethritis is the rule with both gonococcal and non-gonococcal sexually acquired arthritis. It may also occur in reactive arthritis secondary to gastrointestinal infections. Balanitis occurs and may reoccur in Reiter's syndrome without fresh exposure to infectious agents (Fisk 1982).

Impotence is a rare presenting feature of large vessel vasculitis (mimicking the Le Riche syndrome) and is not uncommon in systemic sclerosis. Peyronies' disease associates with other disorders of collagen deposition (e.g. Dupuytren's contracture, knuckle pads, and systemic sclerosis).

Female urogenital tract

Probably the most common presenting disorders are dyspareunia (with primary Sjögren's syndrome), recurrent spontaneous abortion (with the antiphospholipid syndrome) (Hughes 1983; Hughes *et al*. 1989), vaginal ulceration (Behçet's), and vaginal prolapse (hypermobility syndrome and its variants) (Al Rawi and Al Rawi 1982).

Neurological system

Disorders affecting this system are frequently the presenting complaint in rheumatological disease.

The wide diversity of neurological features occurring in systemic lupus erythematosus is well recognized. Entrapment neuropathies, most commonly carpal tunnel, is a frequent presenting complaint in inflammatory joint disease (e.g. rheumatoid arthritis). Mononeuritis is common in polyarteritis. Headaches are a feature of temporal arteritis and cervical spondylosis. Migraine occurs more commonly in systemic lupus erythematosus and Sjögren's syndrome (Pal *et al*. 1989). Spastic quadriplegia may be the presenting feature of cervical spondylosis and ossification of the posterior longitudinal ligament (Griffiths and Fitzjohn 1987).

Clinical patterns

The enormously varied nature of clinical medicine means that every physician will have anecdotes about bizarre presentations of relatively common disorders (e.g. the patient who complained 'she waddled when she ran' due to a proximal myopathy, due to osteomalacia, due to renal tubular acidosis, due to primary Sjögren's syndrome; or the patient with gout who snored, both of which resolved when his hypothyroidism was treated). While diagnosis in these patients may provide an intellectual challenge, the successes tend to be matched by the failure to recognize, for example, the patient with polymyalgia who had a hypernephroma.

Certain patterns of disease occur sufficiently frequently, however, that the systemic confirmation or exclusion of an underlying 'rheumatological' disorder is essential.

The commonest problems encountered in the United Kingdom are described below.

Children

The ill, febrile child with a rash suggests the need to exclude systemic juvenile chronic arthritis. Recognition of significant articular pathology is not easy in a small, unhappy, and irritable child.

Young women

Multisystem disease in young women should always raise the possibility of systemic lupus erythematosus. It is in this group that serological tests are most likely to be helpful.

Young men

Reactive arthritis, either sexually or gastrointestinally acquired, frequently causes a systemic illness of greater severity than is usually recognized by the medical profession. In an ill, febrile young man with weight loss and a subacute history, a careful search for clues (e.g. mild back pain and stiffness, a single swollen metatarsal phalangeal joint or toe, or tenderness under the calcanum may provide clinical pointers).

Middle-aged women

Three main diagnostic groups may present in a 'non-rheumatological guise'.

Rheumatoid arthritis not infrequently may present in a non-articular fashion. Pulmonary or pleuropericardial involvement tends to be the most frequent manifestation.

Two disorders may present as malaise or lethargy often associated with other non-specific symptoms. The symptom complex of 'fibromyalgia' is suggested by the finding of multiple tender points and normal screening laboratory investigations. Primary Sjögren's syndrome often also presents in this non-specific fashion. Specific questions about ocular and oral dryness suggest the diagnosis and laboratory investigations (particularly erythrocyte sedimentation rate and serology) are often surprisingly abnormal against a relatively normal physical examination.

Middle-aged men

Rheumatoid arthritis presenting either pleuropericardial disease or pulmonary involvement is the commonest clinical situation. Polyarteritis nodosum may also cause diagnostic confusion in this group with its variety of clinical features ranging from malaise through to mononeuritis multiplex.

Elderly women

Two disorders commonly cause diagnostic confusion—polymyalgia rheumatica and temporal arteritis, as both the non-musculoskeletal features (malaise, weight loss, fever of unknown origin) are often attributed to other conditions and the more specific features (e.g. proximal myalgias and headaches) may be misinterpreted unless a low diagnostic threshold for this treatable disorder is maintained.

Tophaceous gout in the elderly female taking diuretics is increasingly common. The lesions on the distal interphalangeal joints are often diagnosed as Heberden's nodes and ulcerating lesions adjacent to the first metatarsal phalangeal joint attributed to either vascular insufficiency or trophic changes.

Elderly men

Extra-articular presentation of rheumatic disorders is less common in this age group. More commonly one encounters the reverse situation of non-rheumatological disorders (particularly neoplastic) presenting under a rheumatological guise of either joint or bone pain of a paraneoplastic syndrome.

Conclusions

The recognition of an underlying rheumatological diagnosis as the cause of non-articular disease is not merely an academic exercise. To overlook certain diagnoses (e.g. systemic lupus erythematosus or temporal arteritis) may have severe consequences in terms of morbidity and mortality. Furthermore, to pursue increasingly invasive investigations in a patient with fibromyalgia exposes them to all the risks of iatrogenesis.

Ultimately, the establishment of the underlying cause usually depends upon clinical acumen, often helped by maintaining an 'open mind' when diagnostic uncertainty exists, and the passage of time.

Chapter References

Al Rawi, Z.S. and Al Rawi, Z.T. (1982). Joint hypermobility in women with genital prolapse. *Lancet*, **i**, 1439–41.

Bissenden, J.G. and Hall, S. (1990). Kawasaki syndrome: lessons for Britain. *British Medical Journal*, **300**, 1025–6.

Doherty, M., Abawi, J., and Pattrick, M. (1990). Audit of medical inpatient examination: a cry from the joint. *Journal of the Royal College of Physicians*, **24**, 115–8.

Fisk, P. (1982). Reiter's disease. *British Medical Journal*, **284**, 3–4.

Goldenberg, D.L. (1993). Fibromyalgia, chronic fatigue syndrome and mycofacial pain syndrome. *Current Opinion in Rheumatology*, **5**, 199–208.

Griffiths, I.D. and Fitzjohn, T.D. (1987). Cervical myelopathy, ossification of posterior longitudinal ligament and DISH. *Annals of the Rheumatic Diseases*, **46**, 166–8.

Hughes, G.R.V. (1983). Thrombosis, abortion, cerebral disease and the lupus anti-coagulant. *British Medical Journal*, **287**, 1088–9.

Hughes, G.R.V., Asherson, R.A., and Khamashta, M.A. (1989). Anti-phospholipid syndrome: linking many specialties. *Annals of the Rheumatic Diseases*, **48**, 355–6.

Kelly, C.A., Bourke, J.P., Malcolm, A., and Griffiths, I.D. (1987). Lymphadenopathy in rheumatic patients. *Annals of the Rheumatic Diseases*, **46**, 224–7.

Kelly, C.A., Malcolm, A., and Griffiths, I.D. (1990). Chronic pericardial disease in patients with rheumatoid arthritis. *Quarterly Journal of Medicine*, **277**, 461–70.

Knockaert, D.C., Vanneste, L.J., and Bobbaers, H.J. (1993). Recurrent or episodic fever of unknown origin. *Medicine*, **72**, 184–96.

Knockaert, D.C., Vanneste, L.J., Vanneste, S.B., and Bobbaers, H.J. (1992). Fever of unknown origin in the 1980s. *Archives of Internal Medicine*, **152**, 51–5.

Maher, E.R. (1989). Renal tubular acidosis. *British Journal of Hospital Medicine*, **42**, 116–19.

Martin, K., Davies, E.G., and Axford, J.S. (1994). Fever of unknown origin in childhood: Difficulties in diagnosis. *Annals of the Rheumatic Diseases*, **53**, 429–33.

McCredie, M., Celermajor, J., Sholler, G., Kelly, D., Chivers, T., Wang, Y., and Schrieber, L. (1990). Case control study of C.C.H.B.: Assocation with maternal antibodies to Ro and La. *British Journal of Rheumatology*, **29**, 10–14.

O'Connell, S. (1995). Lyme disease in the United Kingdom. *British Medical Journal*, **310**, 303–8.

Pal, B., Gibson, C., Passmore, J., Griffiths, I.D., and Dick, W.C. (1989). A study of headaches and migraine in Sjögren's syndrome and other rheumatic disorders. *Annals of the Rheumatic Diseases*,

48, 312–16.

Pinals, R.S. (1994). Polyarthritis and fever. *New England Journal of Medicine*, **330**, 769–74.

Pouchot, J., Sampalis, J.S., Beaudet, F., *et al*. (1991). Adult Still's disease: manifestations, disease course and outcome in 62 patients. *Medicine*, **70**, 118–36.

Roubenoff, R., Roubenoff, R.A., Ward, L., Hollands, S., and Hellman, D. (1992). Rheumatoid cachexia: depletion of lean body mass in rheumatoid arthritis. Possible association with tumour necrosis factor. *Journal of Rheumatology*, **19**, 1505–10.

Terkeltaub, R., Esdaile, J., De'Cary, F., and Tannenbaum, H. (1983). A clinical study of older age rheumatoid arthritis with comparison to a younger onset group. *Journal of Rheumatology*, **10**, 418–24.

Wolfe, F. (1989). The fibromyalgia syndrome. *Rheumatic Disease Clinics of North America*, **15**, 1–18.

Wouter, J. and Van de Putte, L. (1986). Adult onset Still's disease; clinical and laboratory features, treatment and progress of 45 cases. *Quarterly Journal of Medicine*, **61**, 1055–65.

1.3.1.1 Pregnancy

W. Watson Buchanan and Walter F. Kean

Three events made 1948 the *annus mirabilis* of rheumatology. First, the description of agglutination of sensitized sheep erythrocytes by sera of patients with rheumatoid arthritis (Rose *et al*. 1948) (although this phenomenon had previously been described by Waaler 1940). Second, there was the discovery by Hargraves *et al*. (1948) at the Mayo Clinic of the 'LE' cell. Third, at the same institution, Hench and his co-workers (Hench *et al*. 1949a; Hench *et al*. 1949b) described the effects of the adrenal cortical hormone, 17-hydroxy-11-dehydrocorticosterone (compound E), and pituitary adrenocorticotrophic hormone in rheumatoid arthritis and rheumatic fever. The discovery of cortisone by Hench, which was to earn him a Nobel Prize in Medicine in 1950, had been prompted by his earlier observation of the ameliorative effects of pregnancy in rheumatoid arthritis, which he reasoned was due to a substance released from the adrenal glands (Hench 1938).

This chapter deals with the effect of pregnancy in rheumatic disease, and the effects of rheumatic disease on pregnancy and the fetus.

Rheumatoid arthritis

There have been several reviews on the effects of pregnancy on rheumatoid arthritis (Klipple and Cecere 1989; Kean and Buchanan 1990; Bellamy *et al*. 1991; Spector and Da Silva 1992). Despite the fact that many studies include only a relatively small number of patients and provide only global assessments of improvement, there seems to be overwhelming evidence that pregnancy has a beneficial effect on rheumatoid arthritis. Overall it has been estimated that 75 per cent of patients improve during pregnancy, especially during the last trimester. Improvement frequently begins in the first trimester, and, with few exceptions, is maintained throughout pregnancy. Improvement may be dramatic with complete remission of symptoms, and in some patients even a reduction in the number and size of subcutaneous nodules, so allowing complete discontinuation of all antirheumatic medication. Fluctuation in disease activity is, however, common, even in patients whose arthritis improves or goes into remission. Improvement in one pregnancy usually predicts the same in future pregnancies; likewise, failure to improve or deterioration postpartum is equally predictive. Currently there is no method of predicting the clinical outcome of rheumatoid arthritis during pregnancy, although a role for fetal paternally-inherited class II HLA antigens in amelioration of symptoms has been suggested (Nelson *et al*. 1992a; Nelson *et al*. 1992b; Nelson *et al*. 1993). When symptoms do not improve, progressive erosive changes may be evident on joint radiographs.

Arthritic symptoms usually recur within 1 to 2 months of delivery. Although the degree of joint inflammation may become more severe postpartum, disease activity usually returns to that prior to conception. Lactation has no effect on either timing or severity of postpartum relapse.

Patients with rheumatoid arthritis develop no more side-effects of pregnancy, such as pre-eclampsia, than do healthy subjects. Problems with delivery, likewise, are no greater, although hip involvement may make vaginal delivery more difficult and painful. Rheumatoid arthritis is, therefore, generally not considered a reason for therapeutic abortion (Felbo and Snorrason 1961). Some studies, but not all, have suggested that spontaneous abortion may be more common in patients with rheumatoid arthritis. Still birth rates may also be slightly increased (Spector and Silman 1990). Rheumatoid arthritis otherwise appears to carry no risk for the fetus, and does not affect lactation or resumption of menses. No difference has been observed in pregnancy before the onset of rheumatoid arthritis (Nelson *et al*. 1992).

The reason why so many patients with rheumatoid arthritis improve during pregnancy remains elusive. (Rook *et al*. 1991; Da Silva and Spector 1992). Klipple and Cecere (1989) have summarized the circulating factors with anti-inflammatory and immunoregulatory activity (Table 1), and the potential factors which might operate in lessening of disease activity during pregnancy (Table 2). Although both the total and free plasma cortisol concentrations rise during pregnancy, this is not of sufficient degree to explain remission in rheumatoid arthritis (Persellin 1981). Thus, one of the great quantum leaps in rheumatology was based on a false hypothesis! The placenta may play a significant role as an immune-modifying organ (Klipple and Cecere 1989), but none of the many immunological changes (Pope 1990; Klipple and Cecere 1989), including pregnancy-associated a-glycoprotein (Ostensen *et al*. 1983; Unger 1983), provide an adequate explanation. Sany *et al*. (1986) have reported improvement in patients with rheumatoid arthritis treated with IgG eluted from placental tissue. This was an open study, and will require confirmation by a double-blind, placebo-controlled trial.

Fetal factors
α-Fetoprotein
Other factors in maternal sera
Trophoblast hormones
Oestrogens
Progesterone
Human chorionic gonadotrophin
Human placental lactogen
Maternal factors
Corticosteroids
Blocking antibodies including antitrophoblastic antibodies
Trophoblastic-decidual placental proteins
Pregnancy-associated plasma protein A
Pregnancy-associated α2 globulin
Placental protein 14
Pregnancy-specific β$_1$-glycoprotein
Other pregnancy serum proteins

Reproduced by kind permission of Klipple and Cecere (1989).

Table 1 Circulating factors with immunoregulatory or anti-inflammatory activity in pregnancy sera

Humoral factors
Immunomodulation by circulating trophoblast-antitrophoblast immune complexes
Removal of circulating aggregated immunoglobulins by the placental 'sponge'
Alteration in the carbohydrate side-chain composition of immunoglobulins leading to decreased 'stickiness'
Decreased IgG levels (mild)
Fetal suppression of the maternal immune system
Soluble factors including α-fetoprotein
Fetal suppressor cells in maternal circulation
Depression of cell-mediated immunity
Depression of delayed hypersensitivity by skin testing
Delayed rejection of skin allografts
Increased virulence of recurrence of intracellular pathogens
Decreased lymphocyte counts
Decreased helper T cells with decreased helper/suppressor ratio
Depressed in vitro lymphocyte responses
Decreased natural killer activity
Circulating suppressors of cell-mediated immunity including blocking antibodies
Suppression of inflammatory reactions
Impaired chemotaxis, adherence, phagocytosis, normal breast tissue reduction
Decreased cathepsin B activity
Suppressive serum factor(s)

Table 2 Potential factors in the amelioration of rheumatoid arthritis during pregnancy

Pregnancy has been shown to have the same beneficial effects on experimental arthritis as in humans (Hirahara *et al*. 1986). In mice with collagen-induced arthritis, improvement did not correlate with serum anticollagen antibodies (Williams and Whyte 1989), but a delay in antibody production was observed when the mice were immunized during pregnancy (Whyte *et al*. 1988). Oestrogens when administered to non-pregnant animals so as to maintain pregnancy serum concentrations of oestrogen have been shown to suppress experimental arthritis (Mattson *et al*. 1991). However, progesterone alone had no such effect, but enhanced the effect of oestrogen (Jansson and Holmdahl 1989). Bromocriptine has been shown to suppress postpartum exacerbation of collagen-induced arthritis, probably due to suppression of prolactin release that normally occurs postpartum (Whyte and Williams 1988). Female sex hormones, especially oestradiol, affect the immune status (Ahmed and Talal 1990; Lahita 1990), and the dramatic hormonal changes which occur during pregnancy might be expected to account for improvement in rheumatoid arthritis. It is therefore disappointing that controlled clinical trials of ethinyloestradiol in non-pregnant female patients with rheumatoid arthritis have shown no benefit (Bijlsma *et al*. 1987; Bijlsma and Van den Brink 1992; Da Silva and Hall 1992). Whether oral contraceptives have a protective effect on the development and course of rheumatoid arthritis remains controversial (Hazes *et al*. 1990). Likewise, whether women who develop rheumatoid arthritis are subfertile before the onset of their disease (Yoshino and Uchida 1981; Spector and Silman 1990) or whether multiparity is associated with an increased risk of developing rheumatoid arthritis (Klipple and Cecere 1989) also remain controversial. Although the onset of rheumatoid arthritis is lower than expected during pregnancy while increased after delivery (Silman *et al*. 1992; Lansink *et al*. 1993), female sex hormones probably do not predispose females to development of rheumatoid arthritis (Pritchard 1992). No evidence has been found in twin studies of a strong interaction between genetic and environmental factors (Brennan and Silman 1994). James (1993) has suggested a low androgen level may predispose to the development of rheumatoid arthritis, and it is of interest that males with rheumatoid arthritis tend to have a reduction in testosterone and dihydroepiandrosterone (Catula *et al*. 1984).

Rheumatoid arthritis frequently adversely affects female sexuality with a loss of libido and reduction in frequency of sexual intercourse (Yoshino and Uchida 1981). Nothing has been reported to the authors' knowledge on sexual disturbances in the male patient with rheumatoid arthritis.

Systemic lupus erythematosus

There are a number of difficulties in interpreting the literature on the effects of pregnancy on systemic lupus erythematosus. There is an absence of a generally agreed definition of disease activity, and difficulty in differentiating between systemic lupus erythematosus and pregnancy complications, such as pre-eclampsia. Most studies do not include control data, and differences in hospital referral patterns lead to differences in severity of disease between different groups. In addition, the racial make-up between studies differs: for instance the study reported by Wong *et al*. (1991) from Hong Kong consisted entirely of Chinese patients, whereas in the study of Pistiner *et al*. (1991) from California, 75 per cent of the patients were white, 11 per cent black, 8 per cent Hispanic, and only 6 per cent Asian. Flares in disease activity during pregnancy may be the result of discontinuing drug therapy, as in the study reported by Lockshin *et al*. (1984), in which two patients had exacerbation of cutaneous lupus following withdrawal of hydroxychloroquine therapy. Despite all of these problems, it is possible to conclude from the large number of studies that exacerbation of systemic lupus erythematosus is more likely in pregnancy when the disease is inactive at conception (Lockshin *et al*. 1984; Hayslett 1991; Petri *et al*. 1991; Hayslett 1992; Urowitz *et al*. 1993). Previous pregnancy outcome (Ramsay-Goldman *et al*. 1992; Ramsay-Goldman *et al*. 1993b), low socio-economic status (McAlindon *et al*. 1993), and pre-existing renal disease (Nicklin 1991; Rubbert *et al*. 1992) are also important determinants. Transient impairment in renal function is usually the result of superimposed pre-eclampsia in patients whose systemic lupus erythematosus is otherwise quiescent (Burkett 1989). Pregnancy does not appear to worsen the long-term prognosis of systemic lupus erythematosus (Burkett 1989), and flares in disease activity do not necessarily influence the outcome of pregnancy (Petri *et al*. 1991).

Fertility does not appear to be impaired in women with systemic lupus erythematosus (Tozman *et al*. 1980). However, fertility is probably reduced when the disease activity is severe (Cecere and Persillin 1981; Pistiner *et al*. 1991) since amenorrhoea is common (Martinez-Cordero *et al*. 1982), and sexual intercourse less frequent, if at all (Bellamy *et al*. 1991). The role of drug therapy, such as corticosteroids, on fertility is unknown.

A major pregnancy-associated complication for the mother is pre-eclampsia. This may be difficult to distinguish from an exacerbation of systemic lupus erythematosus (Bellamy *et al*. 1991). The presence of haematuria and red cell casts suggests lupus nephritis, whereas a rapid increase in proteinuria is more suggestive of pre-eclampsia (Bellamy *et al*. 1991). Both the classical and alternate complement pathways are activated during flares in activity of systemic lupus erythematosus during pregnancy (Buyon *et al*. 1992; Hopkinson and Powell 1992). Activation of both pathways with depression of C3 and C4 levels also occurs in pre-eclampsia, but the total CH5O remains normal and may help in the differential diagnosis (Abramson and Buyon 1992). Delivery by caesarean section is common in patients with systemic lupus erythematosus as a result of concern for both mother and child (Urowitz and Gladman 1982). Avascular necrosis of the hips may give rise to mechanical difficulties during vaginal delivery. A case of breast vasculitis and gigantism controlled with high doses of corticosteroids has been described (Propper *et al*. 1991).

Fetal mortality is increased in women with systemic lupus erythematosus (Bellamy *et al*. 1991; Ferro *et al*. 1992; Kutteh *et al*. 1993; Petri 1994). Although earlier literature suggested that early abortions were more common in patients with this disease, recent reports have indicated a rate not significantly different from the expected rate of spontaneous abortion in the normal population (Cecere and Persellin 1981; Mintz and Rodriguez-Alvarez 1989). In contrast, fetal death is markedly increased in the later stages of pregnancy with a still birth rate of approximately 10 per cent (Cecere and Persellin 1981; Mintz and Rodriguez-Alvarez 1989), and especially in those patients with hypocomplementaemia (Shibata *et al*. 1992). Although Jara-Quezada *et al*. (1991) have suggested that this increased late fetal loss might be due to gonadal hormone and prolactin increase, it is much more likely to be the result of antiphospholipid antibodies (Lockshin *et al*. 1987; Harris 1990; Birdsall *et al*. 1992; Ferro *et al*. 1992; Ginsberg *et al*. 1992; Kerslake *et al*. 1992; Out *et al*. 1992; Vianna *et al*. 1992). Delivery of a term infant in a patient with systemic lupus erythematosus and antiphospholipid antibodies is extremely uncommon (Khamashta and Wallington 1991). Lockshin *et al*. (1987) has suggested that antiphospholipid antibodies are the most sensitive marker of pregnancy loss, with a sensitivity of 0.85, and specificity of 0.92. Antiphospholipid antibodies are either IgG, IgA, or IgM, or all three, with a wide spectrum of antibody specificities (Triplett 1989; Chamley *et al*. 1991; Reyes *et al*. 1994). No single test is sufficient because of this heterogeneity and a panel of screening procedures, inhibitor identification, and confirmatory procedures has recently been recommended (Barna and Triplett 1991). IgA antiphospholipid antibodies (Lopez *et al*. 1992) and the lupus anticoagulant (Julkunen *et al*. 1993) appear most likely to result in fetal death. However, although the finding of antiphospholipid antibodies is a major risk factor in pregnancy, several normal pregnancies have been reported (McHugh and Laurent 1989; McHugh *et al*. 1989). Antiphospholipid antibodies have a low incidence in certain races, especially Malays, Chinese, and Indians (Jones *et al*. 1991). They have also been reported to have an increased prevalence among first-degree relatives of patients (Molta *et al*. 1993). Tests for antiphospholipid antibodies need only be performed if the patient has systemic lupus erythematosus or other suggestive clinical features (Harris and Spinato 1991; Out *et al*. 1991a). Fetal death is essentially the result of thromboses causing placental infarction (Out *et al*. 1991b). Chamley *et al*. (1993) have demonstrated that thrombosis may be due to inhibition of heparin-dependent antithrombin III activation by anticardiolipin antibodies. Antiphospholipid antibodies have also been shown to trigger thromboxane production, which can be eliminated by small doses of aspirin (Kaaja *et al*. 1993a). Inflammatory vascular changes also occur in the vascular bed and may lead to thrombosis, which explains the beneficial effects of corticosteroids (Erlendsson *et al*. 1993). Antiphospholipid antibodies do not interfere with the natural inhibitors of blood coagulation and fibrinolysis (Kordich *et al*. 1992). Acute audiovestibular failure has been attributed to antiphospholipid antibodies (Vyse *et al*. 1994). Regular assessment of intrauterine growth to predict fetuses at risk is recommended, and in the later stages of pregnancy this can be supplemented with Doppler flow studies of uteroplacental and umbilical artery circulation (Trudinger *et al*. 1988; Benifla *et al*. 1992; Guzman *et al*. 1992; Walkinshaw *et al*. 1994).

There is still no agreement regarding the medication that patients with antiphospholipid antibodies should be prescribed, and whether this should be given only when there is an indication of danger to the fetus (Derksen 1991; Buchanan *et al*. 1992; Many *et al*. 1992; Tambyraja 1993). High doses of prednisone have been recommended but their beneficial effects have been questioned (Lockshin *et al*. 1989). Low-dose aspirin, used to inhibit thromboxane A_2 synthesis and thrombosis (Branch *et al*. 1985; Kaaja *et al*. 1993a), is currently undergoing clinical trial (Barton and Sibai 1991). Subcutaneous heparin does not cross the placenta and may prove useful in preventing abortion, although maternal bone loss may prove a problem. Other therapies include cytotoxic drugs, dipyridamole, and warfarin, plasma exchange, and high-dose immunoglobulins (Khamashta and Wallington 1991; Ramsay-Goldman *et al*. 1993a). Human IgG has been shown to neutralize lupus anticoagulant activity in vitro, (Said *et al*. 1992), and an intravenous dose of 1 g/kg body weight has been suggested as a prophylactic measure (Kaaja *et al*. 1993b). Immunoadsorbent plasmapheresis has been reported to decrease significantly antiphospholipid antibody titres, but only in a single patient (Kobayashi *et al*. 1992). Further research on antiphospholipid antibodies may lead to more targeted approaches to treatment (Gonzalez-Buritica *et al*. 1988). Antiphospholipid antibodies are not only found in systemic lupus erythematosus but also in several connective tissue diseases, where they are associated with high fetal loss (Buchanan *et al*. 1989; Imai *et al*. 1991). Patients who are discovered to have antiphosopholipid antibodies but have no symptoms should be followed carefully as many will later develop systemic lupus erythematosus (Bagger *et al*. 1993).

The neonatal lupus syndrome is discussed in Chapter 5.7.2 and further discussion of antiphospholipid antibodies in Chapter 5.7.3.

Systemic sclerosis

There is no reliable prospective statistical information on the outcome of pregnancy in progressive systemic sclerosis for either mother or fetus (Black 1990). However, case–control studies have been recently published. An increased rate of infertility in patients destined to develop systemic sclerosis was observed (Silman and Black 1988; Englert *et al*. 1992; Silman 1992). Patients with the disease had three times the rate of fertility problems compared with controls. The study by Giordano *et al*. (1985) showed no evidence of infertility in patients with progressive systemic sclerosis. Steen *et al*. (1989) did not address this problem. Prematurity and low birth weights have been reported (Steen *et al*. 1989; Silman 1992), and a possible increased rate of spontaneous abortions (Silman and Black 1988; Steen *et al*. 1989; Silman 1992).

Scleroderma involving the uterine cervix and perineal skin may cause difficulty in delivery, but is rare (Bellamy *et al*. 1991). Renal crisis is the most serious complication to develop during pregnancy, and appears to occur especially in patients with rapidly progressive diffuse skin involvement (Steen *et al*. 1989). At present there are no absolute prognosticators regarding the outcome of either mother or fetus (Avrech *et al*. 1992; Englert *et al*. 1992). However, patients should be tested for antiphospholipid, anti-Ro (anti-SSA), and anti-La (anti-SSB) antibodies.

Juvenile chronic (rheumatoid) arthritis

Ostensen (Ostensen 1991; Ostensen 1992) has recently reviewed the outcome of juvenile rheumatoid arthritis during pregnancy. Some 60 per cent of patients experienced improvement or total remission of their arthritis, especially during the second half of gestation. No improvement in uveitis was noted. Functional impairment was the main limiting factor in having children, and delivery by caesarean section is frequently required. Postpartum flare occurred in approximately 50 per cent of patients. Attempts to control such flares with infusions of autologous plasma have been unsuccessful (Wallace *et al*. 1987). Adult-onset Still's disease has been reported both during and after pregnancies (Stein *et al*. 1980; de Miguel *et al*. 1992; Le Loet *et al*. 1993). Vasculitis resulting in coronary thrombosis has also been described postpartum (Parry *et al*. 1992). Ostensen (1992) noted that fetal outcome was not affected.

Ankylosing spondylitis

Ankylosing spondylitis in women is milder and often atypical when compared with men (Marks *et al*. 1983). There is no evidence that patients with the disease are less fertile or have an increased frequency of spontaneous abortion, premature labour, or still birth (Husby *et al*. 1988). Ostensen (1992) reported that 80 per cent of disease activity remained unaltered or was aggravated during pregnancy. Gran and Husby (1992) considered ibuprofen to be the preferred drug for control of pain during pregnancy. Postpartum flares unrelated to lactation or return of the menses have been observed (Ostensen 1992). Whether pregnancy causes worsening of the radiological changes, as has been reported, awaits further study. Amyloid-related serum protein concentration has been found to be a useful indicator to monitor disease activity during pregnancy (Ostensen *et al*. 1985). No laboratory test has been found useful in predicting gestational activity of disease (Ostensen 1984). There is a risk that children born to an HLA B27-positive mother with ankylosing spondylitis will develop the disease in later life: this has been estimated at 10 per cent if the child is HLA-B27 positive and virtually zero if the child is HLA-B27 negative (Van der Linden and Khan 1984).

Psoriatic arthritis

There have been few reports of the effects of pregnancy on psoriatic arthritis. In general, less amelioration of arthritis symptoms has been reported in patients with psoriatic arthritis compared with rheumatoid arthritis (Ostensen 1988). However, McNeill (1988) has reported multiple pregnancy-induced remissions in a patient with psoriatic arthritis, and a recent study by Ostensen (1992) suggests that improvement or remission may occur in as many as 80 per cent of patients. No change has been reported in the skin lesions during pregnancy. McHugh and Laurent (1989) noted that 18 per cent of 33 patients had the onset of their arthritis within three months postpartum. The recent studies of Ostensen (1992) put the figure of postpartum flares as high as 70 per cent. No such relation was noted with psoriasis. Further studies with larger numbers of patients require to be performed to confirm the assertion of McHugh and Laurent (1989) that pregnancy may be a risk factor for psoriatic arthritis. Fetal outcome does not appear to be affected (Ostensen 1992).

Gonococcal arthritis

Pregnancy predisposes to disseminated gonococcal infection, especially during the third trimester (Taylor *et al*. 1966; Einstein and Masi 1981). Destructive arthropathy may result if antibiotic therapy is not promptly instituted (Einstein and Masi 1981). No unfavourable effects on either mother or child will occur with proper care.

Gout

Hyperuricaemia and gouty arthritis are rare in premenopausal women. Serum urate concentrations fall to exceeding low levels early in pregnancy, and the finding of hypouricaemia in a young woman should alert the physician to the possibility of pregnancy (Boyle *et al*. 1966). The serum urate concentration gradually rises to within the normal range towards the third trimester. Hyperuricaemia occurs with pre-eclampsia. Lee and Loeffler (1962) reported five spontaneous abortions in 22 pregnancies in patients with gout, and Talbott (1957) suggested that gouty patients might be subfertile. Colchicine is absolutely contraindicated in pregnancy (Levy *et al*. 1991).

Relapsing polychondritis, Wegener's granulomatosis, Sjögren's syndrome, dermatomyositis, and polyarteritis nodosa

Bellamy and Dewar (1990) have reported a favourable maternal and fetal outcome in a patient with relapsing polychondritis: this appears to be the only reference to pregnancy in this condition in the literature.

Patients with Wegener's granulomatosis have been described who had healthy full-term babies (Biesenbach *et al*. 1991) despite treatment with cyclophosphamide (Fields *et al*. 1991).

Sjögren's syndrome predominantly affects middle-aged and elderly women but also occasionally women in the child-bearing years and even children. Patients with primary Sjögren's syndrome have normal pregnancies and healthy babies (Takaya *et al*. 1991; Manthorpe and Manthorpe 1992). However, patients with anti-Ro (anti-SSA) and anti-La (anti-SSB) have the risk of having infants with congenital heart block (Imai *et al*. 1991).

Favourable maternal and fetal outcomes have been reported in dermatomyositis (Ishii *et al*. 1991; Ohno *et al*. 1992), but no series of patients has been published to our knowledge.

An infant born of a mother with polyarteritis nodosa has been described with cutaneous lesions of the disease (Stone *et al*. 1993).

The successful outcome of a pregnancy has been reported in a mother with multicentric reticulohistiocytosis (Conaghan *et al*. 1993).

Miscellaneous musculoskeletal complaints

During the later stages of pregnancy there is an increase in lumbar lordosis, angulation of the lumbosacral junction, and obliquity of the pelvic bones, due to softening of ligaments as a result of hormonal changes. In addition, sliding movement occurs in the sacroiliac joints and radiographs frequently demonstrate gas present as a result of movement. Likewise, hypermobility of the symphysis pubis results in the development of osteitis pubis. It is therefore not surprising that almost 50 per cent of women complain of lumbar backache during the later stages of their pregnancy (Mantle *et al*. 1977). The pain radiates into the buttocks and occasionally down the legs, and may be sufficiently severe to interfere with sleep. The pain is increased during labour, but usually disappears soon after delivery. Patients who experience low-back pain during pregnancy are said to suffer later from lumbar disc disease, but this has not been the subject of a properly controlled study. Acute lumbar disc prolapse has been reported during pregnancy, but whether the incidence of this complication is any higher than in non-pregnant age-matched females has not been studied.

Patients with hypermobility of the sacroiliac joints may go on to develop osteitis condensans ilii, which is usually asymptomatic or associated with only mild symptoms. Patients who develop osteitis condensans ilii are HLA-B27 negative, and the radiological changes usually disappear with time. Wilbur *et al*. (1988) have reported the

use of magnetic resonance imaging to diagnose sacroiliac joint infection during pregnancy. Pain and tenderness over the symphysis pubis may cause difficulty in walking.

Another rarer complication which may give rise to hip pain and difficulty with walking is transient osteoporosis (Longstreth *et al*. 1973). This is usually unilateral and full recovery occurs 3 months to 1 year after delivery. The cause of this complication is unknown (Resnick 1988). Ischaemic necrosis of bone, especially affecting the femoral heads, has been reported in a few patients during pregnancy (Resnick 1988). Generalized osteopenia associated with pregnancy and lactation has been reported (Nordin and Roper 1955). Another extremely rare complication of pregnancy is osteitis condensans of the medial end of the clavicle. Bone sclerosis is apparent on radiography, and the lesion is 'hot' on an isotope bone scan (Resnick 1988). The lesion is quite distinct from sternoclavicular hyperostosis associated with pustulosis palmaris et plantaris.

Carpal tunnel syndrome occurs in approximately one-third of patients during pregnancy, especially in the third trimester. Patients who develop pre-eclampsia are particularly likely to develop this complication. Compression of both median and ulnar nerves is common, with classical median nerve compression alone less common, and ulnar nerve compression alone relatively rare (McLennan *et al*. 1987). Carpal tunnel symptoms spontaneously resolve after delivery, and surgical treatment is only rarely required (Melvin *et al*. 1969).

We dedicate this chapter to the late Nobelist Dr Philip S. Hench (1896–1965) whose discovery of cortisone did so much to promote the development of the medical subspeciality of rheumatology.

Chapter References

Abramson, S.B. and Buyon J.P. (1992). Activation of the complement pathway: comparison of normal pregnancy, pre-eclampsia, and systemic lupus erythematosus. *American Journal of Reproductive Immunology*, **28**, 183–7.

Ahmed, S.A. and Talal, N. (1990). Sex hormones and the immune system—part 2. Animal data. *Baillière's Clinical Rheumatology*, **4**, 13–21.

Avrech, O.M. *et al*. (1992). Raynaud's phenomenon and peripheral gangrene complicating scleroderma in pregnancy—diagnosis and management. *British Journal of Obstetrics and Gynaecology*, **99**, 850–1.

Bagger, P.V. *et al*. (1993). Anticardiolipin antibodies (IgG and IgA) in women with recurrent fetal loss correlate to clinical and serological characteristics of SLE. *Acta Obstetrica et Gynecologica Scandinavica*, **72**, 465–9.

Barna, L.K. and Triplett, D.A. (1991). A report on the First International Workshop for Lupus Anticoagulant Identification. *Clinical and Experimental Rheumatology*, **9**, 557–67.

Barton, J.R. and Sibai, B.M. (1991). Low-dose aspirin to improve perinatal outcome. *Clinical Obstetrics and Gynecology*, **34**, 251–61.

Bellamy, N. and Dewar, C.L. (1990). Relapsing polychrondritis in pregnancy. *Journal of Rheumatology*, **17**, 1525–6.

Bellamy, N., Grigor, R.R., and Naden, R.P. (1991). Arthritis in pregnancy. In *Prognosis in the rheumatic diseases* (ed. N. Bellamy), Ch. 14, pp. 279–319. Kluwer Academic Publishers, Dordrecht.

Benifla, J.L. *et al*. (1992). Predictive value of uterine artery velocity waveforms in pregnancies complicated by systemic lupus erythematosus and the antiphospholipid syndrome. *Fetal Diagnostic Therapeutics*, **7**, 195–202.

Biesenbach, G., Stoger, H., and Zazgornik, J. (1991). Successful pregnancy of twins in a renal transplant patient with Wegener's granulomatosis. *Nephrology Dialysis and Transplantation*, **6**, 139–40.

Bijlsma, W.J. and Van den Brink, H.R. (1992). Estrogens and rheumatoid arthritis. *American Journal of Reproductive Immunology*, **28**, 231–4.

Bijlsma, W.J., Huber-Brunning, O., and Thijssen, J.H.H. (1987). Effect of estrogen treatment on clinical and laboratory manifestations of rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **46**, 777–9.

Birdsall, M., Pattison, N., and Chamley, L. (1992). Antiphospholipid antibodies in pregnancy. *Australian and New Zealand Journal of Obstetrics and Gynecology*, **32**, 328–30.

Black, C.M. (1990). Systemic sclerosis and pregnancy. *Baillière's Clinical Rheumatology*, **4**, 105–24.

Boyle, J.A. *et al*. (1966). Serum uric acid and levels in normal pregnancy with observations on the renal excretion of urate in pregnancy. *Journal of Clinical Pathology*, **19**, 501–3.

Branch, D.W. *et al*. (1985). Obstetric complications associated with the lupus anticoagulant. *New England Journal of Medicine*, **313**, 1322–6.

Brennan, P. and Silman, A. (1994). Breast-feeding and the onset of rheumatoid arthritis. *Arthritis and Rheumatism*, **37**, 808–13.

Buchanan, R.R. *et al*. (1989). Antiphospholipid antibodies in the connective tissue diseases: their relation to the antiphospholipid syndrome and forme fruste disease. *Journal of Rheumatology*, **16**, 757–61.

Buchanan N.M. *et al*. (1992). A study of 100 high risk lupus pregnancies. *American Journal of Reproductive Immunology*, **28**, 192–4.

Burkett, G. (1989). Lupus nephropathy and pregnancy. *Clinics in Obstetrics and Gynecology*, **28**, 310–23.

Buyon, J.P. *et al*. (1992). Activation of the alternative complement pathway accompanies disease flares in systemic lupus erythematosus during pregnancy. *Arthritis and Rheumatism*, **35**, 55–61.

Catulo, M. *et al*. (1984). Preliminary results of serum androgen level testing in men with rheumatoid arthritis. (Letter.) *Arthritis and Rheumatism*, **27**, 958.

Cecere, F.A. and Persellin, R.H. (1981). The interaction of pregnancy and the rheumatic diseases. *Clinics in the Rheumatic Diseases*, **7**, 747–68.

Chamley, L.W., Pattison, N.S., and McKay, E.J. (1991). Separation of lupus anticoagulant from anticardiolipin antibodies by ion-exchange and gel filtration chromatography. *Haemostasis*, **21**, 25–9.

Chamley, L.W., McKay, E.J., and Pattison, N.S. (1993). Inhibition of heparin/antithrombin III cofactor activity by anticardiolipin antibodies: a mechanism for thrombosis. *Thrombosis Research*, **71**, 103–11.

Conaghan, P. *et al*. (1993). A uniquie presentation of multicentre reticulohistiocytosis in pregnancy. *Arthritis and Rheumatism*, **36**, 269–72.

Da Silva, J.A. and Hall, G.M. (1992). The effects of gender and sex hormones on outcome in rheumatoid arthritis. *Baillière's Clinical Rheumatology*, **6**, 196–219.

Da Silva, J.A. and Spector, T.D. (1992). The role of pregnancy in the course and aetiology of rheumatoid arthritis. *Clinical Rheumatology*, **11**, 189–94.

de Miguel, E. *et al*. (1992). Adult Still's disease and pregnancy. *Journal of Rheumatology*, **19**, 498.

Derksen, R.H. (1991). Systemic lupus erythematosus and pregnancy. *Rheumatology International*, **11**, 121–5.

Einstein, B.I. and Masi, A.T. (1981). Disseminated gonococcal infection (DGI) and gonococcal arthritis (GCA): 1. Bacteriology, epidemiology, host factors, pathogen factors and pathology. *Seminars in Arthritis and Rheumatism*, **10**, 155–72.

Englert, H. *et al*. (1992). Reproductive function prior to disease onset in women with scleroderma. *Journal of Rheumatology*, **19**, 1575–9.

Erlendsson, K. *et al*. (1993). Relation of antiphospholipid antibody and placental bed inflammatory vascular changes to the outcome of pregnancy in successive pregnancies of 2 women with systemic lupus erythematosus. *Journal of Rheumatology*, **20**, 1779–85.

Felbo, M. and Snorrason, E. (1961). Pregnancy and the place of therapeutic abortion in rheumatoid arthritis. *Acta Obstetrica et Gynecologica Scandinavica*, **40**, 116–26.

Ferro, D. *et al*. (1992). Methods for detecting lupus anticoagulants and their relation to thrombosis and miscarriage in patients with systemic lupus erythematosus. *Journal of Clinical Pathology*, **45**, 332–8.

Fields, C.L. *et al*. (1991). Wegener's granulomatosis complicated by pregnancy. A case report. *Journal of Reproductive Medicine*, **36**, 463–6.

Ginsberg, J.S. *et al*. (1992). Relationship of antiphospholipid antibodies to pregnancy loss in patients with systemic lupus erythematosus: a cross-sectional study. *Blood*, **80**, 975–80.

Giordano, M. *et al*. (1985). Pregnancy and systemic sclerosis. *Arthritis and Rheumatism*, **28**, 237–8.

Gonzales-Buritica, H., Khamashta, M.D., and Hughes, G.R.V. (1988). Hypothesis: is there a role for phospholipases in the anti-cardiolipin syndrome? *Clinical and Experimental Rheumatology*, **6**, 341–2.

Gran, J.T. and Husby, G. (1992). Ankylosing spondylitis. Current drug treatment. *Drugs*, **44**, 585–603.

Guzman, E. *et al*. (1992). Uterine-umbilical artery Doppler velocimetry in pregnant women with systemic lupus erythematosus. *Journal of Ultrasound Medicine*, **11**, 275–81.

Hargraves, M.N., Richmond, M., and Morton, R. (1948). Presentation of two bone marrow elements: the 'tart-cell' and the 'L.E.' cell. *Proceedings of Staff Meetings of the Mayo Clinic*, **23**, 25–8.

Harris, E.N. (1990). Maternal autoantibodies and pregnancy—1: the antiphospholipid antibody syndrome. *Baillière's Clinical Rheumatology*, **4**, 53–68.

Harris, E.N. and Spinato J.A. (1991). Should anticardiolipin tests be performed in otherwise healthy pregnant women? *American Journal of Obstetrics and Gynecology*, **165**, 1271–7.

Hayslett, J.P. (1991). Maternal and fetal complications in pregnant women with systemic lupus erythematosus. *American Journal of Kidney Disease*, **17**, 123–6.

Hayslett, J.P. (1992). The effect of systemic lupus erythematosus on pregnancy and pregnancy outcome. *American Journal of Reproductive Immunology*, **28**, 199–204.

Hazes, J.M. *et al*. (1990). Pregnancy and the risk of developing rheumatoid arthritis. *Arthritis and Rheumatism*, **33**, 1770–5.

Hench, P.S. (1938). Amelioration effect of pregnancy on chronic atrophic (infectious) rheumatoid arthritis, fibrositis and intermittent hydroarthrosis. *Proceedings of Staff Meetings of the Mayo Clinic*, **13**, 161–7.

Hench, P.S. *et al*. (1949*a*). The effect of a hormone of the adrenal cortex (17-hydroxy-11-dehydrocorticosterone: compound E) and of pituitary adrenocorticotrophic hormone on rheumatoid arthritis. *Proceedings of Staff Meetings of the Mayo Clinic*, **24**, 181–97.

Hench, P.S. *et al*. (1949*b*). The effects of the adrenal cortical hormone 17-hydroxy-11-dehydrocorticosterone (compound E) in the acute phase of rheumatic fever: preliminary report. *Proceedings of Staff Meetings of the Mayo Clinic*, **24**, 277–97.

Hirahara, F. *et al*. (1986). Collagen-induced arthritis and pregnancy in mice: the effects of pregnancy on collagen-induced arthritis and the high incidence of infertility in arthritic female mice. *American Journal of Reproductive Immunology and Microbiology*, **11**, 44–54.

Hopkinson, N.D. and Powell, R.J. (1992). Classical complement activation induced by pregnancy: implications for management of connective tissue diseases. *Journal of Clinical Pathology*, **45**, 66–7.

Husby, G., Ostensen, M., and Gran, J.T. (1988). Ankylosing spondylitis and pregnancy. *Clinical and Experimental Rheumatology*, **6**, 165–7.

Imai, F. *et al*. (1991). A case of maternal subclinical Sjögren's syndrome associated with congenital heart block. *Japanese Journal of Medicine*, **30**, 39–42.

Ishii, N. *et al*. (1991). Dermatomyositis and pregnancy. Case report and review of the literature. *Dermatologica*, **183**, 146–9.

James, W.H. (1993). Rheumatoid arthritis, the contraceptive pill, and androgens. *Annals of the Rheumatic Diseases*, **52**, 470–4.

Jansson, L. and Holmdahl, R. (1989). Oestrogen suppression of collagen arthritis IV. Progesterone alone does not affect the course of arthritis but enhances the oestrogen-mediated therapeutic effect. *Journal of Reproductive Immunology*, **15**, 141–50.

Jara-Quezada, L., Graef, A., and Lavalle, C. (1991). Prolactin and gonadal hormones during pregnancy in systemic lupus erythematosus. *Journal of Rheumatology*, **18**, 349–53.

Jones, H.W. *et al*. (1991). Anticardiolipin antibodies in patients from Malaysia with systemic lupus erythematosus. *Annals of the Rheumatic Diseases*, **50**, 173–5.

Julkunen, H. *et al*. (1993). Fetal outcome in lupus pregnancy: a retrospective case–control study of 242 pregnancies in 112 patients. *Lupus*, **2**, 125–31.

Kaaja, R., Julkunen, H., and Viinikka, L. (1993 *a*). Production of prostacyclin and thromboxane in lupus pregnancies: effect of small dose of aspirin. *Obstetrics and Gynecology*, **81**, 327–31.

Kaaja, R. *et al*. (1993*b*). Intravenous immunoglobulin treatment of pregnant patients with recurrent pregnancy losses associated with antiphospholipid antibodies. *Acta Obstetrica et Gynecologica Scandinavica*, **72**, 63–6.

Kean, W.F. and Buchanan, W.W. (1990). Pregnancy and rheumatoid disease. In *Pregnancy and the rheumatic diseases* (ed. A.L. Parke), pp. 125–40. *Baillière's Clinical Rheumatology*, Chapter 8. Baillière Tindall, London.

Kerslake, S. *et al*. (1992). Early Doppler studies in lupus pregnancy. *American Journal of Reproductive Immunology*, **28**, 172–5.

Khamashta, M.A. and Wallington, T. (1991). Management of the antiphospholipid syndrome. *Annals of the Rheumatic Diseases*, **50**, 959–62.

Klipple, G.L. and Cecere, F.A. (1989). Rheumatoid arthritis and pregnancy. *Rheumatic Disease Clinics of North America*, **15**, 213–39.

Kobayashi, S. *et al*. (1992). Immunoadsorbent plasmapheresis for a patient with antiphospholipid syndrome during pregnancy. *Annals of the Rheumatic Diseases*, **51**, 339–401.

Kordich, L.C. *et al*. (1992). Natural inhibitors of blood coagulation and fibrinolysis in patients with lupus anticoagulant. *Blood Coagulation and Fibrinolysis*, **3**, 765–71.

Kutteh, W.H. *et al*. (1993). Association of anticardiolipin antibodies and pregnancy loss in women with systemic lupus erythematosus. *Fertility and Sterilization*, **60**, 449–55.

Lahita, R.G. (1990). Sex hormones and the immune system—part I. Human data. *Baillière's Clinical Rheumatology*, **4**, 1–12.

Lansink, M., de Boer, A., Dijkmans, B.A., and vanden Brouke, J.P. (1993). The onset of rheumatoid arthritis in relation to pregnancy and childbirth. *Clinical and Experimental Rheumatology*, **11**, 171–4.

Le Loet, X. *et al*. (1993). Adult onset Still's disease and pregnancy. *Journal of Rheumatology*, **20**, 1158–61.

Lee, F.I. and Loeffler, E.E. (1962). Gout and pregnancy. *British Journal of Obstetrics and Gynaecology*, **69**, 299–304.

Levy, M., Spino, M., and Read, S.E. (1991). Colchicine: a state-of-the-art review. *Pharmacotherapy*, **11**, 196–211.

Lockshin, M.D. *et al*. (1984). Lupus pregnancy: case control study demonstrating absence of lupus exacerbation during and after pregnancy. *American Journal of Medicine*, **77**, 893–8.

Lockshin, M.D. *et al*. (1987). Antibody to cardiolipin, lupus anticoagulant and fetal death. *Journal of Rheumatology*, **14**, 259–62.

Lockshin, M.D., Druzin, M.L., and Qamar, T. (1989). Prednisone does not prevent recurrent fetal death in women with antiphospholipid antibody. *American Journal of Obstetrics and Gynecology*, **60**, 439–43.

Longstreth, P.L., Malinak, L.R., and Hill, C.S. (1973). Transient osteoporosis of the hip in pregnancy. *Obstetrics and Gynecology*, **41**, 563.

Lopez, L.R. *et al*. (1992). Clinical significance of immunoglobulin A versus immunoglobulins G and M anticardiolipin antibodies in patients with systemic lupus erythematosus. Correlation with thrombosis, thrombocytopenia, and recurrent abortion. *American Journal of Clinical Pathology*, **98**, 449–54.

Manthorpe, T. and Manthorpe, R. (1992). Congenital complete heart block in children of mothers with primary Sjögren's syndrome. (Letter.) *Lancet*, **340**, 1359–60.

Mantle, M.J., Greenwood, R.M., and Currey, H.L.F. (1977). Backache in pregnancy. *Rheumatology and Rehabilitation*, **16**, 95–101.

Many, A. *et al*. (1992). Treatment of patients with antiphospholipid antibodies during pregnancy. *American Journal of Reproductive Immunology*, **28**, 216–18.

Marks, S.H., Barnett, M., and Calin, A. (1983). Ankylosing spondylitis in women and men: a case–control study. *Journal of Rheumatology*, **10**, 624–8.

Martinez-Cordero, E. *et al*. (1982). Systemic lupus erythematosus and amenorrhea. Hormonal profile and pathogenic significance. (Abstract.) *Arthritis and Rheumatism*, **25**, S81.

Mattson, R. *et al*. (1991). Maintained pregnancy levels of oestrogen afford complete protection from post-partum exacerbation of collagen-induced arthritis. *Clinical and Experimental Immunology*, **85**, 41–7.

McAlindon, T. *et al*. (1993). Environmental factors predicting nephritis in systemic lupus erythematosus. *Annals of the Rheumatic Diseases*, **52**, 720–4.

McHugh, N.J. and Laurent, M.R. (1989). The effect of pregnancy on the onset of psoriatic arthritis. *British Journal of Rheumatology*, **28**, 50–2.

McHugh, N.J., Reilly, P.A., and McHugh, L.A. (1989). Pregnancy outcome and autoantibodies in connective tissue disease. *Journal of Rheumatology*, **16**, 2–6.

McLennan, G., Oats, J.N., and Walstab, J.E. (1987). Survey of hand symptoms in pregnancy. *Medical Journal of Australia*, **147**, 542–4.

McNeill, M.E. (1988). Multiple pregnancy-induced remissions of psoriatic arthritis: case report. *American Journal of Obstetrics and Gynecology*, **159**, 896–7.

Melvin, J.L., Burnett, C.M., and Johnson, E.W. (1969). Median nerve conduction in pregnancy. *Archives of Physical Medicine*, **50**, 75–80.

Mintz, G. and Rodriguez-Alvarez, E. (1989). Systemic lupus erythematosus. *Rheumatic Disease Clinics of North America*, **15**, 255–72.

Molta, C. *et al*. (1993). Childhood-onset systemic lupus erythematosus: antiphospholipid antibodies in 37 patients and their first-degree relatives. *Pediatrics*, **92**, 849–53.

Nelson, J.L. *et al*. (1992*a*). Pregnancy outcome in women with rheumatoid arthritis before disease onset. *Journal of Rheumatology*, **19**, 18–21.

Nelson, J.L. *et al*. (1992*b*). Remission of rheumatoid arthritis during pregnancy and maternal–fetal class II alloantigen disparity. *American Journal of Reproductive Immunology*, **28**, 226–7.

Nelson, J.L. *et al*. (1993). Maternal–fetal disparity in HLA class II alloantigens and the pregnancy-induced amelioration of rheumatoid arthritis. *New England Journal of Medicine*, **329**, 466–71.

Nicklin, J.L. (1991). Systemic lupus erythematosus and pregnancy at the Royal Women's Hospital, Brisbane 1979–1989. *Australian and New Zealand Journal of Obstetrics and Gynecology*, **31**, 128–33.

Nordin, B.E.C. and Roper, A. (1955). Post-pregnancy osteoporosis—a syndrome? *Lancet*, **i**, 431.

Ohno, T., Imai, A., and Tamaya, T. (1992). Successful outcomes of pregnancy complicated with dermatomyositis. Case reports. *Gynecologic and Obstretic Investigation*, **35**, 187–9.

Ostensen, M. (1984). The influence of pregnancy and blood parameters in patients with rheumatoid disease. *Scandinavian Journal of Rheumatology*, **17**, 67–70.

Ostensen, M. (1988). Pregnancy in psoriatic arthritis. *Scandinavian Journal of Rheumatology*, **17**, 67–70.

Ostensen, M. (1991). Pregnancy in patients with a history of juvenile rheumatoid arthritis. *Arthritis and Rheumatism*, **34**, 881–7.

Ostensen, M. (1992). The effect of pregnancy on ankylosing spondylitis, psoriatic arthritis, and juvenile rheumatoid arthritis. *American Journal of Reproductive Immunology*, **28**, 235–7.

Ostensen, M., von Schoultz, B., and Husby, G. (1983). Comparison between serum alpha 2-pregnancy-associated globulin and activity of rheumatoid arthritis and ankylosing spondylitis during pregnancy. *Scandinavian Journal of Rheumatology*, **12**, 315–18.

Ostensen, M., Marhaug, G., and Husby, G. (1985). Amyloid-related serum protein (SAA) during and after pregnancy in healthy women and women with rheumatoid disease. *Acta Pathologica, Microbiolgica et Immunologica Scandinavica, Section C, Immunology*, **93**, 1–5.

Out, H.J. *et al*. (1991*a*). Prevalence of antiphospholipid antibodies in patients with fetal loss. *Annals of the Rheumatic Diseases*, **50**, 553–7.

Out, H.J. *et al*. (1991*b*). Histopathological findings in placentae from patients with intra-uterine fetal death and anti-phospholipid antibodies. *European Journal of Obstetrics and Gynecology, Reproductive Biology*, **41**, 179–86

Out, H.J. *et al*. (1992). A prospective, controlled multicenter study on the obstetric risks of pregnancy in women with antiphospholipid antibodies. *American Journal of Obstetrics and Gynecology*, **167**, 26–32.

Parry, G., Goudevenos, J., and Williams, D.O. (1992). Coronary thrombosis postpartum in a young woman with Still's disease. *Clinical Cardiology*, **15**, 305–7.

Persellin, R.H. (1981). Inhibitions of inflammatory and immune responses in pregnancy serum. *Clinics in the Rheumatic Diseases*, **7**, 769–80.

Petri, M. (1994). Systemic lupus erythematosus and pregnancy. *Rheumatic Disease Clinics of North America*, **20**, 87–118.

Petri, M., Howard, D., and Repke, J. (1991). Frequency of lupus flare in pregnancy. The Hopkins Lupus Pregnancy Centre experience. *Arthritis and Rheumatism*, **34**, 1538–45.

Pistiner, M. *et al*. (1991). Lupus erythematosus in the 1980's: a survey of 570 patients. *Arthritis and Rheumatism*, **21**, 55–64.

Pope, R.M. (1990). Immunoregulatory mechanisms present in the maternal circulation during pregnancy. In *Baillière's Clinical Rheumatology* (ed. A.L. Parke), pp. 33–52. Baillière Tindall, London.

Pritchard, M.H. (1992). An examination of the role of female hormones and pregnancy as risk factors for rheumatoid arthritis, using a male population as control group. *British Journal of Rheumatology*, **31**, 395–9.

Propper, D.J. *et al*. (1991). Breast vasculitis in association with breast gigantism in a pregnant patient with systemic lupus erythematosus. *Annals of the Rheumatic Diseases*, **50**, 577–8.

Ramsey-Goldman, R. *et al*. (1992). Previous pregnancy outcome is an important determinant of subsequent pregnancy outcome in women with systemic lupus erythematosus. *American Journal of Reproductive Immunology*, **28**, 195–6.

Ramsey-Goldman, R. *et al*. (1993*a*). Pregnancy outcome in women with systemic lupus erythematosus treated with immunosuppressive drugs. *Journal of Rheumatology*, **20**, 1152–7.

Ramsey-Goldman, R. *et al*. (1993*b*). Pregnancy outcome and anticardiolipin antibody in women with systemic lupus erythematosus. *American Journal of Epidemiology*, **138**, 1057–69.

Resnick, D. (1988). Disorders of other endocrine glands and of pregnancy. In *Diagnosis of bone and joint disorder* (2nd edn) (ed. D. Resnick and G. Niwayama), pp. 2287–317. W.B. Saunders Co., Philadelphia.

Reyes, H. *et al*. (1994). Antiphospholipid antibodies: a critique of their heterogeneity and hegemony. *Seminars in Thrombosis and Hemostasis*, **20**, 89–100.

Rook, G.A. *et al*. (1991). Changes in IgG glycoform levels are associated with remission of arthritis during pregnancy. *Journal of Autoimmunity*, **4**, 779–94.

Rose, H.M. *et al*. (1948). Differential agglutination of normal and sensitized sheep erythrocytes by sera of patients with rheumatoid arthritis. *Proceedings of the Society of Experimental Biology*, **68**, 1–6.

Rubbert, A. *et al*. (1992). Pregnancy course and complications in patients with systemic lupus erythematosus. *American Journal of Reproductive Immunology*, **28**, 205–7.

Said, P.B., Martinuzzo, M.E., and Carreras, L.O. (1992). Neutralization of lupus anticoagulant activity by human immunoglobulin in vitro. *Nouvelle Revue Francaise Hematologie*, **34**, 37–42.

Sany, J. *et al*. (1986). Treatment of rheumatoid polyarthritis with IgG eluted from the placenta. Results of an open study on 31 patients. *Revue du Rhumatisme et des Maladies Osteo-Articulaires*, **53**, 473–9.

Shibata, S. *et al*. (1992). Risk factors in the pregnancy of patients with systemic lupus erythematosus: association of hypocomplementaemia with poor prognosis. *Annals of the Rheumatic Diseases*, **51**, 619–23.

Silman, A.J. (1992). Pregnancy and scleroderma. *American Journal of Reproductive Immunology*, **28**, 238–40.

Silman, A. and Black, C.M. (1988). Increased incidence of spontaneous abortion and infertility in women with scleroderma before disease onset: a controlled study. *Annals of the Rheumatic Diseases*, **47**, 441–4.

Silman, A., Kay, A., and Brennan, P. (1992). Timing of pregnancy in relation to the onset of rheumatoid arthritis. *Arthritis and Rheumatism*, **35**, 152–5.

Spector, T.D. and Da Silva, J.A. (1992). Pregnancy and rheumatoid arthritis: an overview. *American Journal of Reproductive Immunology*, **28**, 222–5.

Spector, T.D. and Silman, A. (1990). Is poor pregnancy outcome a risk factor in rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **49**, 12–14.

Steen, V.D. *et al*. (1989). Pregnancy in woman with systemic sclerosis. *Arthritis and Rheumatism*, **32**, 151–7.

Stein, G.H., Cantor, B., and Panush, R.S. (1980). Adult Still's disease associated with pregnancy. *Arthritis and Rheumatism*, **23**, 248–50.

Stone, M.S. *et al*. (1993). Cutaneous vasculitis in the newborn of a mother with cutaneous polyarteritis nodosa. *Journal of the American Academy of Dermatology*, **28**, 101–5.

Takaya, M. *et al*. (1991). Sjögren's syndrome and pregnancy. *Tokai Journal of Experimental and Clinical Medicine*, **16**, 83–8.

Talbott, J.H. (1957). Remarks on gout and gouty arthritis. *Bulletin of the Hospital for Joint Diseases*, **18**, 3–18.

Tambyraja, R.L. (1993). Fetal salvage in maternal systemic lupus erythematosus. *Annals of the Academy of Medicine* (*Singapore*), **22**, 634–7

Taylor, H.A., Bradford, S.A., and Patterson, S.P. (1966). Gonococcal arthritis in pregnancy. *Obstetrics and Gynecology*, **27**, 776–82.

Tozman, E.C.S., Urowitz, M.D., and Gladman, D.D. (1980). Systemic lupus erythematosus and pregnancy. *Journal of Rheumatology*, **1**, 624–32.

Triplett, D.A. (1989). Antiphospholipid antibodies and reproduction. *American Journal of Reproductive Immunology*, **21**, 123–31.

Trudinger, B.J. *et al*. (1988). Monitoring lupus anticoagulant-positive pregnancies with umbilical artery flow velocity wave forms. *Obstetrics and Gynecology*, **72**, 215–18.

Unger, A. *et al*. (1983). Disease activity and pregnancy associated alpha-2-glycoprotein in rheumatoid arthritis during pregnancy. *British Medical Journal*, **286**, 750–1.

Urowitz, M.B. and Gladman, D.D. (1982). Rheumatic diseases. In *Medical complications during pregnancy* (ed. G.M. Burrow and T.F. Farris), pp. 474–97. W.B. Saunders Co., Philadelphia.

Urowitz, M.B. *et al*. (1993). Lupus and pregnancy studies. *Arthritis and Rheumatism*, **36**, 1392–7.

Van der Linden, S.M. and Khan, M.A. (1984). The risk of ankylosing spondylitis in a HLA-B27 positive individual: a reappraisal. *Journal of Rheumatology*, **11**, 727–8.

Vianna, J.L. *et al*. (1992). Reassessing the status of antiphospholipid syndrome in systemic lupus erythematosus. *Annals of the Rheumatic Diseases*, **51**, 160–1.

Vyse, T., Luxon, L.M., and Walport, M.J. (1994). Audiovestibular manifestations of the antiphospholipid syndrome. *Journal of Laryngology and Otology*, **108**, 57–9

Waaler, E. (1940). On the occurrence of a factor in human serum activating the specific agglutination of sheep blood corpuscles. *Acta Pathologica et Microbiologica Scandinavica*, **17**, 172–88.

Walkinshaw, S.A. *et al*. (1994). *In utero* pacing for fetal congenital heart block. *Fetal Diagnostic Therapeutics*, **9**, 183–5.

Wallace, D.J., Goldfinger, D., and Klinenberg, J.R. (1987). Use of autologous pregnancy plasma to treat a flare of juvenile rheumatoid arthritis: case report and literature review. *Journal of Clinical Apheresis*, **3**, 216–18.

Whyte, A. and Williams, R.D. (1988). Bromocriptine suppresses postpartum exacerbation of collagen-induced arthritis. *Arthritis and Rheumatism*, **31**, 937–8.

Whyte, A., Hasan, Z., and Rutter, F. (1988). Anti-collagen antibody responses in DBA/I(H-29) mice associated with type II collagen-induced arthritis and the effect of pregnancy. *Journal of Reproductive Immunology*, **13**, 277–86.

Wilbur, A.C., Langer, B.G., and Spigos, D.G. (1988). Diagnosis of sacroiliac joint infection in pregnancy by magnetic resonance imaging. *Magnetic Resonance Imaging*, **6**, 341–3.

Williams, R.D. and Whyte, A. (1989). A comparison of anti-type II collagen antibody titers and isotype profiles in pregnant and virgin DBA/1 mice with collagen-induced arthritis. *Journal of Reproductive Immunology*, **15**, 229–39.

Wong, K.L., Chan, K.Y., and Lee, C.P. (1991). Outcome of pregnancy in patients with systemic lupus erythematosus. A prospective study. *Archives of Internal Medicine*, **151**, 269–73.

Yoshino, S. and Uchida, S. (1981). Sexual problems of women with rheumatoid arthritis. *Archives of Physical Medicine and Rehabilitation*, **62**, 122–3.

1.3.1.2 Antirheumatic drugs in pregnancy and lactation

W. Watson Buchanan and Walter F. Kean

The placenta was considered until the middle of this century as a protective barrier for the fetus against harmful environmental toxins. It became only too evident that this was not the case with the reports that maternal rubella infection (Gregg 1941) and thalidomide (McBride 1961) could seriously damage the developing fetus. Unfortunately the results of all experimental findings in animals cannot be extrapolated to humans (thalidomide had produced problems in animals). The fetus is at greatest risk during the first 3 months of gestation, and most drugs have the ability to cross the placenta into the fetal blood stream. There is also concern regarding the delayed effects of drugs following intrauterine exposure: the classic example being the increased risk of adenocarcinoma of the vagina in female fetuses exposed to diethylstilboestrol (Herbst 1981). Thus both short-term and long-term effects of maternal use of drugs is a matter of considerable concern. Unfortunately clinical data are woefully lacking with most drugs. Pregnancy has effects on the pharmacokinetics of antirheumatic drugs (Witter 1993), but, on the whole, these are not of clinical significance. Several excellent reviews of antirheumatic and antigout drugs in pregnancy are available (Berkowitz *et al*. 1986; Brooks and Needs 1989; Brooks and Needs 1990; Kean and Buchanan 1990; Bellamy *et al*. 1991; Buchanan *et al*. 1992; Briggs *et al*. 1994) and we acknowledge and refer to the works of these authors throughout this text.

Pharmacokinetics

The disposition of antirheumatic drugs can be influenced by the physiological changes of pregnancy. Gastric pH, alterations in mucus production, and delay in gastric emptying will affect absorption rates. The concomitant use of antacids and so-called gastric protective agents have the potential to influence the absorption of non-steroidal anti-inflammatory drugs (**NSAIDs**) and other analgesics (Kean and Buchanan 1987). Supplementary mineral compounds and iron preparations may influence the absorption of a NSAID (Day *et al*. 1984) or D-penicillamine (Osman *et al*. 1983). However, the evidence in the literature does not suggest that these influences on absorption are clinically significant (Harkness and Blake 1982; Kean and Buchanan 1987). The increased blood volume of pregnancy results in an apparent increase in the volume of distribution. There is a resultant lower serum albumin level which theoretically produces an increased free fraction of active drug. However, since both type I and type II hepatic biotransformation pathways and renal excretory pathways are intact, these physiological changes are unlikely to have a significant clinical influence on the pregnant patient (Buchanan and Kean 1987). Increased protein binding of salicylates has been recorded in neonatal plasma (84.4 per cent) compared with maternal plasma (77.9 per cent) (Levy 1975). Increase in hepatic biotransformation will influence drugs which are essentially metabolized by the liver. Serum cholinesterase activity is reduced during pregnancy, and this could influence the metabolism of aspirin and methylprednisolone. Drugs such as indomethacin, piroxicam, sulindac, and sulphasalazine, which undergo significant enterohepatic circulation, could have altered kinetic activity related to the relative cholestasis of pregnancy. Renal blood flow and creatinine clearance increase during early pregnancy, therefore drugs such as azapropazone which are primarily excreted by the kidneys could be cleared more rapidly in pregnant patients. However, although this is a theoretical issue, there is no evidence that this occurs in practice. (Krauer and Krauer 1983)

The pharmacokinetic changes in absorption, distribution, metabolism, and excretion by the mother have the potential to influence drug transfer to the fetus. The integrity of the maternal circulation and the rate of transfer across the placenta have the potential to influence the fetus. Levy (1975) has stated that the pharmacodynamics and pharmacokinetics, especially drug elimination by the fetus, have to be considered as important issues when investigating the effects of drugs on the pregnant patient and the fetus.

Analgesics

Notwithstanding the extensive media warnings with respect to the potential side-effects of drugs on the fetus, simple analgesics are frequently prescribed or purchased over the counter during pregnancy. (Barrow and Souder 1971; Levy *et al*. 1975a; Heinonen *et al*. 1977; Golden and Perman 1980; Mangurten and Benawra 1980; Golden *et al*. 1982; Aselton *et al*. 1985) Codeine and acetaminophen do not cross the placenta; it has not been confirmed if they cause fetal malformation (Tyson 1974; Mangurten and Benawra 1980; Golden *et al*. 1982).

Acetaminophen appears safe in short-term use, but long-term high doses have been reported as a possible cause of fatal kidney disease in the newborn (Char *et al*. 1975). The drug is probably not teratogenic, but a possible association with congenital dislocation of the hip and club foot has been described (Heinonen *et al*. 1977). Although acetaminophen has been used extensively by pregnant mothers, there is no evidence that this compound or its metabolites cause congenital abnormalities or fetal morbidity (Collins 1981; Golden *et al*. 1982; Lee 1985). Acetaminophen, unlike aspirin and NSAIDs poses no problem with increase in haemorrhage.

Dextropropoxyphene may cause congenital abnormalities. Dextropropoxyphene and codeine have been implicated in the cause of a neonatal drug withdrawal syndrome (Tyson 1974).

Aspirin

Chronic aspirin use during pregnancy increases the risk of anaemia, intra- and post-partum haemorrhage, and prolongs gestation and labour (Slone *et al*. 1976; Collins 1981; Lee 1985; Benigni *et al*. 1989). Anaemia during pregnancy is significantly increased in mothers receiving aspirin compared with controls (Bleyer and Breckenridge 1970). Low-dose aspirin has been employed to prevent pregnancy-induced hypertension and pre-eclampsia (Romero *et al*. 1988; Schiff *et al*. 1989) although recent studies have failed to prove its efficacy (Italian Study of Aspirin in Pregnancy 1993 ; Sibai *et al*. 1993; CLASP [Collaborative Low-dose Aspirin Study in Pregnancy] Group 1994). No toxic effects in the fetus have been observed with this therapy (Benigni *et al*. 1989; Sibai *et al*. 1989), or with the occasional use of aspirin (Collins 1981 ; Barton and Sibai 1991), although more studies are required to be absolutely certain on this point (Romero *et al*. 1988). Low-dose aspirin is also currently used to prevent fetal loss due to antiphospholipid antibodies in patients with systemic lupus erythematosus (Branch *et al*. 1985; Elder *et al*. 1988), perhaps by preventing antiphospholipid antibodies triggering the production of thromboxane over prostacyclin (Kaaja *et al*. 1993).

There is concern that high doses of aspirin may lead to intrauterine growth retardation, congenital salicylate intoxication, depressed albumin-binding capacity (but no increase in jaundice), an increase in still birth associated with antepartum haemorrhage, and an increase in perinatal mortality (Earle 1961; Arcilla *et al*. 1969; Turner and Collins 1975; Lynd *et al*. 1976; Shapiro *et al*. 1976; Aterman *et al*. 1980). Newborn infants have been described with diminished clotting ability, haemorrhages and a purpuric rash, and increased bleeding at circumcision (Haslam 1975; Rumack *et al*. 1981; Soller and Stander 1981; Stuart *et al*. 1982; Stuart 1983). Maternal aspirin ingestion 1 week prior to delivery has been associated with an increased incidence of intracranial haemorrhage in premature infants (Rumack *et al*. 1981). Aspirin readily crosses the placenta and at term has a higher concentration in the neonate than in the mother (Levy *et al*. 1975b). Non-acetylated salicylates are equally as effective as aspirin in rheumatoid arthritis (Preston *et al*. 1989) and are not associated with bleeding, and should be used routinely to control pain and inflammation in inflammatory joint disease (Bleyer 1974).

Laboratory animal research, with acetylsalicylic acid concentrations in excess of therapeutic doses used in humans, has resulted in fetal abnormalities, including learning difficulties, in the offspring (Butcher *et al*. 1972).

Whether aspirin causes congenital defects in humans remains controversial (Briggs *et al*. 1994). The fact that the Collaborative Perinatal Project which monitored 50 282 mother–child pairs, 14 864 who used aspirin during the first trimester, found no evidence of an increase in congenital defects (Slone *et al*. 1976) does not exclude the possibility that high doses may be teratogenic. Streissguth *et al*. (1987) reported lower intelligence quotients in children born of mothers who had been prescribed aspirin during their pregnancies, but this has not been confirmed by other workers (Klebanoff and Berendes 1988).

Failure of intrauterine devices to prevent conception has been attributed to the anti-inflammatory action of aspirin (Buhler and Papiernik 1983).

Indomethacin and sulindac

No ill effects on the fetus have been reported with the use of sulindac (Carlan *et al*. 1992).

Indomethacin rapidly crosses the placental barrier with resultant similar metabolite concentrations in the plasma of the mother and child. Serious fetal effects have been recorded with the use of indomethacin in the treatment of premature labour, especially primary pulmonary hypertension (Demandt *et al*. 1990; Besinger *et al*. 1991), patent ductus arteriosus constriction and tricuspid regurgitation (Moise *et al*. 1988; Eronen *et al*. 1991; Evans *et al*. 1992), and intracranial haemorrhage and necrotizing enterocolitis (Norton *et al*. 1993). Zuckerman *et al*. (1974) reported on 5 of 12 infants, whose mother had received indomethacin for premature labour, who died within 48 h of birth from respiratory distress. Wiquest *et al*. (1975) did not identify an increase in fetal mortality when indomethacin was used to treat premature labour. It has been suggested that the respiratory failure in the newborn might be a reflection of the prematurity and not necessarily the indomethacin-induced pulmonary hypertension from duct closure. Nevertheless, histological studies have identified increased amounts of smooth muscle in the pulmonary arterial wall of some of the children with pulmonary hypertension whose mothers received indomethacin (Manchester *et al*. 1976). Treatment of polyhydramnios has been reported to cause premature ductal closure (Kirshon *et al*. 1990a; Kirshon *et al*. 1990b; Buderus *et al*. 1993), and also renal failure (Moise 1991). A variety of serious complications has been recorded in fetuses of mothers treated with high doses of indomethacin, including hydrops, oliguric renal failure, gastrointestinal and intraventricular bleeding, and perforation of the terminal ilium (van Haesebrouck *et al*. 1988). However, other workers have reported no congenital fetal abnormalities in mothers who have been treated with indomethacin (Niebyl and Witter 1986), and patent ductus arteriosus has been described in infants born of mothers treated with indomethacin (Atad *et al*. 1987). Indomethacin, like aspirin, can both delay and prolong labour (Morales and Madhav 1993) and increase the risk of post-partum haemorrhage (Reiss *et al*. 1976). Indomethacin can also inhibit the effects of b-blocking drugs and induce maternal hypertension (Schoenfeld *et al*. 1989).

Diflunisal

Fetal abnormalities have not been reported with diflunisal use in humans but the drug has similar effects on parturition and closure of the ductus arteriosus as other NSAIDs. When used in higher doses than those recommended in humans, the drug has been shown to be teratogenic in laboratory animals (van Winzum and Verhaest 1979).

Phenylbutazone

Use of phenylbutazone in clinical practice is now restricted to management of patients with refractory seronegative spondylarthropathies who have failed to respond to other available NSAIDs. Phenylbutazone is not recommended in pregnancy because of its potent prostaglandin inhibitory properties and the potential for serious haematological adverse effects (Buchanan and Kean 1994).

Although congenital fetal abnormalities have been reported with maternal use of phenylbutazone, there is doubt that the relationship is causal (Kullander and Kallen 1976) and the evidence is currently in favour of both phenylbutazone and its metabolite, oxyphenbutazone, being non-teratogenic (Ostensen and Husby 1985).

Diclofenac

Brooks and Needs (1989) have stated that there are no reported fetal abnormalities in humans related to diclofenac; however minor fetal abnormalities in rats and in mice treated with high-dose diclofenac have been reported (Medoricawa *et al*. 1972).

Proprionic acid drugs

No ill effects on the fetus have been reported with the use of ibuprofen (Barry *et al*. 1984). There is no evidence that naproxen, fenoprofen, or ketoprofen are teratogenic in animals or in man but as with all NSAIDs there is the potential for inhibition of labour, prolongation of pregnancy, increase in pre- and post-partum haemorrhage, and premature closure of the ductus arteriosus with resultant pulmonary hypertension (Wilkinson 1980). Wilkinson *et al*. (1979) reported on a case of triplets delivered at 30 weeks who suffered from severe hypoxaemia. The mother had been taking naproxen to suppress labour. The ductus arteriosus was closed in the infants. High concentrations of naproxen (60 µg/ml) were found in the neonatal plasma. Fluid retention and hyponatraemia have been reported in an neonate whose mother took 5 g of naproxen 8 h prior to delivery. The child recovered completely with no observable sequelae (Alun-Jones *et al*. 1986).

Oxicam

Oxicams have been shown to delay parturition in laboratory animals (Wiseman 1985) but piroxicam and tenoxicam have not been shown to have any teratogenic effects in animals and there are no reports of fetal abnormalities in man.

Corticosteroids

No ill effects have been recorded in children born of mothers who received betamethasone during their pregnancies, at least in terms of cognitive and psychological development, motor development and school achievement, and socio-emotional function (MacArthur *et al*. 1981; MacArthur *et al*. 1982; Schmand *et al*. 1990;

Smolders-de Haas *et al*. 1990), although more infections were noted during the first few years of life (Smolders-de Haas *et al*. 1990). High maternal doses of prednisone have been reported to cause immunosuppression in newborn infants (Cote *et al*. 1974), although most had no ill effects (Cederqvist *et al*. 1977). Long-term studies of children whose mothers were treated with dexamethasone showed no ill effects (Wong *et al*. 1982; Collaborative Group on Antenatal Steroid Therapy 1984). A reduction in the incidence of premature closure of the ductus arteriosus has been claimed with the administration of betamethasone to mothers with premature babies (Waffarn *et al*. 1983; Morales *et al*. 1989), but no such effect has been observed in normal weight infants (Wasserstrum *et al*. 1989). An increase in still births has been reported with prednisone treatment of the mother by Warrell and Taylor (1968) , but has not been confirmed by other investigators. An increased risk of fetal death has been reported in mothers treated with betamethasone to prevent respiratory distress syndrome (Liggins and Howie 1972), but refuted by others (Nochimson and Petrie 1979). Low levels of cortisol have been observed in premature infants of mothers prescribed betamethasone within 48 h of delivery, although this corrected within 2 h of birth (Dorr *et al*. 1986). An increased incidence of hypoglycaemia has been noted in newborn infants whose mothers received betamethasone (Papageorgiou *et al*. 1979), and congenital cataracts have been attributed to maternal prednisone therapy (Kraus 1975). Leukaemoid reactions in newborn infants may occur with maternal betamethasone (Bielawski *et al*. 1978) and dexamethasone therapy (Otero *et al*. 1981; Anday and Harris 1982). Large doses of corticosteroid administered to the mother during the first trimester have been associated with cleft palate in the offspring (Bongiovani and McPadden 1960). Popert (1962) in a study of 10 patients treated with corticosteroids for rheumatic disease during pregnancy, found that there were no maternal or fetal abnormalities in 20 pregnancies. Grigor *et al*. (1977) reported that fetal immaturity occurred in the offspring of pregnant women with systemic lupus treated with corticosteroids.

Although there have been case reports linking congenital malformations with cortisone therapy in the mother, there are no such reports with prednisone, dexamethasone, and betamethasone. Prednisone therapy does not affect either the number or morphology of chromosomes (Jensen 1967) but high doses may cause damage to spermatogenesis, which, however, is reversible (Mancini *et al*. 1966). Betamethasone has been reported as a potentiating factor in a ritodrine-induced crisis in a mother (Gonen *et al*. 1982). There is scant information on the use of small doses of prednisone (10 mg daily or equivalent) in pregnancy, and although probably safe, should be avoided if at all possible. Corticosteroids should only be administered in pregnancy when it is considered absolutely necessary for the treatment of the mother.

Gold complexes

Gold complexes have been extensively used in pregnant mothers. There are reports of fetal abnormalities in infants born of mothers who received gold sodium thiomalate during pregnancy (Miyamoto *et al*. 1974; Rogers *et al*. 1980; Fuchs and Lippert 1986). Other authors (Cohen *et al*. 1981; Tarp and Graudal 1985) did not find that gold complexes were associated with fetal abnormalities. Gold concentrations have been found in infants at delivery that are similar to gold concentrations found in the mothers (Rocker and Henderson 1976; Richards 1977; Cohen *et al*. 1981), but no obvious adverse effects were identified in the infants. A small number of pregnant mothers who received the oral gold complex, auranofin, have been followed. No adverse effects were identified either in the mother or in the child, but in view of the potential immunosuppressant effects of the triethylphosphine ligand of the auranofin, it is our opinion that this compound should be avoided during pregnancy. Ostensen and Husby (1985) have suggested that intramuscular gold may be given with caution during pregnancy; but how does one monitor the effects on the fetus?

Sulphasalazine

Sulphasalazine is a conjugate of 5-aminosalicylic acid, a salicylate analogue, linked by an azo bond to sulphapyridine. Sulphasalazine results in a reduced sperm count in males but does not appear to affect female fertility. Sulphasalazine and its metabolites, including sulphapyridine, cross the placenta, and case reports of children born with congenital malformation whose mothers received the drug during pregnancy have been recorded (Jarnerot *et al*. 1981; Newman and Correy 1983). There does not appear to be an increased incidence of fetal abnormality when mothers have received sulphasalazine for the treatment of inflammatory bowel disease (Mogadam 1981; Vender and Spirol 1982; Newman and Correy 1983).

Although several reports have suggested that there might be a fetal risk with the use of sulphasalazine in pregnancy (Craxi and Pagliarello 1980; Newman and Correy 1983; Hoo *et al*. 1988) on the whole this does not appear to be clinically significant (Brooks and Needs 1989; Briggs *et al*. 1994). It should be noted that all available data are for inflammatory bowel disease not rheumatoid arthritis.

Aminoquinolines

Chloroquine phosphate and hydroxychloroqine sulphate are 4-amino-7-chloroquinoline compounds that were first synthesized in the earlier part of the twentieth century and used during the Second World War as a treatment for malaria. Chloroquine differs from hydroxychloroqine in that the n-ethyl portion of the chloroquine is b-hydroxylated, otherwise the structure is equivalent to the parent molecule. Antimalarial compounds cross the placenta and can accumulate in the fetus, although the concentrations are much lower than in maternal tissue (Ullberg *et al*. 1970). Chromosomal damage has been reported *in vitro* with antimalarial drugs (Neill *et al*. 1973) and fetal sensory neural hearing loss has been reported after exposure to chloroquine phosphate prescribed during the first trimester (Hart and Naunton 1964). Dr Anne Parke (1993) has reported that fetal loss in systemic lupus is not due to the antimalarial drug therapy, but to the disease process.

Some workers have reported no significant risk to the fetus (Ross and Garatsos 1974; Suhonen 1983) or increase in congenital abnormalities (Roubenoff *et al*. 1988), but the potential for accumulation of the drugs in the uveal tract (Ullberg *et al*. 1970) and to cause chromosomal damage (Neill *et al*. 1973) have led to almost uniform avoidance of their use in pregnancy (Roubenoff *et al*. 1988).

D-Penicillamine

D-Penicillamine can cross the placenta and is therefore potentially teratogenic (Lyle 1978; Rosa 1986). The reaction of penicillamine with an aldehyde or ketone results in water elimination and formation of a five-membered thiazolidine ring with the carbon atom of the carbonyl group and the attached organic group incorporated into the ring, bonding to the sulphur and nitrogen atoms. During collagen synthesis, penicillamine can inhibit or impair tropocollagen cross-linking between the allysine and hydroxyl lysine residues by thiazolidine formation between the aldehyde group of allysine and the amino and sulphur functions of D-penicillamine; the result is the formation of a thiazolidine ring.

Endres (1981) reported on 87 cases of patients who have used D-penicillamine during pregnancy. Two of the children had severe connective tissue defects. In one case, the mother had rheumatoid disease and received D-penicillamine at a dosage of 900 mg per day for 5 months during the pregnancy (Soloman *et al*. 1977), and in the other case, a patient with cystinuria received 2000 mg per day of D-penicillamine for 9 months of pregnancy (Mjolnerod *et al*. 1971). Toxicity appears to be dose related, i.e. less than 500 mg per day is unlikely to be associated with fetal abnormalities (Endres 1981). In general D-penicillamine taken in pregnancy has proven safe (Roubenoff *et al*. 1988) but there are case reports of cutis laxa and other connective tissue impairment in infants whose mothers received the drug during pregnancy (Mjolnerod *et al*. 1971; Soloman *et al*. 1977; Lyle 1978; Endres 1981; Harpey *et al*. 1983).

It is not possible from the above reports to establish a relative risk; however, the potential for teratogenicities is sufficiently serious to suggest that in a non-fatal disease such as rheumatoid arthritis, D-penicillamine should be withheld during pregnancy, especially since the majority of patients will improve during the gestation period. D-Penicillamine should also be withheld in cystinuric patients but high fluid volume and increased urinary pH should be maintained. In patients with Wilson's disease, the dosage of D-penicillamine should not be reduced during pregnancy as the disease process increases the risk of spontaneous abortion. In 1976, Lakatos and colleagues reported that D-penicillamine was of benefit in the treatment of hyperbilinrubinaemia due to haemolytic disease of the newborn (Lakatos *et al*. 1976). The author stated that the bilirubin was destroyed by a copper–penicillamine complex. This finding was not supported by biochemical evidence in the report and to our knowledge this study has neither been confirmed nor refuted.

The controversy whether D-penicillamine should be continued during pregnancy is ongoing, but probably should be avoided (Ostensen and Husby 1985; Miehle 1988).

Methotrexate

There is a considerable body of evidence that treatment with methotrexate in pregnancy may cause congenital abnormalities (Milunsky *et al*. 1968; Powell and Ekert 1971; Warkany 1981) and myelosuppression in the fetus (Okun *et al*. 1979; Pizzuto *et al*. 1980). The drug can cause chromosomal abnormalities (Schleuning and Clemm 1987), but whether these predict abnormalities in the fetus has not been determined. Methotrexate can reduce ovarian and testicular function, but is reversible, and subsequent successful pregnancies have been reported (Barnes and Link 1983) with no higher than expected frequency of congenital abnormalities in the fetus (Rustin *et al*. 1984). A patient treated for Reiter's syndrome over 5 years, subsequently fathered a normal infant (Perry 1983). Methotrexate should, if

possible, be avoided in pregnancy (Ostensen and Husby 1985).

Azathioprine

Immunosuppression of the newborn has been reported with doses prescribed for rheumatoid arthritis (Cote *et al*. 1974; De Witte *et al*. 1984), although reduction in dosage at 32 weeks of pregnancy reduces the risk (Davison *et al*. 1985). Infants that are small for their gestational age have been reported in mothers receiving azathioprine, but this was probably due to underlying disease. Chromosomal abnormalities have been described with the use of azathioprine but the clinical significance has not been determined (Registration Committee of the European Dialysis and Transplant Association 1980). Azathioprine appears to interfere with the effectiveness of intrauterine contraceptives (Zerner *et al*. 1981; Davison and Lindheimer 1982). Surprisingly most workers have found azathioprine to be relatively safe in pregnancy, even in large doses, with no ill effects on the fetus (Kirk 1991). Ostensen and Husby (1985) recommended that azathioprine might be given in pregnancy, but advised caution.

Cyclophosphamide

Although malformed newborns have been described as a result of the use of cyclophosphamide, healthy infants have also been born (Coates 1970; Sweet and Kinzie 1976; Kirshon *et al*. 1988). Many of the mothers who had malformed newborns also had received other antineoplastic drugs and radiation. Less damage to the fetus appears to occur when treatment is confined to the second and third trimester (Nicholson 1968; Dobbing 1977). Congenital abnormalities have also been reported with paternal use of cyclophosphamide (Russell *et al*. 1976; Evenson *et al*. 1984). Permanent amenorrhoea may result from the use of cyclophosphamide (Uldall *et al*. 1972), but normal pregnancies have been reported in patients who had had high doses of cyclophosphamide for the treatment of neoplasms (Lee *et al*. 1989). Azoospermia can occur in the male (Hinkes and Plotkin 1973), but is reversible (Watson *et al*. 1985). Congenital abnormalities have been reported in infants fathered by patients who had had azoospermia as a result of cyclophosphamide, but had recovered (Russell *et al*. 1976; Evenson *et al*. 1984). Chromosomal abnormalities have been described with the use of cyclophosphamide, but the clinical significance is not known (Tolchin *et al*. 1974). Nurses should not handle cyclophosphamide when pregnant. The consensus of opinion is that cyclophosphamide should be avoided in pregnancy (Ostensen and Husby 1985).

Cyclosporin

Abortions and fetal abnormalities (Pujals *et al*. 1989) and growth retardation (Lau and Scott 1985) have been described with the use of cyclosporin in pregnancy. However, successful pregnancies have also been reported (Doria *et al*. 1992). Although cyclosporin does not appear to pose a major risk to the fetus or to be a human teratogen (Briggs *et al*. 1994), the numbers of reports to date are few. Cyclosporin should therefore probably be avoided in pregnancy until more data are available.

Allopurinol

No adverse fetal effects have been attributable to allopurinol (Briggs *et al*. 1994). However, gout is uncommon in women of child-bearing age, and the drug, if it has to be used, should be prescribed in as low a dose as possible.

Probenecid

No congenital defects have been reported with the use of probenecid in pregnancy (Lee and Loeffler 1962). The comments above regarding allopurinol pertain also to probenecid.

Colchicine

Colchicine is used in gout, chondrocalcinosis, and familial Mediterranean fever. In a small number of human fetuses exposed to colchicine, no congenital abnormalities were reported (Sokal and Lessmann 1960; Zemer *et al*. 1976; Cohen *et al*. 1977; Levy and Yaffe 1978). Colchicine may produce azoospermia (Merlin 1972), but not in all men (Bremer and Paulsen 1976). There has been concern of the mutagenic effects on sperm causing a higher risk of trisomic offspring (Cestari *et al*. 1965; Ferreira and Buoniconti 1968). Levy *et al*. (1991a) in a recent state-of-the-art review conclude that colchicine should be avoided in pregnancy because of its cytogenic effects and reported association with Down's syndrome.

Lactation

It has been estimated that approximately 50 per cent of babies discharged from hospitals in the United States are breast fed, and this number appears to be increasing (Briggs *et al*. 1994). Despite the fact that the potential for harm to the suckling infant from maternal ingestion of drugs has long been recognized (Soranus of Ephesus in the second century AD warned nursing mothers to refrain from taking medicines and alcohol), there is at present a relative paucity of knowledge on the safety of maternally-ingested medicaments on the nursing infant. A nursing mother produces on average 600 ml of milk per day. The milk has a pH greater than 7.1, and consists of a suspension of proteins, mostly casein and lactalbumin, and fat, in a carbohydrate—mineral solution. Drugs excreted in milk may be bound either to the proteins or adsorbed or sequestered in the milk fat globules. Drugs pass through the ductal cells primarily by passive diffusion, the concentration achieved being dependent on the concentration gradient, lipid solubility, and degree of ionization of the drug, and on its binding to protein and milk fat globules. Despite the large number of antirheumatic drugs and their extensive use, few data are available concerning drug concentrations in human milk or the effects of these drugs on the suckling infant. Drugs for which no clinical or biochemical data are available include: cortisone, betamethasone, and dexamethasone; diflunisal, meclofenamate, and sulindac; penicillamine; and probenecid (Briggs *et al*. 1994). The use of antirheumatic drugs during lactation has been the subject of several reviews (Ostensen and Husby 1985; Roubenoff *et al*. 1988; Brooks and Needs 1989; Kean and Buchanan 1990; Bellamy *et al*. 1991; Briggs *et al*. 1994). The following is a brief summary of the various antirheumatic and antigout drugs in lactation.

Analgesics

Acetaminophen is excreted in breast milk in very small concentrations but has been considered a safe analgesic for use in nursing mothers (Bitzen *et al*. 1981). A single case of rash in a breast-feeding infant has been described (Matheson *et al*. 1985). Codeine is lipid soluble and could be present in breast milk, however like acetaminophen, it is considered by the Committee on Drugs of the American Academy of Pediatrics (1994) as compatible with breast feeding.

Aspirin and NSAIDs

Aspirin has a short half-life of 10 to 15 min and is readily converted to salicylate after ingestion. The latter is excreted into breast milk in low concentrations. Kinetic studies have shown that after administration of 450 to 650 mg of aspirin to the mother, 21 per cent of the maternal dose appears in the breast milk over a 24-h period (Berlin *et al*. 1980; Findlay *et al*. 1981). Infants have the ability to cleave salicyl phenolic glucuronide and absorb the free salicylic acid. Maximum salicylate concentrations are found in breast milk 2 h after the maternal maximal concentrations (Findlay *et al*. 1981). Since the elimination of salicylates does not follow first-order kinetics, but is dose dependent, the plasma half-life increases as the dose increases. Thus, with small oral doses the plasma half-life may be only 1 to 2 h, whereas with doses exceeding 3 g per day the half-life may be as long as 18 h. With low dosage of aspirin or other salicylates, i.e. less than 3 g per day, there is probably little risk to the suckling infant (Ostensen and Husby 1985; Committee on Drugs of the American Academy of Pediatrics 1994); but with doses in excess of 3 g per day, salicylate intoxication is a distinct possibility. There is, however, only one report of an infant developing salicylate toxicity (Clark and Wilson 1981). Kwan *et al*. (1978) have recommended that because of the long half-life of sulindac and its active metabolites, this medication should not be given to nursing mothers because of the risk of transfer of this drug to the fetal circulation. Similarly, drugs such as diflunisal that have a long half-life are not appropriate for pregnant mothers during lactation.

As previously mentioned no clinical or biochemical data exist on a number of NSAIDs. However, ibuprofen, naproxen, phenylbutazone and oxyphenbutazone, piroxicam, and tolmetin have all been considered compatible with breast feeding by the Committee on Drugs of the American Academy of Pediatrics (1994). Brooks and Needs (1990) have stated that diclofenac sodium has not been detected in human milk after single doses of 50 mg or after 1 week of treatment at 100 mg per day. The authors have stated that since only a small amount of diclofenac is present in breast milk and since the drug has a short half-life of 2 h, that it is suitable for use in lactating mothers. They also state that some of the metabolites have longer half-lives and may be present in the breast milk, but in small quantities. Ibuprofen was not detected in the breast milk of 12 pregnant patients receiving 1600 mg of ibuprofen over 24 h (Townsend *et al*. 1984). Small amounts of naproxen were recovered in breast milk after administration of 250 mg of naproxen to a pregnant mother, but less than 0.26 per cent of the maternal dose was identified in the infant

(Jamali and Stevens 1983). Minimal concentrations of fenoprofen and ketoprofen have been identified in breast milk (Rubin *et al*. 1984). Piroxicam use during lactation results in 1 to 3 per cent of the maternal plasma concentration being present in the breast milk (Ostensen 1983; Ostensen *et al*. 1988). Ostensen and Husby (1985) have also considered diclofenac, flufenamic acid, ketoprofen, and mefenamic acid to pose no risk. Seizures in a breast-fed infant whose mother was receiving indomethacin have been described (Eeg-Olofsson *et al*. 1978), although whether these could be attributed to the drug has been questioned (Fairhead 1978). Nevertheless, indomethacin has been considered compatible with breast feeding (Committee on Drugs of the American Academy of Pediatrics 1994).

Corticosteroids

Systemic corticosteroids, e.g. prednisone, enter breast milk in small quantities. Administration of 5 mg of prednisolone to the mother results in less than 0.23 per cent of the dose being present in the breast milk (McKenzie *et al*. 1975). Surprisingly there is no data on the effect of these drugs on the suckling infant. However, in the standard doses used to treat rheumatoid arthritis, i.e. 10 mg/day or less of prednisone or equivalent, there seems little chance of an infant receiving significant amounts (Brooks and Needs 1989; Committee on Drugs of the American Academy of Pediatrics 1994). There also seems little chance of an infant being breast fed by a mother ill with serious systemic lupus erythematosus or vasculitis who requires large doses of corticosteroids.

Gold complexes

Ostensen *et al*. (1986) estimated that suckling infants might receive 20 per cent of the maternal dose of gold complexes. Trace amounts of aurothioglucose have been identified in breast milk (Blau 1973; Ostensen *et al*. 1986). Rashes, hepatitis, nephritis, and blood disorders have been reported in infants (Blau 1973) and Ostensen *et al*. (1986) concluded that probably nursing is best avoided. However, the Committee on Drugs of the American Academy of Pediatrics (1994) concluded that breast feeding was compatible with treatment with gold sodium thiomalate.

Sulphasalazine

Unmetabolized sulphasalazine and 5-aminosalicylic acid are either present in milk in minute amounts or are undetectable, but sulphapyridine is present in 30 to 60 per cent of maternal serum concentrations (Janerot and Into-Malmberg 1979). One infant has been described who developed bloody diarrhoea while the mother, a slow acetylator, was receiving sulphasalazine (Branski *et al*. 1986). The acetylator status of the infant was not determined. Sulphasalazine should therefore be given to nursing mothers with caution (Committee on Drugs of the American Academy of Pediatrics 1994).

Hydroxychloroquine and chloroquine

Nation *et al*. (1984) estimated that a suckling infant would receive approximately one-third of the mother's daily dose of these antimalarial drugs, but other workers found very small amounts in milk (Ostensen *et al*. 1985). In a study of 27 pregnancies in which the mothers received antimalarial drugs during the first trimester, there was no evidence that the infants were in any way affected (Levy *et al*. 1991b). Despite these findings, it is best to avoid the use of antimalarial drugs during breast feeding if at all possible (Anderson 1991; Levy *et al*. 1991b; Parke 1993; Committee on Drugs of the American Academy of Pediatrics 1994).

Penicillamine

No data is available on the effects of maternal ingestion of penicillamine on the suckling infant. Since congenital abnormalities have been reported in infants of mothers receiving penicillamine (Lyle 1978), the drug is best avoided during breast feeding (Ostensen and Husby 1985).

Methotrexate

Only small amounts of methotrexate are excreted in milk (Johns *et al*. 1972). However, in view of the potential for immunosuppression, neutropenia, effects on growth, and carcinogenesis in the suckling infant, it would seem prudent to avoid prescribing the drug to nursing mothers (Committee on Drugs of the American Academy of Pediatrics 1994).

Azathioprine

Little information is available on the use azathioprine during breast feeding. Only small amounts of azathioprine are present in milk (Anderson 1977) and, as a consequence, the drug has been considered safe but should be given with caution (Bellamy *et al*. 1991).

Cyclophosphamide

Cyclophosphamide is only prescribed in rheumatic practice for two conditions: lupus glomerulonephritis and vasculitis. It is excreted in large quantities in breast milk (Wiernik and Duncan 1971). It is absolutely contraindicated during breast feeding (Committee on Drugs of the American Academy of Pediatrics 1994).

Cyclosporin

There are no reports on the use of cyclosporin during breast feeding, but it is not recommended (Committee on Drugs of the American Academy of Pediatrics 1994).

Allopurinol

Gout is relatively rare in premenopausal women, and no clinical data are available. Allopurinol is present only in very small amounts in milk, and only slightly higher concentrations of the metabolite, oxypurinol, are found. The drug is considered safe during breast feeding (Committee on Drugs of the American Academy of Pediatrics 1994).

Probenecid

No data is available on the use of probenecid during breast feeding.

Colchicine

Colchicine is present in very small amounts in breast milk (Milunsky and Milunsky 1991). The Committee on Drugs of the American Academy of Pediatrics (1994) considers colchicine compatible with breast feeding.

Box 1 Drugs and pregnancy

DO'S	DON'TS
Do not take any drug or medicinal agent when pregnant or if pregnancy is suspected, unless that medication has been recommended by a physician In general the only medications which can be permitted are those which are considered to be essential, e.g., insulin and other products for diabetes, thyroid replacement hormone, antiseizure medication, etc. Under appropriate circumstances drugs which are essential to cardiovascular function, e.g., digoxin, and to respiratory function, e.g., inhalers for asthma, may be permitted when considered necessary by a physician Antirheumatic drugs should not be taken (see don'ts) except under exceptional circumstances when musculoskeletal pain is severe and cannot be controlled by the use of local physical measures, such as hot packs, ice packs, and physical therapy	Antirheumatic medications should not be taken during pregnancy or if pregnancy is suspected It is the opinion of the authors that there is an inadequate evidence base on which to recommend the use of anti-inflammatory drugs, except in exceptional circumstances

Box 2 Antirheumatic drug use by lactating mothers

It is the opinion of the authors that there is insufficient evidence upon which to recommend the use of antirheumatic drugs during lactation. While the case reports and studies referrred to in this chapter suggest that the effects on the child would be minimal, it is the opinion of the authors that there is insufficient quantity of data to support unrestricted use of antirheumatic drugs for the lactating mother

By convention, most rheumatologists will prescribe NSAIDs to lactating mothers, if the clinical situation warrants the use of these drugs

Conclusion

It can be appreciated that we are woefully lacking in data on the effects of antirheumatic drugs on the fetus and the suckling infant. Many of the new drugs have no data either in pregnancy or during lactation. In general most reports are anecdotal or at best on small series of patients.

The decision to use these various antirheumatic and antigout drugs depends on the severity of the patient's illness and the availability of therapeutic alternatives. It is useful to remember that although alcohol has been considered dangerous to the breast-fed infant since antiquity, it was considered to be no more than a superstition in this century until two decades ago (Shaywitz 1978). In general it is wise to avoid all drugs or medicaments during lactation. None of the drugs recommended as compatible with breast feeding have adequate clinical data on adverse effects on breast-fed infants, and rely entirely on comparison of the concentrations of the drugs found in milk samples in relation to maternal serum.

A helpful simple guide is the 'rule of three' (Sackett *et al*. 1986), which is useful when no events of a particular kind are observed. The 'rule of three' states that if no events of a particular kind are observed in a study of *x* individuals, one can be 95 per cent certain that the event occurs no more than 3 divided by *x*. Thus, if 3000 patients are studied then one can be certain that any event that does not occur in any of these patients occurs less often than 3 times in 3000 exposed subjects, or that it has an incidence rate of less than 0.001. In most studies of antirheumatic drugs in pregnancy and lactation the number of patients studied is less than 100. Needless to say it has proved extremely difficult to estimate risk, especially as patients may have more than one disease and be taking other medication (Roubenoff *et al*. 1988). If there is a place for pharmacoepidemiological studies then it is in this area of clinical therapeutics. Large population registers and population databases are necessary to provide accurate numerator/denominator ratios. The answer can only come with a database of drugs used in pregnancy and lactation (Bellamy *et al*. 1991). Meanwhile it seems prudent to use the lowest possible dose of a drug during pregnancy and lactation to control symptoms, and to avoid completely those drugs which are not recommended. Important points to remember when prescribing drugs during pregnancy and lactation are summarized in Box 1 and Box 2.

Chapter References

Alun-Jones, E., Williams, J., and Clwyd, G. (1986). Hyponatremia and fluid retention in a neonatal associated with maternal naproxen overdosage. *Clinical Toxicology*, **24**, 257–60.

Anday, E.K. and Harris, M.C. (1982). Leukemoid reaction associated with antenatal dexamethasone administration. *Journal of Pediatrics*, **101**, 614–16.

Anderson, P.O. (1977). Drugs and breast feeding—a review. *Drug Intelligence and Clinical Pharmacology*, **11**, 208

Anderson, P.O. (1991). Drug use during breast feeding. *Clinical Pharmacology*, **10**, 594–624.

Arcilla, R.A., Thilenius O.G., and Ranniger K. (1969). Congestive heart failure from suspected ductal closure *in utero*. *Journal of Pediatrics*, **75**, 74–8.

Aselton, P. *et al*. (1985). First trimester drug use and congenital disorders. *Obstetrics and Gynecology*, **65**, 451–5.

Atad, J., Lissak, A., Rofe, A., and Abramovici, H. (1987). Patent ductus arteriosus after prolonged treatment with indomethacin during pregnancy: case report. *International Journal of Obstetrics*, **25**, 73–6.

Aterman, K., Holzbecker, M., and Ellenberger, H.A. (1980). Salicylate levels in a stillborn infant born to a drug-addicted mother, with comments on pathology and analytical methodology. *Clinical Toxicology*, **16**, 263–8.

Barnes, A.B. and Link, D.A. (1983). Childhood dermatomyositis and pregnancy. *American Journal of Obstetrics and Gynecology*, **46**, 335–6.

Barrow, M.V. and Souder, D.E. (1971). Propoxyphene and congenital malformations. *Journal of the American Medical Association*, **217**, 1551–2.

Barry, W.S., Meinzinger, M.M., and Howse, C.R. (1984). Ibuprofen overdose and exposure *in utero*: results from a postmarketing voluntary reporting system. *American Journal of Medicine*, **77**, 35–9.

Barton, J.R. and Sibai, B.M. (1991). Low-dose aspirin to improve perinatal outcome. *Clinical Obstetrics and Gynecology*, **34**, 251–61.

Bellamy, N., Grigor, R.R., and Naden, R.P. (1991). Arthritis in pregnancy. In *Prognosis in the rheumatic diseases* (ed. N. Bellamy), pp. 279–319. Kluwer Academic Publishers, Dordrecht.

Benigni, A. *et al*. (1989). Effect of low-dose aspirin on fetal and maternal generation of thromboxane by platelets in women at risk for pregnancy-induced hypertension. *New England Journal of Medicine*, **321**, 357–62.

Berkowitz, R.L., Coustan, D.R., and Mochizuki, T.K. (1986). *Handbook for prescribing medications during pregnancy* (2nd edn). Little, Brown, and Company, Boston.

Berlin, C.M., Pascuzzi, M.J., and Yaffe, S.J. (1980). Excretion of salicylate in human milk. *Clinical Pharmacology and Therapeutics*, **27**, 245.

Besinger, R.E., Niebyl, J.R., Keyes, W.G., and Johnson, T.R.B. (1991). Randomized comparative trial of indomethacin and ritodrine for the long-term treatment of preterm labour. *American Journal of Obstetrics and Gynecology*, **164**, 981–8.

Bielawski, D., Hiatt, I.M., and Hegyi, T. (1978). Betamethasone-induced leukaemoid reaction in pre-term infant. *Lancet*, **i**, 218–9.

Bitzen, P.O. *et al*. (1981). Excretion of paracetamol in human breast milk. *European Journal of Clinical Pharmacology*, **20**, 123–5.

Blau, S.P. (1973). Metabolism of gold during lactation. *Arthritis and Rheumatism*, **16**, 777–8.

Bleyer, W.A. (1974). Maternal ingested salicylates as a cause of neonatal hemorrhage. *Journal of Pediatrics*, **85**, 736–7.

Bleyer, W.A. and Breckenridge, R.T. (1970). Studies on the detection of adverse drug reactions in the newborn. II. The effects of prenatal aspirin on newborn hemostasis. *Journal of the American*

Medical Association, **213**, 2049–53.

Bongiovanni, A.M. and McPadden, A.J. (1960). Steroids during pregnancy and possible fetal consequences. *Fertility and Sterility*, **2**, 181–6.

Branch, D.W., Scott, J.R., Kochenour N.K., and Hershgold, E. (1985). Obstetric complications associated with the lupus anticoagulant. *New England Journal of Medicine*, **13**, 1322–6.

Branski, D. *et al*. (1986). Bloody diarrhea—a possible complication of sulfasalazine transferred through human breast milk. *Journal of Pediatric Gastroenterological Nutrition*, **5**, 317–7.

Bremer, W.J. and Paulsen, C.A. (1976). Colchicine and testicular function in man. *New England Journal of Medicine*, **294**, 1384–5.

Briggs, G.G., Freeman, R.K., and Yaffe, S.J. (1994). *Drugs in pregnancy and lactation* (4th edn). Williams and Wilkins, Baltimore.

Brooks, P.M. and Needs, C.J. (1989). The use of anti-rheumatic medication during pregnancy and in the puerperium. *Rheumatic Disease Clinics of North America*, **15**, 789–806.

Brooks, P.M. and Needs, C.J. (1990). Anti-rheumatic drugs in pregnancy and lactation. *Ballières Clinical Rheumatology*, **4**, 157–71.

Buchanan, W.W. and Kean, W.F. (1987). Implication of anti-rheumatic drug therapy in elderly patients with osteoarthritis. *Journal of Rheumatology*, **14** (Suppl. 4), 98–100.

Buchanan, W.W. and Kean W.F. (1994). Rare side effects in nonsteroidal anti-inflammatory drugs. In *Nonsteroidal anti-inflammatory drugs. Mechanisms and clinical uses* (2nd edn) (ed. A.J. Lewis and D.E. Furst), pp. 217–43. Marcel Dekker, New York.

Buchanan, W.W., Needs, C.J., and Brooks, P.M. (1992). Rheumatic diseases: the arthropathies. In *Principles and practice of medical therapy in pregnancy* (2nd edn) (ed. N. Gleicher), pp. 428–34. Appleton and Lange, Norwalk, Connecticut.

Buderus, S., Thomas, B., Fahnenstich, H., and Kowalewski, S. (1993). Renal failure in two preterm infants: toxic effect of prenatal maternal indomethacin treatment? *British Journal of Obstetrics and Gynaecology*, **100**, 97–8.

Buhler, M. and Papiernik, E. (1983). Successive pregnancies in women fitted with intrauterine devices who take antiinflammatory drugs. *Lancet*, **i**, 483.

Butcher, R.E., Vorhees, C.V., and Kimmmel, C.A. (1972). Learning impairment from maternal salicylate treatment in rats. *Nature*, **236**, 211.

Carlan, S.J., O'Brien, W.F., O'Leary, T.D., and Mastrogiannis, D. (1992). Randomized comparative trial of indomethacin and sulindac for the treatment of refractory preterm labour. *Obstetrics and Gynecology*, **79**, 223–8.

Cederqvist, L.L., Merkatz, I.R., and Litwin, S.D. (1977). Fetal immunoglobulin synthesis following maternal immunosuppression. *American Journal of Obstetrics and Gynecology*, **129**, 687–90.

Cestari, A.N. *et al*. (1965). A case of human reproductive abnormalities possibly induced by colchicine treatment. *Revista Brasileira de Biologia* (*Rio de Janeiro*), **25**, 253–6.

Char, V.C., Chandra, R., Fletcher, A.B., and Avery, G.B. (1975). Polyhydramnios and neonatal renal failure—a possible association with maternal acetaminophen ingestion. *Journal of Pediatrics*, **86**, 638–9.

Clark, J.H. and Wilson, W.G. (1981). A 16-day-old breast-fed infant with metabolic acidosis caused by salicylate. *Clinical Pediatrics* (*Philadelphia*), **20**, 53–4.

CLASP (Collaborative Low-dose Aspirin Study in Pregnancy) Group (1994). CLASP: a randomized trial of low-dose aspirin for the prevention and treatment of pre-eclampsia among 9364 pregnant women. *Lancet*, **343**, 619–29.

Coates, A. (1970). Cyclophosphamide in pregnancy. *Australian and New Zealand Journal of Obstetrics and Gynecology*, **10**, 33–4.

Cohen, D.L., Orzel, J., and Taylor, A. (1981). Infants of mothers receiving gold therapy. *Arthritis and Rheumatism*, **24**, 104–5.

Cohen, M.M., Levy, M., and Eliakim, M. (1977). A cytogenetic evaluation of long-term colchicine therapy in the treatment of familial Mediterranean fever (FMF). *American Journal of Medical Science*, **274**, 147–52.

Collaborative Group on Antenatal Steroid Therapy (1984). Effects of antenatal dexamethasone administration in the infant: long-term follow-up. *Journal of Pediatrics*, **104**, 259–67.

Collins, E. (1981). Maternal and fetal effects of acetaminophen and salicylates in pregnancy. *Obstetrics and Gynecology*, **58** (Suppl.), 57S–62S.

Committee on Drugs of the American Academy of Pediatrics (1994). The transfer of drugs and other chemicals into human milk. *Pediatrics*, **93**, 137–50.

Cote, C.J., Meuwissen, H.J., and Pickering, R.J. (1974). Effects on the neonate of prednisone and azathioprine administered to the mother during pregnancy. *Journal of Pediatrics*, **85**, 324–8.

Craxi, A. and Pagliarello, F. (1980). Possible embryotoxicity of sulfasalazine. *Archives of Internal Medicine*, **140**, 1674.

Davison, J.M. and Lindheimer, M.D. (1982). Pregnancy in renal transplant recipients. *Journal of Reproductive Medicine*, **27**, 613–21.

Davison, J.M., Dellagrammatikas, H., and Parkin, J.M. (1985). Maternal azathioprine therapy and depressed haemopoiesis in the babies of renal allograft patients. *British Journal of Obstetrics and Gynecology*, **92**, 233–9.

Day R.O. *et al*. (1984). Anti-rheumatic drug interactions. *Clinics in the Rheumatic Diseases*, **10**, 251–75.

Demandt, E. *et al*. (1990). Prenatal indomethacin toxicity in one member of monozygous twins; a case report. *European Journal of Obstetrics, Gynecology, and Reproductive Biology*, **35**, 267–9.

De Witte, D.B. *et al*. (1984). Neonatal pancytopenia and severe combined immunodeficiency associated with antenatal administration of azathiprine and prednisone. *Journal of Pediatrics*, **105**, 625–8.

Dobbing, J. (1977). Pregnancy and leukaemia. *Lancet*, **i**, 1155.

Doria, A. *et al*. (1992). Cyclosporin A in a pregnant patient affected with systemic lupus erythematosus. *Rheumatology International*, **12**, 77–8.

Dorr, H.G. *et al*. (1986). Antenatal betamethasone therapy: effects on maternal, fetal, and neonatal mineralocorticoids, glucocorticoids, and progestins. *Journal of Pediatrics*, **108**, 990–3.

Earle, R., Jr. (1961). Congenital salicylate intoxication—report of a case. *New England Journal of Medicine*, **265**, 1003–4.

Eeg-Olofsson, O., Malmros, I. Elwin, C.E., and Steen, B. (1978). Convulsions in a breast-fed infant after maternal indomethacin. *Lancet*, **ii**, 215.

Elder, M.G. *et al*. (1988). Low-dose aspirin in pregnancy. *Lancet*, **i**, 410.

Endres, W. (1981). D-Penicillamine in pregnancy—to ban or not to ban? *Klinische Wochenschrift*, **59**, 535–7.

Eronen, M. *et al*. (1991). The effects of indomethacin and a b-sympathomimetic agent on the fetal ductus arteriosus during treatment of premature labour: a randomized double-blind study. *American Journal of Obstetrics and Gynecology*, **164**, 141–6.

Evans, D.J., Kofinas, A.D., and King, K. (1992). Intraoperative amniocentesis and indomethacin treatment in the management of an immature pregnancy with completely dilated cervix. *Obstetrics and Gynecology*, **79**, 881–2.

Evenson, D.P. *et al*. (1984). Male reproductive capacity may recover following drug treatment with the L-10 protocol for acute lymphocytic leukemia. *Cancer*, **53**, 30–6.

Fairhead, F.W. (1978). Convulsions in a breast-fed infant after maternal indomethacin. *Lancet*, **ii**, 576.

Ferreira, N.R. and Buoniconti, A. (1968). Trisomy after colchicine therapy. *Lancet*, **ii**, 1304.

Findlay, J.W. *et al*. (1981). Analgesic drugs in breast milk and plasma. *Clinical Pharmacology and Therapeutics*, **29**, 625–33.

Fuchs, U. and Lippert, T.H. (1986). Gold therapy and pregnancy. *Deutsche Medizine Wochenschrift*, **111**, 31–4.

Golden, N.L., King, K.C., and Sokol, R.J. (1982). Propoxyphene and acetaminophen: possible side effects on the fetus. *Clinical Pediatrics*, **21**, 752–4.

Golden, S.M. and Perman, K.I. (1980). Bilateral clinical anophthalmia: drugs as potential factors. *Southern Medical Journal*, **73**, 1404–7.

Gonen, R., Samberg, I., and Sharf, M. (1982). Hypertensive crisis associated with ritodrine infusion and betamethasone administration in premature labour. *European Journal of Obstetrics, Gynecology, and Reproductive Biology*, **13**, 129–32.

Gregg, N. McA. (1941). Congenital cataracts following German measles in the mother. *Transactions of the Ophthalmological Society of Australia*, **3**, 35–46.

Grigor, R.R. *et al*. (1977). Outcome of pregnancy in systemic lupus erythematosus. *Proceedings of the Royal Society of Medicine*, **70**, 99–100.

Harkness, J.A.L. and Blake, L. (1982). Penicillamine nephropathy and iron. *Lancet*, **ii**, 2368–9.

Harpey, J.P. *et al*. (1983). Cutis laxa and low serum zinc after antenatal exposure to penicillamine. *Lancet*, **ii**, 144–5.

Hart, C.N. and Naunton, R.F. (1964). The ototoxicity of chloroquine phosphate. *Archives of Otolaryngology and Head and Neck Surgery*, **80**, 407–12.

Haslam, R.R. (1975). Neonatal purpura secondary to maternal salicylism. *Journal of Pediatrics*, **86**, 653.

Heinonen, O.P., Slone, D., and Shapiro, S. (1977). *Birth defects and drugs in pregnancy*, pp. 286–95. Publishing Sciences Group, Littleton, MA.

Herbst, A.L. (1981). Diethylstilbestrol and other sex hormones during pregnancy. *Obstetrics and Gynecology*, **58** (Suppl), 35–40.

Hinkes, E. and Plotkin, D. (1973). Reversible drug-induced sterility in a patient with acute leukemia. *Journal of the American Medical Association*, **223**, 1490–1.

Hoo, J.J., Hadro, T.A., and Von Behren, P. (1988). Possible teratogenicity of sulfasalazine. *New England Journal of Medicine*, **318**, 1128.

Italian Study of Aspirin in Pregnancy (1993). Low dose aspirin in prevention and treatment of intrauterine growth retardation and pregnancy-induced hypertension. *Lancet*, **341**, 396–400.

Jamali, F. and Stevens, D.R.S. (1983). Naproxen excretion in milk and its uptake by the infant. *Drug Intelligence and Clinical Pharmacology*, **17**, 910–11.

Jarnerot, G. and Into-Malmberg, M.B. (1979). Sulphasalazine treatment during breast feeding. *Scandinavian Journal of Gastroenterology*, **14**, 869–71.

Jarnerot, G., Into-Malmberg, M.B., and Esbjorner, E. (1981). Placental transfer of sulphasalazine and sulphapyridine and some of its metabolites. *Scandinavian Journal of Gastroenterology*, **16**, 693–7.

Jensen, M.K. (1967). Chromosome studies in patients treated with azathioprine and amethopterin. *Acta Medica Scandinavica*, **182**, 445–55.

Johns, D.G. *et al*. (1972). Secretion of methotrexate into human milk. *American Journal of Obstetrics and Gynecology*, **112**, 978–80.

Kaaja, R., Julkunen, H., Viinikka, L., and Ylikorkala, O. (1993). Production of prostacyclin and thromboxane in lupus pregnancies: effect of small dose of aspirin. *Obstetrics and Gynecology*, **81**, 327–31.

Kean, W.F. and Buchanan, W.W. (1987). Variables affecting the absorption of non-steroidal anti-inflammatory drugs from the gastro-intestinal tract. *Japanese Journal of Rheumatology*, **1**, 159–70.

Kean, W.F. and Buchanan, W.W. (1990). Pregnancy and rheumatoid disease. *Baillière's Clinical Rheumatology*, **4**, 125–40.

Kirk, E.P. (1991). Organ transplantation and pregnancy. A case report and review. *American Journal of Obstetrics and Gynecology*, **164**, 1629–34.

Kirshon, B. *et al*. (1988). Teratogenic effects of first-trimester cyclophosphamide therapy. *Obstetrics and Gynecology*, **72**, 462–4.

Kirshon, B., Mari, G., and Moise, K.J., Jr., (1990 *a*). Indomethacin therapy in the treatment of symptomatic polyhydramnios. *Obstetrics and Gynecology*, **75**, 202–5.

Kirshon, B., Mari, G., Moise, K.J., Jr., and Wasserstrum, N. (1990 *b*). Effect of indomethacin on the fetal ductus arteriosus during treatment of symptomatic polyhydramnios. *Journal of Reproductive Medicine*, **35**, 529–32.

Klebanoff, M.A. and Berendes, H.W., (1988). Aspirin exposure during the first 20 weeks of gestation and IQ at four years of age. *Teratology*, **37**, 249–55.

Krauer, B. and Krauer, F. (1983). Drug kinetics in pregnancy. In *Handbook of clinical pharmacokinetics* (ed. M. Gibaldi and L. Prescott), pp. 1–17. ADIS Press, Sydney.

Kraus, A.M. (1975). Congenital cataract and maternal steroid injection. *Journal of Pediatric Ophthalmology*, **12**, 107–8.

Kullander, S. and Kallen, B. (1976). A prospective study of drugs in pregnancy. *Acta Obstetrica et Gynecologica Scandinavica*, **55**, 289–95.

Kwan, K.C. *et al*. (1978). Sulindac: chemistry, pharmacology and pharmacokinetics. *European Journal of Rheumatology and Inflammation*, **1**, 9–11.

Lakatos, L., Kover, B., Vekerdy, S., and Dvoracsek, E. (1976). D-penicillamine therapy of neonatal jaundice: comparison with phototherapy. *Acta Paediatrica Scientiarum Hungaricae*, **17**, 93–102.

Lau, R.J. and Scott, J. R. (1985). Pregnancy following renal transplantation. *Clinics in Obstetrics and Gynecology*, **23**, 339–50.

Lee, F.I. and Loeffler, F.E. (1962). Gout and pregnancy. *Journal of Obstetrics and Gynecology of the British Commonwealth*, **69**, 299.

Lee, P. (1985). Anti-inflammatory therapy during pregnancy and lactation. *Clinical Investigative Medicine*, **8**, 328–33.

Lee, R.B., Kelly, J., Elg, S.A., and Benson, W.L. (1989). Pregnancy following conservative surgery and adjunctive chemotherapy for stage III immature teratoma of the ovary. *Obstetrics and Gynecology*, **73**, 853–5.

Levy, G. (1975). Salicylate pharmacokinetics in the human neonate. In *Basic and therapeutic aspects of perinatal pharmacology* (ed. P.L. Morselli, F. Garattini, and F. Sereni), pp. 319–30. Raven Press, New York.

Levy, G., Garretson, L.K., and Soda, D.M. (1975 *a*). Evidence of placental transfer of acetaminophen. *Pediatrics*, **55**, 895.

Levy, G., Procknal, J.A., and Garrettson, L.K. (1975 *b*). Distribution of salicylate between neonatal and maternal serum at diffusion equilibrium. *Clinical Pharmacology and Therapeutics*, **18**, 210–4.

Levy, M. and Yaffe, C. (1978). Testicular function in patients with familial Mediterranean fever on long-term colchicine treatment. *Fertility and Sterility*, **29**, 667–8.

Levy, M., Spino, M., and Read, S.E. (1991 *a*). Colchicine: a state-of-the-art review. *Pharmacotherapy*, **11**, 196–211.

Levy, M. *et al*. (1991*b*). Pregnancy outcome following first trimester exposure to chloroquine. *American Journal of Perinatology*, **8**, 174–8.

Liggins, G.C. and Howie, R.N. (1972). A controlled trial of antepartum glucocorticoid treatment for prevention of the respiratory distress syndrome in premature infants. *Pediatrics*, **50**, 515–25.

Lyle, W.H. (1978). Penicillamine in pregnancy. *Lancet*, **i**, 606–7.

Lynd, P.A., Andreasen, A.C., and Wyatt, R.J. (1976). Intrauterine salicylate intoxication in a newborn. A case report. *Clinical Pediatrics* (*Philadelphia*), **15**, 912–13.

MacArthur, B.A. *et al*. (1981). Cognitive and psychosocial development of 4-year-old children whose mothers were treated antenatally with betamethasone. *Pediatrics*, **68**, 638–43.

MacArthur, B.A. *et al*. (1982). School progress and cognitive development of 6-year-old children whose mothers were treated antenatally with betamethasone. *Pediatrics*, **70**, 99–105.

McBride, W.G. (1961). Thalidomide and congenital abnormalities. *Lancet*, **ii**, 1358.

McKenzie, S.A., Selley, J.A., and Agnew, J.E. (1975). Secretion of prednisone into breast milk. *Archives of Disease in Childhood*, **50**, 894–6.

Manchester, D., Margolis, H.S., and Sheldon, R.E. (1976). Possible association between maternal indomethacin therapy and primary pulmonary hypertension of the newborn. *American Journal of*

Obstetrics and Gynecology, **126**, 467–9.

Mancini, R.E. *et al*. (1966). Effect of prednisolone upon normal and pathologic human spermatogenesis. *Fertility and Sterility*, **17**, 500–13.

Mangurten, H.H. and Benawra, R. (1980). Neonatal codeine withdrawal in infants of non-addicted mothers. *Pediatrics*, **65**, 159–60.

Matheson, I., Lunde, P.K.M., and Notarianni, L. (1985). Infant rash caused by paracetamol in breast milk? *Pediatrics*, **76**, 651–2.

Medorikawa, O. *et al*. (1972). Teratological studies of CGP45840 in mice and rats. *Clinical Report Journal*, **6**, 41–9.

Merlin, H.E. (1972). Azoospermia caused by colchicine—a case report. *Fertility and Sterility*, **23**, 180–1.

Miehle, W. (1988). Current aspects of D-penicillamine and pregnancy. *Zeitschrift für Rheumatologie*, **47** (Suppl. 1), 20–3.

Milunsky, J.M. and Milunsky, A. (1991). Breast-feeding during colchicine therapy for familial Mediterranean fever. *Journal of Pediatrics*, **119**, 164.

Milunsky, G., Graef, J.W., and Gaynor, M.F. (1968). Methotrexate-induced congenital malformation. *Journal of Pediatrics*, **72**, 790–5.

Miyamoto, T. *et al*. (1974). Gold therapy in bronchial asthma-special emphasis upon blood level of gold and its teratogenicity. *Journal of the Japanese Society of Internal Medicine*, **63**, 1190–7.

Mjolnerod, O.K., Pasmussen, K., Dommerud, S.A., and Gjeruldsen, S.T. (1971). Congenital connective tissue defect probably due to D-penicillamine treatment in pregnancy. *Lancet*, **i**, 673–5.

Mogadam, M., Dobbins, W.O., Korelitz, B.I., and Ahmed, S.W. (1981). Pregnancy in inflammatory bowel disease: effects of sulfasalazine and corticosteroids on fetal outcome. *Gastroenterology*, **80**, 72–6.

Moise, K.J., Jr. (1991). Indomethacin therapy in the treatment of symptomatic polyhydramnios. *Clinical Obstetrics and Gynecology*, **34**, 310–8.

Moise, K.J., Jr., *et al*. (1988). Indomethacin in the treatment of premature labour: effects on the fetal ductus arteriosus. *New England Journal of Medicine*, **319**, 327–31.

Morales, W.J. and Madhav, H. (1993). Efficacy and safety of indomethacin compared with magnesium sulfate in the management of preterm labour. A randomized study. *American Journal of Obstetrics and Gynecology*, **169**, 97–102.

Morales, W.J. *et al*. (1989). Use of ampicillin and corticosteroids in premature rupture of membranes: a randomized study. *Obstetrics and Gynecology*, **73**, 721–6.

Nation, R.L., Hackett, L.P., Dusci, L.J., and Ilett, K.F. (1984). Excretion of hydroxychloroquine in human milk. *British Journal of Clinical Pharmacology*, **17**, 368–9.

Neill, W.A. *et al*. (1973). Action of chloroquine phosphate in rheumatoid arthritis II chromosome damaging effect. *Annals of the Rheumatic Diseases*, **32**, 547–50.

Newman, N.M. and Correy, J.F. (1983). Possible teratogenicity of sulphasalazine. *Medical Journal of Australia*, **1**, 528–9.

Nicholson, H.O. (1968). Cytotoxic drugs in pregnancy: review of reported cases. *Journal of Obstetrics and Gynaecology of the British Commonwealth*, **75**, 307–12.

Niebyl, J.R. and Witter, F.R. (1986). Neonatal outcome after indomethacin treatment for preterm labour. *American Journal of Obstetrics and Gynecology*, **155**, 747–9.

Nochimson, D.J. and Petrie, R.H. (1979). Glucocorticoid therapy for the induction of pulmonary maturity in severely hypertensive gravid women. *American Journal of Obstetrics and Gynecology*, **133**, 449–51.

Norton, M.E. *et al*. (1993). Neonatal complications after the administration of indomethacin for preterm labour. *New England Journal of Medicine*, **329**, 1602–7.

Okun, D.B., Groncy, P.K., Sieger, L., and Tanaka, K.R. (1979). Acute leukemia in pregnancy: transient neonatal myelosuppression after combination chemotherapy in the mother. *Medical Pediatric Oncology*, **7**, 315–9.

Osman, M.A. *et al*. (1983). Reduction of oral penicillamine absorption by food, antacid and ferrous sulphate. *Clinical Pharmacology and Therapeutics*, **33**, 465–70.

Ostensen, M. (1983) Piroxicam in human breast milk. *European Journal of Clinical Pharmacology*, **25**, 829–30.

Ostensen, M. and Husby, G. (1985). Antirheumatic drug treatment during pregnancy and lactation. *Scandinavian Journal of Rheumatology*, **14**, 1–7.

Ostensen, M., Brown, N.D., Chiang, P.K, and Aarbakke, J. (1985). Hydroxychloroquine in human breast milk. *European Journal of Clinical Pharmacology*, **28**, 357.

Ostensen, M. *et al*. (1986). Excretion of gold into human breast milk. *European Journal of Clinical Pharmacology*, **32**, 251–2.

Ostensen, M., Matheson, I., and Laufen, H. (1988). Piroxicam in breast milk after long term treatment. *European Journal of Clinical Pharmacology*, **35**, 567–9.

Otero, L. *et al*. (1981). Neonatal leukocytosis associated with prenatal administration of dexamethasone. *Pediatrics*, **68**, 778–80.

Papageorgiou, A.N. *et al*. (1979). The antenatal use of betamethasone in the prevention of respiratory distress syndrome: a controlled double-blind study. *Pediatrics*, **63**, 73–9.

Parke, A.L. (1993). Antimalarial drugs, pregnancy and lactation. *Lupus*, **2** (Suppl. 1), 821–3.

Perry, W.H. (1983). Methotrexate and teratogenesis. *Archives of Dermatology*, **119**, 874.

Pizzuto, J. *et al*. (1980). Treatment of acute leukemia during pregnancy: presentation of nine cases. *Cancer Treatment Reports*, **64**, 679–83.

Popert, A.J. (1962). Pregnancy and adrenocortical hormones, some aspects of their interactions in rheumatic disease. *British Medical Journal*, **i**, 967.

Powell, H.R. and Ekert, H. (1971). Methotrexate-induced congenital malformations. *Medical Journal of Australia*, **2**, 1076–7.

Preston, S.J. *et al*. (1989). Comparative analgesic and anti-inflammatory properties of sodium salicylate and acetylsalicylic acid (aspirin) in rheumatoid arthritis. *British Journal of Clinical Pharmacology*, **27**, 607–11.

Pujals, J.M. *et al*. (1989). Osseous malformation in baby born to woman on cyclosporin. *Lancet*, **i**, 667.

Registration Committee of the European Dialysis and Transplant Association (1980). Successful pregnancies in women treated by dialysis and kidney transplantation. *British Journal of Obstetrics and Gynaecology*, **839**, 45.

Reiss, U., Atad, J. Rubinstein, I., and Zuckerman, H. (1976). The effect of indomethacin in labour at term. *International Journal of Gynecology and Obstetrics*, **14**, 369–74.

Richards, A.J. (1977). Transfer of gold from mother to foetus. *Lancet*, **i**, 99.

Rocker, I. and Henderson, M.J.H. (1976). Medical innumeracy. *Lancet*, **ii**, 1246.

Rogers, J.G. *et al*. (1980). Possible teratogenic effect of gold. *Australian Paediatric Journal*, **16**, 194–5.

Romero, R., Lockwood, C., Oyarfzun, E., and Hobbins, J.C. (1988). Toxemia: new concepts in an old disease. *Seminars in Perinatology*, **12**, 302–23.

Rosa, F.W. (1986). Teratogen update: penicillamine. *Teratology*, **33**, 127–31.

Ross, J.B. and Garatsos, S. (1974). Absence of chloroquine-induced ototoxicity in a fetus. *Archives of Dermatology*, **109**, 573.

Roubenoff, R. *et al*. (1988). Effects of antiinflammatory and immunosuppressive drugs on pregnancy and fertility. *Seminars in Arthritis and Rheumatism*, **18**, 88–110.

Rubin, A. *et al*. (1984). A profile of the physiological disposition and gastrointestinal effects of fenoprofen in man. *Current Medical Research and Opinion*, **2**, 529–44.

Rumack, C.M. *et al*. (1981). Neonatal intracranial hemorrhage and maternal use of aspirin. *Obstetrics and Gynecology*, **58** (Suppl.), 52S–56S.

Russell, J.A., Powles, R.L., and Oliver, R.T.D. (1976). Conception and congenital abnormalities after chemotherapy of acute myelogenous leukaemia in two men. *British Medical Journal*, **ii**, 1508.

Rustin, G.J.S. *et al*. (1984). Pregnancy after cytotoxic chemotherapy for gestational trophoblastic tumours. *British Medical Journal*, **288**, 103–6.

Sackett, D.L., Haynes, R.B., Gent, M., and Taylor, D.W. (1986). Compliance. In *Monitoring for drug safety* (2nd edn) (ed. W.H.W. Inman), pp. 471–83. MTP Press, Lancaster, England.

Schiff, E. *et al*. (1989). The use of aspirin to prevent pregnancy-induced hypertension and lower the ratio of thromboxane A2 to prostacyclin in relatively high risk pregnancies. *New England Journal of Medicine*, **321**, 351–6.

Schleuning, M. and Clemm, C. (1987). Chromosomal aberrations in a newborn whose mother received cytotoxic treatment during pregnancy. *New England Journal of Medicine*, **317**, 1666–7.

Schmand, B. *et al*. (1990). Psychological development of children who were treated antenatally with corticosteroids to prevent respiratory distress syndrome. *Pediatrics*, **86**, 58–64.

Schoenfeld, A., Freedman, S., Hod, M., and Ovadia, Y. (1989). Antagonism of antihypertensive drug therapy in pregnancy by indomethacin? *American Journal of Obstetrics and Gynecology*, **161**, 1204–5.

Shapiro, S. *et al*. (1976). Perinatal mortality and birthweight in relation to aspirin taken during pregnancy. *Lancet*, **i**, 1375–6.

Shaywitz, B.A. (1978). Fetal alcohol syndrome: an ancient problem rediscovered. *Drug Therapy*, **8**, 15–108.

Sibai, B.M., Mirro, R., Chesney, C.M., and Leffler, C. (1989). Low-dose aspirin in pregnancy. *Obstetrics and Gynecology*, **74**, 551–7.

Sibai, B.M. *et al*. (1993). Prevention of pre-eclampsia with low-dose aspirin in healthy, nulliparous pregnant women. *New England Journal of Medicine*, **329**, 1213–18.

Slone, D. *et al*. (1976). Aspirin and congential malformations. *Lancet*, **i**, 1373–5.

Smolders-de Haas, H. *et al*. (1990). Physical development and medical history of children who were treated antenatally with corticosteroids to prevent respiratory distress syndrome: a 10 to 12 year follow-up. *Pediatrics*, **85**, 65–70.

Sokal, J.E. and Lessmann, E.M. (1960). Effects of cancer chemotherapeutic agents on the human fetus. *Journal of the American Medical Association*, **172**, 1765–72.

Soller, R.W. and Stander, H. (1981). Maternal drug exposure and perinatal intracranial hemorrhage. *Obstetrics and Gynecology*, **58**, 735–7.

Solomon, L. *et al*. (1977). Neonatal abnormalities associated with D-penicillamine treatment during pregnancy. *New England Journal of Medicine*, **296**, 54–5.

Streissguth, A.P. *et al*. (1987). Aspirin and acetaminophen use by pregnant women and subsequent child IQ and attention decrements. *Teratology*, **35**, 211–9.

Stuart, M.J. (1983). Aspirin and maternal or neonatal hemostasis. *New England Journal of Medicine*, **308**, 281.

Stuart, M.J., Gross, S.J., Elrad, H., and Graeber, J.E. (1982). Effects of acetylsalicylic-acid ingestion on maternal and neonatal hemostasis. *New England Journal of Medicine*, **307**, 909–12.

Suhonen, R. (1983). Hydroxychloroquine administration in pregnancy. *Archives of Dermatology*, **199**, 185–6.

Sweet, D.L. and Kinzie, J. (1976). Consequences of radiotherapy and antineoplastic therapy for the fetus. *Journal of Reproductive Medicine*, **17**, 241–6.

Tarp, U. and Graudal, H. (1985). A follow up study of children exposed to gold *in utero*. *Arthritis and Rheumatism*, **28**, 235–6.

Tolchin, S.F. *et al*. (1974). Chromosome abnormalities from cyclophosphamide therapy in rheumatoid arthritis and progressive systemic sclerosis (scleroderma). *Arthritis and Rheumatism*, **17**, 375–82.

Townsend, R.J. *et al*. (1984). Excretion of ibuprofen in breast milk. *American Journal of Obstetrics and Gynecology*, **149**, 184–5.

Turner, G. and Collins, E. (1975). Fetal effects of regular salicylate ingestion in pregnancy. *Lancet*, **ii**, 338–9.

Tyson, H.K. (1974). Neonatal withdrawal symptoms associated with maternal use of propoxyphene hydrochloride (Darvon). *Journal of Pediatrics*, **85**, 684–5.

Uldall, P.R., Feest, T.G., Morley, A.R., and Tomlinson, B.E. (1972). Cyclophosphamide therapy in adults with minimal-change nephrotic syndrome. *Lancet*, **ii**, 1250–3.

Ullberg, S., Lindquist, N.J., and Sjostrand, S.E. (1970). Accumulation of chororetinotoxic drugs in fetal eye. *Nature*, **227**, 1257–8.

van Haesebrouck, P. *et al*. (1988). Oligohydramnios, renal insufficiency, and ileal perforation in preterm infants after intrauterine exposure to indomethacine. *Journal of Pediatrics*, **113**, 738–43.

van Winzum, C. and Verhaest, L. (1979). Diflunisal. *Clinics in the Rheumatic Diseases*, **5**, 707–31.

Vender, R.J. and Spiro, H.W. (1982). Inflammatory bowel disease and pregnancy. *Journal of Clinical Gastroenterology*, **4**, 231–49.

Waffarn, F. *et al*. (1983). Effect of antenatal glucocorticoids on clinical closure of the ductus arteriosus. *Americal Journal of Diseases of Children*, **137**, 336–8.

Warkany, J. (1981). Teratogenicity of folic and antagonists. *Cancer Bulletin*, **33**, 76–7.

Warrell, D.W. and Taylor, R. (1968). Outcome for the foetus of mothers receiving prednisolone during pregnancy. *Lancet*, **i**, 117–8.

Wasserstrum, N. *et al*. (1989). Betamethasone and the human fetal ductus arteriosus. *Obstetrics and Gynecology*, **74**, 897–900.

Watson, A.R., Rance, C.P., and Bain, J. (1985). Long term effects of cyclophosphamide on testicular function. *British Medical Journal*, **291**, 1457–60.

Wiernik, P.H. and Duncan, J.H. (1971). Cyclophosphamide in human milk. *Lancet*, **i**, 912.

Wilkinson, A.R. (1980). Naproxen levels in preterm infants after maternal treatment. *Lancet*, **ii**, 591–2.

Wilkinson, A.R., Aynsley-Green, A., and Mitchell, M.D. (1979). Persistent pulmonary hypertension and abnormal prostaglandin E levels in preterm infants after maternal treatment with naproxen. *Archives of Disease in Childhood*, **54**, 942–5.

Wiquest, N., Lundstrom, V., and Green, K. (1975). Premature labour and indomethacin. *Prostaglandins*, **10**, 515–26.

Wiseman, E.H. (1985). Piroxicam and related oxicams. In *Anti-inflammatory and anti-rheumatic drugs* (ed. K.D. Rainsford), Vol. 2, pp. 218–9. CRC Press, Boca Raton, FL.

Witter, F.R. (1993). Clinical pharmacokinetics in the treatment of rheumatoid arthritis in pregnancy. *Clinical Pharmacokinetics*, **25**, 444–9.

Wong, Y.C., Beardsmore, C.S., and Silverman, M. (1982). Antenatal dexamethasone and subsequent lung growth. *Archives of Disease in Childhood*, **57**, 536–8.

Zemer, D., Pras, M., Sohar, E., and Gafni, J. (1976). Colchicine in familial Mediterranean fever. *New England Journal of Medicine*, **294**, 170–1.

Zerner, J. *et al*. (1981). Intrauterine contraceptive device failures in renal transplant patients. *Journal of Reproductive Medicine*, **26**, 99–102.

Zuckerman, H., Reiss, V., and Rubinstein, I. (1974). Inhibition of human premature labour by indomethacin. *Obstetrics and Gynecology*, **44**, 787–9.

1.3.2 The skin

C. R. Lovell and P. J. Maddison

Introduction

Why is it important to examine the skin in a patient with arthritis? The skin is the most accessible organ for clinical examination and investigation. Pattern recognition of skin lesions is valuable in the diagnosis and prognosis of underlying joint disease. Often, the skin bears the brunt of toxicity from antirheumatic drugs.

Joint abnormalities may be mirrored by similar changes in the skin. For example, joint hypermobility and skin hyperextensibility coexist in heritable disorders of connective tissue, such as the Ehlers–Danlos syndrome. Cutaneous erythema and hyperpigmentation may overlie an inflamed joint. Furthermore, many patterns of cutaneous involvement are highly specific to certain arthropathies, such as the characteristic erythematous eruption of erythema marginatum in rheumatic fever (Fig. 1), the 'butterfly' facial erythema of acute lupus erythematosus and erythema chronicum migrans following infection with *Borrelia burgdorferi*, which can result in Lyme disease.

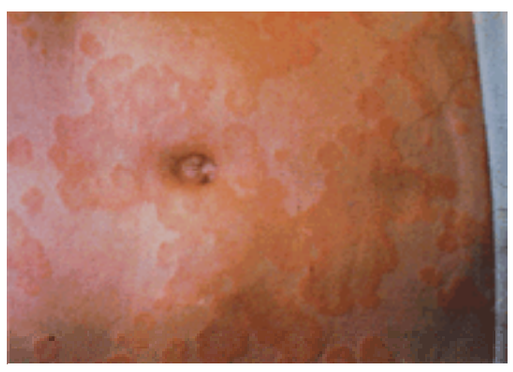


Fig. 1 Erythema marginatum in rheumatic fever: reddish annular lesions may expand rapidly to form large figurate patches; the eruption is typically evanescent, lasting 2 to 3 days at most (photograph by courtesy of Dr E. Pascual).

The degree and extent of skin lesions can be of prognostic importance, for example the cutaneous vasculitis and nodules in rheumatoid disease.

Finally, several antirheumatic drugs produce highly specific, and potentially serious, cutaneous reactions that require accurate diagnosis and prompt management.

This chapter comprises:

1. A brief outline of the major descriptive terms used in dermatology and the simple investigations required to make a firm dermatological diagnosis.
2. The differential diagnoses of important cutaneous abnormalities.
3. Aspects of treatment of the skin in rheumatic disorders.
4. Cutaneous side-effects of commonly used antirheumatic drugs.

The approach to the patient with skin lesions

History

A careful history should include the following features.

1. duration of the eruption;
2. duration of individual lesions (e.g. weals in simple urticaria individually last less than 24 h, whereas they persist for two or more days in urticarial vasculitis);
3. timing of lesions (e.g. the eruption of Still's disease is typically worse in the early afternoon, and evening);
4. seasonal variation of lesions (e.g. 'summer' vasculitis);
5. the relation of skin lesions to joint disease (e.g. scattered gonococcal pustules occurring with flitting arthritis in the bacteraemic phase);
6. whether the skin lesions occur singly or in crops, as in Henoch–Schönlein purpura;
7. where the lesions occur;
8. do lesions blister or crust? (Remember that many infiltrated or hyperkeratotic lesions may be described as 'blisters' by the patient. True blisters are seen in some drug reactions, notably fixed drug eruption and erythema multiforme. Superficial blisters rupture early, leaving eroded areas, as in penicillamine-induced pemphigus.);
9. whether lesions itch;
10. whether lesions are tender;
11. are lesions caused or aggravated by light? (Several drugs cause photosensitivity and cutaneous lupus typically occurs on areas of sun exposure.);
12. whether there is cold sensitivity—enquire about features of Raynaud's phenomenon;
13. whether there is orogenital ulceration of recent onset;
14. a careful history of present and recent drug therapy.

Examination

The skin should be examined in a good, preferably natural, light. A torch and pocket magnifying glass are invaluable. Since some skin lesions are transient, repeated examination of the skin may be necessary.

Note the distribution and morphology of individual lesions. Knowledge of a few technical terms is essential (Table 1).

Term	Definition	Examples
Macule	A flat, impalpable area of altered skin colour	A cafe-au-lait spot
Papule	A lump smaller than 0.5–1 cm diameter	A viral wart
Nodule	A lump greater than 0.5–1 cm diameter	An epidermoid ('sebaceous') cyst
Plaque	A slightly raised, circumscribed area, often disc shaped	Plaque psoriasis
Vesicle	A small blister (fluid-filled lesion), less than 0.5 cm diameter	Herpes simplex
Bulla	A blister greater than 0.5 cm diameter	A friction blister
Pustule	An accumulation of pus in the skin	Folliculitis
Erosion	A loss of epidermis that heals without scarring	An abrasion
Ulcer	A defect or loss of dermis and epidermis produced by sloughing of necrotic tissue	Pyoderma gangrenosum
Telangiectasia	Visibly dilated, small vascular lesions that blanch on pressure	'Spider naevi'
Purpura	Extravasation of blood, forming coalescent areas of discoloration, which does not blanch on pressure; small lesions are termed petechiae (1–2 mm diameter)	
	Larger areas of extravasation of blood are termed ecchymoses	Traumatic bruising
Erythema	Redness	
Violaceous	Purplish hue	Lichen planus
Weal	A transient area of dermal oedema which produces a white or skin-coloured papule or nodule surrounded by a 'flare' of erythema	Urticaria
Atrophy	Skin thinning with reduction of recognizable structures and perhaps translucency	
Poikilodermia	Atrophy with telangiectasia	Radiotherapy scars, dermatomyositis
Alopecia	Loss of hair. It may be localized (e.g. alopecia areata) or diffuse (e.g. systemic lupus erythematosus or iron deficiency). Scarring alopecia, with destruction of hair follicles, is a feature of discoid lupus affecting the scalp	
Onycholysis	Increased free edge of the nail	Psoriasis, thyroid's phenomenon

Table 1 Important descriptive terms in dermatology

The regional distribution of lesions

Scalp

Look for alopecia. Is it diffuse or localized? Is it non-scarring (the scalp itself looks normal) or scarring (Fig. 2) (when there are inflammatory changes around hair follicles and evidence of cutaneous atrophy or sclerosis). Look for the well-demarcated areas of scaling typical of psoriasis.

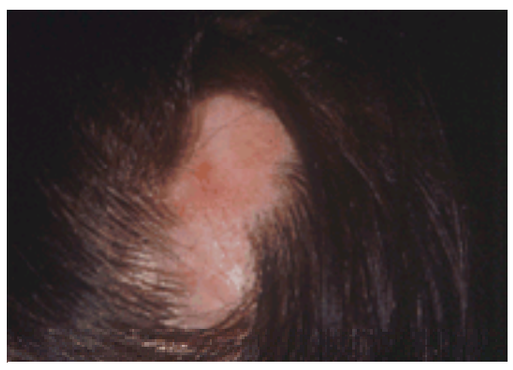


Fig. 2 Scarring alopecia of the scalp in discoid lupus erythematosus.

Face

1. Look for evidence of photosensitivity (which classically spares the 'shaded' areas such as the nasolabial folds, eyelids, chin, and Wilkinson's triangle (behind the ears)).
2. Is there periorbital oedema or the 'heliotrope' violaceous erythema characteristic of dermatomyositis?
3. Examine the nose for infiltration (e.g. lupus pernio in sarcoidosis). Inability to evert the lower eyelid is an early feature of systemic sclerosis.
4. Look inside the mouth for ulcers and the lace-like white streaks of oral lichen planus (Fig. 3 and Fig. 4).

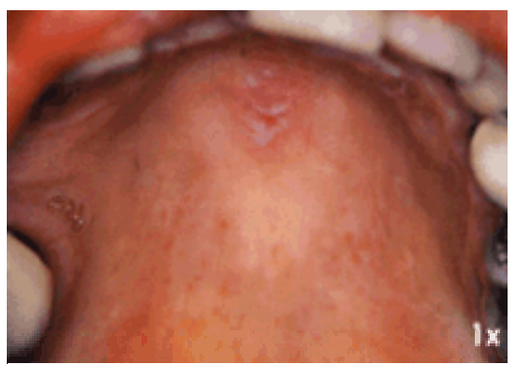


Fig. 3 Involvement of hard palate in subacute lupus erythematosus.

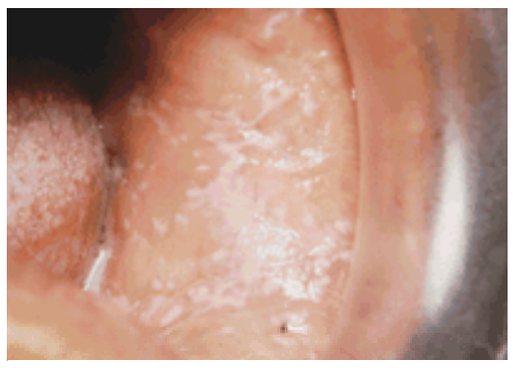


Fig. 4 Oral lichen planus—white streaks on buccal mucosa.

Ears

These are an important site for discoid lupus erythematosus (Fig. 5), gouty tophi, and, occasionally, rheumatoid nodules.

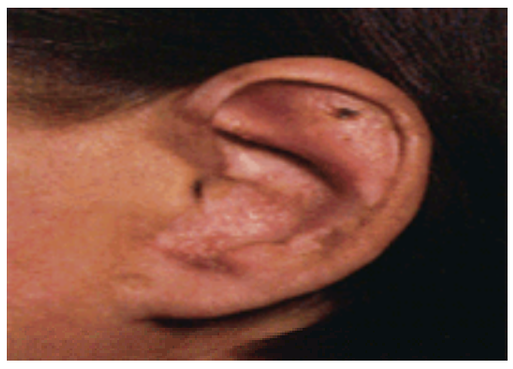


Fig. 5 Discoid lupus erythematosus affecting the external ear.

Hands and feet

1. A photosensitive eruption spares the finger webs and palms. The distribution of erythema on the backs of the fingers helps to distinguish systemic lupus erythematosus from dermatomyositis (Fig. 6).

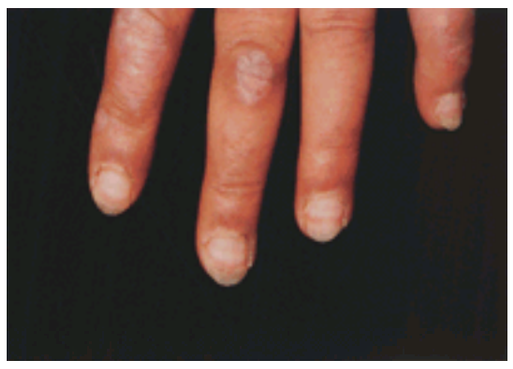


Fig. 6 Plaques of discoid lupus erythematosus on the fingers.

2. In a patient with Raynaud's phenomenon, evidence of ulceration of fingertips, atrophy of finger pulps, induration, and tethering of skin indicates systemic sclerosis. Later features include telangiectasia and calcinosis.
3. Look for palmar erythema, especially seen in rheumatoid disease.
4. Examine the soles for keratoderma blennorrhagica or localized pustular psoriasis.

Nails

1. Look for onycholysis, pitting, salmon patches, and subungual hyperkeratosis (typical of psoriasis) (Fig. 7, Fig. 8 and Fig. 9).

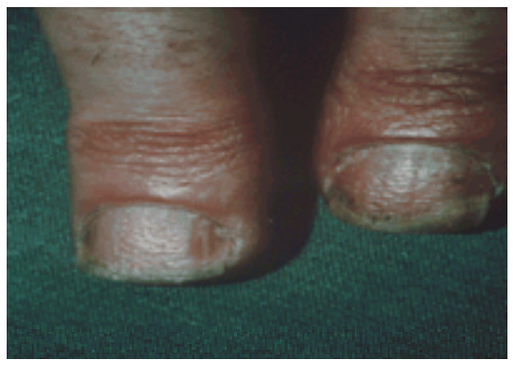


Fig. 7 Psoriasis: nail pits with early onycholysis.

Fig. 8 Psoriasis: onycholysis with 'salmon patch'.

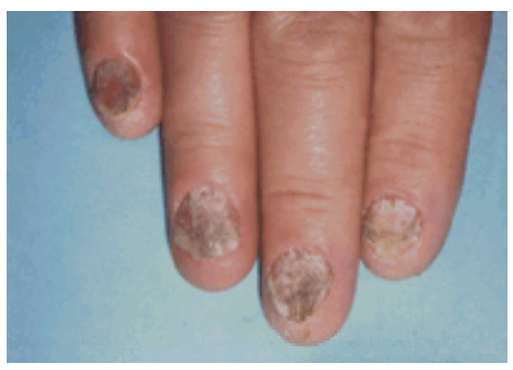


Fig. 9 Psoriasis: onycholysis, gross pitting, and nail deformity.

2. Examine the nailfolds for periungual erythema, digital vasculitis, periungual papules (e.g. in multicentric reticulohistiocytosis), and paronychia (Fig. 10).

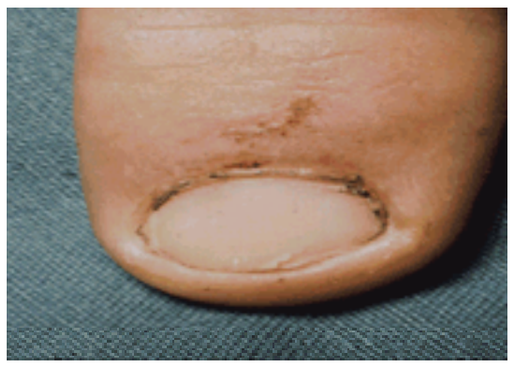


Fig. 10 Dermatomyositis: nailfold vasculopathy (photograph by courtesy of Dr R. S-H. Tan).

3. It is helpful to assess the nailfold capillaries with an ophthalmoscope set at 40 dioptres (10), applying a drop of oil to the cuticle. Capillaroscopy (see below under Investigations) is a useful investigation.
4. Examine the whole body, including the natal cleft and umbilicus (common sites for psoriasis). Look for penile or vulval ulceration, which, like oral ulceration, may be asymptomatic.

For further reading, see Ashton (1995), Burton (1985), and Lovell *et al*. (1990).

Investigation

Routine histological examination of a formalin-fixed skin biopsy may provide considerable useful information (Harrison 1980). The histological features of many lesions are diagnostic, e.g. the palisading granulomas seen in rheumatoid nodules and 'liquefaction degeneration' in cutaneous lupus. Where possible, take an elliptical biopsy across the edge of an early lesion, following skin folds. The biopsy should be at least 10 × 5 mm, unless this is undesirable cosmetically. Include subcutaneous fat in the biopsy. (This is especially important if a panniculitis is suspected, and allows easier closure of the wound.) Use a skin hook or the tip of a needle to lift the end of the biopsy, avoiding forceps trauma. Close the wound with 4/0 or 5/0 silk or prolene. A 5-mm punch biopsy may be adequate for small lesions. Mount the biopsy on filter paper to ensure correct orientation for sectioning. Special stains may be indicated, for example a mucin stain for dermatomyositis. An elastin stain may show the depletion of normal connective tissue structures in localized scleroderma or systemic sclerosis.

Direct skin immunofluorescence is of particular value in lupus erythematosus and in the investigation of blistering disorders. The lupus band test (granular deposition of immunoglobulins at the dermoepidermal junction) is positive in lesional skin in discoid lupus erythematosus and negative in normal skin. It may be negative in early subacute cutaneous lupus erythematosus. In systemic lupus erythematosus, the lupus band test is positive in clinically non-involved skin in 70 per cent of cases, particularly if the biopsy is taken from a sun-exposed site such as the extensor forearm. A punch biopsy is usually adequate for direct immunofluorescence and should be snap-frozen immediately in liquid nitrogen or despatched to the laboratory in appropriate transport medium.

A Kveim test may help in the diagnosis of sarcoidosis, although it is now used less frequently because of the variable reliability of the antigen source. The technique is to inject the antigen intradermally at an identifiable site such as near a mole, or an area tattooed with permanent ink. The site should be biopsied after 4 weeks, whether or not a papule has developed.

Several non-invasive techniques can be used to measure tissue perfusion in the skin, including laser-Doppler flowmetry, transcutaneous oxygen-pressure measurement, thermography, and capillary microscopy. Thermography, using an infrared radiometer or liquid crystal contact, is a useful non-invasive technique in the assessment of Raynaud's phenomenon. Digital blood flow can be measured in a resting state or following a mild thermal stress. Thermography can be used to measure the degree of severity of Raynaud's phenomenon and its response to therapy; it may help to distinguish primary Raynaud's phenomenon from that secondary to connective tissue disease (Love 1980; Ring and Phillips 1984 ; Will *et al*. 1992).

In connective tissue diseases, local vascular changes can be studied by capillary microscopy of the nailfold. This area is easily accessible for examination and the arrangement of the capillaries in the nailfold allows the visualization of the complete capillary loop. In addition, vascular abnormalities appear earlier in the course of the disease than at other sites of the skin of the finger. The morphology of skin capillaries can be studied directly with an ordinary light microscope. Recently, capillaroscopy has been coupled with a videophotometric system and used with computer software to analyse the number and morphology of capillary loops as well as blood-cell velocity under various physiological conditions including local cold exposure (Mahler *et al*. 1987; Fagrell *et al*. 1988).

Morphological changes in capillaries seen in primary Raynaud's approximate those seen in normal individual but the blood-cell velocity is decreased markedly after exposure to cold (Jacohs *et al*. 1987). Characteristic structural abnormalities are found to a much greater extent in connective tissue diseases such as systemic sclerosis, overlap syndromes (including 'mixed connective tissue disease'), and dermatomyositis. They include enlargement of capillary loops (a non-specific finding in idiopathic Raynaud's phenomenon) and, particularly in systemic sclerosis, deformed capillary loops, segmental (or rarely diffuse) loss of capillaries, new capillary formation leading to excessive branching (arborization), and haemorrhage into the nailfold. Their presence in patients presenting with Raynaud's has prognostic significance (Houtman *et al*. 1986).

These changes are dynamic, especially in dermatomyositis. Rapid evolution of structural changes, especially loss of capillaries, is linked with a poor prognosis in systemic sclerosis (Maricq 1988).

Differential diagnosis of common physical signs

Erythema

The red face

Photosensitivity, for example drug-induced
Subacute and acute cutaneous lupus erythematosus may cause diffuse facial erythema. An important cause of photosensitivity is polymorphic light eruption (Fig. 11), which affects 10 per cent of Caucasian females but typically spares the face and hands (Ros and Wennersten 1986).

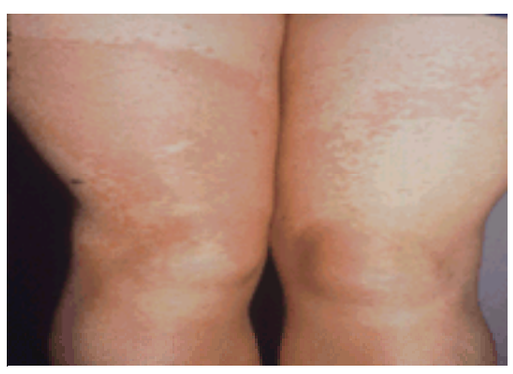


Fig. 11 Polymorphic light eruption: itchy maculopapular lesions occur within hours of sun exposure to sites habitually covered (in this patient the eruption extends to

the line of her Bermuda shorts).

Papules
Telangiectatic papules are found in rosacea and papular lupus erythematosus. These may need to be distinguished by skin biopsy. Papules occur around the alae nasi and lips in sarcoid (Fig. 12), particularly in the Afro-Caribbean patient, and in multicentric reticulohistiocytosis.

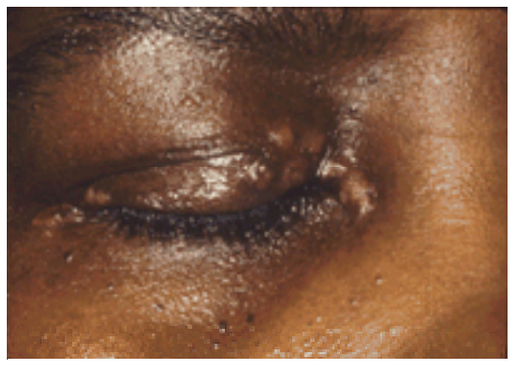


Fig. 12 Cutaneous sarcoid in an Afro-Caribbean patient: yellowish papules around the eyelids.

'Heliotrope'
Violaceous erythema and oedema affect both eyelids in dermatomyositis. Episodic oedema of eyelids, and swollen tongue and lips are characteristic of angioedema, which may be a presenting feature of systemic lupus erythematosus.

'Slapped cheek' erythema
This is characteristic of erythema infectiosum (due to parvovirus infection) (Anderson *et al*. 1983).

Strawberry erythema of the tongue and lips
This occurs in streptococcal infections and in Kawasaki disease, where it is associated with red palms and soles, followed a few days later by a generalized eruption.

Exanthem

1. Maculopapular eruptions are typical of viral infections. The circumscribed pink macules of rubella tend to start on the face, and later become confluent and spread to the limbs. In erythema infectiosum, the 'slapped cheek' erythema is followed by a reticulate eruption starting on the buttocks and spreading distally.
2. The 'raindrop' erythematous lesions of guttate psoriasis occur on the trunk and limbs and may follow a streptococcal sore throat. Scaling may be minimal.
3. Widespread erythema may be a feature of a connective tissue disease, particularly dermatomyositis, in which dusky erythema occurs on light exposed areas and extensor surfaces, later developing telangiectasia and atrophy (poikiloderma).
4. An evanescent, non-pruritic, pinkish maculopapular eruption occurs on the trunk and limbs in Still's disease. Often it is most marked in the late afternoon, and evening at the height of fever.
5. In rheumatic fever, annular erythematous lesions enlarge to form large figurate patches within hours.
6. Erythema chronicum migrans, the characteristic skin lesion of Lyme disease, is a solitary, annular lesion gradually expanding around the site of a tick bite (Fig. 13) (Steere 1989).

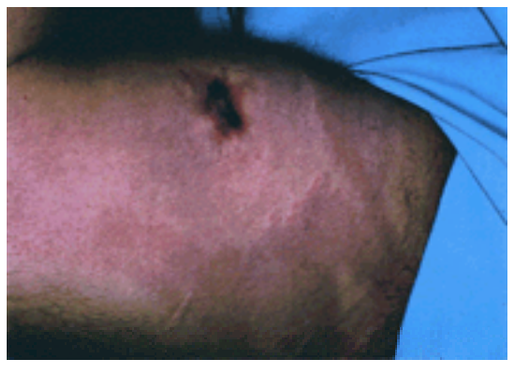


Fig. 13 Erythema chronicum migrans: annular erythema is developing around the site of the tick bite on the thigh (photograph by courtesy of Dr E. Pascual).

7. Many drugs cause a generalized maculopapular eruption (see below).

Localized erythema

Periarticular erythema may be seen around inflamed joints in rheumatoid disease and is a feature of gout and septic arthritis. Erythema may occur over early Heberden's nodes.

Erythroderma

Generalized erythema may be a presenting feature of psoriasis, particularly in the elderly male (Fig. 14). It may also be caused by drugs.

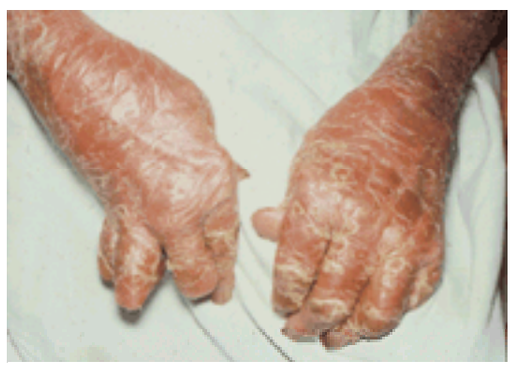


Fig. 14 Erythrodermic psoriasis involving the hands in an elderly man with psoriatic arthritis.

Scaling skin

Scaling is caused by easily detachable keratin, and is due to disordered regulation of epidermal turnover. Any acute inflammation may cause scaling related to the intensity of inflammation, e.g. after gout and drug-induced exanthems. Scaling is characteristic of the different forms of psoriasis and is seen also in lupus erythematosus, particularly discoid forms (Fig. 15 and Fig. 16). In the 'psoriasiform' subset of subacute lupus erythematosus, there is a rim of scale at the periphery of the lesion.

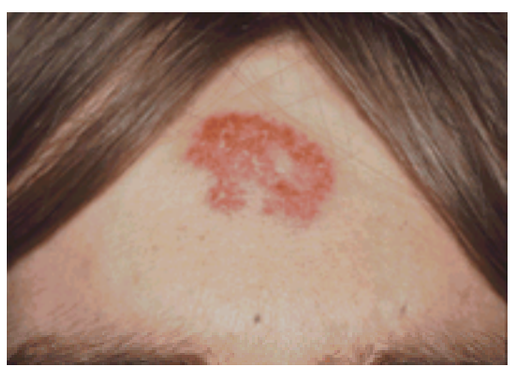


Fig. 15 Scaling plaque of discoid lupus erythematosus on a light-exposed area of forehead.

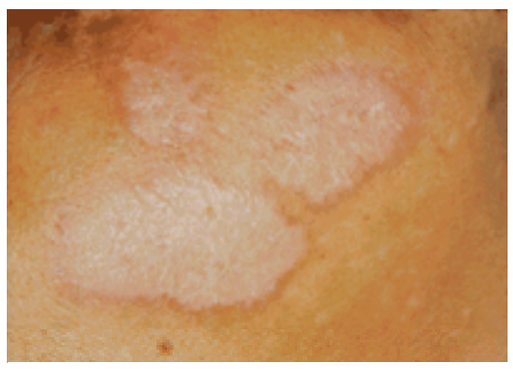


Fig. 16 Discoid lupus erythematosus on the trunk.

Blisters and pustules

Blisters are a feature of pemphigus and pemphigoid, which may be drug-induced. Bullous lesions occur rarely in morphoea, lichen sclerosus et atrophicus, and lupus erythematosus. Epidermolysis bullosa acquisita, characterized by skin fragility, blisters, and milia, may presage the development of systemic lupus erythematosus (Boh *et al*. 1990). Similar lesions are seen in porphyria cutanea tarda and 'pseudoporphyrias' associated with non-steroidal anti-inflammatory drugs (Fig. 17). Widespread bullae are seen in drug-induced erythema multiforme; typical 'target lesions' are seen, particularly on the hands and feet (Fig. 18). Drug-induced toxic epidermal necrolysis is characterized by widespread shedding of skin. Fixed drug eruptions may be bullous; they tend to recur at the same site after repeated exposure to the drug.

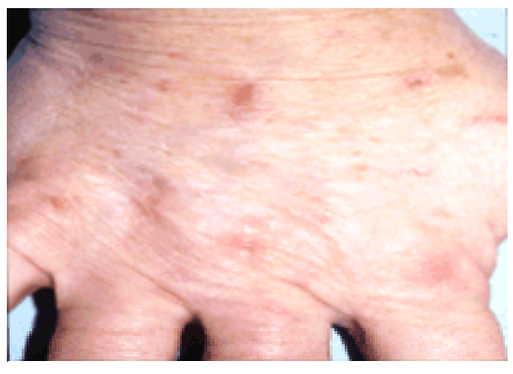


Fig. 17 'Pseudoporphyria' due to naproxen: blisters, milia and scars.

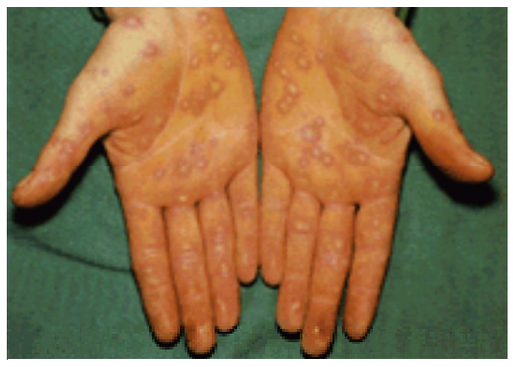


Fig. 18 Erythema multiforme: characteristic 'target lesions' on the palms.

Pustules may be localized to the hands and feet, as in Reiter's syndrome. Localized forms of pustular psoriasis can be indistinguishable. Generalized pustulation may be seen in pustular vasculitis, Behçet's syndrome, intestinal bypass syndromes, and gonococcal bacteraemia, where they may be scanty. Severe pustular acne and hidradenitis suppurativa, a chronic pustular eruption of the flexures, may be associated with inflammatory arthritis. Early lesions of pyoderma gangrenosum may be pustular. A characteristic neutrophilic eruption has been reported as typical of rheumatoid disease (Scherbenske *et al*. 1989), as has subcorneal pustular dermatosis (Sneddon–Wilkinson disease) (Butt and Burge 1995).

Plaques

Plaques on extensor surfaces and the scalp are characteristic of chronic psoriasis. The infiltrated plaques of sarcoidosis occur typically at the anterior scalp margin. Xanthelasmata are a feature of multicentric reticulohistiocytosis and hyperlipidaemias. Plaques of discoid lupus erythematosus are characterized by hyperkeratosis and scarring with 'tintack' scales. Plum-coloured plaques (Fig. 19) occur, especially around the neck and upper trunk, in Sweet's syndrome (acute neutrophilic dermatitis) (van den Driesch 1994). Indurated, tender areas are a feature of panniculitis and erythema nodosum. Localized plaques of morphoea are initially typically violaceous and later become ivory-white and sclerotic. Lesions resembling linear scleroderma are seen in melorheostosis (Fig. 20) (Wagers *et al*. 1972). Urticarial plaques may be a feature of vasculitis and familial Mediterranean fever.

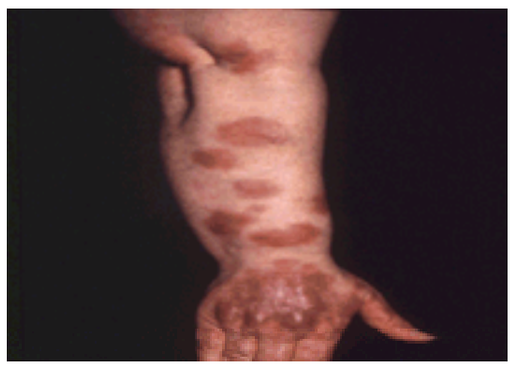


Fig. 19 Sweet's syndrome: plum-coloured plaques on forearm and hand.

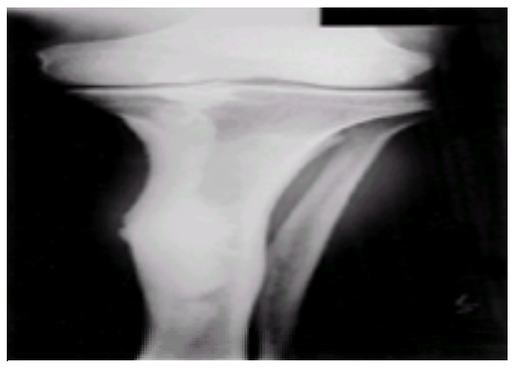


Fig. 20 Melorheostosis: the linear sclerotic streaks resemble dripping candle wax (photograph by courtesy of Dr G. Evison).

Connective tissue naevi present in childhood as firm, skin-coloured or yellowish nodules or plaques distributed asymmetrically on the trunk and proximal limbs. In the Buschke–Ollendorf syndrome they are associated with radiological features of osteopoikilosis (Fig. 21) (Verbov and Graham 1972).

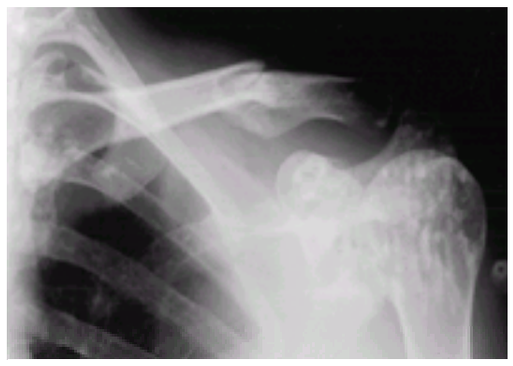


Fig. 21 Osteopoikilosis: note stippling of bone, especially adjacent to joint (photograph by courtesy of Dr G. Evison).

Papules and nodules

Important causes are summarized in Table 2.

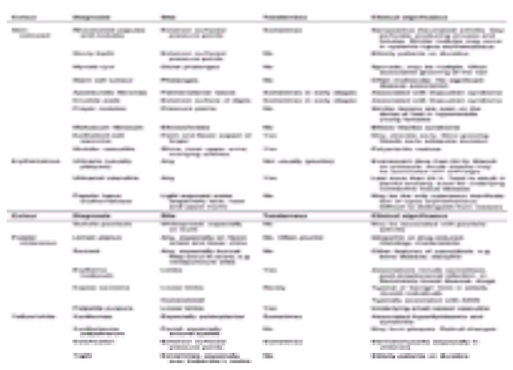


Table 2 Papules and nodules

Purpura and telangiectasia

Impalpable purpuric lesions may be due to diminished numbers of platelets or their dysfunction, for example idiopathic thrombocytopenic purpura or drug-induced marrow aplasia. Trauma, including battering and sports injuries, causes ecchymoses. The 'talon noir', a black area of subcutaneous haemorrhage, occurs on the heel typically in tennis and squash players and may be misdiagnosed as a malignant melanoma. Similarly, splinter haemorrhages and ecchymoses under nails may be due to trauma. The 'crescent sign' occurs typically over the medial or lateral malleoli after a ruptured Baker's cyst. Purpura and ecchymoses are also a feature of corticosteroid toxicity, Ehlers–Danlos syndrome, and haemophilia. Perifollicular haemorrhages occur in scurvy.

Palpable purpura reflects vasculitis, including drug-induced vasculitis and the hypergammaglobulinaemic purpura associated with Sjögren's syndrome. The matt

telangiectasia of hereditary haemorrhagic telangiectasia is indistinguishable from those seen in systemic sclerosis. Telangiectasia and atrophy (poikiloderma) are characteristic of dermatomyositis.

Ulcers

Cutaneous ulceration may have more than one cause in connective tissue diseases. Thus, in rheumatoid arthritis, punched-out ulcers due to vasculitis are common, particularly on the lower limbs. Venous hypertension may play a part in the immobile patient and ulceration may occur over nodules on pressure points. Rheumatoid arthritis is an important cause of pyoderma gangrenosum, which presents as an indurated, expanding, plum-coloured plaque or an acne-like pustule; this ulcerates early and the resulting crater has irregular, 'undermined', thickened, bluish margins. The ulcers may be multiple and heal with considerable scarring. Pyoderma gangrenosum can occur in association with other inflammatory disorders such as Behçet's syndrome.

Neurotropic ulceration occurs rarely in rheumatoid disease, in association with mononeuritis multiplex and severe vasculitis. Foot ulceration may occur with Charcot joints in the patient with syphilis. Similar changes occur in the upper limb in syringomyelia. Cranial arteritis may cause ulceration on the scalp and tongue.

Superficial erosions may be due to pemphigus or more commonly to scratching; in the absence of an obvious dermatosis, pruritus may be due to metabolic causes such as low serum iron or bile-salt retention.

Fingertip ulceration occurs in systemic sclerosis, leaving brownish scars. Severe widespread ulceration may develop rapidly in a child with active dermatomyositis, and reflects the underlying vasculopathy.

Orogenital ulceration

It is important to remember that aphthous ulcers are common in the general population.

In Behçet's syndrome, oral and genital ulcers tend to be more severe and persistent, each lesion often lasting several weeks. They tend to heal with scarring. The ulceration in Reiter's syndrome is typically superficial and painless. Oral lesions of discoid lupus erythematosus occur on the vermilion border and oral mucosa. They are dull red and may have radiating white streaks similar to lichen planus. Oral and palatal ulceration is common in systemic lupus erythematosus and reflects disease activity. Bullous disorders, particularly erythema multiforme, are associated with orogenital involvement. Oral ulceration is common in patients on methotrexate and also occurs following therapy with non-steroidal anti-inflammatory drugs. Gold therapy causes a punctate stomatitis.

Ulceration of part of the tongue may occur in cranial arteritis. A 'geographical' tongue is generally a normal variant, but can occur in psoriasis and Reiter's syndrome.

Pigmentary changes

Relevant causes of hyper- and hypopigmentation are listed in Table 3. Yellow–brown infiltrated lesions include xanthomas and xanthelasmas. Granulomas, such as sarcoid nodules, are dull yellowish on pressure. Rippled yellow streaks occur around the neck and inner arms in pseudoxanthoma elasticum.

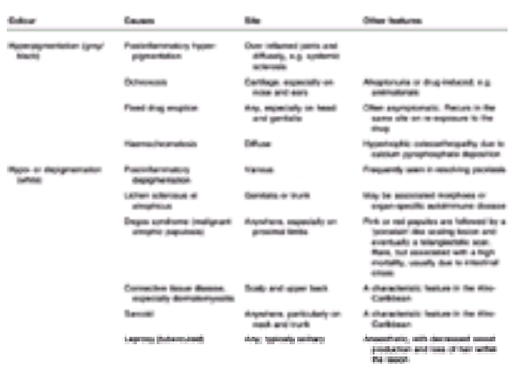


Table 3 Cutaneous hyper- and hypopigmentation

Induration or thickening of skin or subcutis (Jablonska 1975; Jayson and Black 1988)

Diffuse cutaneous sclerosis is the hallmark of systemic sclerosis, where it typically affects the extremities and face, spreading proximally. In generalized morphoea the sclerosis tends to be truncal, with sparing of the areolae; sometimes a typical violaceous margin is seen.

Scleroderma-like changes occur in several metabolic disorders, including phenylketonuria, scurvy, and porphyria cutanea tarda, where cutaneous sclerosis occurs in light-exposed areas. Several industrial chemicals trigger scleroderma-like syndromes (Table 4).

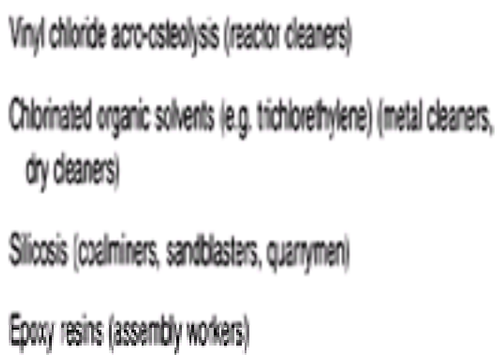


Table 4 Scleroderma-like syndromes due to industrial chemicals

Solid oedema and sclerosis of the limbs occur in the eosinophilia–myalgia syndrome associated with l-tryptophan ingestion (Connolly *et al*. 1990). Stiff swollen fingers, resembling acrosclerosis, occur in rheumatoid arthritis and in overlap connective tissue syndromes. Stiff skin and joint contractures occur rarely in diabetes mellitus. Scleroedema of Buschke affects the upper trunk and neck; the skin can be 'wrinkled' over the indurated subcutis. Fibrotic induration of the lower legs occurs in lipodermatosclerosis ('champagne-bottle legs'), associated with chronic venous stasis. Malignant infiltration of the skin, overlying lung or breast carcinoma, causes extensive erythema and induration (carcinoma erysipelatoides). Gross skin thickening, forming deep folds, is associated with thickened digits in pachydermoperiostosis (Fam *et al*. 1983).

Hair abnormalities

Increased hair growth (hypertrichosis) occurs over sites of hyperaemia such as arteriovenous fistulas and chronically inflamed joints. Drugs, such as corticosteroids,

penicillamine and cyclosporin, and metabolic disorders, such as porphyria cutanea tarda, cause hypertrichosis.

Treatment of cutaneous manifestations of rheumatic diseases

The skin has the advantage of being accessible to topical therapy; high concentrations of otherwise toxic agents can be achieved in the epidermis and superficial dermis. Percutaneous absorption of drugs is affected by many factors, including inflammation of the skin and site variation in the density of dermal connective tissue and epidermal thickness. The most important topical agents in this context are corticosteroids, which are available in bewildering variety. They can be classified in four main groups: mild, moderately potent, potent, and very potent (Table 5).

Potency	Pharmacological name	Trade name
Mild	Hydrocortisone 0.1–2.5%	Various
	Fluocinolone acetonide 0.0025%	Synalar 1:10
Moderate	Clobetasone butyrate 0.05%	Eumovate
	Desoximethasone 0.05%	Stiedex LP
	Flurandrenolone 0.0125%	Haelan
Potent	Betamethasone valerate 0.025%	Betnovate RD
	Betamethasone valerate 0.05%	Betnovate
	Desoximethasone 0.25%	Stiedex
	Fluocinonide 0.05%	Metosyn
	Hydrocortisone 17-butyrate 0.1%	Locoid
	Mometasone furoate 0.1%	Elocon*
	Fluticasone propionate 0.05%	Cutivate*
	Betamethasone dipropionate 0.05%	Diprosalic*
	Fludrocortisone acetonide 0.025%	Topilar
	Beclomethasone dipropionate 0.025%	Propaderm
	Budesonide 0.025%	Pulmicort
Very potent	Clobetasol propionate 0.05%	Dermovate
	Mometasone furoate 0.1%	Elocon
	Flurandrenolone-impregnated tape	Haelan tape — for hospital use only

*Licensed for one whole daily treatment.
†Licensed for nocturnal dressing of hyperkeratinous eczematous skin.
‡Addition of salicylic acid to formulation makes it useful in hyperkeratosis lesions, e.g. corticosteroid tape, psoriasis tape, pharmaceuticals.
Common immunotherapeutic preparations are available for infected lesions and combined preparations for use in infected or suspected psoriasis.

Table 5 Relative potencies of some commonly used topical corticosteroids

Topical sunscreens are also important, particularly in individuals with connective tissue disease. Several are freely available and a few may be had on prescription. The sun protection factor is the ratio of the dose of a particular wavelength of ultraviolet (**UV**)B irradiation that causes minimum erythema with sunscreen over that without sunscreen. Most sunscreens protect the skin adequately from short wavelengths of UV (UVC and UVB) but are inadequate against long-wavelength UV (UVA) and visible light. There is no internationally agreed standard for measuring the degree of protection against UVA or visible light. In recent years, sunscreens are becoming available that contain micromized particles of titanium dioxide or zinc oxide. These preparations block a wider spectrum of light than those that rely on chemical absorption at a specific wavelength. However, individual preference is important; it is essential that the patient is willing to use the preparation regularly.

Systemic agents, such as b-carotene, are somewhat photoprotective.

Cosmetic camouflage is particularly effective for vascular lesions, such as matt telangiectasia in systemic sclerosis.

Treatment of cutaneous ulceration requires careful assessment of the causes and correction of contributory factors such as severe anaemia.

Large, punched-out ulcers reflect necrotizing vasculitis and justify aggressive immunosuppressive therapy. Most vasculitic ulcers are clean and require minimal local interference. Excessive slough can be removed by surgical debridement or the application of a preparation such as Varidase (streptokinase/streptodornase powder for preparing topical solutions). Clean indolent ulcers should respond to hydrocolloid dressings. Sometimes admission to hospital is necessary, often for long periods. Grafting has been found to be necessary on occasion. There are reports of the beneficial use of methotrexate in intractable rheumatoid ulcers.

Rapidly expanding ulcers with a bluish 'undermined' edge should raise the possibility of pyoderma gangrenosum. This requires specific therapy; usually prednisolone (40–60 mg/day) results in rapid healing. Stubborn lesions may require intralesional triamcinolone. Milder forms of pyoderma gangrenosum sometimes respond to minocycline (100 mg twice daily). Other therapeutic agents include clofazimine, dapsone, cyclosporin, and topical disodium cromoglycate.

Digital ulceration occurs with Raynaud's phenomenon, particularly in systemic sclerosis. Patients should avoid excessive manicuring of the cuticles. Digital ulcers can be treated with a topical antimicrobial. Hydrocolloids or occlusive dressings can be helpful.

Management of specific disorders

Cutaneous lupus

A moderate to very potent steroid ointment is generally necessary for chronic discoid lesions, even on the face. Stubborn lesions may even require polythene occlusion. For small lesions, flurandrenolone tape can be cut to size. Prolonged use of topical corticosteroids produces telangiectasia and cutaneous atrophy. Rarely, intralesional triamcinolone (10 mg/ml) is necessary, particularly for lesions on the ears and nose. Injections should be directed as superficially as possible, but even so there is still a risk of atrophic scarring. Disfiguring lesions may respond well to excision and grafting or even dermabrasion. CO_2 laser therapy is an alternative destructive technique (Henderson and Odom 1986)

Sunscreens should be used regularly and, in some cases of subacute cutaneous lupus (Fig. 22), may be the only therapy required.

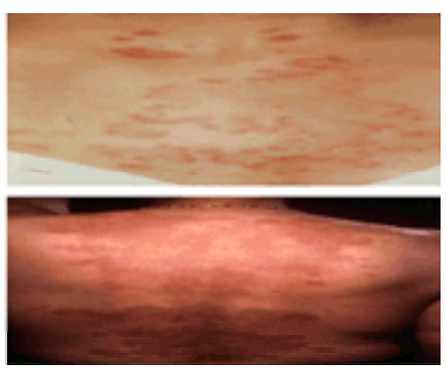


Fig. 22 Subacute cutaneous lupus erythematosus affecting (a) the neck and upper chest, (b) the back. This condition flared after sun exposure and lesions cleared totally with the use of a sunscreen.

In many patients, systemic therapy is necessary. Antimalarials are the drugs of choice, either hydroxychloroquine (400 mg daily) or chloroquine (200 mg daily). Mepacrine may be preferred, although it causes yellow discoloration of the skin and sclerae in therapeutic doses (Fig. 23). (This may enhance the photoprotective effect of the drug.) A combination of two antimalarials may have a synergistic effect without necessarily increased toxicity, although this observation has not yet been confirmed in a controlled study. Although the retinal risk is minimal, patients on antimalarials continuously should have annual ophthalmological assessments. Other drugs used in cutaneous lupus include clofazimine (100 mg daily), dapsone (50–100 mg daily), azathioprine (Askinoff *et al*. 1988), and oral gold (6 mg daily). Aromatic retinoids such as acetretin (50–75 mg daily) (Ruzicka *et al*. 1988) decrease hyperkeratosis and can be used together with antimalarials. Some patients with chilblain lupus or lupus profundus respond dramatically to thalidomide (Knop *et al*. 1983). It is unusual for systemic steroids to be necessary for cutaneous lesions alone.

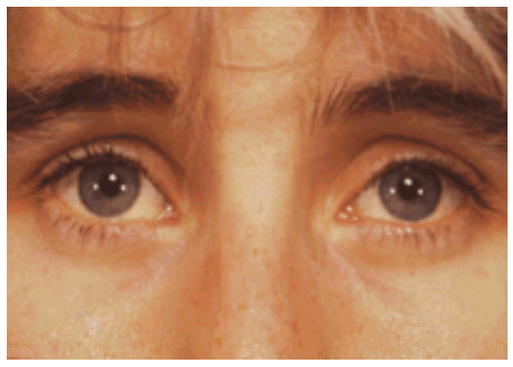


Fig. 23 Yellow pigmentation of the skin and sclerae due to mepacrine.

Dermatomyositis

Extensive skin lesions can be treated with a moderately potent topical steroid. Sunscreens are important. Systemic treatment is usually required. Mild cases respond well to antimalarials but most require corticosteroids, initially in moderately high doses (1 mg/kg body wt) with an immunosuppressive agent, such as methotrexate, as a steroid-sparing agent.

Scleroderma

Localized forms of scleroderma (e.g. morphoea) are often self-limiting, although a topical corticosteroid or occasionally intralesional triamcinolone may help. Extensive lesions may require systemic therapy. Penicillamine (300–600 mg daily) appears to 'soften' skin lesions and may reduce contractures (Fig. 24). Sulphasalazine has been used in generalized morphoea. Severe hemifacial atrophy and ossifying lesions may respond well to plastic surgical reconstruction, sometimes with tissue expansion (Handfield-Jones *et al*. 1988).

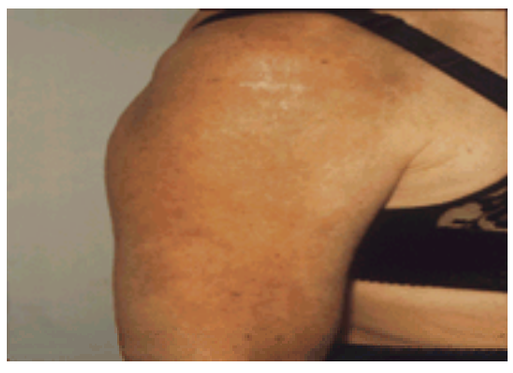


Fig. 24 Linear morphoea: this patient observed increased joint mobility and softening of the skin following penicillamine therapy.

Treatment of the skin lesions in systemic sclerosis is generally supportive, and includes emollients, and treatment of digital ulcers.

Itch is often a major symptom in patients with systemic sclerosis. Its cause is unclear. Symptomatic treatments include the use of emollients such as aqueous cream or emulsifying ointment, applied directly to the skin or as soap substitutes; these also counteract dryness of the skin. Sedative antihistamines may help, as may topical preparations such as 1 to 2 per cent menthol in aqueous cream, or even topical corticosteroids.

Digital ulcers may respond to topical nitroglycerine or stable prostaglandins. In severe digital ulceration, or incipient gangrene of a digit, intravenous infusions of vasodilators such as prostaglandin E_1 or prostacyclin may be helpful.

Psoriasis

Superficial lesions (e.g. guttate psoriasis) respond to topical tar preparations, the topical vitamin D analogue, calcipotriol, or dilute corticosteroids. Potent steroids should not be used. Dithranol is used for thicker plaques. Its use should be supervised carefully as it is irritant and stains skin and clothing. Several different concentrations and bases are available. Initially a weak preparation (e.g. 0.1 per cent) can be used as short-contact therapy and washed off after 30 to 120 min. The strength and duration of therapy can be increased gradually. Dithranol must not be used on the flexures or face. Often admission or day-centre therapy is necessary. Other treatments include tar baths and ultraviolet light.

A minority of patients with extensive or unstable disease (e.g. generalized pustular psoriasis) will require second-line therapy. This includes drugs such as etretinate and acetretin, aromatic retinoids related to vitamin A. They are teratogenic and may cause hypercholesterolaemia and hypertriglyceridaemia. Methotrexate, in low weekly dosage (5–15 mg) is especially useful in erythrodermic or generalized pustular psoriasis, particularly in elderly people. Since the drug appears to be more hepatotoxic in patients with psoriasis than those with rheumatoid disease, most dermatologists would carry out routine liver biopsies in psoriatic patients under 65 years on long-term methotrexate therapy. Hydroxyurea (0.5–1.5 g) is sometimes of value in a proportion of patients; although it may cause short-term marrow toxicity, fatalities are rare (Layton *et al*. 1989). There are encouraging results from cyclosporin, initially 3 to 5 mg/kg per day, in patients with severe disease (Ellis *et al*. 1991), although the long-term renal toxicity of the drug is uncertain. Regular monitoring should include measurement of blood pressure, serum creatinine and measurement of glomerular filtration rate.

PUVA (psoralen + UVA) is valuable in recalcitrant psoriasis of hands and feet or in some individuals with widespread disease. It is particularly effective in combination with low-dose retinoids (Saurat *et al*. 1988). PUVA is potentially carcinogenic to the skin, particularly in patients who have previously received other carcinogens such as methotrexate.

Antimalarials should not be used in psoriatic arthritis, as they may exacerbate cutaneous psoriasis.

Cutaneous side-effects of antirheumatic therapy

Skin lesions are common in patients receiving antirheumatic therapy. They are often non-specific and trivial, although sometimes they may be severe enough to warrant stopping the drug. In the following we list important cutaneous adverse effects of the different forms of antirheumatic therapy (Breathnach and Hintner 1992; Lovell and Maddison 1992).

Non-steroidal anti-inflammatory drugs

1. Non-specific exanthems/morbilliform eruptions are common.
2. Generalized exfoliative dermatitis is rare (has been caused by phenylbutazone, oxyphenbutazone, and indomethacin).
3. Photoallergic reactions (e.g. caused by fenbufen, piroxicam).
4. Photo-onycholysis, skin fragility in light exposed sites, and features resembling porphyria cutanea tarda appear to be a direct phototoxic effect of drugs. These

effects were seen typically with benoxaprofen, but can occur with other non-steroidal anti-inflammatory drugs.
5. Acneiform lesions (e.g. caused by naproxen).
6. Stomatitis has been reported with indomethacin.
7. Urticaria may be immunologically mediated (e.g. caused by diclofenac, mefanamic acid) or may be a pharmacological effect of the drug (typically aspirin in patients with nasal polyposis).
8. Erythema multiforme and toxic epidermal necrolysis (Fig. 25) (Korstanje 1995) (e.g. caused by fenbufen, benoxaprofen, piroxicam, sulindac, and especially phenylbutazone, oxyphenbutazone, and diclofenac, where it has been fatal).

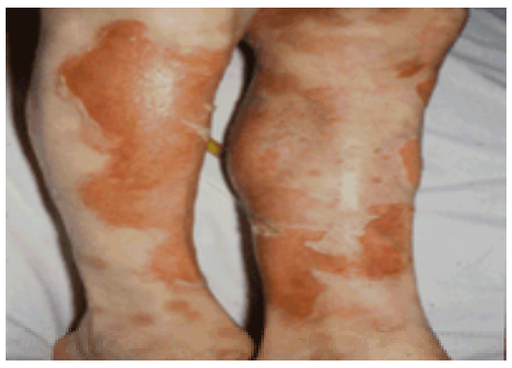


Fig. 25 Toxic epidermal necrolysis.

9. Fixed drug eruptions (Fig. 26). One or more well-demarcated erythematous lesions, sometimes with blistering, resolve, leaving persistent, dusky grey-black pigmented areas. The eruption occurs at the same site, with or without further lesions, on re-exposure to the drug. Systemic symptoms are few, if any, and re-exposure to the drug is justifiable if diagnostically necessary. Oxyphenbutazone is a classical cause, although a non-pigmenting fixed drug eruption has been reported due to piroxicam.

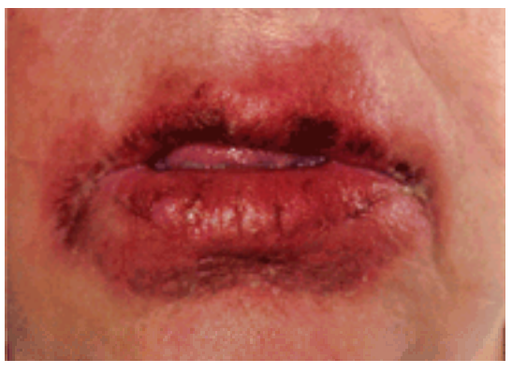


Fig. 26 Fixed drug eruption. A well-demarcated, often ovoid, dusky area of erythema, later followed by slatey hyperpigmentation.

10. Several topically applied non-steroidals (e.g. ketoprofen) cause allergic contact dermatitis and photoallergic contact dermatitis.
11. Purpura may be due to thrombocytopenia (e.g. caused by fenoprofen, phenylbutazone) or allergic vasculitis (e.g. caused by phenylbutazone, naproxen).

Gold

Cutaneous reactions occur at any stage of treatment in 25 per cent more of patients on intramuscular gold salts; they may require cessation of therapy. A high incidence of mucocutaneous side-effects appears to be linked with therapeutic efficacy. Gold-salt dermatitis appears to be related to HLA class I B35 antigen. Oral gold (auranofin) causes similar, although less severe, toxic reactions to intramuscular gold. Important side-effects include the following.

1. Pruritus.
2. Non-specific maculopapular/morbilliform eruptions—these may progress to fatal erythroderma if the drug is continued.
3. An eczematous or pityriasis rosea-like eruption. Even in relatively banal eruptions, histology may reveal striking features resembling lichen planus. Typical lichen planus (Fig. 27) may develop, with hyperkeratotic changes, especially in the scalp, sometimes causing scarring alopecia.

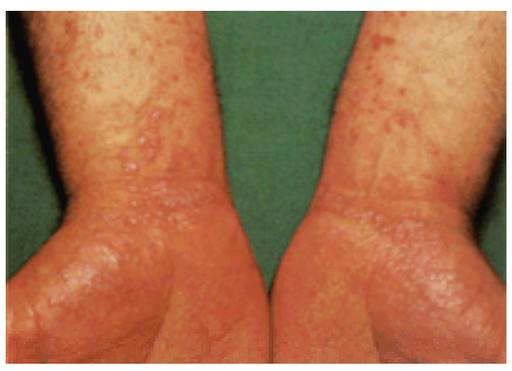


Fig. 27 Lichen planus: typical violaceous papules. Drug-induced lichen planus may be indistinguishable from the idiopathic form or may also exhibit eczematous features.

4. Intramuscular gold therapy may reactivate pre-existing nickel sensitivity (e.g. to jewellery).
5. Punctuate or aphthous stomatitis (common, may be the only adverse effect).
6. Purpura due to thrombocytopenia or (rarely) vasculitis.
7. Erythema multiforme and toxic epidermal necrolysis (rare).
8. Prolonged gold therapy may cause slate-grey pigmentation (chrysiasis), especially in light-exposed sites.

D-Penicillamine

Many adverse effects from this drug are similar to those seen in gold therapy. An urticarial or morbilliform eruption is common early in therapy, although usually it can be avoided by a 'go low, go slow' dose regimen.

Immunologically mediated skin reactions may mimic inflammatory dermatoses such as lichen planus. The eruptions may persist for several years after the drug is discontinued. Typical cicatricial pemphigoid may occur, with the development of fibrous conjunctival bands (synechiae). Pemphigus foliaceus is the most common pattern induced by penicillamine; it may be mild and the crusted erythematous truncal lesions may mimic eczema (Fig. 28). However, direct immunofluorescence of uninvolved skin reveals the characteristic intercellular deposition of IgG. More rarely, the drug may induce pemphigus vulgaris, in which flaccid bullae rupture to form extensive eroded areas of skin. Oral lesions and vulvovaginitis may occur. Penicillamine may precipitate lesions closely resembling morphoea, systemic sclerosis, dermatomyositis, and both discoid and systemic lupus erythematosus. Antibodies to native DNA are found, unlike in other forms of drug-induced lupus. Other

inflammatory adverse effects include thrombocytopenic purpura, stomatitis, glossitis, and alopecia.

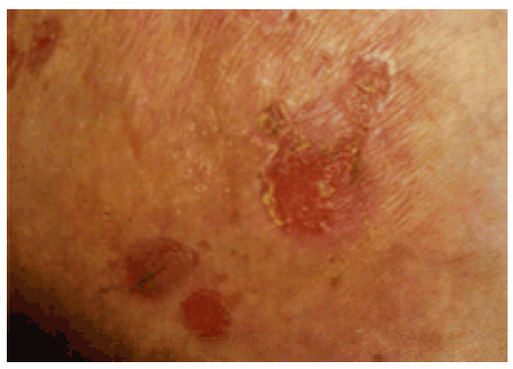


Fig. 28 Pemphigus foliaceus of the trunk, induced by D-penicillamine. The superficial, eroded eruption can be confused with eczema.

Penicillamine has a profound effect on connective tissue proteins: it inhibits collagen synthesis and cleaves newly formed intermolecular cross-links; it also stimulates elastin synthesis. Although effects on connective tissue are generally associated with the high doses used in metabolic disorders such as Wilson's disease, elastosis perforans serpiginosa has been reported in one child receiving penicillamine for rheumatoid arthritis. Typical features include serpiginous and annular crusted plaques on the limbs. Histological changes include transepidermal elimination of elastin. Pseudoxanthoma elasticum-like changes and elastosis may occur in relatively low doses (e.g. 750 mg/day).

Sulphasalazine

Eruptions caused by the sulphapyridine moiety are typical of those induced by sulphonamides in general. They include:

1. leucocytoclastic vasculitis;
2. photosensitivity;
3. a pruritic, scaly, maculopapular eruption (common)—desensitization may be possible;
4. a lupus-like syndrome;
5. erythema multiforme/toxic epidermal necrolysis (rare).

Antimalarials

Reactions include:

1. diffuse grey itchy pigmentation, especially in light-exposed areas;
2. blackish pigmentation (especially nose and ears) due to acquired ochronosis—nails may also be affected;
3. bright yellow pigmentation (mepacrine);
4. bleaching of red or blonde hair;
5. photosensitivity or exacerbation of porphyria cutanea tarda;
6. exacerbation of psoriasis;
7. toxic epidermal necrolysis (rare);
8. lichenoid eruptions (rare);
9. a generalized pustular eruption (hydroxychloroquine).

Corticosteroids

Reactions include:

1. atrophy, telangiectasia, and easy bruising;
2. increased risk of cutaneous sepsis;
3. dermal atrophy, causing dimpled scars (Fig. 29)—this may follow injudicious local infiltration with corticosteroids.

Fig. 29 Dermal atrophy from corticosteroid injection.

Immunosuppressive drugs

Reactions include:

1. Opportunistic infections (e.g. multiple warts and molluscum contagiosum lesions) may follow chronic immunosuppressive therapy (e.g. with azathioprine and methotrexate).
2. Cytotoxic agents cause alopecia.
3. Pigmentation of skin and nails occurs with cyclophosphamide.
4. Hypersensitivity reactions to azathioprine may present as an exanthem or generalized urticarial eruption.
5. Orogenital ulcers occur typically with methotrexate.

Some key points for skin manifestations of rheumatic diseases (Table 6)

1. Examination of the skin in good light may clarify a rheumatological diagnosis. Examine skin thoroughly, including nailfolds, genitalia, and feet. Examine oral mucosa.
2. Describe and record skin lesions accurately
3. Skin biopsy for conventional histology or immunofluorescence is often helpful. Try to biopsy an early lesion and orientate correctly.
4. Cutaneous manifestations of rheumatic diseases require specific, often topical, therapy. Don't be afraid to use potent topical corticosteroids in cutaneous lupus—but keep under review. Don't overtreat cutaneous lupus with systemic therapy.
5. Antirheumatic drugs produce a wide range of skin side-effects; even after several months of therapy, cutaneous intolerance may develop. In any rash that doesn't fit a precise pattern—consider a drug cause.
6. Stress the importance of photoprotection.

Table 6 Important clinical points

Chapter References

Anderson, M.J. *et al*. (1983). Human parvovirus, the cause of erythema infectiosum (fifth disease). *Lancet*, **ii**, 1378–81.

Ashton, R.E. (1995). Teaching non-dermatologists to examine the skin: a review of the literature and some recommendations. *British Journal of Dermatology*, **132**, 221–5.

Askinoff, R., Werth, V.P., and Franks, A.G., Jr (1988). Resistant discoid erythematosus of palms and soles: Successful treatment with azathioprine. *Journal of the American Academy of Dermatology*, **19**, 961–5.

Boh, E., Roberts, L.J., Lieu, T-S., Gammon, W.R., and Sontheimer, R.D. (1990). Epidermolysis bullosa acquisita preceding the development of systemic lupus erythematosus. *Journal of the American Academy of Dermatology*, **22**, 587–93.

Breathnach, S.M. and Hintner, H. (1992). *Adverse drug reactions and the skin*. Blackwell Scientific, Oxford.

Burton, J.L. (1985). *Essentials of dermatology*. Churchill Livingstone, Edinburgh.

Butt, A. and Burze, S.M. (1995). Sneddon-Wilkinson disease in association with rheumatoid arthritis. *British Journal of Dermatology*, **132**, 313–15.

Connolly, S.M. *et al*. (1990). Scleroderma and l-tryptophan: a possible explanation of the eosinophilia–myalgia syndrome. *Journal of the American Academy of Dermatology*, **23**, 451–7.

Ellis, C.N. *et al*. (1991). Cyclosporin for plaque-type psoriasis. Results of a multidose, double-blind trial. *New England Journal of Medicine*, **324**, 277–84.

Fagrell, B., Eriksson, S.E., Malmstrom, S., and Sjolund, A. (1988). Computerised data analysis of capillary blood cell velocity. *International Journal of Microcirculation Clinical and Experimental*, **7**, 276–81.

Fam, A.G., Chin-Sang, H., and Ramsay, C.A. (1983). Pachydermoperiostosis: scintigraphic, thermographic, plethysmographic, and capillaroscopic observations. *Annals of the Rheumatic Diseases*, **42**, 98–102.

Handfield-Jones, S.E. *et al*. (1988). Ossification in linear morphoea with hemifacial atrophy—treatment by surgical excision. *Clinical and Experimental Dermatology*, **13**, 385–8.

Harrison, P.V. (1980). A guide to skin biopsies and excisions. *Clinical and Experimental Dermatology*, **5**, 235–45.

Henderson, D.L. and Odom, J.C. (1986). Laser treatment of discoid lupus erythematosus. *Lasers in Surgical Medicine*, **6**, 12–15.

Houtman, P.M., Kallenberg, C.G.M., Fidler, V., and Wouda, A.A. (1986). Diagnostic significance of nailfold capillary patterns in patients with Raynaud's phenomenon. *Journal of Rheumatology*, **13**, 556–63.

Jablonska, S. (1975). *Scleroderma and pseudoscleroderma*. Polish Medical Publishers, Warsaw.

Jacohs, M.J.H.M., Breslau, P.J., Slaaf, D.W., Reneman, R.S., and Lemmens, J.A.J. (1987). Nomenclature of Raynaud's phenomenon—a capillary microscopic and hemorheologic study. *Surgery*, **101**, 136–45.

Jayson, M.I.V. and Black, C.M. (1988). *Systemic sclerosis: scleroderma*. Wiley, Chichester.

Knop, J. *et al*. (1983). Thalidomide in the treatment of 60 cases of chronic discoid lupus erythematosus. *British Journal of Dermatology*, **108**, 461–6.

Korstanje, M.J. (1995). Drug-induced mouth disorders. *Clinical and Experimental Dermatology*, **20**, 10–18.

Layton, A.M. *et al*. (1989). Hydroxyurea in the management of therapy-resistant psoriasis. *British Journal of Dermatology*, **121**, 647–53.

Love, T.J. (1980). Thermography as an indicator of blood perfusion. *Annals of the New York Academy of Sciences*, **335**, 429–37.

Lovell, C.R. and Maddison, P.J. (1992) Rheumatoid arthritis and the skin. In *Recent advances in dermatology* (ed. R. H. Champion and R. J. Pye), pp. 1–16. Churchill Livingstone, London.

Lovell, C.R., Maddison, P.J., and Campion, G.V. (1990). *The skin in rheumatic disease*. Chapman and Hall Medical, London.

Mahler, F., Saner, H., Bose, C., and Annaheim, M. (1987). Local cold exposure test for capillaroscopic examination of patients with Raynaud's syndrome. *Microvascular Research*, **33**, 422–7.

Maricq, M.R. (1988). Raynaud's phenomenon and microvascular abnormalities in scleroderma (systemic sclerosis). In *Systemic sclerosis: scleroderma* (ed. M. I. V. Jayson and C. M. Black), pp. 151–66. Wiley, Chichester.

Ring, E.F.J. and Phillips, B. (ed.) (1984). *Recent advances in medical thermology*. Plenum, London.

Ros, A. and Wennersten, G. (1986). Current aspects of polymorphous light eruptions in Sweden. *Photodermatology*, **3**, 298–302.

Ruzicka, T., Meurer, M., and Bieber, T. (1988). Efficacy of acetretin in the treatment of cutaneous lupus erythematosus. *Archives of Dermatology*, **124**, 897–902.

Saurat, J-H. *et al*. (1988). Randomised double-blind multicentre study comparing acetretin–PUVA, etretinate–PUVA and placebo–PUVA in the treatment of severe psoriasis. *Dermatologica*, **177**, 218–24.

Scherbenske, J.M. *et al*. (1989). Rheumatoid neutrophilic dermatitis. *Archives of Dermatology*, **125**, 1105–8.

Steere, A.C. (1989). Lyme disease. *New England Journal of Medicine*, **321**, 586–96.

Verbov, J. and Graham, R. (1972). Buschke–Ollendorff syndrome: disseminated dermofibrosis with osteopoikilosis. *Clinical and Experimental Dermatology*, **11**, 17–26.

van den Driesch, P. (1994). Sweet's syndrome (acute febrile neutrophilic dermatosis). *Journal of the American Academy of Dermatology*, **31**, 535–56.

Wagers, L.T., Young, A.W., and Ryan, S.F. (1972). Linear melorheostotic scleroderma. *British Journal of Dermatology*, **86**, 297–301.

Will, R.K., Ring, E.F.J., Clarke, A.K., and Maddison, P.J. (1992). Infrared thermography: what is its place in rheumatology in the 1990's? *British Journal of Rheumatology*, **31**, 337–44.

1.3.3 Neurological complications

I. T. Ferguson and Peter Hollingworth

Introduction

When acute, as in cervical cord compression, neurological complications are among the few rheumatological emergencies. When they cause a gradual functional decline, they may be misconstrued as deterioration of the arthritis. Either way, the addition of weakness or sensory symptoms to an already painful, stiff limb may result in severe disability. Drugs used in the treatment of rheumatic diseases (Blain and Stewart-Wynne 1986 ; Mastaglia 1989) can themselves cause neurological side-effects (Table 1). Coincidental neurological diseases must be considered in the differential diagnosis and some neurological disorders have musculoskeletal manifestations (Table 2). Such is the interplay between rheumatological and neurological disorders that the rheumatologist must be alert to the diagnostic pitfalls they present (Table 3).

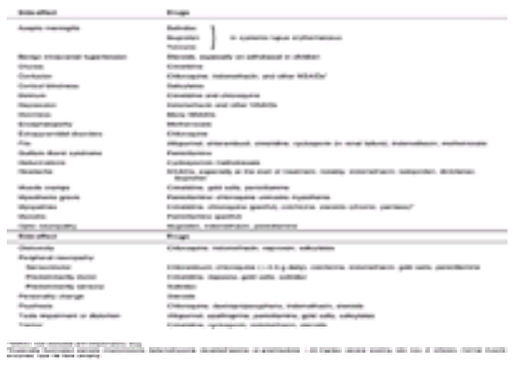


Table 1 Neurological side-effects of drugs used for rheumatic disorders

Neurological disorders	Rheumatic symptoms and signs
Parkinson's disease	Partial, stiff shoulder, arm or hand; less frequently lower limb
Multiple sclerosis	Stiff (spastic) legs may be confused with or coincide with cervical spondylitic myelopathy
Spinal cord tumour	Radicular pain with aching, stiff legs; legs feel swollen with dorsal column involvement
Dorsal root entry-zone lesion e.g. multiple sclerosis, tumour, neuropathies	Lightning limb pains
Motor neurone disease	Wasted hand, confused with T1 root lesion from cervical spondylosis
Syringomyelia, tabes dorsalis, primary sensory neuropathy, diabetes	Neuropathic joints; swollen fingers (*main en casserole*) with syringomyelia
Hereditary sensorimotor neuropathy (Charcot-Marie-Tooth)	Pes cavus
Myotonias	Stiff hands, often worse in cold
Myasthenia	Weak, sometimes aching muscles, may be confused with polymyositis

Table 2 Neurological disorders presenting with rheumatic symptoms (see also Table 3)

Table 3 Some important points to remember

Accurate assessment of power, tone, reflexes, or the plantar response may be impossible in a limb already weak, stiff, or deformed from arthritis, so recognition of

these complications demands a high index of suspicion and help from sophisticated investigations.

The recognition of patterns of symptoms and signs is necessary in forming a diagnosis (Box 1).

Box 1 Important points for the rheumatologist to remember

1. Joint damage often prevents a meaningful clinical assessment of neurological symptoms, so special investigations are usually necessary.
2. Beware of attributing deteriorating function to progression of joint disease in rheumatoid arthritis: consider superimposed neurological complications.
3. Be alert to hysteria, mimicking neurological disease, coexisting with rheumatological or organic neurological diseases.
4. The differential diagnosis of cerebral lupus includes complications of lupus uraemia, sepsis and hypertensive encephalopathy, and side-effects of drugs in its treatment.
5. Cervical myelopathy and rheumatoid arthritis is suggested by deteriorating function, 'electric shocks' in the arms; recent worsening of neck pain.
6. Mononeuritis multiplex may progress rapidly and is a rheumatological emergency.
7. Entrapment neuropathies are the commonest neurological complication of rheumatic diseases.

Weakness patterns

The distribution of weakness is important. Symmetrical, proximal weakness suggests muscle disease but other causes should be considered (Table 4). Difficulty getting out of a chair, climbing stairs, lifting objects from a high shelf, or brushing hair are common.

Table 4 Causes of proximal weakness in adults

In upper motor neurone disease, shoulder abductors and arm extensors are affected more than the flexors, whereas in the legs the flexors are affected more than the extensors. Hyper-reflexia, hypertonia, and an extensor plantar response are usually present.

Myasthenia gravis is characterized by fluctuating weakness with fatiguability. The weakness tends to be proximal and associated with diplopia and dysarthria towards the end of the day. Symptomatic myasthenia responding to intravenous edrophonium can occur in polymyositis, where a reduced number of acetylcholine receptors has been demonstrated. Penicillamine may produce myasthenia responsive to the edrophonium test. Antibodies to acetylcholine receptors are present and recovery usually follows withdrawal of the drug. The Eaton–Lambert syndrome, with painful, proximal muscle weakness, is not associated with antibodies to acetylcholine receptors. Three-quarters of males and 30 per cent of females will have an underlying carcinoma. Electromyography is diagnostic.

Distal weakness usually results from disease of peripheral nerves, either of their roots or a polyneuropathy. Involvement of specific muscle groups may allow precise anatomical localization (Guarantors of Brain 1988). Weakness of one limb is usually neurogenic, except for tendon rupture. The differential diagnosis includes motor neurone disease, neuralgic amyotrophy, and early fascioscapulohumeral dystrophy. Hysterical or non-organic weakness (Table 5) is commonly superimposed on organic disease (Marsden 1986).

Great display of effort to little effect
Intermittent or jerky weakness
Simultaneous contraction of agonist and antagonist muscles
Failure of stabilizing or synergistic muscles to contract when testing limb power
Inconsistencies in performance: for example the patient may have weak quadriceps when lying on the couch yet be able to walk, squat, or jump normally

Table 5 Features of hysterical weakness

Muscle pain

When at rest, this is rarely from primary muscle disease. Exertion pain is commonly non-organic, once peripheral vascular disease has been excluded. However, McArdle's disease (mild myophosphorylase deficiency) and phosphofructokinase deficiency should be considered if the pain is in the calves, increases with further exertion, and leads to myoglobulinuria. The ischaemic lactate test, electromyography, and muscle biopsy are necessary for diagnosis.

Muscle twitching

Fasciculation is painless, visible, muscle contractions of small groups of muscle fibres, often felt by the patient. It is benign in the absence of weakness or wasting which, if present, suggests anterior horn cell disease or motor neuropathy. Fibrillation is contraction of single fibres seen in the tongue or exposed muscle, and may be recorded by electromyography. Myokymia is a slower or coarse contraction of bands of muscle fibres, most often seen in healthy individuals. Facial myokymia can occur in brain-stem multiple sclerosis and Whipple's disease.

Contractures

These occur in both neurogenic and myopathic muscle disease; Duchenne muscular dystrophy, spinal muscular atrophy, and Friedreich's ataxia are common causes. They may develop rapidly in acute polymyositis.

Reflex patterns

Deep tendon reflexes may be absent at joints stiffened by arthritis. The plantar response may be difficult to assess in hallux valgus or after Keller's operation. Absent ankle jerks after the age of 60 is often normal. Exaggerated reflexes, especially with clonus, point to an upper motor-neurone lesion. Absent or depressed reflexes occur in peripheral neuropathies. A combination of an absent segmental reflex and hyper-reflexia below that level, as in the inverted supinator sign (absent biceps and supinator jerk but present triceps jerk, indicating a C5/6 cord lesion), is of considerable localizing value. In polymyositis, the reflexes may be increased, possibly

because of muscle spindle hyperexcitability secondary to inflammation.

Sensory patterns

Sensory loss may be the most important clue to involvement of the nervous system in patients with arthritis. The distribution and type of sensory deficit usually permits anatomical localization (Guarantors of Brain 1988); for example anaesthesia of the little finger and ulnar half of the ring finger and palm, extending no further than the wrist crease, indicates an ulnar nerve lesion rather than a C8 radiculopathy.

Knowledge of key dermatomes (Fig. 3) and the area supplied by peripheral nerves is essential. Loss of pain and temperature with sparing of joint position sense points to a cord or brain-stem lesion. Cervical cord compression from atalantoaxial subluxation may be confused with a peripheral neuropathy, as the sensory loss can be in a glove-and-stocking distribution, perhaps from cord ischaemia. As sensory examination is subjective, electrophysiological testing may be the only way to determine the site of nerve damage.

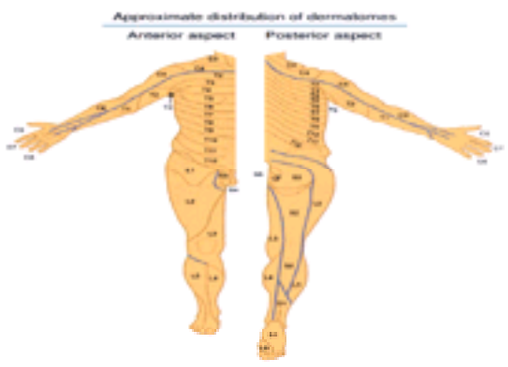


Fig. 3 Variation and overlap occurs between dermatomes. Also applies to fingers, but thumb is usually supplied by C6 and little finger C8. (Adapted from Guarantors of Brain 1988.)

Central nervous system involvement in rheumatic diseases

Fortunately, central nervous system complications are a common manifestation of only the rarer rheumatic diseases. Most are mild and self limiting, but some have devastating consequences. Difficulties abound in understanding these complications as there is a dearth of agreed diagnostic criteria, specific signs, laboratory tests, or imaging techniques. When a patient with a rheumatic disease disorder develops central nervous system disease, one should ask whether it is:

1. a direct complication and, if so, is it residual or progressive;
2. secondary to another complication, such as uraemia, infection, or hypertension;
3. a side-effect of therapy, such as steroid psychosis or drug-induced systemic lupus erythematosus;
4. non-organic, such as hysteria or depression;
5. coincidental, such as a cerebral thrombosis from atheroma.

The most common and closely studied central nervous system disorders that present to the rheumatologist occur in systemic lupus erythematosus (Feinglass *et al*. 1976; Gibson and Myers 1976; Adelman *et al*. 1986) and the allied lupus-like and antiphospholipid syndrome (Hughes 1985; Asherson *et al*. 1989). Details are given in the appropriate chapters but how these concern the neurologist will be considered in depth here.

Neuropsychiatric systemic lupus erythematosus

The incidence of neuropsychiatric systemic lupus erythematosus may be as high as 50 per cent (Hanly *et al*. 1992) but the incidence of serious complications, as opposed to minor psychiatric disturbance, is far less. The most common features are depression, fits, and headaches, although virtually every manifestation of central nervous system disease has been described (Singer and Denburg 1990), including subtle cognitive defects in memory, intellect, and learning (Hanly *et al*. 1994). Most victims are young women. It usually occurs when the systemic lupus erythematosus is active and early in its course, though some patients have had seizures for years before the patient presents with other manifestations. (Details are given in Chapter 5.7.1)

Pathogenic factors

1. Antineuronal antibodies, directed notably against the cytoskeletal neurofilaments, tend to be associated with diffuse neuropsychiatric systemic lupus erythematosus (Robbins *et al*. 1988).
2. IgG anticardiolipin antibodies occur in patients with antiphospholipid syndrome and in 20 to 50 per cent with systemic lupus erythematosus. By their action on platelet and endothelial cells, they predispose to thrombosis of arteries, veins, and venous sinuses as well as cross-reacting with sphingolipids. They tend to be associated focal lesions in neuropsychiatric lupus (McHugh *et al*. 1988).

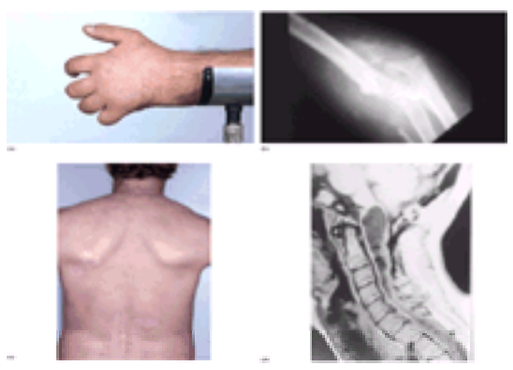


Fig. 1 Syringomyelia.

3. Vegetations on the aortic and mitral valves of some patients carrying anticardiolipin antibodies might be a source of emboli to the brain, causing transient ischaemic attacks, strokes, or dementia (Asherson *et al*. 1987; Asherson and Lubbe 1988).
4. Cerebral vasculitis, causing ischaemia, oedema, and infarction is rare.
5. Antiribosomal P antibodies might be associated with lupus psychosis and severe lupus depression. An initial study (Bonfa *et al*. 1987) found that 90 per cent of those suffering from lupus psychosis carried high levels of these antibodies, and that the levels of these antibodies rose before or during the psychotic episode, suggesting both a diagnostic and predictive value in measuring anti-P antibodies.

However, half of all the lupus patients in this study who carried anti-P antibodies did not suffer from lupus psychosis and six subsequent studies have given conflicting results (Derksen *et al*. 1990; Schneebaum *et al*. 1991; Nojima *et al*. 1992; Van Dam *et al*. 1991; Sato *et al*. 1991; Teh *et al*. 1992), either confirming or refuting the initial findings or suggesting an association of raised levels of these antibodies with severe lupus depression. Until the position is clarified, single or sequential measurement of anti-P antibodies is felt to be of no value (Teh and Isenberg 1992).

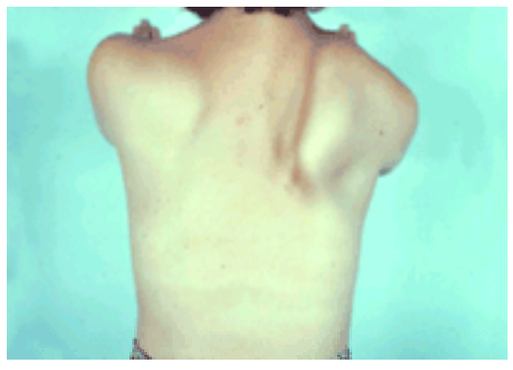


Fig. 2 Neuralgic amyotrophy with winging of the scapula.

Investigations

The diagnosis of neuropsychiatric lupus is based principally on symptoms and signs, but the most useful, generally available investigations (Schrieber *et al*. 1988) are listed below.

1. Formal neuropsychological testing may be the most sensitive and clinically useful measure of brain function in systemic lupus erythematosus; it measures function rather than structural abnormalities, it is non-invasive, objective, and it can be used serially (Ginsburg *et al*. 1992). Shortened versions of these tests are available (Jacobs *et al*. 1977).
2. Magnetic resonance imaging. Three distinct patterns detected by increased intensity of T_2 weighted images are recognized (Aisen *et al*. 1985; Molad *et al*. 1992).
   a. Small, multifocal lesions in the white matter, not detected on computed tomography probably represent microinfarcts or demyelination.
   b. Large lesions, also in the white matter and evident on computed tomography, are infarcts.
   c. Large lesions in the grey matter, not evident on computed tomography, are caused by oedema. The lesions may resolve or evolve to infarction (Sibbitt *et al*. 1989).

   Magnetic resonance imaging also detects intracranial catastrophes such as cerebral venous thrombosis, haematoma, abscess, and tumour. Patients with focal neurological findings or seizures are more likely to have abnormalities detected by magnetic resonance imaging than computed tomography, but it cannot diagnose neuropsychiatric lupus in the absence of clinical features.
3. Electroencephalography. Abnormalities are found in 80 per cent. Focal changes are associated with fits and focal neurological deficits, and diffuse changes with organic brain syndrome. Serial evaluation may monitor disease activity in response to treatment (Ritchlin *et al*. 1992).
4. Anticardiolipin antibody. This may be associated with thromboses or emboli that may warrant anticoagulation (Asherson and Lubbe 1988).
5. Echocardiography. Valvular vegetations as a possible source of emboli may be demonstrated if anticardiolipin antibodies are present.

Analysis of cerebrospinal fluid is important to exclude infectious meningitis. In acute, unstable, catastrophic central nervous system disease, computed tomography scanning will identify haemorrhage, large infarction, abscess, or tumour. Blood levels of anti-DNA antibody, complement, and immune complexes do not correlate with neuropsychiatric lupus. Less frequently available but useful investigations are positron emission tomography, which detects reduced cerebral metabolism (Stoppe *et al*. 1990), and single-photon-emission tomography which detects abnormalities of cerebral blood flow (Nossent *et al*. 1991).

Differential diagnosis

Exclusion of the many complications of systemic lupus erythematosus or its treatment which have central nervous system manifestations (Table 6) is imperative (Futran *et al*. 1986).

Septicaemia
Intracranial sepsis
Accelerated atherosclerosis
Hypertensive encephalopathy
Uraemic encephalopathy
Psychiatric disorders
Emboli from valvular vegetations
Drug side-effects, especially steroid psychosis[1]
Multifocal leucoencephalopathy

[1]See Table 1.

Table 6 Other complications of systemic lupus erythematosus or its treatment with central nervous system manifestations

Treatment

Most patients have mild, self-limiting disease that does not warrant treatment. Corticosteroids and cytotoxic drugs are often given for serious manifestations but their value is unproved. Patients receiving more than 100 mg/day of prednisolone for more than a month are at greater risk of dying from infection than neuropsychiatric lupus (Sargent *et al*. 1975). The evolution of irreversible lesions on serial magnetic resonance imaging could be an indicator for starting treatment, which should be early to prevent progression or new lesions developing (Sibbitt *et al*. 1989).

It is imperative to consider whether other factors are contributing to the clinical picture that would respond to different treatment (Table 7).

Subdural haematoma
Persistent intracerebral haemorrhage causing symptoms or signs
Some patients with chorea from antiphospholipid syndrome have strokes which may be prevented by prophylactic warfarin or aspirin
Multiple emboli associated with valvular vegetations in antiphospholipid syndrome causing transient ischaemic attacks or multi-infarct dementia also respond to the above regimen
Organic depression can respond to antidepressants

Table 7 Features of neuropsychiatric systemic lupus erythematosus requiring specific treatment

Antiphospholipid syndrome (see Chapter 5.7.3)

Anticardiolipin antibody appears to be responsible for the thrombotic and, possibly, embolic manifestations that occur in systemic lupus erythematosus, lupus-like syndromes, and antiphospholipid syndrome. Most patients have a history of a multiplicity of seemingly unrelated disorders, notably livedo reticularis, hypertension, and a history of deep venous thromboses (Asherson *et al*. 1989). Antiplatelet drugs or anticoagulation may be indicated. The principal neurological manifestations are:

1. cerebral thromboses;
2. progressive dementia, either in isolation or associated with repeated cerebrovascular accidents resulting from multiple emboli that arise from cardiac vegetations;
3. transient ischaemic attacks;
4. chorea, sometimes in pregnancy or as a harbinger of a stroke;
5. global amnesia;
6. transverse myelitis (see below).

Other connective tissue diseases

Central nervous system disease occurs in polyarteritis nodosa, Wegener's granulomatosis, mixed connective tissue disease, scleroderma, Sjögren's syndrome (Hietaharju *et al*. 1992), and giant cell arteritis, but far less frequently than in systemic lupus erythematosus or antiphospholipid syndrome. Cranial nerve lesions (see below) may predominate. Vasculitis, anticardiolipin antibodies, and overlap with systemic lupus erythematosus are the principal factors (Shannon and Goetz 1989). Other rheumatic disorders may have central nervous system manifestations (Table 8).

Behçet's syndrome	Recurrent neurological symptoms, principally brainstem (Table 9)
Hyperviscosity syndrome	The rheumatological associations are Sjögren's syndrome and rheumatoid arthritis; diverse symptoms occur often with confusion, ataxia, and visual symptoms; retinopathy may be marked with engorged tortuous veins
Lyme disease	Early: aseptic meningitis
	Late: 15 per cent develop a fluctuating meningoencephalitis with cranial nerve lesions
Sarcoidosis	Cranial nerve lesions, aseptic meningitis, fits, hydrocephalus
Takayasu's arteritis	Carotid or vertebrobasilar ischaemia (see below)
	Retinopathy with arteriovenous anastomoses around the disc
Whipple's disease	Slow progression of apathy, personality change, memory loss, or hypersomnia
	Occasionally, cranial nerve lesions, basal myelopathy, facial, motor, or sensory signs

Table 8 Other rheumatic diseases with central nervous system manifestations

Brain-stem syndromes

The brain-stem comprises the mid-brain, pons, and medulla. It conducts the major motor and sensory tracts and gives rise to all the cranial nerves except the olfactory and optic. The vertebrobasilar arteries supply the brain-stem and the cerebellum, as well as the cerebrum via the circle of Willis. The features of brain-stem disease are listed in Table 9. The principal rheumatological associations are:

Bilateral motor or sensory long-tract signs
Crossed (face/limb) motor or sensory signs
Dissociated pain and temperature loss
Cerebellar signs
Stupor or coma
Dysconjugate eye movements; nystagmus
Horner's syndrome
Unilateral deafness or pharyngeal weakness (these cranial nerves are not affected by single hemispheric disease)
Bulbar weakness

Table 9 Features of brainstem disease

1. vertebrobasilar insufficiency;
2. inflammatory conditions, notably Behçet's disease and connective tissue diseases;
3. platybasia from vertical and atlantoaxial subluxation or Paget's disease.

Vertebrobasilar insufficiency occurs in cervical spondylosis, atlantoaxial subluxation, giant cell arteritis, and Takayasu's arteritis. Its manifestations are myriad but the principal features are diplopia, circumoral numbness, dysarthria, and ataxia. Symptoms are often provoked by moving the head (Brust 1989).

Cranial nerve palsies

These occur in brain-stem disease or as a cranial neuropathy. They may be single or multiple, so only the notable associations are mentioned in (Table 10).

Brainstem disease (see above)
Connective tissue diseases: any nerve may be affected, in any combination, but notably trigeminal neuropathy in scleroderma and mixed connective tissue disease.
Giant-cell arteritis: optic; oculomotor
Lyme disease: multiple
Sarcoidosis: facial; optic chiasm
Whipple's disease: multiple

Table 10 Cranial nerve palsies in rheumatic diseases

Spinal cord disease

Only cervical myelopathy consequent to rheumatoid arthritis or cervical spondylosis presents to the rheumatologist with any frequency. Other causes (Table 11) usually affect the thoracic or lumbar cord. Most result from compression of the cord or its blood supply by vertebral structures, so neck or back pain are usually, but not always, found in association with spastic weakness of the limbs and radicular signs at this level.

Cervical disc prolapse
Cervical spondylosis
Paget's disease
Rheumatoid arthritis
Spinal tuberculosis
Staphylococcal discitis

Table 11 Rheumatic diseases complicated by myelopathy

Magnetic resonance imaging has major advantages over other forms of imaging as it distinguishes the cord from other structures and identifies compression from soft tissues, such as rheumatoid pannus. It can be performed with the neck in flexion, and axial views show irreversible changes such as myelomalacia. Magnetic resonance imaging uses non-ionizing radiation and the contrast medium, gadolinium, is not toxic.

Cervical myelopathy and rheumatoid arthritis

While subluxation of the cervical spine is common in rheumatoid arthritis (25–36 per cent) (Winfield *et al*. 1981) cervical myelopathy is rare and mostly in those with long-standing, crippling disease. Half will die within 5 years, often from unrelated causes (Nakano 1975). Compression is usually at the craniocervical junction from anterior or vertical atlantoaxial subluxation, and less frequently from subaxial subluxation. Compression may not occur with large displacements when the canal is roomy, or it may be caused by pressure from pannus, evident on magnetic resonance imaging (Komusi *et al*. 1985), with only small displacements. In half, the myelopathy progresses unless halted by surgery.

Recognition of this potentially fatal complication is often delayed for months, thereby jeopardizing surgical success, as functional decline is usually so insidious it may be dismissed as deterioration of the arthritis.

A high degree of suspicion is necessary and the principal indicators that should alert the rheumatologist are:

1. recent (less than 18 months) neck pain or occipital neuralgia;
2. paraesthesias or numbness in the limbs or trunk, 'electric-shocks' on neck movement such as lifting the head off the pillow or reading;
3. the patient's own account of diminished motor ability or documented change since the last examination (Crockard *et al*. 1985).

Paraesthesias may be in a glove-and-stocking distribution that could be mistaken for a peripheral neuropathy. Patients with spinothalamic involvement often describe sensations of heat or cold. Posterior column symptoms include sensations of the leg being encased in plaster or walking through quicksand. Spastic weakness may affect arms or legs. Compression of vertebral arteries may cause symptoms of vertebrobasilar insufficiency. Sphincter disturbance, usually urinary retention, occasionally occurs. An acute quadriparesis may follow a fall, whiplash injury, or forcible hyperextension during anaesthesia.

Manubriosternal joint subluxation in rheumatoid arthritis is associated with major deformities of the cervical spine, perhaps from the weight of the head causing chronic flexion of the neck (Khong and Rooney 1982).

Recognition of cord compression and assessment of its progression relies heavily on the history: evaluation of motor signs is difficult with destroyed joints and sensory symptoms are often patchy and transient. Demonstration of obscure 'textbook' signs of atlantoaxial instability may entertain, but do not guide management.

Surgical decompression may halt deterioration of the myelopathy and is recommended even in the relatively symptomless patient (Agarwal *et al*. 1992). Subaxial subluxation may need stabilization at the same time or may develop later (Henderson *et al*. 1993).

Other causes of atlantoaxial subluxation include Jaccoud's arthropathy complicating systemic lupus erythematosus (Babini *et al*. 1990), ankylosing spondylitis (Hunter 1989), juvenile chronic arthritis, Down's syndrome, Klippel-Feil malformation, multiple epiphyseal dysplasia, and rheumatic fever.

Myelopathy caused by cervical spondylosis

Complaints of difficulty climbing stairs, heaviness, stiffness or dragging of the legs, suggesting spastic weakness, in a patient with neck pain, should alert the rheumatologist to this complication. Other symptoms include paraesthesias, numbness or aching of the legs, or weakness of the arms. Sphincter disturbance is rare. A history of recurrent brachial neuralgia is uncommon in these patients.

Neurological signs are always present; 80 per cent have weakness, either of the arms or a para-, hemi-, or quadriparesis, with hyper-reflexia in most and an extensor plantar response in half. Hand muscles may atrophy and mimic syringomyelia. Fasciculation, attributed to interference with the descending blood supply to the lumbar segment, is sometimes present in the legs and can confuse the diagnosis with amyotrophic lateral sclerosis. The most useful clinical sign is sensory loss in the arms in a dermatomal pattern, but the usual picture is patchy sensory loss in the arms, together with a combination of dorsal column and/or spinothalamic tract involvement in the legs, usually without a clear sensory level. Cord damage is chiefly at levels C4–C7 and results from a combination of direct pressure from osteophytes, disc degeneration or ligamentous hypertrophy, trauma during neck movement, and interference with the spinal blood supply. Magnetic resonance imaging and computed tomography myelography are the preferred investigations.

Laminectomy should be recommended with caution; untreated most patients are little changed after many years. Alternative or additional disorders may be present (atheroma of the vertebral and carotid arteries, multiple sclerosis, amyotropic lateral sclerosis, neurosyphilis, spinal tumours). Surgery improves only half (Rowland 1989).

Transverse myelitis in systemic lupus erythematosus or antiphospholipid syndrome

The onset of this rare complication is dramatic with paraesthesias, numbness, and weakness ascending over hours to a thoracic level, and always associated with loss of sphincter function. Magnetic resonance imaging can be useful in the diagnosis and monitoring of these patients (Lavalle *et al*. 1990; Boumpas *et al*. 1990). Even high-dose prednisolone does not affect the poor prognosis, but improvement has been reported following pulse intravenous methylprednisolone followed by repeated pulses of intravenous cyclophosphamide (Barile and Lavalle 1992).

Cauda equina syndrome

This rare complication of ankylosing spondylitis develops after more than two decades, often when the disease is clinically inactive. Arachnoiditis causes adhesions

resulting in an enlargement of the caudal sac and arachnoid cysts causing pressure erosion of the adjacent bone and damage to the cauda equina. The principal presenting features are sensory with paraesthesias, pain or numbness in the perineum, buttocks, or lower limbs. Sometimes there is sphincter disturbance and, later, motor impairment. Computed tomography may be diagnostic. Decompression of the cyst, either at laminectomy or by shunting, has variable success (Tyrrell *et al*. 1994; Mitchell *et al*. 1990).

Peripheral nervous system involvement in rheumatic diseases

Radicular pain

The most common causes of radicular pain in rheumatological practice are cervical spondylosis and lumbar disc disease (Bland 1990), which are described in detail elsewhere. However, several points should be stressed.

1. Arm and leg pains associated with spinal disease are often wrongly attributed to nerve root (radicular) compression. In fact, the pain is commonly referred from muscles, tendons, joint capsules, and ligaments. This pain is perceived away from the site of origin; it radiates widely (from the cervical spine to the hand, arm, chest, and scapula regions; from the lumbar spine to the sacroiliac region, buttocks, and posterior thigh) in a poorly localized, non-dermatomal distribution, more proximally than distal. It is felt as a deep, aching pain and is not associated with neurological signs (Bland 1990).
2. Radicular pain is less common. It arises from compression (or ischaemia) of the dorsal roots or their ganglia from osteophytes, disc herniation, oedema, or fibrous tissue. Radicular pain follows a dermatomal pattern, it is perceived as sharp and lancinating, and is usually exacerbated by movement of the spine, coughing, or sneezing. Paraesthesias in a dermatomal distribution are often associated. Signs of nerve root compression or irritation are usually present (Fig. 4). Dermatomal numbness or myotomal weakness may occur.

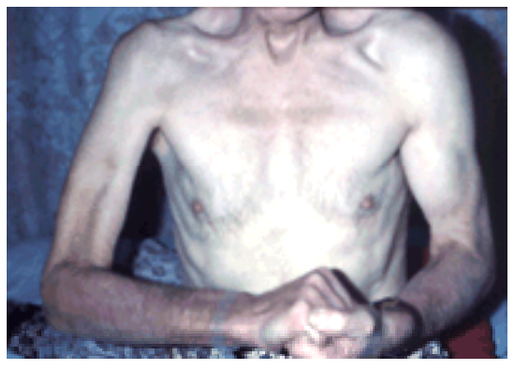


Fig. 4 A right cervical 5th and 6th nerve root lesion with wasted biceps.

3. Soft tissue rheumatism in the arm (shoulder lesions, epicondylitis, carpal tunnel syndrome, de Quervain's tenosynovitis) are strongly associated with cervical spondylosis, possibly unmasked by radicular compression in the neck.
4. Compression of the lumbar nerve root from disc herniation is too frequently diagnosed. The straight leg raising test is specific, but only under the age of 30, and only for L5/S1 lesions; for higher lesions the femoral stretch test is indicated. The straight leg raising test is deemed positive only when it is restricted to less than 40° by the radicular pain. A crossed, straight leg raising test is even more specific; radicular pain is felt on the affected side when the contralateral leg is lifted. Loss, but not depression, of a reflex is a valuable localizing sign.
5. Radicular pain may be mistaken as arising from other structures sharing the same dermatomes; for example C5 as shoulder pain, thoracic nerves as angina or pleurisy, L3 as hip pain, L5 as knee or ankle pain.

Peripheral neuropathies

Several kinds of peripheral neuropathy (Asbury and Gilliatt 1984) complicate rheumatic diseases (Table 12, Table 13) but may be overlooked if the resulting pain or weakness is taken as reflecting joint disease. Neuropathies may also be manifestations of another complication, for example uraemic neuropathy in systemic lupus erythematosus, a side-effect of therapy such as gold salts, or coincidental such as diabetic or alcoholic neuropathy (Fig. 5). Subclinical neuropathy may be unmasked by a second insult, so dual pathology, such as entrapment neuropathy together with polyneuropathy, should be considered. Confirmation is by electrodiagnostic studies. A nerve biopsy occasionally aids diagnosis.

Type	Clinical features
Polyneuropathy (see Table 13)	Distal sensory (glove-and-stocking) and motor impairment; areflexia
Mononeuropathy including entrapment neuropathies	Loss of single peripheral nerve function
Mononeuritis multiplex	Involvement of two or more nerves; usually indicative of a systemic disease, notably vasculitis
Autonomic neuropathy	Postural hypotension, impotence, flushing, sweating, loss of sinus arrhythmia, loss of reactive tachycardia following the Valsalva manoeuvre, abnormal gut motility

Table 12 Classification of peripheral neuropathies

Sensory	Glove-and-stocking distribution of pain, paraesthesias or numbness; typically starts in the feet and only affects the hands when it has risen to the level of the knees
	The pain is spontaneous or touch-induced; often the threshold to pain or touch is delayed but has an exaggerated response
	Paraesthesias are usually perceived as tingling or as light bands
	Sensory loss may be all modalities, large fibres only (position and vibration), or small fibres only (pain and temperature)
Motor	Distal weakness, with depressed reflexes, wasting, and occasionally fasciculation

Table 13 Features of polyneuropathy (the deficit may be sensory, motor, or mixed)

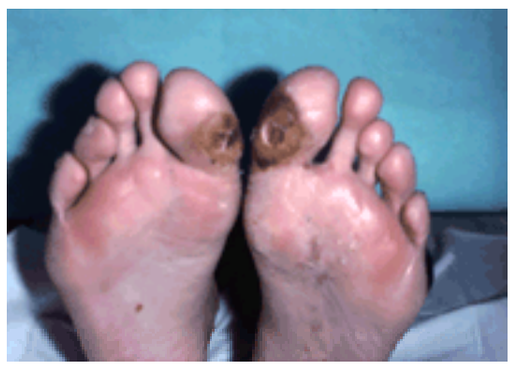


Fig. 5 Painless ulcers in hereditary sensorineuropathy.

Entrapment neuropathies

In patients with joint disease, peripheral nerves may be damaged by pressure from deformed joints or periarticular structures, by callipers, or from the operating table. The onset is usually insidious, with pain, paraesthesias, or weakness. The differential diagnosis includes other causes of regional pain, such as brachial neuralgia, mononeuritis multiplex complicating vasculitis, and reflex sympathetic dystrophy. Confirmation is by electrodiagnosis.

Conservative treatment includes removal of the offending callipers, rest, splints, and steroid injections. Surgical decompression or transposition may be necessary for persistent pain or if motor signs develop.

Only those entrapment neuropathies relevant to the rheumatologist will be considered here (Dawson *et al*. 1983).

Median nerve

Pressure on the median nerve in the carpal tunnel causes tingling, pain, or numbness in the radial three and a half digits, sometimes radiating up to, or even above, the elbow. It is commonly present at night, which can be relieved by exercising the hand or hanging it out of bed. Symptoms sometimes persist through the day so the patient cannot manipulate small objects and drops things. Signs may be absent. When present, sensory loss is confined to the median nerve distribution, with blunting of touch and pinprick in the radial three and a half digits. Symptoms may be provoked by Tinel's sign (tapping the nerve at the volar surface of the wrist) or Phalen's sign (sustained flexion at the wrist). Thenar weakness or wasting is a late feature.

Common causes include obesity, fluid retention from pregnancy, arthritis, Colles, or scaphoid fractures. Acromegaly, hypothyroidism, amyloidosis, or chronic haemodialysis are rare.

Anterior interosseous nerve

Pressure at the elbow of this branch of the median nerve results in weakness of the deep flexors of the thumb, index, and middle fingers, with pain in the forearm but without sensory signs. The characteristic sign is loss of normal thumb-index pincer movement.

Ulnar nerve

Pressure behind the medial epicondyle of the elbow produces numbness in the ulnar one and a half digits and the corresponding side of the hand, but not proximal to the wrist, which differentiates it from a C8 radiculopathy. The intrinsic muscles of the hand and the deep flexors of the ring and little fingers may be weak (Fig. 6). Occasionally the ulnar nerve may be trapped in Guyon's canal at the wrist. Motor loss is then confined to the intrinsic muscles of the hand, and sensory loss to the palmar aspect of the ulnar side of the hand.

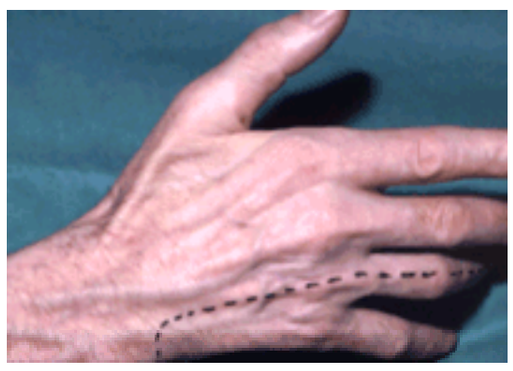


Fig. 6 Slight flexion of the 4th and 5th fingers in an ulnar nerve lesion at the elbow. (The sensory loss is indicated by the dotted line.)

Posterior interosseous nerve

Entrapment at the elbow causes weakness of extension of all fingers and may be mistaken for tendon ruptures. Sensory symptoms are absent.

Thoracic outlet compression syndrome

The lower elements of the brachial plexus, derived from the C8 and T1 roots, are compressed in isolation or together with the brachial artery and vein at the thoracic outlet. Various structures in this region have been implicated—cervical rib, scalenus anterior, pectoralis minor, and a costoclavicular band—each with its own name and operation. In reality, the evidence is usually lacking and the syndrome is far less common than is diagnosed (Bland 1990).

Presentation is with pain, paraesthesias, and numbness in the ulnar side of the hand and forearm associated with vasomotor changes. The various provocative tests described are now discounted. Confirmation requires arteriography and electrodiagnostic studies.

Sciatic nerve

This may be trapped by the piriformus muscle causing pain in the lateral thigh, foot drop, and an absent ankle jerk. Compression by a Baker's cyst gives only the motor signs.

Lateral cutaneous nerve of thigh

Tight clothes or obesity can compress this nerve as it passes under the inguinal ligament causing pain, paraesthesias, and numbness in the outer side of the thigh (meralgia paraesthetica) which may be mistaken for hip pain, an L3 root lesion, or trochanteric bursitis.

Common peroneal nerve

Where it winds around the head of the fibula, this nerve is vulnerable to pressure from the joint, callipers, or the operating table. Foot drop and weakness of eversion are associated with sensory loss on the dorsum of the foot.

Posterior tibial nerve

Compression occurs in the tarsal tunnel behind the medial malleolus causing pain and tingling in the sole on standing or at night.

Autonomic neuropathy

A mild and possibly clinically irrelevant autonomic neuropathy has been demonstrated in rheumatoid arthritis (Toussirot *et al*. 1993), systemic lupus erythematosus (Liote and Osterland 1994), CREST syndrome (Hermosillo *et al*. 1994), and systemic sclerosis (Dessein *et al*. 1992).

Peripheral neuropathies in particular diseases

Rheumatoid arthritis

Carpal tunnel syndrome may be the presenting feature of rheumatoid arthritis, but other neuropathies usually develop in established disease (Good *et al*. 1965). Entrapment neuropathies are the most frequent and may be multiple. An indolent pure sensory neuropathy is also common. The mechanism is unknown. The sensory loss in atlantoaxial subluxation may mimic a polyneuropathy in the arms. A rare but serious problem is mononeuritis multiplex, usually where there is also clinical and serological evidence of vasculitis. The nerves of the upper arm and thigh are vulnerable because of their sparse supply of nutrient arteries. The onset is usually sudden, with dermatomal pain and paraesthesias followed within hours by weakness. Additional nerves are involved asymmetrically over days or weeks.

Polyarteritis nodosa

Mononeuritis multiplex (see above) occurs in 50 to 70 per cent of patients, sometimes as a presenting feature, but typically in the systemically ill patient with multisystem disease (Conn and Dyck 1975). The kidneys are usually involved. Lesions may summate to resemble a polyneuropathy. Occasionally a peripheral sensory neuropathy slowly develops (see Table 13).

Sjögren's syndrome

Mild or subclinical, motor, or sensory neuropathy occurs in one-fifth of patients. Mononeuritis multiplex is rare and usually associated with frank vasculitis; anti-Ro is over represented in this group (Binder *et al*. 1988).

Scleroderma

A mild sensorineuropathy has been described (Schady *et al*. 1991). A parasympathetic autonomic neuropathy, evident on cardiovascular tests, is reported to be common and arguably related to the Raynaud's syndrome, oesophageal dysmotility, and bowel disturbance (Hermosillo *et al*. 1994). Carpal tunnel syndrome is the only other peripheral neuropathy reported with any frequency; fibrosis following surgical decompression may worsen the symptoms. Subacute combined degeneration of the cord is reported in vitamin B_{12} deficiency consequent to sclerodermatous bowel involvement (Lee *et al*. 1984).

Systemic lupus erythematosus

Mild peripheral neuropathies, mostly sensory, occur particularly during active disease. Mononeuritis multiplex is uncommon.

L-*tryptoph an toxiticy*

Before this drug was withdrawn, several cases were described of a vasculitis associated with eosinophilic myositis and neuropathy.

Other rheumatological disorders

Mixed connective tissue disease, Churg-Strauss syndrome, amyloidosis, sarcoidosis, Wegener's granulomatosis, cryoglobulinaemia, Lyme disease, and giant cell arteritis may all be complicated by peripheral neuropathies.

Chapter References

Adelman, D.C., Saltiel, E., and Klinenberg, J.R. (1986). The neuropsychiatric manifestations of systemic lupus: an overview. *Seminars in Arthritis and Rheumatism*, **15**, 185–99.

Agarwal, A.K., Peppelman, W.C., Kraus, D.R., Pollock, B.H., Stolzer, B.L., Eisenbeis, C.H. and Donaldson, W.F. (1992). Recurrence of cervical spine instability in rheumatoid arthritis following previous fusion: can disease progress be prevented by early surgery? *Journal of Rheumatology*, **19**, 1364–70.

Aisen, A.M., Grabrielson, T.O., and McCune, W.J. (1985). MR imaging of systemic lupus erythematosus involving the brain. *American Journal of Neuroradiology*, **6**, 197–201.

Asbury, A.K. and Gilliatt, R.W. (1984). *Peripheral nerve disorders: a practical approach*. Butterworths, London.

Asherson, R.A. and Lubbe, W.F. (1988). Cerebral and valve lesions in systemic lupus erythematosus: association with antiphospholipid antibodies (editorial). *Journal of Rheumatology*, **15**, 539–43.

Asherson, R.A., Derkston, R.W.H.M., Harris, E.N., and Hughes, G.R.V. (1987). Chorea in systemic lupus erythematosus and lupus-like disease: association with antiphospholipid antibodies. *Seminars in Arthritis and Rheumatism*, **16**, 253–9.

Asherson, R.A., Khamashta, M.A., and Ordis-Ros, J. (1989). The 'primary' antiphospholipid syndrome: major clinical and serological features. *Medicine*, **68**, 366–77.

Babini, S.M., Maldonado Cocco, J.A., Babini, J.C., de la Sota, M., Arturi, A., and Marcos, J.C. (1990). Atlantoaxial subluxation in systemic lupus erythematosus: further evidence of tendinous alterations. *Journal of Rheumatology*, **17**, 173–7.

Barile, L. and Lavalle, C. (1992). Transverse myelitis in systemic lupus erythematosus—the effect of I.V. pulse methylprednisolone and cyclophosphamide. *Journal of Rheumatology*, **19**, 370–2.

Binder, A., Snaith, M.C., and Isenberg, D. (1988). Sjögren's syndrome: a study of its neurological complications. *British Journal of Rheumatology*, **27**, 275–80.

Blain, P.G. and Stewart-Wynne, E. (1986). Neurological disorders. In *Textbook of adverse drug reactions*, (ed. D.M. Davies), pp. 494–509. Oxford University Press, New York.

Bland, J.H. (1990). Cervical and thoracic pain including thoracic outlet syndrome and brachial neuritis. *Current Opinion in Rheumatology*, **2**, 242–52.

Bonfa, E., Golombek, S.J., Kaufman, L.D., Skelly, S., Weissbach, H., Brot, N., and Elkon, K.B. (1987). Association between lupus psychosis and anti-ribosomal P protein antibodies. *New England Journal of Medicine*, **317**, 265–71.

Boumpas, D.T., Patronas, M.J., Dalakas, M.C., Hakim, C.A., Klippel, J.H., and Balow, J.E. (1990). Acute transverse myelitis in systemic lupus erythematosus: magnetic resonance imaging a review of the literature. *Journal of Rheumatology*, **17**, 89–92.

Brust, J.C.M. (1989). Cerebral infarction. In *Merritt's textbook of neurology*, (ed. L.P. Rowland), pp. 206–14. Lea and Febiger, Philadelphia.

Conn, D.C. and Dyck, P.J. (1975). Angiopathic neuropathy in connective tissue diseases. In *Peripheral neuropathy* (ed. P.J. Dyck, P.K. Thomas, and E.H. Leubert), pp. 1149–65. W.B. Saunders,

Philadelphia.

Crockard, H.A., Essigman, W.K., Stevens, J.M., Pozo, J.L., Ransford, A.O., and Kendall, B.E. (1985). Surgical treatments of cervical cord compression in rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **44**, 809–16.

Dawson, D.M., Hallett, M., and Millender, L.H. (1983). *Entrapment neuropathies*. Little Brown, Boston.

Denburg, J.A. and Behmann, S.A. (1994). Lymphocytes and neural antigens in neuropsychiatric lupus: presence of an elutable, immunoprecipitable lymphocyte/neuronal 52 kDa reactivity. *Annals of the Rheumatic Diseases*, **53**, 304–8.

Derksen, R.H.W.M., van Dam, A.P., Gmelig Meyling, F.H.J., Bijlsma, J.W.J., and Smeenk, R.J.T. (1990). A prospective study on anti-ribosomal P proteins in two cases of familial lupus and recurrent psychosis. *Annals of the Rheumatic Diseases*, **49**, 779–82.

Dessein, P.H., Joffe, B.I., Metz, R.M., Millar, D.L., Lawson, M., and Stanwix, A.E. (1992). Autonomic dysfunction in systemic sclerosis: sympathetic activity and instability. *American Journal of Medicine*, **93**, 143–50.

Feinglass, E.J., Arnett, F.C., Porsch, L.A., Zizic, T.M., and Stevens, M.B. (1976). Neuropsychiatric manifestations of systemic lupus erythematosus: diagnosis, clinical spectrum and relationship to other features of the disease. *Medicine*, **55**, 323–6.

Futran, J., Shore, A., Murray, B., Urowitz, M.B., and Grossman, H. (1986). Subdural haematoma in systemic lupus erythematosus. Report and review of the literature. *Journal of Rheumatology*, **14**, 387–1.

Gibson, T. and Myers, A.R. (1976). Nervous system involvement in systemic lupus erythematosus. *Annals of the Rheumatic Diseases*, **35**, 398–406.

Ginsburg, K.S., Wright, E.A., Larson, M.C., Fossel, A.H., Albert, M., Schur, P.H., and Liang, M.H. (1992). A controlled study of the prevalence of cognitive dysfunction in randomly selected patients with systemic lupus erythematosus. *Arthritis and Rheumatism*, **35**, 776–82.

Good, A.F., Christopher, R.P., Koepke, G.H., Bender, L.F., and Tarter, M. (1965). Peripheral neuropathies associated with rheumatoid arthritis: a clinical and electrodiagnostic study of 70 cases in rheumatoid arthritis. *Annals of Internal Medicine*, **63**, 87.

Guarantors of Brain (1988). *Aids to the investigation of peripheral nerve injuries*. Ballière Tindall, London.

Hanly, J.G., Walsh, N.M.G., and Sangalang, V. (1992). Brain pathology in systemic lupus erythematosus. *Journal of Rheumatology*, **19**, 732–41.

Hanly, J.G., Fisk, J.D., Sherwood, G., and Eastwood, B. (1994). Clinical course of cognitive dysfunction in systemic lupus erythematosus. *Journal of Rheumatology*, **21**, 1825–31.

Henderson, F.C., Geddes, J.F., and Crockard, H.A. (1993). Neuropathology of the brain stem and spinal cord in end-stage rheumatoid arthritis: implications for treatment. *Annals of the Rheumatic Diseases*, **52**, 629–37.

Hermosillo, A.G., Ortiz, R., Dayaque, J., Casanova, J.M., and Martinez-Lavin, M. (1994). Autonomic dysfunction in diffuse scleroderma vs CREST: an assessment of computerised heart rate variabilities. *Journal of Rheumatology*, **21**, 1849–54.

Hietaharju, A., Korpela, M., Ilonen, J., and Frey, H. (1992). Nervous system disease, immunological features and HLA phenotype in Sjögren's syndrome. *Annals of the Rheumatic Diseases*, **51**, 506–9.

Hughes, G.R.V. (1985). The anticardiolipin syndrome. *Clinical and Experimental Rheumatology*, **3**, 285–6.

Hunter, T. (1989). The spinal complications of ankylosing spondylitis. *Seminars in Arthritis and Rheumatism*, **19**, 172–82.

Jacobs, J.W., Bernard, M.R., Delgado, A., and Strain, J.J. (1977). Screening for organic mental syndrome in the mentally ill. *Annals of Internal Medicine*, **86**, 40–6.

Khong, T.K. and Rooney, P.J. (1982). Manubrio-sternal joint subluxation in rheumatoid arthritis. *Journal of Rheumatology*, **9**, 712–15.

Komusi, T., Munro, T., and Harth, M. (1985). Radiologic review: the rheumatoid cervical spine. *Seminars in Arthritis and Rheumatism*, **14**, 187–95.

Lavalle, C., Pizarro, S., Drenkard, L., Sanchez-Guerrero, J., and Alarcon-Segovia, D. (1990). Transverse myelitis: a manifestation of systemic lupus erythematosus strongly associated with antiphospholipid antibodies. *Journal of Rheumatology*, **17**, 34–7.

Lee, P., Bruni, J., and Sukenik, S. (1984). Neurological manifestations of systemic sclerosis (scleroderma). *Journal of Rheumatology*, **11**, 480–3.

Liote, F. and Osterland, C.K. (1994). Autonomic neuropathy in systemic lupus erythematosus: Cardiovascular autonomic fuction assessment. *Annals of the Rheumatic Diseases*, **53**, 671–4.

McHugh, N.J., Maymo, J., Skinner, R.P., James, I., and Maddison, P.J. (1988). Anticardiolipin antibodies, livedo reticularis and major cerebrovascular and renal disease in systemic lupus erythematosus. *Annals of the Rheumatic Diseases*, **47**, 110–5.

Marsden, C.D. (1986). Hysteria — a neurologist's view. *Psychological Medicine*, **16**, 277–88.

Mastaglia, F.L. (1989). Iatrogenic (drug-induced) disorders of the nervous system. In *Neurology and general medicine* (ed. M.J. Aminoff), pp. 505–32. Churchill Livingstone, New York.

Mitchell, M.J., Sartoris, D.J., Moody, D., and Resnick, D. (1990). Cauda equina syndrome complicating ankylosing spondylitis. *Radiology*, **175**, 521–5.

Molad, Y., Siddi, Y., Gornish, M., Lerner, M., Pinkhas, J., and Weinberger, A. (1992). Lupus anticoagulant: correlation with magnetic resonance imaging of brain lesions. *Journal of Rheumatology*, **19**, 556–61.

Nakano, K.K. (1975). Neurological complications of rheumatoid arthritis. *Orthopaedic Clinics of North America*, **6**, 861–81.

Nojima, Y., Minota, S., Yamada, A., Takaku, F., Aotsuka, S., and Yokohari, R. (1992). Correlation of antibodies of ribosomal P protein with psychosis in patients with systemic lupus erythematosus. *Annals of the Rheumatic Diseases*, **51**, 1053–5.

Nossent, J.C., Hovestadt, A., Schonfeld, D.H.W., and Swaak, A.J.E. (1991). Single-photon-emission computed tomography of the brain in the evaluation of cerebral lupus. *Arthritis and Rheumatism*, **34**, 1397–403.

Ritchlin, C.T., Chabot, R.J., Alper, K., Buyon, J., Belmont, H.M., Roueby, R., and Abramson, S.B. (1992). Quantatative elecroencephalograpy. *Arthritis and Rheumatism*, **35**, 1330–42.

Robbins, M.L., Kornguth, S.E., Bell, C.L., Kalinke, D.E., England, D., Turski, P., and Graziazo, F.M. (1988). Anti-neurofilament antibody evaluation in neuropsychiatric systemic lupus erythematosus. *Arthritis and Rheumatism*, **31**, 623–31.

Rowland, L.P. (1989). Myelopathy caused by cervical spondylosis. In *Merritt's textbook of neurology*, (ed. L.P. Rowland), pp. 409–11. Lea and Febiger, Philadelphia.

Sargent, J.S., Lockshin, M.D., Klempner, M.S., and Lipsky, B.A. (1975). Central nervous system disease in systemic lupus erythematosus. *American Journal of Medicine*, **58**, 644–54.

Sato, T., Uchiumi, T., Azawa, T., Kikuchi, M., Nakano, M., Kominami, R., and Arakawa, M. (1991). Auto-antibodies against ribosomal proteins found with high frequency in patients with systemic lupus erythematosus with active disease. *Journal of Rheumatology*, **18**, 1681–4.

Schady, W., Sheard, A., Hassell, A., Holt, L., Jason, M.I.V., and Klimiuk, P. (1991). Peripheral nerve dysfunction in scleroderma. *Quarterly Journal of Medicine*, **80**, 661–75.

Schneebaum, A.B., Singleton, J.D., Sterling, G., West, S.G., Blodgett, J.K., Allen, L.G., Cheronis, J.C., and Kotzin, B.L. (1991). Association of psychiatric manifestations with antibodies to ribosomal P proteins in systemic lupus erythematosus. *Americal Journal of Medicine*, **90**, 54–62.

Schrieber, L., McCredie, M., Tugwell, P., and Brook, M. (1988). An evaluation of the role of laboratory investigations in establishing a diagnosis of CNS lupus. *British Journal of Rheumatology*, **27**, 138–42.

Shannon, K.M. and Goetz, G.G. (1989). Connective tissue diseases and the nervous system. In *Neurology and general medicine* (ed. M.J. Aminoff), pp. 389–413. Churchill Livingstone, New York.

Sibbitt, W.L., Sibbitt, R.R., Griffey, R.H., Eckel, C., and Bankhurst, A.D. (1989). Magnetic resonance and CT imaging in the evaluation of acute neuropsychiatric disease in systemic lupus erythematosus. *Annals of the Rheumatic Diseases*, **48**, 1014–22.

Singer, J. and Denburg, J.A. (1990). Diagnostic criteria for neuropsychiatric systemic lupus erythematosus: the results of a consensus meeting. *Journal of Rheumatology*, **17**, 1397–402.

Stoppe, G., Wildhagen, K., Seidel, J.W., Meyer, G.J., Schober, O., Heintz, P., Kunkel, H., Deicher, H., and Hundeshagen, H. (1990). Positron emission tomography in neuropsychiatric lupus erythematosus. *Neurology*, **40**, 304–8.

Teh, L.S., Bedwell, A.E., Isenberg, D.A., Gordon, C., Emery, P., Charles, P.J., Harper, M., Amos, N., and Williams, B.D. (1992). Antibodies to protein P in systemic lupus erythematosus. *Annals of the Rheumatic Diseases*, **51**, 489–94.

Teh, L.S. and Isenberg, D.A. (1994). Anti ribosomal P protein antibodies in systemic lupus erythematosus. *Arthritis and Rheumatism*, **37**, 307–15.

Toussirot, E., Serratrice, G., and Valentin, P. (1993). Autonomic nervous system involvement in rheumatoid arthritis. 50 cases. *Journal of Rheumatology*, **20**, 1508–14.

Tyrrell, P.N.M., Davies, A.M., and Evans, N. (1994). Neurological disturbances in ankylosing spondylitis. *Annals of the Rheumatic Diseases*, **53**, 714–17.

Van Dam A., Nossent, H., de Jong, J., Meilof, J., ter Borg, E-J., Swaak, T., and Smeenk, R. (1991). Diagnostic value of antibodies against ribosomal phosphoproteins: a cross-section and longitudinal study. *Journal of Rheumatology*, **18**, 1026–34.

Winfield, J., Cook, D., Brook, A.S., and Corbett, M. (1981). A progressive study of the radiological changes in the cervical spine in rheumatoid disease. *Annals of the Rheumatic Diseases*, **40**, 109–14.

1.3.4 The cardiovascular system

Laura F. Wexler and Michael E. Luggen

Chest pain

Cardiac disorders that present as chest pain are almost invariably serious or potentially serious, requiring prompt diagnosis, close monitoring, and rapid initiation of treatment. Thus, it is critical that any patient with rheumatic disease who complains of chest pain be carefully and thoroughly evaluated. Chest pain of cardiac origin may be difficult to distinguish from the pain associated with costochondritis, arthritis, bursitis, or any other musculoskeletal disorder involving the chest and shoulders. More commonly, gastrointestinal symptoms exacerbated by medications used to treat rheumatic disorders may mimic cardiac pain.

Angina pectoris

Angina refers to the chest pain that results from imbalance of myocardial oxygen supply and demand; it occurs most commonly in patients with myocardial ischaemia secondary to coronary arterial disease but is also experienced by patients with acute myocarditis, chronic cardiomyopathy, severe systemic hypertension, or certain types of valve disease. The description of angina published by William Heberden (Heberden 1772) remains worthy of attention.

'... a disorder of the breast marked with strong and peculiar symptoms.

The seat of it, and sense of strangling, and anxiety with which it is attended, may make it not improperly called angina pectoris. They who are afflicted with it, are seized while they are walking, (more especially if it be up hill, and soon after eating) with a painful and most disagreeable sensation in the breast, which seems as if it would extinguish life, if it were to increase or to continue, but the moment they stand still, all this uneasiness vanishes.'

Patients frequently describe anginal pain as a band-like sensation around the chest or as pressure, tightness, squeezing, burning, or aching rather than as pain; the anxiety and sense of impending doom it elicits may seem out of proportion to the actual severity of the discomfort described. It is usually substernal but may radiate to either side of the chest or be isolated to the shoulders, arms (left, right or both), lower (but not upper) jaw, or epigastrium. It may thus be misinterpreted by the patient as bursitis, arthritis, toothache, or indigestion. An important feature distinguishing angina from rheumatic or gastrointestinal pain is that angina is almost never a sharp, localized pain that the patient can point to with one finger and it is not associated with tenderness over the area of pain. Unlike pericarditis (see below), angina is neither pleuritic nor positional. Although angina is often precipitated by exertion or stress (physical or emotional) and relieved by rest, it may occur solely or predominantly at rest. An episode of angina usually lasts several minutes and may recur many times in a day. However, it does not persist unchanged for hours or days. Angina is very often associated with dyspnoea; in some patients, dyspnoea may be the sole or predominant symptom associated with myocardial ischaemia. Although the quality and location of angina varies greatly from patient to patient, chronic, stable angina usually has a repetitive pattern in a particular patient.

Box 1 Cardiovascular complications of rheumatic diseases

Rheumatic disease	Coronary disease	Pericarditis	Myocardial disease	Conduction system disease	Valvular disease
Amyloidosis	+	–	++	++	–
Ankylosing spondylitis	–	–	+	++	++
Behçet's disease	–	+	–	–	+
Ehlers-Danlos syndrome	–	–	–	+	++
Hypersensitivity vasculitis	–	+	–	–	–
Juvenile rheumatoid arthritis	–	+	+	–	+
Kawasaki's disease	++	+	++	–	+
Lyme disease	–	+	+	++	–
Mixed connective tissue disease	–	++	–	–	++
Polyarteritis nodosa	++	+	–	–	–
Polymyositis dermatomyositis	+	+	++	++	–
Reiter's syndrome	–	+	+	+	+
Rheumatic fever	–	++	++	+++	+
Rheumatoid arthritis	+	+	+	+	+
Sarcoidosis	–	–	++	++	–
Systemic lupus erythematosus	+	++	+	+	+
Systemic sclerosis	+	++	++	++	–
Takayasu's disease	+	–	–	–	–
Temporal arteritis	–	–	–	–	+
Wegener's granulomatosis	+	+	–	–	–

+++ Occurs in >50% patients; ++ Occurs in 5–49% patients; + Occurs in <5% patients; – Occurs rarely if ever

Sublingual nitroglycerine almost invariably relieves angina within seconds or a few minutes and the response to nitroglycerine is sometimes considered a diagnostic test to help differentiate cardiac from non-cardiac pain. It should be noted that nitroglycerine is an effective smooth-muscle vasodilator and will also relieve epigastric pain associated with oesophageal spasm, which often mimics angina.

New-onset angina or angina that abruptly increases in severity or frequency, or that occurs at rest for the first time in a patient with previously stable exertional angina, is called unstable angina and is of particular significance because it is associated with a high risk of progression to acute myocardial infarction.

Acute myocardial infarction

Severe, unremitting anginal pain that lasts 30 min or more should suggest the possibility of an acute myocardial infarction, especially if the pain is associated with nausea, vomiting, diaphoresis, dyspnoea, light-headedness, or syncope. Contrary to common wisdom, an acute myocardial infarction is usually not precipitated by sudden exertion or stress (as musculoskeletal pain may be). Most often it occurs at rest, without warning. The mechanism of acute myocardial infarction is usually acute coronary thrombosis induced by endothelial injury or spasm in a coronary artery previously damaged by atherosclerosis or other inflammatory process. Rarely, spasm or embolus in a previously normal coronary artery can lead to acute infarction.

Pericarditis

The pain associated with pericarditis may closely mimic angina in quality and radiation pattern but it is not brought on by exertion or emotional stress and is not relieved by rest. It is usually located in the mid-anterior chest and is often pleuritic (exacerbated by deep inspiration or cough), positional, or aggravated by swallowing. Non-steroidal anti-inflammatory drugs frequently alleviate pericardial pain but are ineffective for anginal pain.

Acute aortic dissection

The pain associated with acute aortic dissection may mimic an acute myocardial infarction, but is often distinguished by its extreme severity and radiation to the back. The pain is usually so catastrophic as to bring patients to the emergency room promptly; in other words, aortic dissection would not present as a mild to moderate, repetitive anginal pain syndrome.

Dyspnoea

Dyspnoea is the cardinal symptom of congestive heart failure. Congestive heart failure associated with the rheumatic disorders can be the result of myocardial damage (from either primary involvement of cardiac muscle or from ischaemic damage resulting from coronary arterial disease) or valvular heart disease. However, dyspnoea may also represent an anginal equivalent, i.e., dyspnoea rather than chest pain may be the sole or predominant symptom of angina (myocardial ischaemia) in patients with obstructive coronary arterial disease.

Dyspnoea of cardiac origin may be difficult to distinguish from dyspnoea related to obstructive or restrictive lung disease but certain clues can suggest its origin. Paroxysmal nocturnal dyspnoea (nocturnal episodes of a sudden sense of suffocation that is relieved by sitting up) is more common in congestive heart failure than in pulmonary disease, as is orthopnoea. Dyspnoea on exertion is a feature of both cardiac and intrinsic pulmonary disease but cardiac dyspnoea is rarely precipitated by hot, humid weather or exposure to specific irritants, is not relieved by productive coughing or by bronchodilators, and, with the exception of severe pulmonary oedema, does not cause wheezing.

Dizzy spells and syncope

Dizzy spells are exceedingly common and may or may not indicate a significant cardiac arrhythmia, the most common cause being enhanced vagal tone leading to hypotension. Syncope or near-syncope may be a sign of cardiovascular involvement in patients with rheumatic disorders that affect the conduction system disease and cause heart block (see Table 4). Syncope may also occur in patients with coronary arteritis in whom myocardial ischaemia induces ventricular arrhythmias, sinus bradycardia, or heart block. Autonomic dysfunction, resulting in symptomatic postural hypotension (as occurs with amyloid), is another cause.

Rheumatic disease	Frequency[a]
Amyloidosis	++
Ankylosing spondylitis	++
Ehlers–Danlos syndrome	+
Lyme disease	++
Polymyositis/dermatomyositis	++
Reiter's syndrome	+
Rheumatoid arthritis	+
Rheumatic fever	+++
Sarcoidosis	++
Systemic lupus erythematosus	+
Systemic sclerosis	++

[a]Approximate frequency of clinically significant atrioventricular block in patients who have rheumatic disease. Atrioventricular block occurs rarely if ever in rheumatic diseases not listed.
+++ Occurs in >50% of patients; ++ occurs in 5–49% of patients; + occurs in <5% of patients.

Table 4 Rheumatic diseases causing conduction system disease

It is usually possible to distinguish cardiogenic syncope from transient neurological events, such as seizures or transient ischaemic attacks, on the basis of a careful history. Tonic–clonic movements and loss of sphincter tone, which characterize seizures, rarely occur with syncope. Recovery from a syncopal episode usually takes place within seconds or minutes and the patient returns promptly to an awake and alert state without the prolonged drowsiness and confusion characteristic of the postictal state. In contrast to cerebral ischaemic events, focal neurological deficits do not usually occur with syncope of cardiac origin.

Palpitations may indicate the presence of a significant arrhythmia but more often reflect anxiety and increased cardiac awareness, as there is a very poor correlation between the perception of palpitations and the presence of significant arrhythmias.

Specific cardiovascular syndromes associated with the rheumatic disorders

Coronary arterial disease

Several distinct patterns of coronary arterial disease occur in association with the rheumatic disorders (Fauci *et al*. 1978; Lie 1987) (Table 1): focal obstructive disease of the large extramural coronary arteries or ostial stenosis of the right or left main coronary artery in patients with giant cell or other forms of aortitis may mimic atherosclerotic coronary disease. Diffuse vasculitis and/or vasospastic disease of small intramural arteries is often asymptomatic but may lead to myocardial damage. Coronary artery disease may occasionally be the initial or sole manifestation of a rheumatic disorder; more commonly, clinical evidence of coronary artery disease will occur in the setting of diffuse systemic vasculitis or other evidence of the underlying rheumatic disorder.

	Size of artery[a]	
Rheumatic disease	Large	Small
Amyloidosis	–	+
Kawasaki's disease	++	–
Polyarteritis nodosa	++	+
Polydermatomyositis	+	+
Rheumatoid arthritis	+	+
Systemic lupus erythematosus	+	+
Systemic sclerosis	–	+
Takayasu's disease	+	–
Temporal arteritis	+	–
Wegener's granulomatosis	–	+

[a]Approximate frequency of clinically significant coronary arterial disease in patients with the rheumatic disorder. Coronary arterial disease of significance occurs rarely if ever in those rheumatic disorders not listed.
+++ Occurs in >50% of patients; ++ occurs in 5–49% of patients; + occurs in <5% of patients; – occurs rarely if at all.

Table 1 Rheumatic diseases involving the coronary arteries

Acute vasculitis involving the large epicardial coronary arteries is uncommon; it has been most often reported in association with systemic lupus erythematosus (Doherty and Siegel 1985; Mandel 1987), polyarteritis nodosa (Przybojewski 1981; Cassling *et al*. 1985), and rheumatoid arthritis (Pizzarello and Goldberg 1985; Van Albada-Kuipers *et al*. 1986). Clinically, it may result in acute obstruction of a coronary artery and exactly mimic obstructive, atherosclerotic, coronary arterial disease, presenting as unstable angina, acute myocardial infarction, congestive heart failure, or sudden death from an arrhythmia.

In addition to causing a true vasculitis, systemic lupus erythematosus has been associated with the development of premature atherosclerosis of coronary arteries. Doherty and Siegel (1985) reviewed the literature on 33 patients less than 35 years old who had systemic lupus erythematosus with evidence of coronary arterial

disease (angina, acute myocardial infarction, congestive heart failure, or sudden death); atherosclerosis was reported in twelve, coronary vasculitis in seven, and occlusive disease without clear distinction in six. The remainder had either embolic or thrombotic lesions or small-vessel hyaline arteriolar disease presenting as myocardial infarction.

Coronary atherosclerosis in young women with systemic lupus was rarely reported in the era before steroids. In 1975, Bulkley and Roberts (1975) reported finding coronary atherosclerosis at autopsy in 8 of 36 young patients with systemic lupus erythematosus and suggested that corticosteroids might be an aetiological factor in accelerated atherosclerosis. However, a subsequent review (Haider and Roberts 1981) of 22 patients with systemic lupus (all of whom had been treated with prednisone) submitted to autopsy revealed that only a subset (10 of 22) had truly accelerated coronary atherosclerosis, i.e., disease that was more severe than that found in age- and sex-matched controls. The subgroup with accelerated coronary disease had more severe hypertension and higher cholesterol concentrations, which are potent risk factors for coronary atherosclerosis and which are exacerbated by corticosteroids. However, they also had more severe lupus-associated pericardial and mitral valvular disease, suggesting that immunological factors related to the systemic lupus itself might predispose susceptible patients to accelerated atherosclerosis.

Coronary vasculitis and accelerated atherosclerosis in systemic lupus may be related, rather than distinct, processes since various types of vascular injury and/or immunological mediators have been implicated in the development of atherosclerosis in the general population. Acute vasculitis, as a form of vascular injury, may serve as an initiating factor for the development of atherosclerosis in susceptible patients, in whom the presence of other risk factors (hypertension, increased cholesterol levels) accelerates the process. In this schema, very early steroid therapy might be expected to reduce vascular injury and thus suppress or delay the development of atherosclerosis. Support for this hypothesis can be inferred from a report by Fukumoto *et al*. (1987), in which patients with systemic lupus treated with steroids had less intimal thickening in coronary vessels than patients with systemic lupus who had never been on steroids. In this autopsy study, age and the duration of the systemic lupus erythematosus correlated best with severity of intimal thickening.

In a patient with systemic lupus presenting with chest pain suggestive of acute unstable angina or possible myocardial infarction, the clinical or angiographic differentiation between atherosclerosis and pure acute vasculitis is difficult, but may be of considerable importance because of conflicting indications for and against the use of corticosteroids. In isolated cases, patients with systemic lupus and unstable angina reportedly responded to high-dose steroids, with resolution of chest pain and the disappearance of angiographically documented coronary stenoses (although this is not unequivocal evidence of vasculitis since coronary stenoses caused by coronary spasm or thrombosis in patients with atherosclerosis may also resolve spontaneously). On the other hand, steroids are generally considered to be contraindicated in patients with acute myocardial infarction, based on studies of experimental myocardial infarction (Hammerman *et al*. 1983; Vivaldi *et al*. 1987) in which high-dose steroids led to delayed healing of the infarct and thinning of the scar. The few, small, controlled clinical trials of steroid therapy (Madias and Hood 1982) in acute myocardial infarction have not demonstrated deleterious effects of steroids (none clearly demonstrates beneficial effects either) but numerous case reports (Bulkley and Roberts 1974; Silverman and Pfeifer 1987), most involving rather prolonged use of high-dose steroids for Dressler syndrome, have suggested a relation between early steroid therapy, the formation of ventricular aneurysms, and an increased risk of myocardial rupture. Patients who present with unstable angina are at high risk of progression to acute myocardial infarction; thus, the decision to try a course of steroids for possible coronary vasculitis in a patient with systemic lupus erythematosus with unstable angina or other evidence of acute myocardial ischaemia entails some potential risk. The duration of risk associated with steroid therapy after a myocardial infarction is not certain; in experimental studies, even a short course of steroids administered immediately after the infarct impaired scar formation for weeks after the initial event.

Coronary vasculitis can occur with or without evidence of extracardiac vasculitis so that absence of other evidence of systemic lupus activity is not helpful in excluding acute vasculitis involving a major coronary artery. Alternatively, absence of standard coronary risk factors does not exclude atherosclerosis in a young patient with systemic lupus erythematosus. While hypertension and hyperlipidaemia tend to be more prevalent in patients with systemic lupus who have accelerated atherosclerosis, severe atherosclerosis and acute myocardial infarction have been well documented in young female patients with few, if any, of the standard atherosclerotic risk factors (Bulkley and Roberts 1975).

Apart from biopsy of coronary arteries, there is no definitive way to differentiate coronary atherosclerosis from coronary arteritis. There is limited information on coronary angiography in patients with systemic lupus erythematosus (Heibel *et al*. 1976) and polyarteritis nodosa (Cassling *et al*. 1985) who have coronary disease. Unlike in Kawasaki disease (Suzuki *et al*. 1986) (see Chapter 5.11.8), there are no diagnostic radiographic features of vasculitis in coronary arteries in systemic lupus. Serial angiograms showing the abrupt appearance of an obstructive lesion in a previously normal vessel, aneurysmal dilation of the coronary arteries, and elongated smooth lesions have been suggested as radiographic features of coronary vasculitis, but all of these phenomena can also occur in atherosclerosis.

Treatment strategies in unstable angina and acute myocardial infarction

In a patient presenting with acute, unstable angina without evidence of infarction in whom there is a strong suspicion of acute coronary vasculitis, a short course of high-dose steroids may be appropriate in conjunction with standard treatment (antianginal drugs and aspirin and/or heparin) moving to early angiography in those who do not respond promptly to medical therapy.

The initial treatment of an acute myocardial infarction is the same regardless of the underlying coronary pathology (Suzuki *et al*. 1986). The patient should be under intensive care and treated with the usual medical regimen (nitrates, aspirin and/or heparin, and, if not contraindicated, a b-blocker). The current practice of administering thrombolytic therapy promptly to patients who present within the first few hours of a clearly documented acute myocardial infarction presents a dilemma in patients with systemic lupus as there are no data on the risk:benefit ratio of giving acute thrombolytic therapy to a patient with acute myocardial infarction caused by coronary vasculitis. The cases reviewed by Doherty and Siegel (1985) suggest that patients with systemic lupus and acute coronary syndromes have a fair likelihood of having intracoronary thrombus; anticardiolipin antibodies have been implicated as a contributory factor (Greisman *et al*. 1991). Whether or not the haemorrhagic risk associated with thrombolytic therapy (especially that of intracerebral haemorrhage) is in any way altered in patients with systemic vasculitis is unknown. Steroids should probably not be used if there is clear evidence of evolving myocardial infarction unless there are compelling reasons to do so, i.e., severe vasculitis involving other organs, in which case doses and duration of therapy should be minimal.

Vasculitis affecting the ascending aorta (giant-cell arteritis, Takayasu arteritis, temporal arteritis, non-specific aortoarteritis) will occasionally involve the proximal segment (ostium) of the right or left coronary artery, presenting as acute myocardial infarction, unstable angina, congestive heart failure, or sudden death (Holsinger *et al*. 1962; Cipriano *et al*. 1977; Save Soderbergh *et al*. 1986; Mitnick *et al*. 1990). In most reported cases other evidence of disease activity implicated arteritis as the cause of the coronary disease, but in several cases aortitis was a previously unsuspected finding discovered at the time of coronary bypass surgery for severe angina or at autopsy in a patient who died of an acute myocardial infarction. In a case reported by Mitnick *et al*. (1990), a patient with temporal arteritis developed ostial proximal coronary involvement despite low-dose steroid therapy that had been effective in controlling the systemic symptoms of the disease.

Vasculitis involving the small, intramural coronary arteries is commonly observed at autopsy in polyarteritis nodosa (Holsinger *et al*. 1962; Przybojewski 1981) and rheumatoid arthritis (15–20 per cent) (Haider and Roberts 1981; Lie 1987), but it is usually asymptomatic and only rarely presents as chest pain or an acute myocardial infarction (Morris *et al*. 1986). Small-vessel coronary arteritis may be responsible for diffuse, patchy areas of myocardial fibrosis leading to congestive heart failure or to clinical syndromes reflecting the specific area of damage, such as arrhythmia or heart block related to fibrosis of the specialized conduction tissues. In systemic sclerosis, diffuse narrowing of intramural coronary arteries secondary to intimal thickening and periadventitial sclerosis is occasionally found, and may be associated with myocardial infarction and sudden death (James 1974; Owens and Follansbee 1987; Nair *et al*. 1988). Diffuse coronary vasospasm is also thought to play an important part in this disorder and may be responsible for the fact that diffuse myocardial fibrosis, out of proportion to frankly obstructive coronary disease, is commonly observed in this condition. (Follansbee *et al*. 1993). Amyloidosis may involve small coronary arteries, and cause angina and myocardial infarction (Saffitz *et al*. 1983).

Evaluation of patients with suspected coronary disease

Patients with rheumatic disorders who develop angina, dyspnoea on exertion, or heart failure should be referred for a complete cardiovascular evaluation. Evaluation should include screening for risk factors for atherosclerosis and a stress test to aid in diagnosis and also to screen for high-risk profiles (early or profoundly positive tests implying the presence of severe or extensive disease). Patients with severe arthritis who cannot perform a bicycle or treadmill exercise test can be evaluated with pharmacological stress tests (dipyridamole, adenosine, or dobutamine thallium scintigraphy or echocardiography), which have comparable sensitivity and specificity for the detection of significant coronary arterial disease. Most young or middle-aged patients with objective evidence (abnormal electrocardiogram, positive stress test) or strong clinical evidence (classic symptoms, history of myocardial infarction, positive risk factors) of coronary arterial disease should undergo coronary angiography to define the severity of involvement and determine the appropriateness of medical as compared to revascularization therapy (angioplasty or bypass surgery). The decision in elderly or inactive patients who are limited by their rheumatological disease is more complex and should be based on the severity of their

cardiac symptoms and the efficacy of standard medical therapy.

Pericardial disease

Diffuse or focal inflammatory changes in the pericardium are found commonly at autopsy in many of the rheumatic diseases but clinically evident pericardial disease occurs less often (Table 2). Acute pericarditis, with or without effusions, occurs most often in systemic lupus erythematosus (Doherty and Siegel 1985; Mandel 1987), rheumatoid arthritis (John *et al*. 1979; Pizzarello and Goldberg 1985) and juvenile rheumatoid arthritis (Brewer 1977; Svantesson *et al*. 1983), and mixed connective tissue disease (Alpert *et al*. 1983) as well as in acute rheumatic fever (Stollerman 1988) (see also Chapter 5.3.12); systemic sclerosis (Owens and Follansbee 1987) is more often associated with chronic effusions. The clinical syndrome of pericarditis or pericardial effusion progressing to tamponade or constrictive pericarditis is rare, but has been reported in rheumatoid arthritis, systemic sclerosis, and systemic lupus.

Rheumatic disease	Frequency[a]
Behçet's disease	+
Hypersensitivity vasculitis	+
Juvenile rheumatoid arthritis	+
Kawasaki's disease	+
Lyme disease	+
Mixed connective tissue disease	++
Polyarteritis nodosa	+
Polymyositis/dermatomyositis	+
Reiter's syndrome	+
Rheumatic fever	++
Rheumatoid arthritis	+
Systemic lupus erythematosus	++
Systemic sclerosis	++
Wegener's granulomatosis	+

[a]Approximate frequency of clinical pericarditis in patients who have rheumatic disease. Pericarditis occurs rarely if ever in those rheumatic diseases not listed. +++Occurs in >50% of patients; ++occurs in 5–49% of patients; +occurs in <5% of patients.

Table 2 Rheumatic diseases causing pericardial involvement

Acute pericarditis

The pericardium has two layers, the visceral pericardium, which is adherent to the epicardial surface of the heart, and the parietal pericardium, which is separated from the visceral layer by a small amount of pericardial fluid (25–50 ml). Acute pericardial inflammation, which can result from a wide spectrum of pathological processes, usually results in a syndrome characterized by chest pain, a pericardial friction rub, and electrocardiographic changes. Two of the above are usually adequate to make the diagnosis. Pericarditis may or may not be associated with increased accumulation of fluid in the pericardial space.

Pericardial pain will be worsened by actions that increase the contact between the pericardial layers or between the parietal pericardium and the adjacent lungs or mediastinal contents. Thus, patients may complain of pain when changing position, coughing, deep breathing, or swallowing. The pericardial friction rub is produced by the rubbing together of the roughened, inflamed surfaces of the visceral and parietal pericardium during the cardiac cycle. It is a harsh, scratching, crunching or grating sound heard over the anterior precordium or left chest, and may be heard best with the patient leaning forward and holding forced expiration. Classically, the rub is triphasic, i.e., composed of three discrete sounds, corresponding to the motion of the heart in late diastole (atrial systole), early ventricular systole, and early diastole (rapid diastolic filling). The rub may also be biphasic with single late diastolic and systolic components or uniphasic with a single systolic component. The rub may be pleuropericardial, i.e., related to the cardiac cycle but present only during inspiration.

Electrocardiographic changes associated with pericarditis include depression of the PR segment, diffuse elevation of the ST segment, and diffuse inversion of the T wave (Fig. 1), which usually follow each other sequentially and resolve within a few days. Atrial arrhythmias, most commonly atrial fibrillation or flutter, are common. Ventricular ectopy is unusual unless there is associated myocardial involvement. The echocardiogram is not specifically helpful in making the diagnosis of pericarditis unless there is an associated pericardial effusion.

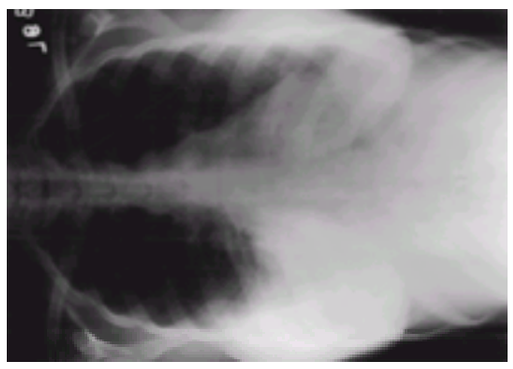


Fig. 1 A chest radiograph of a 23-year-old female with systemic lupus erythematosus and pleuropericardial disease.

Pericardial effusion

The signs and symptoms associated with a pericardial effusion are dependent on the size of the effusion and the rapidity with which it has accumulated. The pericardium has the capacity to stretch so that large collections of pericardial fluid may be tolerated without symptoms if the fluid has accumulated very slowly (Fig. 2). In contrast, small, rapidly accumulating effusion associated with, for example, an acute flare of systemic lupus erythematosus, can produce pericardial tamponade. When there is a large effusion, the electrocardiogram may show diminution in QRS voltage; changes characteristic of acute pericarditis may or may not be present. Echocardiography is the diagnostic technique of choice for identifying the presence of an effusion, assessing its size, and guiding pericardiocentesis. Diagnostic pericardiocentesis is indicated primarily to exclude treatable infections, particularly tuberculosis, or neoplasm. Purulent pericarditis has been reported (rarely) in rheumatoid arthritis. Characteristics which suggest that the effusion is rheumatic in origin are discussed below.

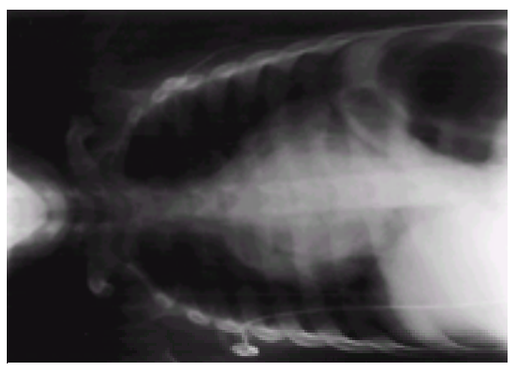


Fig. 2 A child with systemic-onset juvenile rheumatoid arthritis complicated by a large pericardial effusion. Note the 'water-bottle shape' of the cardiac silhouette.

Pericardial tamponade

Large or rapidly accumulating effusions may lead to tamponade, a condition in which diastolic filling of the heart is restricted by the pressure of fluid within the pericardial space. Once the pressure in the pericardial space exceeds that in the systemic and pulmonary veins emptying into the heart, diastolic filling ceases and there is a precipitous drop in cardiac output followed by hypotension, and, if not treated by immediate decompression of the pericardial space, cardiovascular collapse and death.

Signs of impending pericardial tamponade include evidence of obstructed venous inflow into the heart (elevated jugular venous pressure) and decreased cardiac output (sinus tachycardia, narrow pulse pressure, decreased urine output, cool extremities, altered mental status). Pulsus paradoxicus, a drop in systolic pressure of more than 10 mmHg during each inspiratory cycle, is a characteristic sign of impending tamponade. Electrical alternans, an electrocardiographic pattern of alternating increase and decrease in QRS voltage (and/or T-wave height) is another sign suggestive of impending tamponade.

Echocardiographic signs of pericardial tamponade include indenting of the right atrium or right ventricle during diastolic filling. (This sign may be obscured in patients with significant left ventricular dysfunction.)

If the diagnosis of pericardial tamponade is suspected but cannot be confirmed by echocardiogram, the patient should undergo a right heart catheterization. If the diagnosis of tamponade is confirmed, the patient should then undergo pericardiocentesis for relief of tamponade and for diagnostic evaluation of the fluid. Tamponade can occur with any type of inflammatory pericarditis and does not suggest a specific cause.

The risks of percutaneous pericardiocentesis are minimized when there is a large effusion and when echocardiographic or fluoroscopic guidance is available. Alternatively, a pericardial window may be created surgically or percutaneously using a balloon-tipped catheter. If the procedure is done in the setting of impending or suspected tamponade, a catheter should be left in the pericardial space for recurrent drainage since highly inflammatory, rapidly accumulating effusions are likely to recur. Steroid therapy is usually effective, and pericardial sclerosis or pericardial stripping is rarely required.

Constrictive pericarditis

Constrictive pericarditis refers to a state of chronic restriction of cardiac filling by a thickened, rigid pericardium. Recurrent or severe pericarditis associated with the rheumatic disorders very rarely leads to this complication; it has been reported occasionally in systemic sclerosis (Owens and Follansbee 1987) and rheumatoid arthritis (John *et al*. 1979). Restricted inflow into the right heart results in elevation of systemic venous pressure and a positive Kaussmaal sign (failure of the jugular venous pressure to fall or paradoxical elevation of the venous pulse wave with inspiration), peripheral oedema, and hepatic congestion. Inflow from the pulmonary veins into the left side of the heart is also impeded and patients are frequently dyspnoeic and fatigued. Electrocardiographic signs of acute pericarditis are usually absent, and instead of a rub a pericardial knock (a loud third heart sound corresponding to early diastolic inflow into the constricted, non-compliant ventricle) may be heard. Constrictive pericarditis may be difficult to distinguish from restrictive cardiomyopathy, a heart muscle disorder in which a similar pattern of impedance to cardiac filling results from diminished compliance of the myocardium, most often as a result of infiltrative diseases such as amyloid. The distinction can be critical since severe constrictive pericarditis responds poorly to medical therapy and usually requires surgical stripping for definitive relief. CT scanning or echocardiography can demonstrate pericardial thickening but neither establishes whether or not there is physiological constriction. While cardiac catheterization may show haemodynamic patterns suggestive of constrictive pericarditis, thoracotomy may be required to make a definitive diagnosis.

Differentiating pericarditis of rheumatic origin from other causes of pericarditis

The distinction between rheumatic pericarditis and other common causes of pericarditis is difficult to make on the basis of cardiac findings alone. In most large series of patients with clinical evidence of pericarditis, the rheumatic diseases account for less than 2 per cent of cases. Probably the most common cause of pericarditis in young patients is acute viral pericarditis, in which the pattern of chest pain, physical findings, and electrocardiographic changes are similar to those of pericarditis of rheumatic origin; leucocytosis and a raised erythrocyte sedimentation rate are also seen in both. Pericarditis associated with rheumatic diseases often occurs in the setting of an obvious flare involving other non-cardiac manifestations of the rheumatic condition; however, pericarditis is occasionally the sole or presenting finding, especially in systemic lupus erythematosus. In systemic lupus, rheumatoid arthritis, and mixed connective-tissue disease, an associated pleural effusion is common, but this can also occur in up to 60 per cent of patients with acute, non-rheumatic pericarditis.

The echocardiogram may help differentiate acute viral from rheumatic pericarditis in that certain of the rheumatic disorders are associated with echocardiographic findings other than effusion. In patients presenting with acute viral pericarditis, pericardial thickening is uncommon, but 10 to 15 per cent of patients with rheumatoid arthritis (Pizzarello and Goldberg 1985) and a somewhat smaller percentage of patients with systemic sclerosis (Owens and Follansbee 1987) have evidence of chronic pericardial disease; in patients with systemic lupus, echocardiographic abnormalities including pericardial thickening, valvular insufficiency, and sterile vegetation are present in 30 to 50 per cent (Doherty and Siegel 1985).

Analysis of the pericardial fluid is important in excluding infectious or malignant pericarditis, especially in patients on steroids or immunosuppressive agents. Pericardial biopsy may be required to make the diagnosis of tuberculous pericarditis. Pericardial fluid may offer some clues to the diagnosis of specific rheumatic disorders, although there are no pathognomonic features. Pericardial fluid glucose is low in rheumatoid arthritis (Pizzarello and Goldberg 1985) and systemic lupus erythematosus (Doherty and Siegel 1985) but is usually normal in viral pericarditis. Complement has also been reported to be low but difficulties in interpretation arise in that the amounts of complement have often been uncorrected for the protein content of the effusion and systemic complement levels. Moreover, the number of patients who have had these tests is small. The same is true of tests for antinuclear antibody in pericardial fluid; it is unclear whether they provide any significant information over and above that which can be obtained by testing the patients' blood. Furthermore, antinuclear antibodies have been reported in pericardial fluid from non-rheumatic pericardial effusions and it is likely that any condition which can produce a positive antinuclear antibody test in peripheral blood may produce a similarly positive test in an exudative pericardial effusion.

Myocardial disease

Primary myocardial disease includes acute inflammatory myocarditis and chronic cardiomyopathy, which is characterized as either dilated, restrictive, or hypertrophic.

Myocarditis

Autopsy series report histological evidence of myocarditis in many of the chronic rheumatic disorders (Table 3), with perivascular and interstitial inflammatory cells and focal fibrosis in up to 40 per cent of cases of systemic lupus erythematosus (Doherty and Siegel 1985; Mandel 1987) and 20 per cent of cases of rheumatic arthritis (Pizzarello and Goldberg 1985). However, there is generally a poor correlation between histological findings and clinical evidence of myocardial dysfunction. Acute, severe myocarditis presenting as heart failure is recognized clinically far less often. Acute myocarditis is characterized by rapid onset and progression of congestive heart failure, with pulmonary congestion and cardiomegaly on chest radiography, sinus tachycardia, diffuse, non-specific changes in ST and T wave on electrocardiograms, and atrial and ventricular arrhythmias. Dilation of all four cardiac chambers with markedly depressed systolic function of both ventricles is demonstrable on the echocardiogram or radionuclide ventriculogram. Severe, fulminant myocarditis may present as low cardiac-output state and shock. There is enzymatic evidence of myocardial necrosis (elevated serum creatine kinase-MB), and histological evidence of myocyte degeneration, as well as diffuse inflammatory infiltration on endomyocardial biopsy. Occasionally, acute myocarditis may be the sole or presenting sign of systemic lupus or rheumatoid arthritis and will mimic acute viral myocarditis, but, more commonly, other manifestations of rheumatic disease activity are present. Treatment includes bed rest, pharmacological support with positive inotropic agents, afterload-reducing agents, diuretics, and usually systemic anticoagulation to prevent venous thrombosis and systemic embolization from ventricular mural thrombi. There is anecdotal evidence supporting the use of high-dose steroids in severe myocarditis but few data on the use of immunosuppressive agents (Talwar *et al*. 1988).

Rheumatic disease	Frequency[1]
Amyloidosis	++
Ankylosing spondylitis	+
Juvenile rheumatoid arthritis	+
Kawasaki's disease	++
Lyme disease	+
Polymyositis/dermatomyositis	++
Reiter's syndrome	+
Rheumatic fever	++
Rheumatoid arthritis	+
Sarcoidosis	++
Systemic lupus erythematosus	+
Systemic sclerosis	++
Takayasu's disease	+

[1]Approximate frequency of clinically significant myocarditis or cardiomyopathy in patients who have rheumatic disease. Myocardial disease occurs rarely if ever in those rheumatic disorders not listed.
+++ Occurs in >50% of patients. ++ occurs in 5–49% of patients. + occurs in <5% of patients.

Table 3 Rheumatic diseases causing myocardial disease

The role of myocardial biopsy in patients with rheumatic disorders who develop myocardial dysfunction is not clear as there has been a poor correlation between histological evidence of myocarditis and degree of myocardial dysfunction or response to corticosteroids. However, in patients with acute or progressive myocardial dysfunction in whom more common causes of heart failure have been excluded (myocardial infarction, valvular disease, severe anaemia), myocardial biopsy may be helpful in documenting the presence of inflammatory changes and excluding alternative, albeit rare, causes of myocardial failure (such as hydroxychloroquine toxicity, amyloid, sarcoid, or haemochomatosis) before starting a trial of corticosteroid therapy.

Acute viral myocarditis may be indistinguishable from other forms of inflammatory myocarditis. It has been suggested that steroids may be deleterious in the acute stages of viral myocarditis, based on studies in murine myocarditis in which viral replication was enhanced by steroid administration (Reyes and Lerner 1985). This observation has not been confirmed in human studies and there are anecdotal reports of marked improvement in patients with apparent viral myocarditis following administration of steroids and occasionally of cytotoxic agents. While it is not yet clear whether or not corticosteroids and immunosuppressive drugs are helpful in acute viral myocarditis (this is currently under investigation in a large clinical trial in the United States), there is no clear evidence that they are harmful. Thus, it appears reasonable to administer a course of steroids in acute myocarditis suspected of being associated with a rheumatic disorder, even if acute viral myocarditis cannot be ruled out.

Dilated cardiomyopathy

The term cardiomyopathy refers to any chronic disorders of cardiac muscle; cardiomyopathies are currently divided into dilated (congestive), hypertrophic, or restrictive types. Dilated cardiomyopathy is characterized by enlargement the four cardiac chambers with depressed systolic function. Myocardial dysfunction may involve all four cardiac chambers or may be isolated to either the left- or right-sided chambers. Atrial and ventricular arrhythmias are common. Patients usually present with insidious onset of dyspnoea on exertion and diminished exercise tolerance, or with new onset of atrial fibrillation; more rapid onset with progression to pulmonary oedema may also occur. Unlike acute myocarditis, extracardiac signs of acute vasculitis or inflammation are usually not present and there is neither serological nor histological evidence of acute inflammation. The aetiology of dilated cardiomyopathy associated with the rheumatic disorders is multifactorial. A picture of dilated cardiomyopathy may develop secondary to long-standing hypertension in systemic lupus erythematosus (Doherty and Siegel 1985) or in scleroderma (Owens and Follansbee 1987), or after extensive myocardial infarction in patients with coronary arteritis (Lie 1987), as a result of diffuse fibrosis in sarcoid (Shammas and Movahed 1993), or as a consequence of diffuse amyloid infiltration (Hamer *et al*. 1992). Diffuse ischaemic damage may also occur in small-vessel coronary vasculitis from obstruction, in situ thrombosis, or, in the case of scleroderma, coronary vasospasm (Follansbee *et al*. 1993). Long-standing mitral or aortic insufficiency leading to cardiac failure may also present as a dilated cardiomyopathy. Infiltrative myopathy, as in amyloid (Kyle and Bayrd 1975), may also cause a dilated cardiomyopathy.

In the absence of these underlying factors, clinically significant dilated cardiomyopathy as a primary consequence of the rheumatic disorders is uncommon, although many patients have histological evidence of patchy fibrosis and minimal chronic inflammatory reaction. Subclinical abnormalities of myocardial function, insufficient to cause frank congestive heart failure, have been described in several studies of patients with stable systemic lupus in the absence of hypertension or obvious coronary vasculitis (Doherty and Siegel 1985).

Restrictive cardiomyopathy

This is characterized by abnormal compliance of the heart leading to impaired diastolic filling with little or no impairment of systolic function. Clinically, patients present with signs and symptoms of biventricular heart failure (elevated systemic venous pressure and pulmonary congestion) and impaired cardiac reserve, including limited exercise tolerance, but without marked cardiomegaly. The most common cause of restrictive cardiomyopathy associated with the rheumatic disorders is primary or secondary amyloid (Chew *et al*. 1975). Although cardiac amyloid can cause dilated cardiomyopathy, the initial or predominant manifestation of amyloid infiltration of the myocardial walls may be restricted diastolic filling and impaired compliance, leading to depressed cardiac output reserve and elevation of systemic and pulmonary venous pressures (see also Chapter 5.13.1).

Little can be done to improve diastolic function in restrictive cardiomyopathy. Symptomatic pulmonary congestion can be relieved by diuretics, but these must be used with caution since underfilling of the left ventricle easily leads to low cardiac output with fatigue, hypotension, and renal insufficiency. It is essential to recognize that the signs, symptoms, and physiological consequences of restrictive cardiomyopathy may be virtually identical to those of constrictive pericarditis, which can be treated surgically. Clues to the differential diagnosis can be obtained by CT scanning, echocardiography, or cardiac catheterization, but thoracotomy is sometimes required to exclude constrictive pericarditis.

Hypertrophic cardiomyopathy

Hypertrophic cardiomyopathy is a form of cardiomyopathy characterized by massive thickening of the ventricular walls and is not specifically associated with the rheumatic disorders. However, severe left-ventricular hypertrophy developing in response to hypertension in patients with systemic lupus or scleroderma may mimic it. Patients with severe left-ventricular hypertrophy often demonstrate abnormal diastolic properties of the left ventricle, such that inordinately high pressures develop during left ventricular filling, with resultant elevations in left atrial pressure and symptoms of pulmonary congestion. Systolic function is preserved initially, but long-standing untreated hypertension may eventually cause left ventricular dilation and contractile failure.

Conduction system disease

The cardiac conduction system consists of the sinoatrial node located in the high right atrium, the atrioventricular node located at the junction of the aortic root, anterior mitral leaflet and interventricular septum, the intra-atrial conduction fibres, which connect the two nodes, and the His–Perkinje system, which consists of specialized conduction fibres emanating from the atrioventricular node and spreading through the interventricular septum and ventricular walls. The conduction system is subject to damage by numerous mechanisms in the rheumatic disorders (Table 4) as it is intimately associated with, or surrounded by, dense connective tissue and receives its blood supply from small- to medium-sized arteries potentially affected by vasculitis. The proximity of the atrioventricular node to the aortic ring may result in conduction disorders in patients with inflammatory aortitis or connective tissue diseases involving the aorta. Thus, the conduction system of the heart may be damaged by fibrinoid necrosis of connective tissue or vasculitis of the atrioventricular or sinoatrial nodal arteries in systemic lupus erythematosus (Doherty and Siegel 1985; Mandel 1987), infiltration by rheumatoid (Pizzarello and Goldberg 1985) or sarcoid nodules (Shammas and Movahed 1993), or by amyloid infiltration (Mandel 1987), diffuse myocardial fibrosis in scleroderma (James 1974; Owens and Follansbee 1987), inflammatory myocarditis or degenerative myopathy in polydermatomyositis and Lyme disease (Steere *et al*. 1980; Askari and Huettner 1982; Haupt and Hutchins 1982), or extension of aortic root involvement to adjacent structures in ankylosing spondylitis (Nagyhegyi *et al*. 1988), Reiter's disease (Good 1974), and Ehlers–Danlos syndrome (Leier *et al*. 1980).

The presence of a first-degree atrioventricular block, Mobitz I (Wenkebach) atrioventricular block, or bundle branch block on electrocardiography in an asymptomatic patient with a rheumatic disorder suggests the possibility of conduction system involvement but does not necessarily predict progression to complete heart block and is not an indication for prophylactic implantation of a pacemaker. Periodic follow-up electrocardiograms are important and careful monitoring should be undertaken if the patient experiences near or actual syncope. Higher degrees of atrioventricular block (Mobitz II) or complete heart block imply advanced conduction system

disease and are a strong indication for prophylactic implantation of a pacemaker even if the patient is asymptomatic.

Patients who present with syncope or light-headedness in whom second- or third-degree heart block, or severe sinus bradycardia (pulse less than 40), or sinus pauses greater than 5 s can be documented are clearly candidates for the implantation of a permanent transvenous pacemaker. In symptomatic patients with less severe degrees of bradyarrhythmia or heart block, it is important to correlate the patient's symptoms with the occurrence of bradycardia during ambulatory monitoring or electrophysiological testing before implanting a pacemaker. Despite the potential for involvement of the conduction system, symptomatic bradyarrhythmias are uncommon in the rheumatic disorders and unless an arrhythmic cause is clearly documented, patients with syncope or light-headedness should be carefully investigated for other causes before permanent pacing is undertaken.

Aortic insufficiency

Aortic insufficiency can occur as a consequence of dilation or dissection of the aortic root, or as a result of damage or deformation of the aortic valve leaflets themselves (Table 5). Stretching of the aortic root with failure of the leaflets to coapt occurs in inflammatory diseases, especially ankylosing spondylitis (Nagyhegyi *et al*. 1988; O'Neill 1992), Reiter's disease (Good 1974), Takayasu disease (Ishikawa 1978), giant-cell arteritis (Pyeritz and Wappel 1983), Behçet's disease (James and Thomson 1982), and in the connective tissue diseases, especially Marfan syndrome (Marsalese *et al*. 1989), and Ehlers–Danlos syndrome (Leier *et al*. 1980). Disease of the aortic root in Marfan syndrome can be particularly treacherous in that acute aortic dissection can develop, which may present as fulminant aortic insufficiency. Marfan syndrome and the other connective tissue diseases are also associated with myxomatous degeneration of the aortic valve leaflets, which may contribute to aortic insufficiency. Scarring of the leaflets occurs occasionally as a result of Libman–Sachs endocarditis in systemic lupus (Doherty and Siegel 1985; Mandel 1987) (the mitral valve is affected more commonly). Deformation of the aortic valve by rheumatoid nodules is another rare cause of aortic insufficiency. Aortic valve disease is also reported in juvenile rheumatoid arthritis (Heyd and Glaser 1990).

Rheumatic disease	Valve[1]	
	Aortic[2]	Mitral
Ankylosing spondylitis	++	–
Behçet's disease	+	–
Ehlers–Danlos syndrome	++	++
Juvenile rheumatoid arthritis	+	–
Kawasaki's disease	–	+
Marfan syndrome	++	+++
Mixed connective tissue disease	–	++
Reiter's syndrome	+	–
Rheumatic fever	+	+
Rheumatoid arthritis	+	+
Systemic lupus erythematosus	+	+
Takayasu's disease	+	–
Temporal arteritis	+	–

[1] Approximate frequency of clinically significant valvular insufficiency in patients with rheumatic disease. Valvular insufficiency occurs rarely or even in those rheumatic diseases not listed.
[2] Due to either valvular pathology or aortic root dilatation.
+++ Occurs in > 50% of patients, ++ occurs in 15–49% of patients, + occurs in < 5% of patients, – occurs rarely if at all.

Table 5 Rheumatic diseases causing valvular insufficiency

The presentation and management of aortic insufficiency depends on whether the valve insufficiency develops slowly or abruptly. Slowly progressive aortic insufficiency is usually well tolerated and asymptomatic for many years as the left ventricular chamber is able to compensate by dilation and hypertrophy in response to the gradual imposition of the increased volume load. In contrast, acute aortic insufficiency presents the unprepared left ventricle with an acute volume overload, which results in severe heart failure. Acute, severe aortic insufficiency constitutes a medical emergency requiring early surgical replacement of the aortic valve. Chronic aortic insufficiency can usually be managed conservatively for many years, delaying valve replacement until the patient develops signs and symptoms of left ventricular dysfunction.

The characteristic, diastolic blowing decrescendo murmur of aortic insufficiency can be soft and is best detected with the bell of the stethoscope at the mid- to lower-left sternal border with the patient leaning forward and holding his or her breath in full expiration. In patients with aortic insufficiency secondary to dilation of the aortic root, the murmur is often heard best at the right upper sternal border. In addition to the diastolic blowing murmur, there is often a systolic ejection-type murmur at the base, reflecting the increased stroke volume in systole, and a low-frequency mid-diastolic rumble (Austin–Flint murmur) at the lower left sternal border that mimics mitral stenosis and is caused by premature closure of the anterior mitral-valve leaflet by the aortic regurgitant jet. Other characteristic physical findings in aortic insufficiency reflect the high stroke volume: the apical cardiac impulse is displaced to the left and is hyperdynamic, and the peripheral and carotid pulses are bounding. Echo-Doppler studies can best assess the severity and progression of aortic root- or valve-related aortic insufficiency.

The decision when to refer patients for aortic valve replacement is based on development of symptoms and evidence of deterioration of left ventricular function, ideally assessed by serial echocardiography. Asymptomatic patients with good left-ventricular systolic function (normal ejection fraction) do not require valve replacement. Patients who develop symptoms of congestive heart failure (dyspnoea on exertion) can usually obtain symptomatic benefit from medical therapy (digoxin, diuretics, and vasodilator therapy), but valve replacement is usually recommended early in such patients if there is evidence of systolic dysfunction, in order to offset progression to severe, irreversible, left ventricular dysfunction, which greatly increases the morbidity and mortality of valve replacement surgery and limits the benefits achieved.

Although most patients develop obvious symptoms when left ventricular systolic function deteriorates, members of a perplexing subgroup remain asymptomatic despite echocardiographic evidence of a decreasing ejection fraction. When to replace the aortic valve in an asymptomatic patient with left ventricular dysfunction demonstrated on echocardiogram remains controversial. However, there is agreement that careful periodic examination (history and physical examination) and serial studies of left ventricular function by echocardiogram should be routine for all patients with aortic insufficiency, with referral to a specialist for more detailed cardiac evaluation and consideration for valve replacement as soon as symptoms of heart failure or signs of left ventricular dysfunction on echocardiogram appear.

Mitral valve insufficiency

Clinically significant mitral valve insufficiency is uncommon in association with rheumatic disease except as a late consequence of acute rheumatic fever (see Table 5). Rheumatoid nodules may deform the mitral leaflets and result in mitral insufficiency (Pizzarello and Goldberg 1985), as can Libman–Sachs endocarditis and valvulitis in systemic lupus erythematosus (Doherty and Siegel 1985; Mandel 1987). The connective tissue disorders, especially Marfan syndrome, are associated with myxomatous degeneration of the mitral leaflets (Van Albada-Kuipers *et al*. 1986) and mitral valve prolapse, which occasionally causes significant mitral insufficiency but more commonly causes the characteristic mid-systolic murmur without haemodynamically significant valvular insufficiency. Occasionally, mitral valve prolapse results in acute rupture of the chordae tendineae, which causes acute, severe mitral insufficiency and presents as acute pulmonary oedema, requiring urgent valve replacement.

As in aortic insufficiency, slowly developing mitral insufficiency is usually well tolerated for long periods. The left ventricle undergoes gradual dilation and hypertrophy, enabling it to pump the increased stroke volume (normal diastolic filling from the pulmonary veins plus the additional load of blood regurgitation back into the left atrium through the incompetent mitral valve in systole). Dilation of the left atrium enables it to accommodate the mitral regurgitant volume without marked elevation in left atrial pressure. Left ventricular function may remain well compensated for many years. The onset of left ventricular dysfunction, or sudden worsening of the mitral insufficiency, will result in an increase in left atrial pressure and signs and symptoms of pulmonary congestion. Initially, symptoms can be ameliorated by a combination of digoxin, diuretics, and afterload-reduction (vasodilator) therapy, but severe or recurrent bouts of heart failure should prompt referral to a specialist for more detailed evaluation and consideration of valve replacement or repair, as is possible in some instances.

Chapter References

Alpert, M. A. *et al*. (1983). Cardiovascular manifestations of mixed connective tissue disease in adenitis. *Circulation*, **68**, 1182–93.

Askari, A. D. and Huettner, T. L. (1982). Cardiac abnormalities in polymyositis/dermatomyositis. *Seminars in Arthritis and Rheumatism*, **12**, 208–19.

Brewer, E., Jr. (1977). Juvenile rheumatoid arthritis–cardiac involvement. *Arthritis and Rheumatism*, **20**, 231–6.

Bulkley, B. H. and Roberts, W. C. (1974). Steroid therapy during acute myocardial infarction: a cause of delayed healing and ventricular aneurysm. *American Journal of Medicine*, **56**, 244–50.

Bulkley, B. H. and Roberts, W. C. (1975). The heart in systemic lupus erythematosus and the changes induced in it by corticosteroid therapy: a study of 36 necropsy patients. *American Journal of Medicine*, **58**, 243–64.

Cassling, R. S., Loritz, J. B., Olson, D. R., Hubbard, T. F., and McManus, B. M. (1985). Fatal vasculitis (periarteritis nodosa) of the coronary arteries: angiographic ambiguities and absence of aneurysms at autopsy. *Journal of the American College of Cardiology*, **6**, 707–14.

Chew, C., Ziady, G. M., Raphael, M. J., and Oakley, C. M. (1975). The functional defect in amyloid heart disease: the stiff heart syndrome. *American Journal of Cardiology*, **36**, 438.

Cipriano, P. R., Silverman, J. F., Perlroth, M. G., Griepp, R. B., and Wexler, L. (1977). Coronary arterial narrowing in Takayasu's aortitis. *American Journal of Cardiology*, **39**, 744.

Doherty, N. E. and Siegel, R. J. (1985). Cardiovascular manifestations of systemic lupus erythematosus. *American Heart Journal*, **110**, 1257–65.

Fauci, A. S., Haynes, B. F., and Katz, P. (1978). The spectrum of vasculitis: clinical, pathological, immunological and therapeutic considerations. *Annals of Internal Medicine*, **89** (Part 1), 660–76.

Follansbee, W. P., Zerbe, T. R., and Medsger, T. A., Jr. (1993). Cardiac and skeletal muscle disease in systemic sclerosis (scleroderma): a high risk association. *American Heart Journal*, **126**, 194–203.

Fukumoto, S., Tsumagari, T., Kinjo, M., and Tanaka, K. (1987). Coronary atherosclerosis in patients with systemic lupus erythematosus at autopsy. *Acta Pathologica Japonica*, **37**, 1–9.

Good, A. (1974). Reiter's disease: a review with special attention to cardiovascular and neurologic sequelae. *Seminars in Arthritis and Rheumatism*, **3**, 253–86.

Greisman, S. G., Thayaparan, R., Godwin, T. A., and Lockshin, M. D. (1991). Occlusive vasculopathy in systemic lupus erythematosus. Association with anticardiolipin antibody. *Archives of Internal Medicine*, **151**, 389–92.

Haider, Y. S. and Roberts, W. C. (1981). Coronary arterial disease in systemic lupus erythematosus. Quantification of degrees of narrowing in 22 necropsy patients (21 women) aged 16 to 37 years. *American Journal of Medicine*, **70**, 775–81.

Hamer, J. P. M., Janssen, S., van Rijswijk, M. H., and Lie, K. I. (1992). Amyloid cardiomyopathy in systemic non-hereditary amyloidosis. *European Heart Journal*, **13**, 623–7.

Hammerman, H., Kloner, R. A., Hales, S., Schoen, F., and Braunwald, E. (1983). Dose-dependent effects of short-term methylprednisolone on myocardial infarct extent, scar formation and ventricular function. *Circulation*, **68**, 446–53.

Haupt, H. M. and Hutchins, G. M. (1982). The heart and cardiac conduction system in polymyositis-dermatomyositis: a clinicopathologic study of 16 autopsied patients. *American Journal of Cardiology*, **50**, 998–1006.

Heberden, W. (1772). Some account of a disorder of the breast. *Medical Transactions* (Royal College of Physicians, London), **2**, 59.

Heibel, R. H., O'Toole, J. D., Curtiss, E. I., Medsger, T. A., Reddy, S. P., and Shaver, J. A. (1976). Coronary arteritis in systemic lupus erythematosus. *Chest*, **69**, 700–3.

Heyd, J. and Glaser, J. (1990). Early occurrence of aortic valve regurgitation in a youth with systemic-onset juvenile rheumatoid arthritis. *American Journal of Medicine*, **89**, 123–4.

Holsinger, D. R., Osmundson, P. J., and Edwards, J. E. (1962). The heart in periarteritis nodosa. *Circulation*, **26**, 610–18.

Ishikawa, K. (1978). Natural history and classification of occlusive thromboaortopathy (Takayasu's disease). *Circulation*, **57**, 27.

James, T. N. (1974). De Subitaneis Mortibus. VIII. Coronary arteries and conduction system in scleroderma heart disease. *Circulation*, **50**, 844–56.

James, D. G. and Thomson, A. (1982). Recognition of the diverse cardiovascular manifestations in Behcet's disease. *American Heart Journal*, **105**, 457–8.

John, J. T., Jr., Hough, A., and Sergent, J. S. (1979). Pericardial disease in rheumatoid arthritis. *American Journal of Medicine*, **66**, 385–9.

Kyle, R. A. and Bayrd, E. D. (1975). Amyloidosis: review of 236 cases. *Medicine*, **54**, 271.

Leier, C. V., Call, T. D., Fulkerson, P. K., and Wooley, C. F. (1980). The spectrum of cardiac defects in the Ehlers–Danlos syndrome, type I and II. *Annals of Internal Medicine*, **92**, 171–8.

Lie, J. T. (1987). Coronary vasculitis: a review in the current scheme of classification of vasculitis. *Archives of Pathology and Laboratory Medicine*, **111**, 224–33.

Madias, J. E. and Hood, W. B., Jr. (1982). Effects of methylprednisolone on the ischemic damage in patients with acute myocardial infarction. *Circulation*, **65**, 1106–13.

Mandel, B. F. (1987). Cardiovascular involvement in systemic lupus erythematosus. *Seminars in Arthritis and Rheumatism*, **17**, 126–41.

Marsalese, D. L. *et al*. (1989). Marfan's syndrome: natural history and long-term follow-up of cardiovascular involvement. *Journal of the American College of Cardiology*, **14**, 422–8.

Mitnick, A. J., Tunick, P. A., Rotterdam, H., and Esposito, R. (1990). Antemortem diagnosis of giant cell aortitis. *Journal of Rheumatology*, **17**, 708–11.

Morris, P. B., Linder, M. J., Heinsimer, J. A., Alatky, M. A., and Reimer, K. A. (1986). Rheumatoid arthritis and coronary arteritis. *American Journal of Cardiology*, **57**, 689–90.

Nagyhegyi, G. *et al*. (1988). Cardiac and cardiopulmonary disorders in patients with ankylosing spondylitis and rheumatoid arthritis. *Clinical and Experimental Rheumatology*, **6**, 17–26.

Nair, C. K., Goli-Rijanki, R., Lyckholm, L., and Sketch, M. H. (1988). Inferior myocardial infarction complicated by mural thrombus and systemic embolization despite anticoagulation in progressive systemic sclerosis with normal coronary arteriograms. *American Heart Journal*, **116**, 1357–60.

O'Neill, T. W. (1992). The heart in ankylosing spondylitis. *Annals of the Rheumatic Diseases*, **65**, 705–6.

Owens, G. R. and Follansbee, W. P. (1987). Cardiopulmonary manifestations of systemic sclerosis. *Chest*, **91**, 118–27.

Paronen, I. (1948). Reiter's disease: a study of 344 cases observed in Finland. *Acta Medica Scandinavica*, **131** (Suppl. 212), 1–112.

Pizzarello, R. A. and Goldberg, J. (1985). The heart in rheumatoid arthritis. In *Rheumatoid arthritis. etiology, diagnosis, management* (eds. P. D. Utsinger, N. J. Zvaifler, and G. E. Ehrlich), pp. 431–40. Lippincott, Philadelphia.

Przybojewski, J. Z. (1981). Polyarteritis nodosa in the adult. Report of a case with repeated myocardial infarction and a review of cardiac involvement. *South African Medical Journal*, **60**, 512–18.

Pyeritz, R. E. and Wappel, M. A. (1983). Mitral valve dysfunction in the Marfan syndrome. Clinical and echocardiographic study of prevalence and natural history. *American Journal of Medicine*, **74**, 797.

Reyes, M. P. and Lerner, A. M. (1985). Coxsackie virus myocarditis—with special reference to acute and chronic effects. *Progress in Cardiovascular Diseases*, **27**, 373.

Saffitz, J. E., Sazama, K., and Roberts, W. C. (1983). Amyloidosis limited to small arteries causing angina pectoris and sudden death. *American Journal of Cardiology*, **51**, 1234.

Save Soderbergh, J., Malmvall, B. E., Andersson, R, and Bengtsson, B. A. (1986). Giant cell arteritis as a cause of death: report of nine cases. *Journal of the American Medical Association*, **255**, 493–6.

Shammas, R. L. and Movahed, A. (1993). Sarcoidosis of the heart. *Clinical Cardiology*, **16**, 462–72.

Silverman, H. W. and Pfeifer, M. P. (1987). Relation between use of anti-inflammatory agents and left ventricular free wall rupture during acute myocardial infarction. *American Journal of Cardiology*, **59**, 363.

Steere, A. C. *et al*. (1980). Lyme carditis: cardiac abnormalities of Lyme disease. *Annals of Internal Medicine*, **93**, 8–16.

Stollerman, G. H. (1988). Rheumatic and heritable connective tissue disease of the cardiovascular system. In *Heart disease*, (3rd edn) (ed. E. Braunvald), pp. 1706–33. Saunders, Philadelphia.

Suzuki, A. *et al*. (1986). Coronary arterial lesions of Kawasaki disease: cardiac catheterization findings of 1,100 cases. *Pediatric Cardiology*, **7**, 3–9.

Svantesson, H., Bjorkhem, E., and Elborgh, R. (1983). Cardiac involvement in juvenile rheumatoid arthritis. A follow-up study. *Acta Paediatrica Scandinavica*, **72**, 345–50.

Talwar, K. K. *et al*. (1988). Myocardial involvement and its response to immunosuppressive therapy in nonspecific aortoarteritis (Takayasu's disease)—a study of endomyocardial biopsy. *International Journal of Cardiology*, **21**, 323–4.

Van Albada-Kuipers, G. A., Bruijn, J. A., Weostedt, M. L., Breedveld, F. C., and Eulderink F. (1986). Coronary arteries complicating rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **45**, 963–5.

Vivaldi, M. T., Eyre, D. R., Kloner, R. A., and Schoen, F. J. (1987). Effects of methylprednisolone on collagen biosynthesis in healing acute myocardial infarction. *American Journal of Cardiology*, **60**, 424–5.

1.3.5 The respiratory system

Stephen G. Spiro and David A. Isenberg

Introduction

The lungs and chest wall are commonly affected by several of the rheumatological conditions. The involvement may be a direct manifestation of disease or it may be indirect as a result or complication of treatment. This chapter will detail the common symptoms of pulmonary diseases and show how often these are implicated in rheumatological conditions. The role of radiography and physiological tests in the assessment of these conditions will then be described. Finally, there is advice on which investigations should be carried out, together with a discussion on the treatment of pulmonary infections related to rheumatological disorders. A synopsis of how these disorders can affect the different respiratory structures is given in Table 1.

Rheumatoid arthritis	Bronchitis Bronchiectasis Obliterative bronchiolitis	Pneumonia Fibrosing alveolitis Rheumatoid nodules	Hypertension	Pleurisy Effusion Empyema	
Systemic lupus erythematosus		Pneumonia Fibrosing alveolitis Rheumatoid nodules	Hypertension	Pleurisy Effusion	Shrinking lungs with high diaphragms
Systemic sclerosis	Bronchiectasis	Fibrosing alveolitis Aspiration pneumonia	Hypertension		'Crossed' chest
Sjögren's syndrome	Bronchitis	Fibrosing alveolitis Lymphoma			
Dermatomyositis		Aspiration pneumonia			Myopathy
Polymyositis		Fibrosing alveolitis			
Ankylosing spondylitis		Upper lobe fibrosis			Costovertebral joint fixation
Behçet's syndrome		Haemorrhage	Aneurysm		
Pulmonary vasculitides		Nodules			
Relapsing polychondritis	Upper airway narrowing				

Table 1 Respiratory associations of the rheumatic disorders

Respiratory symptoms

Cough

Most rheumatological conditions that affect the lungs cause cough (Fig. 1). The airways can be involved in a variety of ways. It was once considered that asthma was more common in patients with rheumatoid arthritis, but there now seems no particular association except that of two relatively common conditions. Chronic bronchitis is likely to occur in patients who smoke.

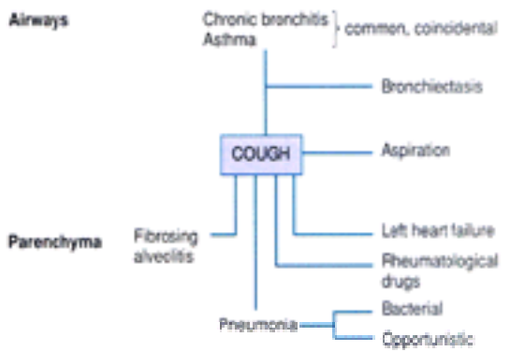


Fig. 1 The causes of cough.

Asthmatics will complain of waking during the night to cough, usually with a tight or wheezy chest. They will often be worse early in the morning and after exercise, exposure to cold air, pungent smells, and other chemical irritants. The diagnosis is seldom obscure and responds promptly to inhaled b_2-sympathomimetic agonists. In adults, asthma is associated more with cough and mucus (often yellow due to an increase in eosinophils) rather than wheezy breathlessness. The patient with chronic bronchitis will, by definition, have morning cough and sputum for 3 consecutive months for at least 2 years. He or she will be prone to acute exacerbations with infection, producing larger than usual quantities of yellow or green sputum, and needs antibiotics for control. Gradually deteriorating exercise tolerance will follow,

especially if smoking does not cease.

Chronic bronchial sepsis (or bronchiectasis) with persistent cough and the production of more than 30 ml of purulent sputum a day is found in some patients with rheumatoid arthritis. There is an increased incidence of bronchiectasis in these patients and the typical radiographic features are usually confined to just one or two lobes. Occasionally, cough with purulent sputum is an immunological reaction to the underlying condition, usually rheumatoid arthritis, but also occurs with inflammatory bowel disease. In these patients the symptoms respond dramatically to corticosteroids and not just antibiotics.

The development of parenchymal disease is almost universally associated with a persistent, dry, irritating cough, worse with exercise.

Fibrosing alveolitis occurs in association with several of the rheumatic disorders including rheumatoid arthritis, systemic sclerosis, Sjögren's syndrome, and systemic lupus (Table 1). There are no clinical, physiological, radiological, or histopathological differences between 'lone' cryptogenic fibrosing alveolitis (with or without associated rheumatoid factor) and the alveolitis in patients with rheumatic diseases. The pathogenesis is uncertain—it is possible that the alveolitis develops as a reaction to immune complexes deposited in the pulmonary capillaries. Limitations on exercise imposed by the joint disease may explain why some of these patients present relatively late with their respiratory symptoms, although florid fibrosing alveolitis with positive rheumatoid factor can herald the onset of rheumatoid arthritis and precede the joint manifestations by months or years. Usually, however, it is the patients with more severe rheumatoid arthritis who develop pulmonary fibrosis. In addition to cough, other clinical features include finger clubbing in up to 70 per cent of patients, bilateral inspiratory basal crackles, and, late in the disease, cyanosis.

The physiological and radiological features of fibrosing alveolitis are described below. However, the symptoms of cough together with infiltrates on the chest radiograph may also indicate a drug-induced pneumonitis (e.g. gold, cyclophosphamide, methotrexate), pneumonia (common in patients with chronic diseases, especially if also on corticosteroid or other immunosuppressive therapy), left ventricular failure (more likely on steroids or some non-steroidal anti-inflammatory drugs; and in old age), and opportunistic infections associated with the primary autoimmune rheumatic disorder or the consequence of its therapy.

The treatment of cough can be difficult but is helped greatly if the cause is identified. Asthma responds promptly and well to inhaled steroids with additional b_2-agonists when necessary. Chronic bronchitis improves on stopping smoking, and most exacerbations respond rapidly to conventional oral antibiotics. The treatment of fibrosing alveolitis is unsatisfactory, although initially oral corticosteroids improve the cough, but may be withdrawn if lung function does not improve in order to prevent long-term steroid side-effects. Cough suppressants such as pholcodeine or codeine linctus in sufficiently large doses are a boon at night. n the most refractory cases, nebulized bupivacaine local anaesthetic solution (0.03 per cent) can be helpful for a few hours, but patients risk aspiration and must avoid eating when their larynx is anaesthetized.

Cough, especially in association with purulent secretions and radiographic pulmonary infiltrations, can be caused by aspiration of stomach and oesophageal contents in systemic sclerosis. The high incidence of oesophageal reflux, especially when the patient is supine, can cause major problems with pulmonary infection, particularly by anaerobic organisms.

Breathlessness

Breathlessness is a common feature of most pulmonary conditions and causes associated with connective tissue disorders can be listed (Fig. 2).

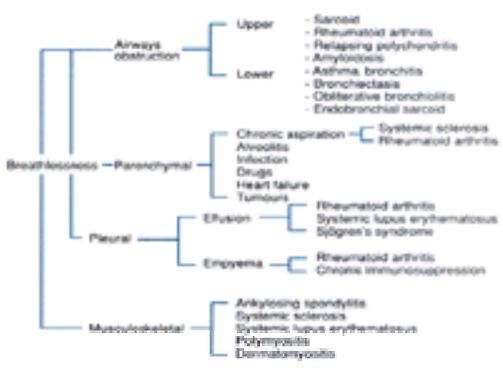


Fig. 2 The causes of breathlessness.

Airways

Asthma, chronic bronchitis, and bronchiectasis, if sufficiently severe, will result in breathlessness on exertion and in asthmatics at rest as well.

Upper airways obstruction will, because of its dominant effect on airways resistance, readily cause dyspnoea. Collapse of the cartilaginous support of the trachea and main bronchi in relapsing polychondritis can cause great difficulty with both inspiration and expiration and will be associated with loud, harsh wheezing and dyspnoea. This rare syndrome is often also associated with a deep, bovine cough. Endobronchial sarcoidosis or laryngeal infiltration by sarcoid can result in severe upper airways obstruction and breathlessness, particularly on exertion. Minor degrees of upper airways obstruction have been reported in up to 18 per cent of autopsies in patients with rheumatoid arthritis. Synovitis of the cricoarytenoid joint may cause supraglottic narrowing and also hoarseness and breathlessness.

Obliterative bronchiolitis has been recognized more recently to complicate rheumatoid arthritis. It is rare and can be rapidly fatal, although in most, whilst the condition seems to be of sudden onset and rapidly progressive, it becomes stable leaving the patient to suffer severe exertional dyspnoea with grossly reduced lung function. Pathologically the small airways (less than 3 mm in diameter) are narrowed or obliterated by a chronic inflammatory process with fibrotic scar formation. Wheezing is not a symptom but on examination high-pitched, end-inspiratory, basal squeaks are often readily audible. The condition responds poorly to corticosteroid or other therapy.

Parenchymal disease

These complications include fibrosing alveolitis, pulmonary infections, aspiration, drug-induced pneumonitis, heart failure, and acute manifestations such as acute lupus pneumonia. The history and the radiographic appearances are sometimes helpful but often frustratingly non-specific. The history may be of more importance, for example drug history, rate of onset of symptoms, and chest pain. Often a diagnosis is made only on lung biopsy (see below).

Sarcoidosis is often associated with widespread pulmonary infiltration, predominantly in the mid and upper zones. However, this is rarely accompanied by physical signs or breathlessness. Late progressive fibrotic changes in sarcoidosis cause shrinkage of the upper lobes with subsequent distortion of the lower lobes, and dyspnoea can develop.

Pleural disease

Pleural effusion

This is the most common respiratory sign in patients with rheumatoid arthritis. It is regarded generally as the only complication of rheumatoid arthritis to be more common in men. It is usually asymptomatic, although there can be pleuritic pain, dyspnoea, and fever. Although the effusions are usually small and unilateral, they can be large and bilateral. Large pleural effusions with dyspnoea are also occasionally seen in systemic lupus, although classically the effusions are small. It is more common in lupus to obtain a history of pleuritic chest pain and breathlessness with no visible effusion. Aspiration of an effusion is required usually to exclude infection or malignancy. The fluid is characteristically clear coloured, an exudate (protein greater than 30 g/l) with, in rheumatoid arthritis, a very low or unrecordable glucose concentration. The concentrations of lactate dehydrogenase and cholesterol are raised. Rheumatoid factor is detected frequently in the fluid. Pleural biopsy is rarely

helpful, with non-specific features of chronic inflammation only. If the effusions fail to resolve, or relapse after drainage, oral corticosteroids may be rapidly effective. It is rare to have to resort to surgical pleurectomy.

Empyema

Empyema is relatively common in patients with rheumatoid arthritis. This usually is secondary to long periods of corticosteroid therapy, but is seen less often in other autoimmune rheumatic disorders that are similarly treated. Breathlessness is common, as is a fever with malaise and weight loss. The diagnosis is established on aspiration of turbid fluid containing pus cells. The fluid should be cultured for aerobic and anaerobic organisms as well as tuberculosis. The empyema should be treated intensively, with local drainage as well as systemic antibiotics. Rib resection should be done early to facilitate drainage if intercostal tube drainage is not effective. A decortication may be necessary when the infection has resolved as a thick residual pleural cortex can cause severe restriction of movement of the underlying lung. Ideally, surgical decortication should be done at least 6 weeks after the infection has resolved, thus allowing the underlying lung tissue to heal and making surgery easier and less traumatic.

Chest wall disease

Breathlessness can be caused by the patient's inability to ventilate the lungs adequately. Gross restriction of the chest wall can occur in advanced systemic sclerosis and in ankylosing spondylosis. In the presence of normal lung fields on a chest radiograph, musculoskeletal causes of dyspnoea are often overlooked and psychogenic reasons may have to be considered. However, clinical examination of the patient and appropriate lung function tests should leave the physician in little doubt as to the correct cause.

Weakness of the respiratory muscles is often insidious and difficult to detect unless there is obvious wasting of other muscle groups such as in the upper limb girdle. Again, lung function tests are extremely helpful but the physician should be alert to the possibility of muscle weakness. In systemic lupus the contractile properties of the diaphragm are sometimes impaired, allowing it to keep its curved shape and resulting in lungs that look small radiographically on full inspiration. In polymyositis and dermatomyositis the well-recognized proximal myopathy often extends to involve the respiratory muscles.

Rarer causes of dyspnoea

These include pulmonary haemorrhage in Behçet's syndrome, usually due to a leaking aneurysm of a pulmonary artery, or bleeding from a pulmonary infiltration subsequent to the pneumonitis occasionally seen in this condition.

Tumours can develop in some autoimmune rheumatic disorders, particularly rheumatoid arthritis and Sjögren's syndrome, and less frequently dermatomyositis. Lymphoma can develop as a result of long-term immunosuppressive therapy.

Radiological aspects

There is a plethora of pulmonary abnormalities providing a large differential diagnosis for radiographic shadowing in association with connective tissue diseases (Fig. 3). Some of these are detailed next.

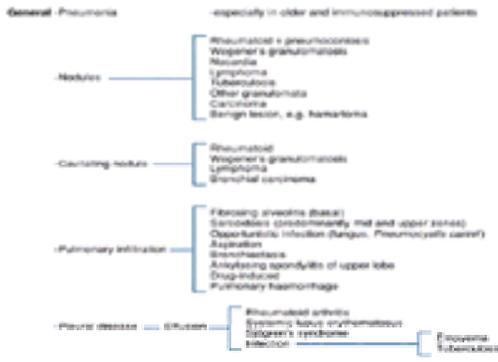


Fig. 3 The radiological abnormalities of rheumatic diseases.

Pulmonary nodules

Although the most common cause of a round lesion on the chest radiograph is a bronchial carcinoma and the next most common a benign lesion such as a hamartoma, there are several rheumatological causes or associations with such a lesion.

In rheumatoid arthritis, nodules histologically identical to subcutaneous necrobiotic nodules can occur. They tend to be peripheral, and can increase in size, remain static, or disappear (Fig. 4). They can also cavitate and they can occasionally cause a pneumothorax. When single, the differential diagnosis will include another benign lesion, a bronchogenic carcinoma, tuberculosis, or another granuloma (Fig. 3). The diagnosis is often elusive, although occasionally fine-needle aspiration of the lesion yields granulation tissue suggestive of a necrobiotic nodule.

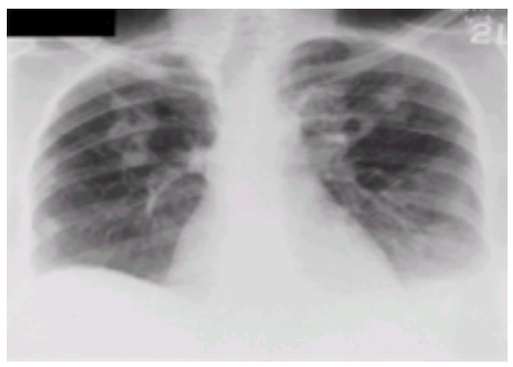


Fig. 4 Necrobiotic nodules with bilateral effusions (the right is subpulmonary) in a case of rheumatoid arthritis.

Caplan's syndrome occurs when simple coal-workers' pneumoconiosis, or another pneumoconiosis, is complicated by the development of rheumatoid nodules, which are visible on the chest radiograph. These often arise in crops and may sometimes enlarge rapidly.

Nodular shadowing can occur in Wegener's granulomatosis, often as the only abnormality associated with malaise, fever, and a high erythrocyte sedimentation rate. The antineutrophil cytoplasmic antibody test is usually positive. The nodules can enlarge rapidly, become multiple, and cavitate producing relatively large, thin-walled lesions (Fig. 5). Diagnosis by fibreoptic transbronchial lung biopsy is unusual as a moderately large artery is required to provide evidence of vasculitis. Open lung biopsy is usually necessary if tissue for diagnosis is required.

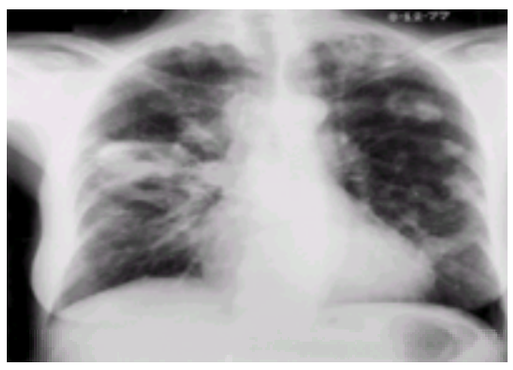


Fig. 5 Multiple nodules of Wegener's granulomatoisis; some are cavitating.

Nocardia, a variety of the Actinomycetaceae, can develop as an opportunistic infection in patients on prolonged immunosuppressive therapy. They form single or multiple nodules, which may cavitate. Usually extrapulmonary lesions soon develop, as for example in joints and the subcutaneous tissues.

It is often impossible to distinguish a malignant from a benign nodule. This is particularly the case for pulmonary metastases, which are often spherical with smooth margins. If they are multiple and a range of sizes, then malignant disease is probable.

The classical malignant primary tumour will have an irregular margin with 'sun ray' spicules radiating out into the surrounding parenchyma. The presence of flecks of calcium cannot guarantee benign disease. Computed tomography (**CT**) often identifies subpleural nodules invisible on routine chest radiographs. The likelihood of malignant origin will increase if the nodules are greater than 10 mm in diameter. Most subpleural nodules are granulomata, that is old tuberculosis or sarcoid, and in the United States may often be due to fungal disease such as histoplasmosis or coccidioidomycosis.

Hyperinflated lungs

These occur only in some cases of obliterative bronchiolitis, unless there is emphysema as an independent feature in a patient who also happens to have a connective tissue disorder. The lung fields in obliterative bronchiolitis are clear, with no other specific features (Fig. 6). They are, however, often of normal size and appearance.

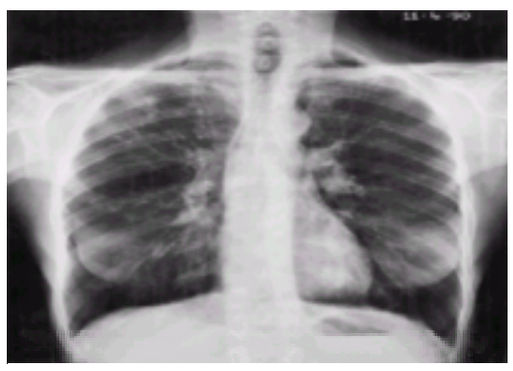


Fig. 6 The lungs are markedly overinflated, and there is a reduction in the vascular markings; these are non-specific findings but consistent with obliterative bronchiolitis. There is a necrobiotic nodule in the left lower lobe, behind the heart.

Small lungs

This appearance occurs in lupus; bilateral diaphragmatic weakness prevents the diaphragms from shortening and 'flattening' on inspiration, and both diaphragms remain elevated and deeply domed. There may be linear atelectatic shadows in the lower lobes, particularly above the diaphragm as a consequence of the basal hypoventilation. There is an apparent mediastinal widening in these patients but only as a consequence of the volume loss and not due to mediastinal disease.

Parenchymal shadowing

Diffuse, predominantly basal, reticulonodular shadowing, possibly with overall reduction in lung volume, is typical of fibrosing alveolitis. This occurs in association with rheumatoid arthritis, systemic lupus, systemic sclerosis, and Sjögren's syndrome (Fig. 7 and Fig. 8). The appearances are non-specific and identical to those cases of lone fibrosing alveolitis. Interstitial lung disease in rheumatoid arthritis correlates with the severity of joint disease, ranging from an incidence of just 1 per cent in mild rheumatoid arthritis to a 30 to 40 per cent association in patients with severe joint deformity. The interstitial infiltrates are depicted radiographically as small, irregular shadows or opacities, predominantly at the lung bases. These shadows are also termed reticulonodular densities, ground-glass densities, or increased interstitial markings. Eventually the infiltrates extend to the upper lung fields. They can occasionally respond remarkably well to treatment with corticosteroids (Fig. 9). If the interstitial fibrosis progresses, it can develop areas of honeycombing comprising round or polygonal air cysts with well-defined walls and measuring less than 5 mm in diameter. The cysts may eventually become confluent, forming large destructive cavities in the subpleural regions. Rupture can cause pneumothorax.

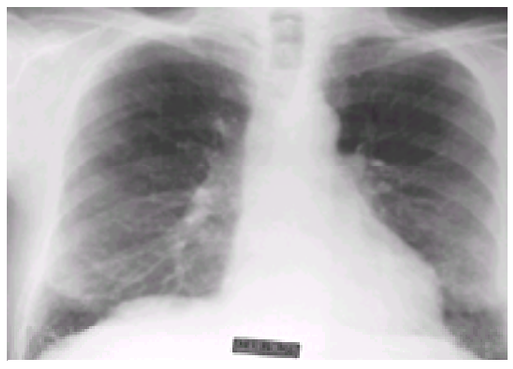


Fig. 7 Typical appearances of established fibrosing alveolitis with a reticulonodular pattern concentrated mainly in the lower lobes.

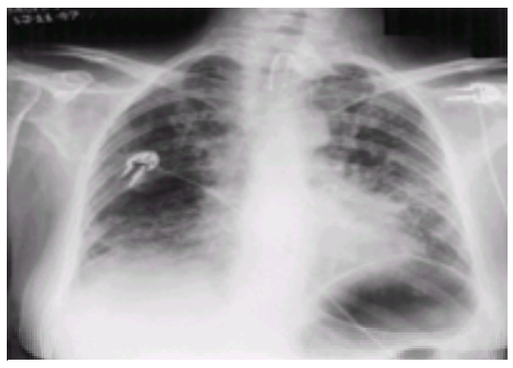


Fig. 8 Widespread coarse reticular shadowing mainly in the lower zones with associated loss of volume in a patient with rheumatoid arthritis. Note erosion of right humeral head. Tracheostomy in place.

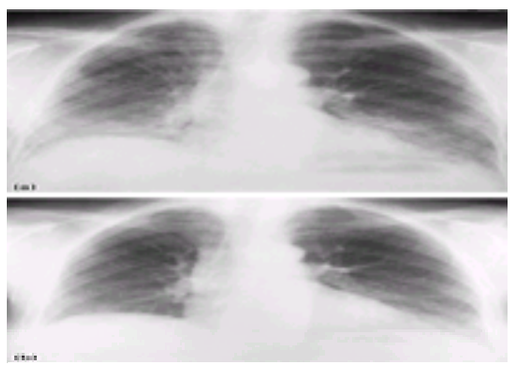


Fig. 9 (a) Reticulonodular and confluent shadowing at the lung bases; (b) 8 months later, almost complete resolution after corticosteroid treatment.

High-resolution CT scanning provides an 'early warning' means of assessing the lungs. In a study of 17 patients with early scleroderma it was shown that 15 (88 per cent) had a variety of abnormalities, especially subpleural cysts and honeycombing (Warrick *et al*. 1991). In contrast the plain chest radiograph was abnormal in only 10 cases (59 per cent). Similarly changes consistent with rheumatoid-associated lung disease, notably interstitial fibrosis, were detected in patients with rheumatoid arthritis at a time when the chest radiograph was often normal (Fewins *et al*. 1991). This form of scanning also facilitates the precise distribution of alveolitis and the extent of any lung fibrosis and destruction (Fig. 10 and Fig. 11).

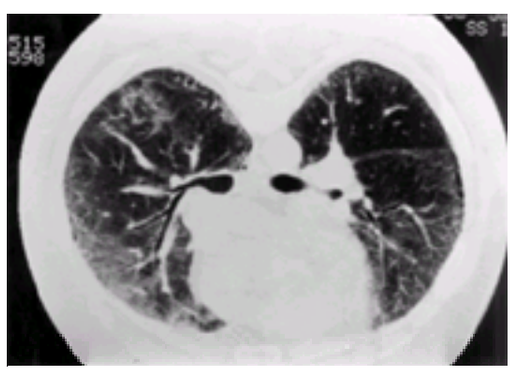


Fig. 10 Prone, high-resolution CT at the level of the carina. Diffuse parenchymal opacification in a subpleural distribution. This pattern is seen during the desquamative phase of fibrosing alveolitis.

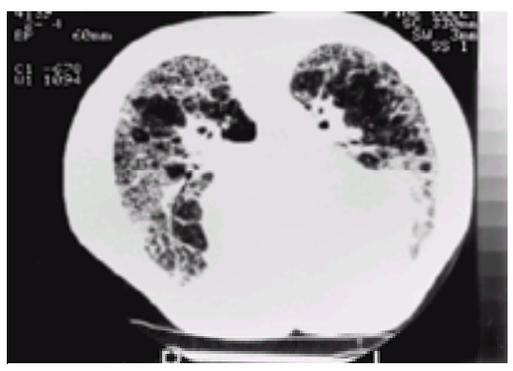


Fig. 11 Prone, high-resolution CT through the lower lobes. There is a reticular pattern in a subpleural distribution representing established fibrosis in a case of systemic sclerosis.

Other causes of diffuse pulmonary infiltrates include:

1. Sarcoidosis. The shadowing is usually more nodular and spares the lung bases. It is often associated with bilateral hilar lymphadenopathy, few symptoms, and commonly no decrease in lung size.
2. Infection. Opportunistic infection with fungi can also cause diffuse shadowing. However, the lesions are often more nodular and, although usually affecting more than one lobe, are often focal.
3. Polyarteritides. In polyarteritis nodosa a peripheral, soft, ill-defined, predominantly apical and/or midzone shadowing can develop, usually in association with an elevated peripheral eosinophil count. This 'pulmonary eosinophilia' can also be associated with pleural or pericardial effusions and is very responsive to corticosteroid therapy. It is similar to a hypersensitivity drug reaction within the lung such as occurs with gold salts (Fig. 12).

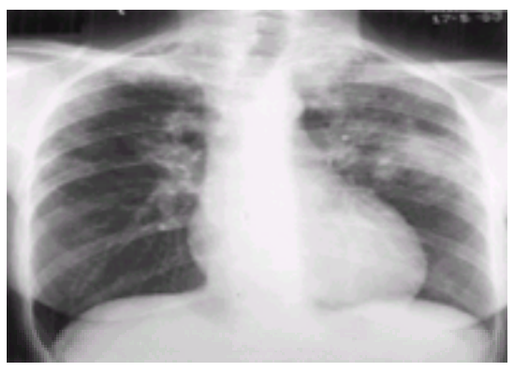


Fig. 12 Mild upper-zone consolidation, predominantly peripheral on the right, typical of pulmonary eosinophilia.

4. Upper lobe fibrosis is seen in ankylosing spondylosis, with contracture of the upper lobes and upward traction of the hilas. Eventually the upper lobes may cavitate. Apical pleural thickening may also develop and be extensive (Fig. 13).

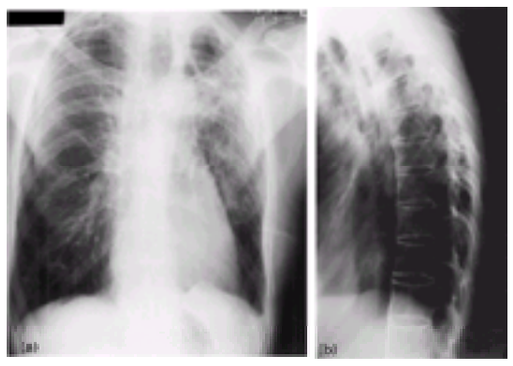


Fig. 13 (a) Linear shadows in upper zones with marked elevation of the hila consistent with upper-zone fibrosis in a case of ankylosing spondylosis. There is fusion of the costotransverse joints. (b) Lateral dorsal spine shadowing anterior syndesmophytes typical of ankylosing spondylosis.

5. Cavitation in the upper lobes, and to a lesser extent the mid zones, amongst florid scarring is also seen in sarcoidosis. This progressive course of sarcoid usually develops after the bilateral hilar lymphadenopathy has regressed. In both sarcoidosis and ankylosing spondylosis the cavities can become colonized by *Aspergillus* spp. to form mycetomata. These semisolid masses appear to 'sit' in the cavity surrounded by a halo of air, a characteristic appearance. Serum aspergillal precipitins will be strongly positive.
6. Focal intrapulmonary shadowing can occur after aspiration of gastric contents. Classically, aspiration happens when the subject is lying prone and the inhaled contents run posteriorly to the posterior segment of the upper lobes or apical segment of the lower lobes. Consolidation will be concurrent with the symptoms and signs of infection. Aspiration of anaerobic organisms is frequent and can cause pulmonary cavities. However, there can also be aspiration into basal segments, especially if it is associated with gross dysphagia as in advanced systemic sclerosis.
7. Bronchiectasis is recognized by peribronchial thickening causing 'tram-lines' on the chest radiograph due to a widened lumen and the thickened bronchial walls. Cystic changes can also be detected in cases where the damage is extensive. The lesions are patchy but usually confined to the middle and lower lobes.

Rare causes of pulmonary infiltrations will include acute exacerbations of lupus, a lymphoid tumour in, for example, Sjögren's syndrome, and intrapulmonary haemorrhage in a vascular disorder such as polyarteritis, Wegener's granulomatosis, or Behçet's syndrome.

Pleural disease

Pleural thickening is very unusual in autoimmune rheumatic disorders but pleural effusions are moderately common, especially in rheumatoid arthritis. They also occur in lupus and Sjögren's syndrome, and can reach considerable size.

Pulmonary function tests

Pulmonary function tests are not diagnostic but provide invaluable information as to the pattern and extent of an abnormality. They should be carried out early in any patient with pulmonary complications so that a baseline can be established and followed in order to evaluate the response to therapy, or assist in making decisions when to start therapy.

Spirometry

This is a simple test for almost any patient. It requires some co-operation for maximal inspiration and forced expiration until 'empty'.

The forced expiratory manoeuvre measures the forced expired volume in 1 s (**FEV_1**) and the forced vital capacity (**FVC**). The FEV_1/FVC ratio, expressed as a percentage, should be greater than 75 per cent, but falls with normal ageing to 65 per cent at 70 years. Figure 14 shows typical obstructive and restrictive patterns of spirometry. The obstructive pattern, with a proportionately greater decrease in FEV_1 than FVC, is found in all types of airflow obstruction. In pulmonary infiltrative conditions with or without fibrosis the FEV_1 is preserved relatively (although reduced) compared to the FVC. This results in an FEV_1/FVC ratio of greater than 80 per cent, but with reduced absolute values, the extent depending on the severity of the lesion.

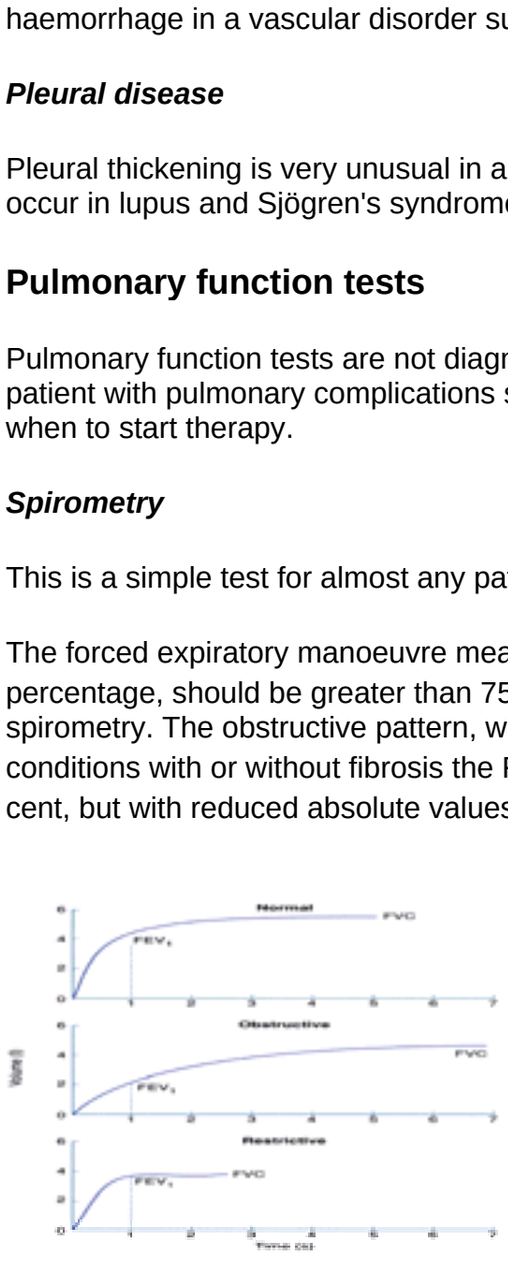


Fig. 14 Normal, obstructive, and restrictive patterns of spirometry.

In patients with stiff chest walls or weak muscles the pattern of spirometry is also restrictive, with reduction in both FEV_1 and FVC. The FEV_1/FVC per cent is not always increased as the elastic recoil pressures within the lung remain normal in contrast to parenchymal diseases.

Peak expiratory flow (PEF)

This is the most commonly tested lung function and yet is not particularly sensitive. The PEF is the instantaneous flow rate during the first 10 ms of expiration after a maximum inspiration. It is a measure of airflow obstruction, dominated by flow from the upper airways. The PEF is of little value in restrictive pulmonary conditions as the airways are normal. However, where upper airways obstruction is suspected the PEF is reduced severely. An indication that there may be obstruction in the trachea, carina, or main bronchi can be deduced if the ratio of FEV_1 (ml) to the PEF (l/min) is greater than 10.

Flow volume curves

The envelope of expiratory flow from total lung capacity (**TLC**) to residual volume (**RV**) and then back in again to TLC can be simply captured as flow versus volume (Fig. 15). This provides a sensitive measure of peak flow and expiratory flow. Flows along subsections of the vital capacity—usually at 25 and 50 per cent above the residual volume (V_{25}:V_{50}) are measured and compared to age- and sex-matched normal values. These flow rates represent flow in the small airways and are reduced especially when generalized obstruction predominates. In rheumatological medicine a good example is found in obliterative bronchiolitis (Fig. 15(b)) where flow at V_{50} and V_{25} is reduced especially compared to the FEV_1 and PEF when expressed as the percentage predicted. Pressure-dependent collapse (as in emphysema), that is decreased recoil causing collapse of major airways on forced expiration (Fig. 15(d)), is especially prominent in relapsing polychondritis where the trachea collapses during expiration. Flow volume curves give little information in patients with parenchymal disease.

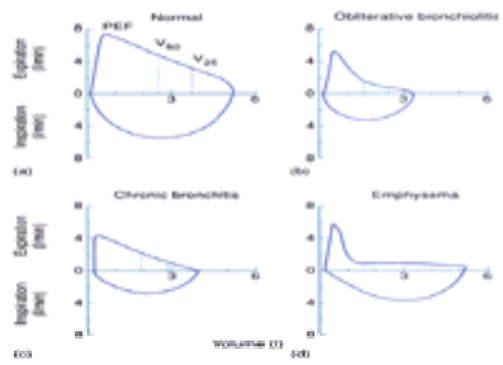


Fig. 15 (a) A normal flow volume curve. During expiration flow reaches an immediate peak (peak expiratory flow) and gradually falls to zero at residual volume. Flow rates at 50 and 25 per cent of the vital capacity are shown. (b) Reduced flows and small vital capacity in obliterative bronchiolitis. (c) 'Volume-dependent' reduction in expiratory flows in chronic bronchitis. (d) 'Pressure-dependent' reduction in expiratory flow in emphysema.

The flow volume loop is also particularly useful for detecting upper airways obstruction. The site of the lesion and whether it is fixed or variable will affect the shape of the flow volume loop (Fig. 16).

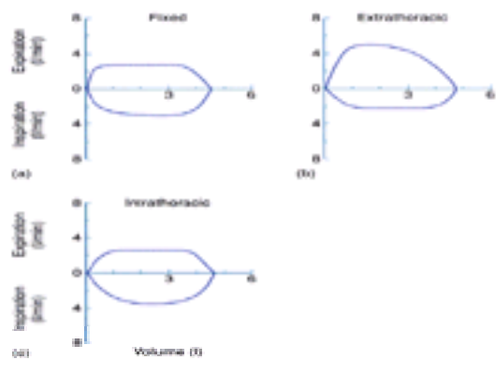


Fig. 16 Upper airways obstruction: (a) fixed airflow obstruction; (b) extrathoracic airflow obstruction, affecting inspiratory flow more than expiratory flow; (c) intrathoracic airflow obstruction, affecting expiratory flow more than inspiratory flow.

Total respiratory pressures

Classically patients with weakness of the respiratory muscles will show reduced spirometric and PEF values. However, this is not diagnostic as it fails to differentiate between muscle weakness and reduced lung volumes secondary to parenchymal disease. The strength of the respiratory muscles is best and simply measured by the maximum inspiratory and expiratory mouth pressures.

If pure diaphragmatic weakness or paralysis is suspected, then spirometry done with the patient first lying and then standing can be most useful. A drop in vital capacity by up to 40 per cent can occur when lying flat if the diaphragm is non-functioning. This is due to the abdominal contents moving towards the chest unopposed when the patient lies flat. This is a remarkably reproducible test.

Thoracic lung volumes

Measurement of TLC and RV in patients with restrictive lung disease is very important before treatment and should be closely followed during therapy. Lung volumes can be measured either by the wash-out (or wash-in) method of using an inert gas as by body plethysmography. In restrictive conditions both RV and TLC will be reduced. However, in patients with muscle weakness, RV may be reduced but TLC can be reasonably normal. Figure 17 illustrates these points.

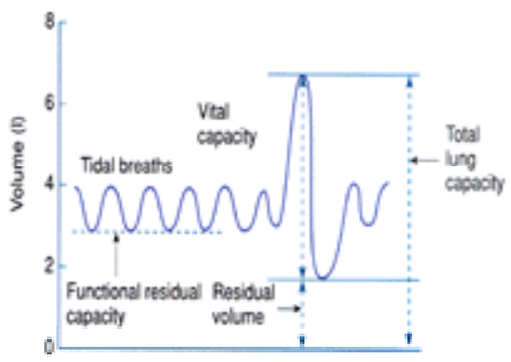


Fig. 17 The thoracic lung volume subdivisions.

Diffusing capacity

The magnitude of carbon monoxide gas transfer (*TL*CO) depends on the volume of the pulmonary capillary bed and the matching between the distribution of ventilation and perfusion within the lungs. The test is done with the subject inspiring a known concentration of carbon monoxide and of an inert gas, helium. As the helium is not absorbed, the change in its concentration in the expired sample will provide a measure of the volume of alveolar gas into which the gas mixture was inspired (alveolar volume, V_A). If the quality of uptake of carbon monoxide per unit of lung is desired, then dividing the *TL*CO by the V_A will provide the carbon monoxide transfer coefficient (*K*CO), a quantitative assessment of the efficacy of uptake by the lung and alveolar capillaries of carbon monoxide per litre of lung volume.

The *TL*CO is reduced when:

1. The alveolar capillaries are reduced in number, for example in emphysema, pulmonary fibrosis, pulmonary embolic disease, and pulmonary hypertension.
2. There is an increased ventilation/perfusion mismatch as occurs in pneumonia, infection, and infiltrates, and sometimes in pulmonary oedema.
3. There is a reduction in haemoglobin content in the blood.
4. The patient is a heavy smoker with high levels of carboxyhaemoglobin.

The *TL*CO is increased when:

1. There is a redistribution of ventilation/perfusion ratios enabling more capillary blood to come into contact with inspired gas as sometimes happens in asthma.
2. There is alveolar haemorrhage increasing the amount of haemoglobin available to carbon monoxide in the alveoli.
3. Polycythaemia increases the level of available haemoglobin in the capillaries.

Walking tests

Assessment of the functional ability of an individual in a manner that uses a normal, habitual activity is very useful. In patients with joint disease, formal treadmill or cycle ergometric exercise testing is often not possible. However, a walking test may be feasible. The subject is asked to walk up and down a corridor of known length or other suitable enclosed space for 2, 6, or 12 min watched by a technician who measures the distance walked and uses standard terms of encouragement at regular intervals. After a practice walk, a second walk distance is recorded. This is often up to 30 per cent longer than the initial distance walked, as the patient develops confidence in the test and his or her ability. Extra information can be obtained, for example, by attaching a portable pulse oximeter to the finger or ear to measure oxygen saturation during the walk. The test is highly reproducible after the test practice walk.

Differences in pulmonary function in the context of rheumatological diseases are summarized in Fig. 18. As indicated earlier, serial performance of these tests provides a useful way of assessing response to therapy. Thus, Akesson and colleagues demonstrated that cyclophosphamides increased vital capacity and static lung compliance in 18 patients with scleroderma who had lung disease (Akesson *et al*. 1994).

Fig. 18 A summary of pulmonary function in rheumatological diseases. PEFR, peak expiratory flow rate; FEV_1, forced expiratory volume in 1 s; FVC, forced vital capacity; TLC, total lung capacity; RV, residual volume; FRC, functional residual capacity; P_1 max, maximal inspiratory mouth pressure; P_E max, maximal expiratory mouth pressure.

Bronchoscopy, lavage, biopsy, and other diagnostic tests

Lung biopsy

Open lung biopsy remains the best method for providing tissue for histopathological diagnosis of chronic infiltrative lung diseases, and a review of several series gave a diagnosis in 94 per cent of cases (Gaensler and Carrington 1980). However, this procedure is by no means routinely performed, particularly when the association between the joint disease and the presence of pulmonary infiltration is both typical and as expected. The ultimate answer will be the response to any attempted therapy. A recent survey in the United Kingdom reviewed the case notes of 200 patients with cryptogenic fibrosing alveolitis, of whom 119 cases had only a clinical diagnosis—based on typical physical signs. Only 15 patients underwent an open lung biopsy, 35 bronchoalveolar lavage, and 66 transbronchial biopsy (Johnson *et al*. 1993).

If open lung biopsy is done, it is usually through a minithoracotomy and at least two areas of lung are sampled. In general the right middle lobe and the lingula are avoided as it is claimed by some that fibrosis occurs more readily in these areas. Similarly, there is little value in removing a piece of grossly fibrotic lung as it will be histologically less helpful. A moderately involved and an apparently uninvolved piece of lung should be sampled. The alternative to open lung biopsy is a transbronchial lung biopsy via the fibreoptic bronchoscope. While this is usually simple, with few complications (e.g. pneumothorax), the biopsies suffer from being very small and probably are not representative of a diffuse, chronic process. Four to six biopsies are recommended, all from one lung to avoid the risk of bilateral pneumothoraces, but from more than one segment whenever possible. Non-specific reactions and non-diagnostic microscopic appearances may be misleading (Wall *et al*. 1981).

Computed tomography

In patients who present with interstitial infiltration together with an obvious autoimmune rheumatic disease the need to obtain pulmonary histology is less urgent or potentially helpful. High resolution CT has been shown to be a good predictor of lung histology in systemic sclerosis, and can discriminate between fibrotic and inflammatory histology (Wells *et al*. 1992). Furthermore, high resolution CT in patients with systemic sclerosis can identify different patterns of the disease; that is ground-glass pattern, mixed, or predominantly reticular patterns—the latter being the most abnormal. These patterns correlate strongly with lung function, especially gas transfer (*TL*CO). Responses to therapy were seen most commonly in patients with a ground-glass appearance on high resolution CT, and least in those with a reticular CT appearance (Wells *et al*. 1993b).

Another potentially useful predictor of prognosis is the rate of clearance of technetium labelled diethylene-triamine-pentacetae (99mTc-DTPA) from the lung. Those patients with interstitial lung disease who had normal clearances had stable disease, whilst those with rapid clearances had deteriorating disease. There was no association between rates of clearance and the likelihood of responding to corticosteroid treatment (Wells *et al*. 1993a), and this was not influenced by any association of the fibrosing alveolitis with an autoimmune rheumatic disease.

Bronchoalveolar lavage

Lavage allows examination of the cellular and non-cellular elements present on the epithelial surface of the alveolar space. It does not truly sample the lung parenchyma; however, in some studies there has been good correlation between the type and the number of inflammatory cells obtained by lavage and those obtained by diagnostic lung biopsy (Reynolds 1987).

The technique does, however, have serious shortcomings both in its diagnostic value and as a tool for following disease activity. There is no satisfactory consensus on what constitutes a 'routine' analysis of fluid from bronchoalveolar lavage. Total and differential cell counts, frequently with identification of T-lymphocyte subsets by monoclonal antibody techniques, are usually made. Fluid from a non-smoker normally contains 85 to 90 per cent macrophages, 7 to 12 per cent lymphocytes, and 0.1 per cent neutrophils (Reynolds 1987). Patients may be classified as having a neutrophil alveolitis or a lymphocyte alveolitis based on an increase in the percentage of neutrophils or lymphocytes. A neutrophil alveolitis is often seen in cryptogenic fibrosing alveolitis, histiocytosis X, and in smokers; a lymphocytic alveolitis is commonly seen in sarcoidosis, hypersensitivity alveolitis, and granulomatous infections. However, there is marked overlap.

The value of bronchoalveolar lavage in the serial follow-up of lesions is not yet established with confidence. However, a study that included a subgroup of patients with rheumatoid arthritis associated with fibrosing alveolitis showed a link between a response to steroids and a decrease in neutrophils in the lavage (Turner-Warwick and Haslam 1987).

In patients suspected of pulmonary infections, especially by opportunistic organisms, bronchoalveolar lavage has a distinct role. On failing to make a diagnosis from sputum, or in the absence of sputum, a lavage via fibreoptic bronchoscopy in patients with pulmonary infiltrates can reveal fungi, *Pneumocystis carinii*, cytomegalovirus, Nocardia, tuberculosis, and other bacteria.

In patients suspected of drug-induced pulmonary shadowing, bronchoalveolar lavage may have a role but often the cell counts are non-specific. Data are too scanty to confirm the use of lavage as a diagnostic tool for drug-induced pathology.

Transbronchial lung biopsy also can provide findings suggestive of drug eruptions—with type I pneumonocyte desquamation and type II cell dysplasia. Hypersensitivity reactions may produce granulomas, interstitial eosinophilic infiltration, or granulation tissue as seen in some drug-induced organizing pneumonias.

Vascular studies

With the identification in the mid-1980s of clinical syndromes associated with the presence of antiphospholipid antibodies, much interest has focused on clotting abnormalities in the lungs of patients with underlying rheumatic disease. The widespread availability of ventilation/perfusion ratio (V/Q) scanning, CT and magnetic resonance imaging scans, and the more specialized pulmonary angiograms and digital subtraction angiography have greatly facilitated the ability to demonstrate the presence of pulmonary emboli and infarction. The association of antiphospholipid antibodies with deep venous thrombosis and thromboembolic pulmonary hypertension seems to be better documented in patients with the primary antiphospholipid antibody syndrome (Asherson *et al*. 1989) than in patients with systemic lupus erythematosus (Gulko *et al*. 1993).

Although much less common, pulmonary arterial aneurysms have been demonstrated in approximately 1 per cent of patients with Behçet's syndrome (Hamuryudan *et al*. 1994). Haemoptysis is the usual presenting symptom and the prognosis is often poor.

Pulmonary infections (Table 2)

Table 2 Pulmonary infections in rheumatology

The rheumatological conditions have a considerable association with pulmonary infection, probably because they induce host-defence defects within the immune system. In systemic lupus, many aspects of the immune system have been reported to be abnormal. These include the antibacterial activity of the alveolar macrophages, defects in the Fc receptor function of the macrophages, and diminished clearance of immune complexes by macrophages.

Defects have also been described in the chemotactic and phagocytic activity of both polymorphonuclear neutrophils and blood monocytes. Lymphopenia is common, and T lymphocytes, especially T suppressors, are decreased.

Patients with rheumatoid arthritis have a predisposition to pleuropulmonary infection. A twofold rise in pneumonia was found when patients with rheumatoid arthritis were compared to those with degenerative joint disease; deaths due to respiratory infection were four times more frequent than expected.

The drugs used to treat rheumatic diseases also adversely affect the host defence system; the complex and many effects of glucocorticoids have been described by Cupps and Fauci (Cupps and Fauci 1982). Cyclophosphamide affects virtually all components of the cellular and humoral immune response. The drug acts primarily during the S phase of the cell cycle and so all rapidly dividing cells are involved. B-lymphocyte depletion occurs early and antibody production can also be inhibited. Azathioprine also inhibits both humoral and cell-mediated immunity, and suppresses the numbers of both T and B lymphocytes. Methotrexate is a folic acid antagonist that interferes with transport of carbon fragments required for thymidine synthesis; DNA synthesis and cellular proliferation are inhibited.

Specific infections

Bacteria

Gram-negative bacteria (*Klebsiella pneumoniae, Escherichia coli*, and *Pseudomonas aeruginosa*) are frequent causes of pneumonias in patients with autoimmune rheumatic disorders. The high frequency of these infections is probably related to the ease with which Gram-negative bacteria colonize the upper respiratory tract in patients with serious or debilitating conditions.

As far as Gram-positive bacteria are concerned, the incidence of pneumococcal infections in this group of patients is probably no higher than in the general population. *Staphylococcus aureus* is a common cause of death in patients with systemic lupus (Hellman *et al*. 1987).

Infection with *Legionella* spp. is more frequent than normal in immunocompromised patients and those on immunosuppressive medication; 95 per cent will be caused by *L. pneumophilia*, and occasional cases by *L. micdadei*. The radiographic features of Legionella pneumonia will usually be indistinguishable from those of the other bacterial pneumonias. However, confusion, prostration, hyponatraemia, and disturbance of liver function are more common. Standard bacterial cultures will be unhelpful but cultures in specific media may provide a diagnosis within 2 to 7 days. Indirect immunofluorescence antibody techniques are positive in 70 to 90 per cent of cases at 10 days.

Fungal infections

Aspergillus

Aspergillus causes two types of infection in patients with autoimmune connective diseases. *Aspergillus* spp. frequently colonize cavities in patients with chronic fibrotic lung disease. In the autoimmune rheumatic diseases, these include sufferers from type III sarcoidosis and, occasionally, patients with upper lobe fibrosis secondary to ankylosing spondylosis. Usually the mycetoma remains asymptomatic but it can cause large haemoptyses. Should haemoptyses develop, resection will be difficult because of the surrounding fibrosis and scarring, with extensive adhesion to the chest wall, and the probability of poor overall lung function. Embolization of bronchial arteries by selective angiography is the treatment of choice and is successful in approximately half of cases.

Invasive aspergillosis is a much more lethal condition. Patients prone to this complication are those with severe defects in both cellular immunity and granulocyte neutrophil count and function. Other factors include cytotoxic or corticosteroid therapy, recent broad-spectrum antibiotic therapy, and severe neutropenia. Over 90 per cent of cases occur in association with the treatment of haematological malignancies or with organ transplantation. The most common initial finding is fever, usually with a patchy bronchopneumonia on the chest radiograph. However, the radiograph can also show nodular infiltration, lobar consolidation, cavitating lesions, or a diffuse miliary or interstitial pattern. Progressive parenchymal necrosis due to invasion of blood vessels by the fungus can cause cavitation with rapid colonization by a mycetoma. The propensity to invade vessels predisposes to death from massive haemoptysis.

The diagnosis of invasive aspergillosis is difficult. Bronchoscopy with lavage is seldom helpful but transbronchial biopsy will produce fungal hyphas in 40 to 70 per cent of cases. The mortality is in excess of 80 per cent, but starting treatment early and restoring the immune function may improve survival. Amphotericin B is the mainstay of therapy and the addition of 5-fluorocytosine is probably helpful. The renal toxicity of amphotericin is a serious potential problem, particularly if there is underlying renal disease. Liposomal amphotericin B with higher local deposition and activity is valuable in patients not tolerant of the conventional drug.

Phycomycetes

These non-septate fungi cause infection much less commonly than aspergillus. They also have a propensity to infect the lung and the central nervous system, and invade the vasculature. The diagnosis is usually by transbronchial lung biopsy. Untreated the mortality rate is nearly 100 per cent. Amphotericin B is the recommended treatment.

Cryptococcus neoformans

This organism usually invades the central nervous system but pulmonary involvement can arise, either in conjunction with infection of the central nervous system or in isolation (Kerkering *et al*. 1981). In up to a quarter of cases there may be no symptoms, only the incidental findings on a chest radiograph. A nodular lesion (single or multiple) is common but alveolar infiltrates are also often found. Sputum culture is positive in up to one-third of cases of cryptococcal pneumonia. Transbronchial lung biopsy is positive in up to half of cases presenting with infiltration, but nodular lesions usually require open lung biopsy for diagnosis.

(Histoplasmosis and coccidiomycosis also occur in endemic areas but are rare.)

Nocardia asteroides

Nocardiosis can complicate autoimmune rheumatic disease, particularly lupus. Prompt diagnosis is critical as the infection disseminates rapidly to cause death. Early treatment can cure up to 80 per cent of cases. Pulmonary involvement usually predominates, with an abnormal chest radiograph in 70 per cent of cases. Multiple nodular densities with or without cavitation are characteristic. A percutaneous fine-needle aspiration biopsy is diagnostic in up to 80 per cent of cases. Treatment with trimethoprim–sulphamethoxazole is highly effective.

Mycobacterial infections

Both typical and atypical mycobacterial infections occur in patients with autoimmune rheumatic disorders. The clinical features are often non-specific and pulmonary infiltration common, usually lacking the classical cavitating appearance found in the non-immunosuppressed. A high degree of suspicion is required. Sputum smears are the investigation of choice but, if negative, alveolar lavage can provide a diagnosis in up to 80 per cent of cases (Stover *et al*. 1984). *Mycobacterium tuberculosis* is sensitive to standard antituberculous therapy—rifampicin, isoniazid, and pyrazinamide—with the prospect of cure in more than 80 per cent. The cure rates for atypical mycobacteria infection are always considerably lower.

Pneumocystis carinii

Although the prevalence of *Pneumocystis carinii* pneumonia has increased dramatically in some populations over the last few years, it is still an unusual opportunistic pathogen in patients with autoimmune rheumatic disorders. Nevertheless, one should be aware of the small possibility of it arising and a bronchoalveolar lavage will almost invariably provide a positive diagnosis if the organism is present. Therapy comprises 2 weeks of intravenous trimethoprim, 20 mg/kg per day, with sulphamethoxazole, 100 mg/kg per day, followed by a week of conventional oral trimethoprim and sulphamethoxazole therapy. If trimethoprim and sulphamethoxazole therapy is not possible, then pentamidine isoethionate, 4 mg/kg per day intravenously for 3 weeks, is effective.

Viral infections

Infections with cytomegalovirus, herpes simplex, and herpes varicella zoster are more of a problem in patients with severe and sustained defects in cellular immunity, that is after organ transplantation, in human immunodeficiency virus-positive patients, and in sufferers from haematological malignancies. Nevertheless, these infections occasionally occur in people with connective tissue disorders, especially if taking immunosuppressive therapy.

Drug-induced pulmonary reactions

The disorders caused by the drugs used in autoimmune rheumatic disorders can be grouped as shown in Table 3. The range of such disorders is clearly wide and they mimic pulmonary reactions to the diseases themselves or other causes of pulmonary damage. There is very little that is helpful in the clinical manifestations, laboratory findings, or radiographic appearances specifically to suggest a drug-induced problem. Several patterns of radiographic abnormality occur, including increased interstitial markings, reticulonodular patterns, and alveolar infiltrates, often bilateral and most common in the lower lung fields.

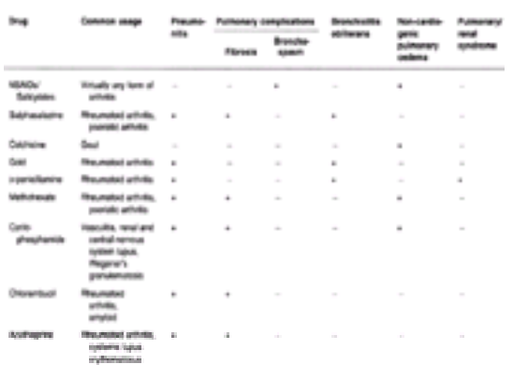


Table 3 Clinical manifestations of drug-induced pulmonary disease

Pulmonary function tests tend to show a restrictive defect with a reduced *TL*CO unless the drug causes a problem that purely affects the airways. Lung biopsy may demonstrate characteristic changes and exclude infection. Interstitial and alveolar inflammation is characterized by mononuclear cell infiltrates. Epithelial abnormalities include proliferation of both type I and type II pneumonocytes, with some dysplastic changes. In those with an acute hypersensitivity reaction, bronchiolitis, giant cells, granuloma formation, and an eosinophil infiltration may develop.

The precise prevalence of clinical manifestations that complicate drug therapy is not established. However, some associations are particularly well founded. For example Chakravarty and Webley reported that 46 per cent of patients treated with gold and 21 per cent treated with D-penicillamine developed a restrictive defect within 2 years of the onset of therapy (Chakravarty and Webley 1992). The chest radiographs often shows dense reticulonodular infiltrates, often basal, and occasionally with pleural fluid. Other features of gold toxicity may be evident elsewhere. Treatment is a combination of stopping therapy and adding corticosteroids and gives excellent results (Israel-Biet *et al*. 1991).

Of particular concern to rheumatologists given its current prominence in treating patients with rheumatoid arthritis and psoriatic arthritis, methotrexate may cause pneumonitis. This complication has been estimated to occur in 7 per cent of patients receiving methotrexate for either neoplastic or non-neoplastic disorders (Cooper *et al*. 1986). In a case–control study, however, no overt risk factors for the development of pneumonitis could be identified (Carroll *et al*. 1994). Even standard-dose regimens of methotrexate can cause pulmonary disease. Presentation of this complication is non-specific with cough, dyspnoea, and pulmonary crackles. Response is usual if the methotrexate is discontinued and high-dose steroids started. The steroids can be withdrawn once maximal improvement is achieved (Hargreaves *et al*. 1992).

Care should be taken if patients develop bronchospasm in relation to salicylates or non-steroidal anti-inflammatory drugs. These episodes usually arise early in treatment with these drugs and can even be seen after a single dose. The episodes can rapidly worsen and death may occur if the drug is continued.

Important points to remember (Table 4)

1. Rheumatoid diseases are associated predominantly with parenchymal infiltration, but airway, pleural, and musculoskeletal involvement are well recognized.
2. Breathlessness may be due to upper airways obstruction or weak muscles or an immobile chest wall. A normal chest radiograph does not eliminate a respiratory cause of dyspnoea.
3. High resolution CT provides clinical patterns of distribution of changes due to fibrosing alveolitis. The grading of those changes correlate to histological appearances and prognosis.
4. Lung function tests should include measurement of gas transfer, assessment of upper airway function (flow-volume curves) and, where indicated, measurement of total respiratory inspiratory and expiratory mouth pressures.
5. Open lung biopsy rarely provides information that is not intuitive from high resolution CT and lung function tests. It is more important where a non-associated cause is suspected, for example infection.
6. Drugs used by rheumatologists can cause pulmonary disease.
7. Infection is relatively common and a diagnosis should be sought, even if invasive tests are necessary.

Table 4 Important points in diagnosis

Chapter References

Akesson, A., Scheja, A., Ludin, A., and Wollheim, F.A. (1994). Improved pulmonary function in systemic sclerosis after treatment with cyclophosphamide. *Arthritis and Rheumatism*, **37**, 729–35.

Asherson, R.A., Khamashta, M., Ordi-Ros, J., Derksen, R.H.W.M., Machin, S.J., Barquinero, J., *et al*. (1989). The 'primary' antiphospholipid syndrome: major clinical and serological features. *Medicine*, **68**, 366–74.

Carroll, G.J., Thomas, R. Phatouros, C.C., Atchison, M.H., Leslie, A.L., Cook, N.J., and D'Souza, I. (1994). Incidence, prevalence and possible risk factors for pneumonitis in patients with rheumatoid arthritis receiving methotrexate. *Journal of Rheumatology*, **21**, 51–4.

Chakravarty, K. and Webley, M.A. (1992). A longitudinal study of pulmonary function in patients with rheumatoid arthritis treated with gold and D-pencillamine. *British Journal of Rheumatology*, **31**, 829–33.

Cooper, J.D., White, D.A., and Matthey, R.A. (1986). Drug-induced pulmonary disease. *American Review of Respiratory Diseases*, **133**, 321–40.

Cupps, T.R. and Fauci, A.S. (1982). Corticosteroid-mediated immunoregulation in man. *Immunology Review*, **65**, 133–55.

Fewins, H.E., McGowan, I. Whitehouse, G.H., Williams, J., and Mallyar, R. (1991). High definition computed tomography in rheumatoid arthritis associated pulmonary disease. *British Journal of Rheumatology*, **30**, 214–16.

Gaensler, E.A. and Carrington, C.B. (1980). Open biopsy for chronic diffuse infiltrative lung disease: clinical, roentgenographic and physiologic correlations in 502 patients. *Annals of Thoracic Surgery*, **30**, 411–26.

Gulko, P.S., Reville, J.D., Koopman, W. Burgurd, S.L., Bartolucci, A.A., and Alarcon Segovia, G.S. (1993). Anticardiolipin antibodies in systemic lupus erythematosus: clinical correlates, HLA associations and impact on survival. *Journal of Rheumatology*, **20**, 1684–93.

Hamuryudan, V., Yurkakul, S., Moral, F., Numan, F., Tüzün, H., Tüzüner, N., Mat, C., Tüzün, Özyazgan, Y., and Yazici, H. (1994). Pulmonary arterial aneurysms in Behçet's syndrome: a report of 24 cases. *British Journal of Rheumatology*, **33**, 48–51.

Hargreaves, M.R., Mowat, A.G., and Benson, M.K. (1992). Acute pneumonitis associated with low dose methotrexate treatment for rheumatoid arthritis: report of five cases and review of published reports. *Thorax*, **47**, 628–33.

Hellman, D.B., Petri, M., and Whiting-O'Keefe, Q. (1987). Fatal infections in systemic lupus erythematosus: the role of opportunistic organisms. *Medicine*, **66**, 341–8.

Israel-Biet, D., Labrune, S., and Huchon, G.J. (1991). Drug induced lung disease. 1990 review. *European Respiratory Journal*, **4**, 465–78.

Johnson, D.A., Gomm, S.A., Kalra, A., Woodcock, A.A., Evans, C.C., and Hind, C.R.K. (1993). The management of cryptogenic fibrosing alveolitis in three regions in the United Kingdom. *European Respiratory Journal*, **6**, 891–3.

Kerkering, T.M., Duma, R.J., and Shadomy, S. (1981). The evolution of pulmonary cryptococcosis. *Annals of Internal Medicine*, **94**, 611–16.

Reynolds, H.Y. (1987). Bronchoalveolar lavage. *American Review of Respiratory Disease*, **135**, 250–63.

Stover, D.E., Zaman, M.B., Hajdu, S.J., Lange, M., Gold, J., and Armstrong, D. (1984). Bronchoalveolar lavage in the diagnosis of diffuse pulmonary infiltrates in the immunosuppressed host. *Annals of Internal Medicine*, **101**, 1–7.

Turner-Warwick, M. and Haslam, P. (1987). The value of serial bronchoalveolar lavages in assessing the clinical progress of patients with cryptogenic fibrosing alveolitis. *American Review of Respiratory Diseases*, **135**, 26–34.

Wall, C.P., Gaensler, E.A., Carrington, C.B., and Hayes, J.A. (1981). Comparison of transbronchial and open biopsies in chronic infiltrative lung diseases. *American Review of Respiratory Diseases*, **123**, 280–5.

Warrick, J.H., Bhalla, M., Schabel, S.I., and Silver, R.M. (1991). High resolution computed tomography in early scleroderma lung disease. *Journal of Rheumatology*, **18**, 1520–7.

Wells, A.U., Hansell, D.M., Corrin, B., Harrison, N.K., Goldstraw, P., and Black, C.M. (1992). High resolution computed tomography as a predictor of lung histology in systemic sclerosis. *Thorax* **47**, 738–42.

Wells, A.U., Hansell, D.M., Harrison, N.K., Lawrence, R., Black, C.M., and Dubois, R.M. (1993 *a*). Clearance of inhaled 99^{m}Tc-DTPA predicts the clinical course of fibrosing alveolitis. *European*

Respiratory Journal, **4**, 797–802.

Wells, A.U., Hansell, D.M., Rubens, M.B., Cullinan, P., Black, C.M., and Dubois, R.M. (1993 *b*). The predictive value of appearance on thin-section computed tomography in fibrosing alveolitis. *American Review of Respiratory Diseases*, **148**, 1076–82.

1.3.6 The gastrointestinal system

Erkki Eerola and Reijo Peltonen

Introduction

Gastroenterology and rheumatology are associated in many ways and historically the idea of the association is not new. In 1672 Sydenham and in 1743 Ives described arthritis in patients with dysentery (Bollet 1981; Midvedt 1987). By the end of the nineteenth century the association between arthritis and shigellosis and inflammatory gut diseases had been recognized. The association can be of aetiological significance, as in enterogenic reactive arthritides, or without a causative link as in many gastrointestinal and rheumatic diseases with manifestations in both systems (Fig. 1 (a and b)). With some diseases it may even be difficult to decide in which category the disease should belong. The cellular and molecular mechanisms of inflammation and tissue damage are basically similar in both chronic inflammatory joint and bowel diseases (Parke 1993). The same drugs can sometimes be used to treat diseases of both systems or gastrointestinal and rheumatic manifestations of the same disease. An important facet is also the large number of gastrointestinal adverse effects of rheumatological medications.

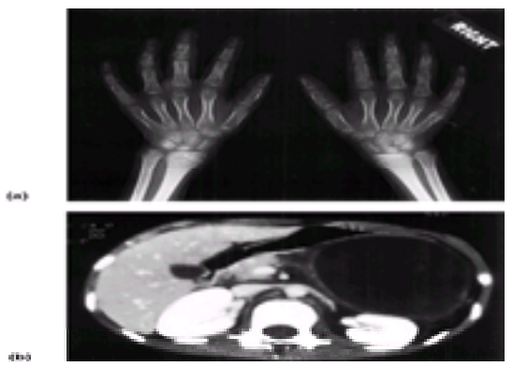


Fig. 1 Post-traumatic pancreatic cyst in a patient with associated lytic bone lesions secondary to lipase release.

Gastroenterological disorders with rheumatic manifestations

Gastrointestinal diseases with rheumatic associations are listed in Table 1. The closest association is in enteropathic arthropathies, which by definition are seronegative axial or peripheral arthropathies associated with gastrointestinal disorders. In patients with inflammatory bowel disease the overall prevalence of arthritic manifestations has been estimated to be approximately 30 per cent, but it seems to be increased (up to 62 per cent) when the patients are studied by rheumatologists (Scarpa *et al*. 1992; Parke 1993). On the other hand, an equally high percentage of patients with spondylitis have subclinical gut inflammation (Mielants *et al*. 1991b); in juvenile spondylarthropathy the prevalence may be even higher (Veys *et al*. 1993). In some of the enteropathic arthropathies rather convincing evidence of an aetiological association does exist. Antigenic material of intestinal microbial origin has been demonstrated in the synovial samples of patients with reactive arthritis after *Yersinia* and *Salmonella* enteric infections (Granfors *et al*. 1989; Granfors *et al*. 1990). Another aetiologically significant finding is the identification of the uncultivable causative agent of Whipple's disease (*Tropheryma whippelii*) from duodenal samples of the patients by using gene amplification techniques (Relman *et al*. 1992). Intestinal bypass surgery is a an example of an anatomical rearrangement which induces microbiological changes in the gut and leads to inflammatory joint symptoms in 6 to 35 per cent or even 52 per cent of the patients (Ross *et al*. 1989). There is a more definite aetiological link, however, in purulent or viral arthritis, when an intestinal pathogen is isolated in an affected joint.

Table 1 Gastrointestinal disorders with rheumatic manifestations

Gastrointestinal manifestations of rheumatic diseases

Gastrointestinal manifestations of rheumatic diseases are listed in Table 2.

Table 2 Gastrointestinal (GI) manifestations of rheumatic diseases (modified from Ryan and Sleisenger 1993)

Rheumatoid arthritis

The most common gastrointestinal manifestations in rheumatoid arthritis are gastric erosions and ulcerations caused by drug therapy. It is very difficult to assess whether the disease itself is an independent risk factor for them because of the widespread use of NSAIDs (Willoughby *et al*. 1986; Doube and Collins 1988). A common manifestation of the disease itself is oesophageal dysmotility, but it is usually benign and rarely causes serious complications or even dysphagia or heartburn. A potentially serious gastrointestinal manifestation is rheumatoid vasculitis, which has been estimated to affect 0.1 per cent of patients, but may be considerably more common. With complications such as ischaemic cholecystitis, appendicitis, pancolitis, bowel ulcerations, infarction, perforation and massive bleeding the mortality is high. In the treatment of vasculitis, cytoxic agents appear to be more effective than corticosteroids (Ryan and Sleisenger 1993). Amyloidosis has also been considered a rare complication although the prevalence appears to be increasing in rheumatoid arthritis. In autopsies, 21 per cent of patients were observed to have amyloidosis and 53 per cent of them had died because of it, although for other than gastrointestinal reasons (Susuki *et al*. 1993). In juvenile rheumatoid arthritis the trend may be the opposite. Amyloidosis as a cause of death is reported to have decreased from 42 per cent in the 1970s to 17 per cent in the 1980s (Savolainen and Isomäki 1993). Aggressive treatment with immunosuppressive agents may be the reason for this decline in mortality (David *et al*. 1993). Rectal biopsy or subcutaneous fat biopsy gives a 75 per cent diagnostic yield for amyloidoisis, but lip biopsy may be even better (86 per cent) (Bocanegra 1995). Interestingly, vasculitis and amyloidosis appear to exclude each other (Susuki *et al*. 1993).

Scleroderma

Gastrointestinal manifestations are common in scleroderma, occurring in 50 to 90 per cent of patients. They may precede the skin involvement, but usually, when the gastrointestinal tract is affected, the disease process is diffuse with multiple levels of involvement. The disease is characterized by mononuclear cell infiltrates, collagen deposition, and small vessel vasculitis in the skin and internal organs (Abu-Shakra *et al*. 1994). The most common symptoms are heartburn and dysphagia, and the most frequent manifestations of the disease include oesophageal dysmotility and lower oesophageal sphincter laxity. Small bowel involvement occurs in 17 to 50 per cent of patients. These manifestations often include segmental dilatation, atony, pseudo-obstruction, and bacterial overgrowth of the small bowel (10 per cent). The reported frequency of colonic involvement varies between 10 and 50 per cent; constipation in about 30 per cent and wide mouth diverticula in over 28 per cent. Other colonic complications such as telangiectasis, pseudo-obstruction, and volvulus are rare and each of them occurs in about 1 per cent of patients (Abu-Shakra *et al*. 1994).

Systemic lupus erythematosus

The most common gastrointestinal manifestation of systemic lupus erythematosus is abdominal pain. Nausea, anorexia, or vomiting affects over 50 per cent of patients. Diarrhoea affects 5 to 25 per cent of patients (Gutierrez *et al*. 1982). In systemic lupus, liver disease is not considered to be one of the most serious problems but it has been reported in 20 per cent and some biochemical evidence of hepatobiliary involvement in over 50 per cent of patients (Runyon *et al*. 1980). Ascites occurs in 10 per cent of patients and can be caused by peritoneal or mesenteric vasculitis, pancreatitis, serositis, nephrotic syndrome, or cardiac failure (Ryan and Sleisenger 1993). Vasculitis is one of the most serious gastrointestinal manifestations of lupus. It affects only 2 per cent of patients, but has a mortality rate of 50 per cent. It may result in colonic perforation, which has been reported to be responsible for about 27 per cent of deaths occurring in patients with lupus (Zizic *et al*. 1975).

Sjögren's syndrome

Xerostomia is one of the two major symptoms of Sjögren's syndrome. Dysphagia is reported in two-thirds of the patients. Abnormalities of oesophageal motility are seen in about 36 per cent and oesophageal webs in up to 10 per cent of patients. Impairment of pancreatic exocrine function is common. Sjögren's syndrome is associated with primary biliary cirrhosis (Tsianos *et al*. 1990) and has even been connected with hepatitis C infection (Haddad *et al*. 1992).

Vasculitides

Vasculitis is one of the pathogenetic mechanisms by which rheumatic diseases, such as rheumatoid arthritis and lupus, cause serious gastrointestinal manifestations. The frequency of gastrointestinal manifestations in diseases which in the first place are regarded as vasculitides varies widely. In Behçet's syndrome, vasculitic oral ulcers constitute a major diagnostic criterium. In Wegener's granulomatosis, the occurrence of gastrointestinal vasculitis is 5 per cent and in some other vasculitic diseases it is even lower. However, gastrointestinal vasculitis is common in polyarteritis nodosa (80 per cent), Henoch–Schönlein purpura (68 per cent), and Churg–Strauss syndrome (42 per cent) (see Fig. 2)

Fig. 2 Perforated lower oesophagus in a patient with dermatomyositis.

Gastrointestinal side-effects of antirheumatic drug treatment

Gastrointestinal side-effects of antirheumatic drugs are listed in Table 3.

Drug	Symptoms
NSAIDs (Hoy and Rooney 1988)	Nausea, oesophagitis, non-ulcer dyspepsia Peptic ulcer Hepatic enzyme abnormalities Steatorrhoea with malabsorption and protein-losing enteropathy Increased intestinal permeability Intestinal ulceration, stricture, obstruction and perforation
D-Penicillamine (Bergamini and Kaye 1988)	Dyspeptic admonitions, of nausea (2–20%) in the first 3–6 months Nausea, anorexia, abdominal pain (5–20%) Diarrhoea (occasional) Hepatic enzyme abnormalities (<5%), toxic hepatitis, hepatic necrosis, cholestasis jaundice Acute haemorrhagic colitis (rare)
Gold (Bergamini and Kaye 1988)	Buccal ulcers (5–20%) Hepatic enzyme abnormalities, cholestatic jaundice, hepatic necrosis Diarrhoea (50%, the most common gastrointestinal side-effect of gold) Gold enterocolitis (enteritis, diarrhoea, fever, eosinophilia, megacolon)
Sulphasalazine (Klotz et al. 1988)	Nausea, vomiting, dyspepsia, malaise, anorexia, abdominal pain, indigestion common, diarrhoea Colonic inflammation, mucosal ulceration, bloody diarrhoea (resembling ulcerative colitis) Hepatic enzyme abnormalities, granulomatous hepatitis, severe hepatitis and necrosis Cholestasis Pancreatitis (Braun and Hassell 1988)
Antimalarials	Anorexia, nausea, vomiting, diarrhoea
Azathioprine (McKendry 1981)	Nausea, vomiting (common), diarrhoea (~ 1%) Stomatitis Liver enzyme abnormalities
Cyclosporin (Kahn et al. 1982)	Nausea, vomiting, diarrhoea Abnormal liver function
Chlorambucil (Clemens 1981)	Nausea, vomiting Stomatitis
Cyclophosphamide (Clemens 1981)	Nausea, vomiting, diarrhoea Stomatitis Hepatic abnormalities
Methotrexate (Kemppainen and Furst 1986)	Nausea, vomiting, diarrhoea, stomatitis Liver enzyme abnormalities Hepatic fibrosis
Corticosteroids (Cooper and Kirwan 1988)	Peptic ulcer, pancreatitis

Table 3 Gastrointestinal side-effects of antirheumatic drug treatment

Non-steroidal anti-inflammatory drugs (NSAIDs)

NSAIDs have improved the quality of life for patients with arthritis for over 90 years. More than 30 million people worldwide consume them daily, and the annual increase in prescriptions is 6 per cent (Giercksky *et al*. 1989; Gabriel and Bombardier 1990; Zeidler 1992). Unfortunately they also cause a higher number of adverse effects than any other class of drug. NSAIDs are able to induce mucosal damage in every part of the gastrointestinal tract from oesophagus to rectum, but most of the serious damage occurs in the gastroduodenal region. The mechanisms involved are believed to be direct topical injury and systemic prostaglandin inhibition (Schoen and Vender 1989). It is the prostaglandin inhibition which is also responsible for the anti-inflammatory effect of the drugs and therefore all NSAIDs have the potential to initiate serious gastrointestinal events. Endoscopic studies have shown that virtually every subject develops gastroduodenal mucosal injury after a single 650 to 1350 mg dose of acetylsalicylic acid. NSAIDs can cause ultrastructural damage of the gastric epithelium within minutes and gross subepithelial haemorrhages and erosions within hours of ingestion (Graham and Smith 1986) (Fig. 3). With long-term treatment 50 to 75 per cent of patients develop mucosal damage, and 10 to 37 per cent gastroduodenal ulcers, a value over 5- to 15-fold greater than would be expected in a healthy population (Larkai *et al*. 1987; Graham *et al*. 1988; Geis *et al*. 1991). The risk is directly dose-related (Griffin *et al*. 1988). However, only approximately 30 to 40 per cent of patients receiving long-term treatment with NSAIDs have symptoms of dyspepsia (Zeidler 1992; Lichtenstein *et al*. 1995) and in as many as 58 per cent of subjects with a life-threatening complication, the complication itself has been the presenting manifestation of NSAID injury without prior dyspepsia as a warning (Langman 1989). The incidence of serious gastrointestinal events is estimated to be 1 to 2/1000 users per year. At special risk are elderly people (nearly 40 per cent of the users, relative risk ratio 5.5), those on corticosteroid therapy (relative risk ratio 1.8), patients using high-dose or multiple NSAID treatment, and also anyone during the first 3 months of treatment (Gabriel *et al*. 1991; Lichtenstein *et al*. 1995). The risk may also vary among different formulations: meclofenemate has the greatest relative risk (8.7) and ibuprofen the lowest (2.3), while those for indomethacin, acetylsalicylic acid, and naproxen lie in between (Gabriel *et al*. 1991; Henry *et al*. 1993). This has some relevance for clinical practice, although the order may reflect mainly the relative anti-inflammatory potency of the drugs (Fries *et al*. 1991; Lichtenstein *et al*. 1995). Newer drugs such as etodolac and nabumetone with selective peripheral and low gastric prostaglandin inhibition may prove to be much safer. With these drugs incidence of ulcers is 0.3 to 0.7 per cent and serious complications occur only in 0.1 per cent of users (Jenner 1990; Schattenkirchner 1990). NSAIDs inhibit prostaglandin synthesis by blocking cyclo-oxygenase. At least two different cyclo-oxygenases exist, Cox-1 and Cox-2. Cox-1 is found mainly in the gastrointestinal tract, while Cox-2 seems to be responsible for peripheral inflammatory responses. By selectively inhibiting Cox-2, the newer NSAIDs may prove to have as good as or even better peripheral anti-inflammatory effect but considerably less gastrointestinal side-effects than the NSAIDs now in use (Hayllar and Bjarnason 1995).

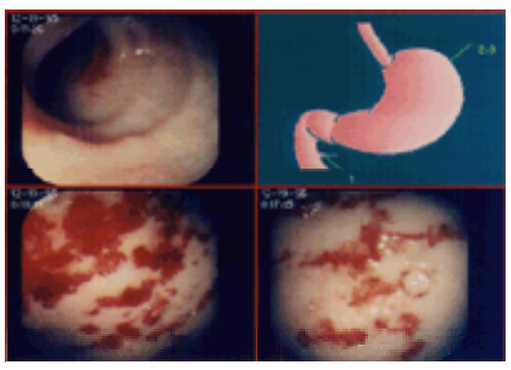


Fig. 3 Gastric lesions in a patient taking non-steroidal anti-inflammatory drugs (NSAIDs).

Oesophageal injury has been reported rarely and seems to come from direct topical irritation of NSAIDs (Minoch and Greenbaum 1991). Asymptomatic, small intestinal enteropathy is seen in 60 to 70 per cent of patients with long-term treatment (Bjarnason *et al*. 1993). It is associated with mild blood loss (1 to 10 ml per day) and increased intestinal permeability. Colonic injury occurs less frequently, but may be serious with perforations and bleeding. NSAIDs have been linked to development of collagenous colitis or colitis resembling inflammatory bowel disease. The drugs are also known to exacerbate pre-existing inflammatory bowel disease. Rectal administration may cause proctitis (Lichtenstein *et al*. 1995). Reversible elevations in liver enzymes are seen in about 5 per cent, and monitoring them may be necessary when NSAID therapy is initiated. Serious complications and deaths from hepatic damage are rare (Rabinovitz and Van Thiel 1992).

It should be realized that neither subjective symptoms nor endoscopic evidence of injury reliably predict the development of serious complications. In order to decrease the incidence of untoward gastrointestinal events induced by NSAIDs high-risk patients must be identified. Risk factors for gastroduodenal injury include age over 60 years, chronic use of NSAIDs, history of peptic ulcer, previous NSAID gastroduodenal intolerance, recent upper abdominal pain due to NSAIDs, multiple NSAID use, cardiovascular or other chronic disease in addition to arthritis, and use of multiple concurrent medications especially corticosteroids (Numo 1992). If high-risk patients must be treated with NSAIDs, either the lowest possible dose, an alternative non-NSAID analgesic, or the non-acetylated salicylate NSAID is recommended. If a NSAID is used, a coadministration of misoprostol should be considered. Misoprostol is a synthetic prostaglandin E_1 analogue and so far the only prophylactic agent that is unequivocally effective in protecting duodenal mucosa as well as the stomach from NSAID-induced injury. Prophylactic use of H_2 blockers and proton pump inhibitors (e.g. omeprazole) seems to offer some protection against duodenal ulcers only. However, even misoprostol cannot be relied upon to prevent more than 40 per cent of serious upper gastrointestinal complications (Silverstein *et al*. 1995), and the drug has not been proven to reduce the number of deaths. The results of the studies on the cost-effectiveness of coadministration of misoprostol with chronic NSAID use have so far been inconclusive. For patients with low risk of gastropathy, routine use of misoprostol is likely to be too costly and the drug is not always well tolerated (Scheiman 1994; Stucki *et al*. 1994). In the prophylaxis of ulcers, misoprostol is the only unequivocally effective agent, but even it has not been proven to reduce serious complications and deaths. The treatment of gastrointestinal injury includes withholding NSAIDs and administering H_2 blockers or omeprazole (Lichtenstein *et al*. 1995).

For a long time it was believed that disease-modifying drugs were many times more toxic than NSAIDs. In fact, because of the high frequency of gastrointestinal problems with NSAIDs, there are no substantial differences in toxicity between the two drug categories and the order in which the drugs are initiated in rheumatic diseases should be carefully considered (Fries *et al*. 1991).

Other antirheumatic drugs

Gastrointestinal side-effects of disease-modifying antirheumatic drugs occur most frequently in treatment with oral gold compounds, over 600 events per 1000 patient-years. With intramuscular gold the incidence is only one-tenth of that, but the complications appear to be more serious (Fries *et al*. 1991). Diarrhoea is the most common gastrointestinal manifestation of oral gold therapy (44 per cent). It responds well to dosage reduction and often declines even with continued use. Although serious gastrointestinal complications are rare, gold enterocolitis has been reported and may be fatal (Langer *et al*. 1987). Diarrhoea should therefore never be overlooked during any kind of gold therapy. In suspicious cases sigmoidoscopy may be required for the correct diagnosis. Stomatitis and mouth ulcers occur in

approximately 20 per cent of patients given intramuscular gold and in 5 to 10 per cent with oral gold compounds (Rangachari and Kean 1989).

Methotrexate and azathioprine have been reported to cause over 200 gastrointestinal adverse effects per 1000 patient-years (Fries *et al*. 1991). In a long-term study, methotrexate caused nausea and/or gastrointestinal distress in 65 per cent of patients, mouth sore or soreness in 55 per cent, and liver enzyme elevation in up to 88 per cent at some time during the treatment. The treatment also tends to lead to hepatic fibrosis verifiable by histology (Zizic *et al*. 1975; Kremer *et al*. 1989). Serious or potentially serious liver abnormalities have been reported in 12 per cent of patients (Augur *et al*. 1990). Hepatic and gastrointestinal toxicity has resulted in discontinuation of therapy in 16 per cent of patients treated with methotrexate (Alarcon *et al*. 1989).

Half of patients given cyclophosphamide complain of nausea, vomiting, or diarrhoea, and stomatitis also is common. Liver enzyme abnormalities are often seen. Serious hepatotoxicity, however, is uncommon; a few cases of hepatic necrosis have been reported when patients had previously been treated with azathioprine (Shaunak *et al*. 1988). Chlorambucil has a low frequency of gastrointestinal side-effects (only 4 per cent) (Clemens 1991).

Penicillamine is responsible for over 100 gastrointestinal adverse events per 1000 patient-years (Fries *et al*. 1991). Gastrointestinal pain or diarrhoea occurs in 8 to 35 per cent of patients. Taste abnormalities are experienced by 2 to 25 per cent of patients within the first 3 to 6 months. The symptom is dose related and usually subsides within a few months of treatment (Rangachari and Kean 1989). Haemorrhagic colitis has been reported (Hickling and Fuller 1979).

Sulphasalazine has a good gastrointestinal safety profile. Dyspepsia, nausea, abdominal discomfort, and mucosal ulceration are the most common side-effects of the therapy. Hepatic injuries, even massive liver necrosis, have been reported. Pancreatitis, cholestasis, and inflammatory diarrhoea are also rare complications (Amos *et al*. 1986).

Antimalarials and corticosteroids have a low number of gastrointestinal side-effects. The main gastrointestinal problems commonly associated with the use of corticosteroids are peptic ulceration (especially when used with NSAIDs), gastric haemorrhage, intestinal perforation, and pancreatitis. The risk of corticosteroids alone causing peptic ulcers is estimated to be very small or absent and this concern should not inhibit the use of small or moderate doses of steroids, for example in rheumatoid arthritis (Cooper and Kirwan 1990; George and Kirwan 1990). However, gastrointestinal problems when caused by corticosteroids tend to be serious and often require admission to hospital. For this reason corticosteroids are considered to be as toxic to the gastrointestinal tract as NSAIDs (Fries *et al*. 1991). Antimalarials appear to be the safest and best-tolerated disease-modifying drugs with respect to the gastrointestinal tract.

Diet and rheumatic diseases

Patients with rheumatoid arthritis often claim that certain foods provoke and certain diets alleviate their symptoms. In fact, the idea that dietary manipulation may alter the symptoms of rheumatoid arthritis is an intrinsic part of the folklore of the disease. In 1980, a large percentage (up to 94 per cent) (Wasner *et al*. 1980; Panush 1984) of patients in Western countries spent nearly one billion American dollars on diet and other similar unconventional therapies (Panush 1987; Buchanan *et al*. 1991; Darlington 1991); 37 to 50 per cent found them beneficial (Brown *et al*. 1980; Darlington *et al*. 1990), although there was no difference in disease progression or outcome between those reporting benefit or failure. In the past, dietary treatment of rheumatoid arthritis was generally considered as quackery (Lockshin 1981), but lately diet has become an important issue in rheumatology (Klinenberg 1985).

Diet therapy may be divided into three types: elimination, supplementation, and oral tolerance (Darlington and Ramsey 1993; Trentham *et al*. 1993). In a survey of 21 popular books (Darlington *et al*. 1990) the following supplements were advocated: vitamins, minerals, cod liver oil, New Zealand green-lipped mussel, kelp, evening primrose oil, devil's claw, yucca, and ginseng. Among other items tried by patients have been garlic, royal jelly, vinegar, honey, and alfalfa (Brown *et al*. 1980), while the following have been forbidden: alcohol, refined cereals, unrefined cereals, rice, acidic foods, fruits in general and especially citrus fruits, tomatoes, red meats, fish, chicken, egg yolks, egg whites, dairy produce, coffee (especially with caffeine), tea, chocolate, soft drinks, sugar (especially white sugar), processed foods, salt, pepper, monosodium glutamate, additives, preservatives, spices, and ice (Darlington *et al*. 1990). Thus, very often diets tried by patients, especially the health food diets, at the same time both supplementand eliminate individual foods. Most studies on pure supplementation diets concentrate on vitamins, minerals, fish oils, and plant oils. The idea that supplementation might be beneficial in rheumatoid arthritis is based on studies which show that patients have low serum levels of certain minerals and on the concept that certain oils, vitamins, and minerals have anti-inflammatory effect (Sköldstam 1989).

Supplementation therapy

Although patients with rheumatoid arthritis are susceptible to a number of dietary deficiencies including pyridoxine (vitamin B_6), histidine, vitamin C, zinc, and folic acid, dietary supplementation with these agents has not been found to improve the disease (Pike 1989). In their reviews of the literature on vitamin and mineral supplements in rheumatoid arthritis, Panush (1984) and Darlington (1991) observed that there is a lack of pertinent information. According to Panush, copper salts, although interesting (Banford *et al*. 1982), will probably not have any important role in the treatment of rheumatic disease. There is no clinical evidence that vitamin C has therapeutic effects in rheumatoid arthritis (Panush 1984; Sköldstam 1989). Tulleken *et al*. (1990) reported that the beneficial effects of fish oil supplementation cannot be ascribed to the antioxidizing properties of vitamin E (10 to 13 mg daily) *per se*. Machtey (1991) reported some benefit in osteoarthritis from vitamin E given at high doses (600 mg/day), and Scherak and Kolarz (1991) found in a multicentre, double-blind, crossover study using even higher doses of vitamin E (1200 mg daily) some benefit to patients with rheumatoid arthritis. There is no difference in the 25-OH-vitamin D concentrations between patients with rheumatoid arthritis and those with osteoarthritis (Bird *et al*. 1980). Although a low serum zinc concentration seems to be a non-specific feature of inflammation (Balogh *et al*. 1980), and the results in a preliminary study (Simkin 1976) did indicate some improvement in the symptoms of patients with rheumatoid arthritis who took oral zinc sulphate, later studies (Job *et al*. 1980; Kennedy *et al*. 1980; Mattingly and Mowat 1982) have not confirmed the result. Honkanen *et al*. (1991) observed that plasma copper and zinc levels correlated with the measures of disease activity, not dietary factors, and concluded that the plasma levels of these trace elements seem to be largely determined by the inflammatory process. A low selenium level was suggested by Tarp *et al*. (1985) as a possible factor in the pathogenesis of rheumatoid arthritis. However, other studies have not shown a correlation between plasma levels of selenium and disease activity (Möttönen *et al*. 1984; Peretz *et al*. 1987). Oral selenium therapy appears to have limited value in the management of rheumatoid arthritis, but glutathione peroxidase may play some role in the complex aetiology of the disease and possibly of other inflammatory disorders (Darlington and Ramsey 1993). Iron (and high concentrations of ferritin) may be associated with worsening of inflammatory joint diseases (Blake and Bacon 1981). All together, the results of therapeutic trials with vitamins and minerals in rheumatoid arthritis are controversial.

Fish oil

Bang *et al*. (1971) noted a low frequency of atherosclerosis among Eskimos, attributed this to the consumption of fish (Bang and Dyerberg 1980), and suggested that -3 polyunsaturated fatty acids, such as eicosapentaenoic acid and docosahexaenoic acid, may be responsible for the health benefit (Dyerberg and Bang 1979). Lucas and Power (1981) demonstrated dietary fat (vegetable oils and animal fat, but not fish oil) in amounts normally consumed in the Western diet to aggravate symptoms of rheumatoid arthritis. Kelley *et al*. (1985) showed that fish oil suppresses autoimmune disease. Lee *et al*. (1985) concluded that fatty acids derived from fish oils may have anti-inflammatory effects by inhibiting the 5-lipo-oxygenase pathway and the leukotriene B_4-mediated functions. A diet rich in fish oil can modulate the humoral and inflammatory components of the allergic response (Lee and Arm 1986) and may also be beneficial in several inflammatory conditions (Lee 1988). Human studies on diets rich in fish or supplemented with fish oil are listed in Table 4.

Table 4 Fish oil in the treatment of rheumatoid arthritis (RA)(a) Change in symptoms (b) Epidemiological evidence for beneficial effect of fish oil in rheumatoid arthritis

Plant oils

Polyunsaturated plant oils, such as g-linolenic acid may also be anti-inflammatory (Pike 1989). This fatty acid is metabolized to dihomo-g-linolenic acid which produces prostaglandins of the E series, such as PGE_1, which have potent anti-inflammatory properties (Mertin and Stackpoole 1981). It also does not lead to the formation of leukotrienes. Acute and chronic inflammation is suppressed by g-linolenic acid in animals (Tate *et al*. 1989). Attention has especially focused on the effects of evening primrose oil (Darlington and Ramsey 1993). Studies on the effects of plant oils on rheumatoid arthritis are listed in Table 5.

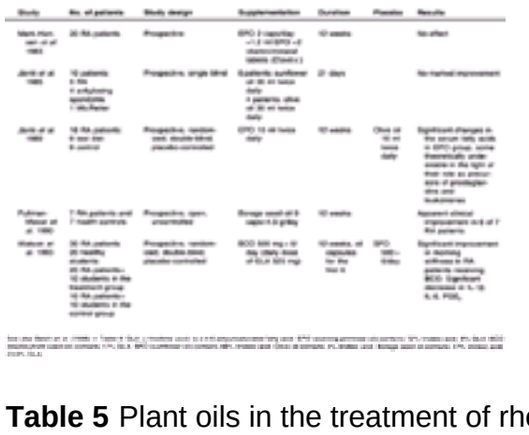


Table 5 Plant oils in the treatment of rheumatoid arthritis (RA)

Conclusions on oil supplementation therapy

Cathcart and Gonnerman (1991), in animal studies, found that fish oil can alleviate the symptoms of arthritis, and that the alteration in the severity of symptoms correlates with the decreased production of prostaglandins by peritoneal macrophages. Angelin and Palmblad (1988) concluded that the effects of fish oils in inflammatory diseases are interesting but, at best, modest, and Lee (1988) agreed that dietary supplementation with fish oil lipids may have anti-inflammatory effects. Kremer (1991) confirmed that fish oil supplementation is associated with clinical benefits in patients with rheumatoid arthritis. Improvement in the condition of the patient is not usually observed until after 12 weeks, but the improvement appears to increase as treatment is continued. Belch (1990) admitted that assessment of the literature does support a real, although modest, effect of fish oil in rheumatoid arthritis, without the production of serious side-effects. Most rheumatologists, however, would expect to achieve better relief of symptoms with NSAIDs. Pike (1989) agreed that some modest, but beneficial, effects of eicosapentaenoic acid consumption have been observed. Sperling (1991) noted that several small clinical trials have demonstrated a modest clinical effect of -3 fatty acids, but that larger, well-designed, randomized, double-blinded, placebo-controlled clinical trials are needed to determine the role, if any, of fatty acids in the treatment of rheumatoid arthritis. Buchanan *et al*. (1991) considered that there is a modest improvement of symptoms, but were of the opinion that it is unlikely that fish oil or evening primrose oil will make a major impact on the treatment of rheumatoid arthritis. Callegari and Zurier (1991) claimed that the potential capacity of particular fatty acids to regulate immune responses is exciting and called for long-term, multicentre, placebo-controlled studies with large numbers of patients to determine whether the plant oil therapy is useful in rheumatoid arthritis. According to Ringertz (1991), the trials on the effects of oil supplements on rheumatoid arthritis are characterized by a lack of improvement of the laboratory signs of inflammation. In a recent review, Darlington and Ramsey (1993) thought that the work on fish oils in the treatment of rheumatoid arthritis shows sufficient promise to justify further studies. The treatment appears to be safe, although long-term toxicity studies are needed, as are studies to determine whether low-dose fish oil is effective. The data on the effects of evening primrose oil for patients with rheumatoid arthritis are also promising, but need confirmation. Dietary supplementation with oils in the treatment of rheumatoid arthritis is of interest, because certain oils have, without doubt, some anti-inflammatory effect.

Elimination therapy

There are numerous reports connecting certain food items and arthritic symptoms (Hicklin *et al*. 1980; Mandell and Conte 1980; Coombs and Oldham 1981; Parke and Hughes 1981; Williams 1981; Bourne *et al*. 1985; Ratner *et al*. 1985; Panush *et al*. 1986; Lunardi *et al*. 1988; O'Farrelly *et al*. 1988; Panush 1990). It is known that food antigens are found in the serum after oral challenge (Paganelli *et al*. 1979). Regardless of the mechanism by which the arthritis is caused, be it called allergy (Williams 1981; Panush 1986), 'allergy' (Panush 1990), pseudo-allergy (Pearson 1991), immune response (Inman 1991), food antigens with molecular mimicry (Perez-Maceda *et al*. 1991), food hypersensitivity (Bahna and Kanuga 1991), or food intolerance (Parke and Hughes 1981), the idea of excluding the provocative item from the diet is delightful in its simplicity. If the item is unknown one may reason as follows: if a patient with constant symptoms fasts and consumes mineral water alone, his or her symptoms should clear completely if they are caused only by food intolerance (exclusion phase). Reintroduction of food items, one at a time, should reveal the item responsible for symptoms (reintroduction phase), and the rechallenge with it will confirm the finding. Some more palatable and nourishing modifications than a total fast also exist (cod and pear, lamb and pear, etc.) (Darlington 1985). Occasionally, diets believed to be beneficial in rheumatoid arthritis, such as certain vegetarian diets, may be combined with this method (Sköldstam 1986). Some diets may modify immunological responsiveness and thus affect the manifestations of rheumatic diseases (Panush 1986). Studies on exclusion and reintroduction of food items in patients with rheumatoid arthritis are listed in Table 6.

Table 6 Diets in the treatment of rheumatoid arthritis (RA)

Oral tolerance therapy

The idea of using oral tolerance in the treatment of rheumatoid arthritis is based on the fact that rheumatoid arthritis is an inflammatory synovial disease thought to involve T cells reacting to an antigen within the joint. The major protein in articular cartilage is type II collagen, and it is a potential autoantigen. Oral tolerance to autoantigens suppresses animal models of autoimmune disease mediated by T cells. In an experiment, oral administration of chicken type II collagen to patients with rheumatoid arthritis decreased the number of swollen and tender joints and some patients even had complete remission of the disease (Trentham *et al*. 1993). This approach appears exciting, but further confirmation from long-term studies is needed (Parke and Parke 1994).

Conclusions on diet therapy in rheumatoid arthritis

Studies on experimental animals have demonstrated that, without any doubt, modification of nutrition is able to delay, prevent, and even reverse the expression of genetically determined autoimmune defects. There is evidence which suggests that nutrition can also modulate the expression of autoimmunity in humans (Homsy *et al*. 1986; Trentham *et al*. 1993). The relationship between nutrition and rheumatic disease could be explained by two mechanisms that are not mutually exclusive:

dietary factors might alter immune and inflammatory responses, or food-related antigens might provoke hypersensitivity responses (Panush 1991). Some patients may also develop a type of allergic reaction to foods, but this is uncommon (O'Driscoll *et al*. 1985). This group of patients should benefit from a specific elimination diet (Sköldstam 1989). It has also been stated that total fasting may represent the most rapid and readily available way of relieving arthritic pain and swelling in rheumatoid arthritis (Palmblad *et al*. 1991). The benefit is usually lost soon after eating is resumed, and therefore total fasting is of little therapeutic significance in a chronic disease like rheumatoid arthritis (Sköldstam and Magnusson 1991). Ringertz (1991) notes that so far there have been no reports that show whether there may be any sustained positive effects in rheumatoid arthritis from dietary manipulations other than total fasting. However, Kjeldsen-Kragh *et al*. (1991) did report such an effect in patients who stayed on a vegetarian diet for 1 year, and Darlington and Ramsey (1991a), Maberly and Anthony (1991), and Seignalet (1992) also recorded similar effects with elimination diets in the treatment of rheumatoid arthritis.

The American College of Rheumatology position statement: diet and arthritis (American College of Rheumatology 1991) concludes that nutritional therapy for arthritis is experimental. Until more data are available, patients should continue to follow balanced and healthy diets, be sceptical of 'miraculous' claims, and avoid elimination diets and fad nutritional practices. Darlington and Ramsey (1993) state that it is essential for orthodox rheumatologists and research workers to continue to investigate the various ways in which patients' diets may influence the course of their disease, not only to find what may be useful in dietary manipulation for the patients' benefit and to keep them from the advice of the unscrupulous, but also so that patients do not adopt useless diets and thus deteriorate, when good medical advice to cease dietary therapy and to return to effective drug treatment would be of benefit. A summary of different diet trials in rheumatoid arthritis is listed in Table 6.

Intestinal microbiology and rheumatic diseases

The intestinal microbial flora is not essential for life (Shahani and Ayebo 1980), although it was once widely believed that Pasteur's work on the microbial nature of fermentation meant that without bacteria the food taken in would be passed unassimilated out of the body (Wilson 1983). Animals free from microbes and on adequate nutrition live more than 30 per cent longer than colonized animals (Drasar 1974; Parsonnet 1992). Although the relationship between host and normal flora is truly symbiotic (Parsonnet 1992), it is evident that intestinal bacteria may have either beneficial or harmful effects on the host, depending on the bacterial species and available substrates for their metabolism (Luckey 1977). It is surprising how well the epithelium of the gastrointestinal tract succeeds in allowing absorption of various nutrients and at the same time serving as a barrier against microbes and harmful antigenic structures. The number of bacteria in a human gut is estimated to be 10^{14}. Viable enteric pathogens may invade the blood stream and cause septic arthritis, but most often the arthritis is reactive (Ahvonen *et al*. 1969; Aho *et al*. 1974; Aho *et al*. 1976). Intestinal bacteria known to induce reactive arthritis include *Shigella, Salmonella, Yersinia, Campylobacter*, and *Clostridium difficile* (Gorbach 1986; Amor 1987; Aho 1989). Certain intestinal parasites such as *Giardia, Cryptosporidium, Strongyloides stercoralis*, and *Taenia saginata* have been associated with reactive arthritis (Bocanegra *et al*. 1981). The association between the gut and arthritis could be as follows: the increased permeability of the intestinal mucosa allows leakage of the antigens, either dietary or microbial in origin, which leads to an inflammatory reaction affecting the joints (Katz and Hollander 1989).

Patients with ankylosing spondylitis have a markedly elevated intestinal permeability (Katz and Hollander 1989), and more than 50 per cent of those with peripheral arthritis have inflammatory gut lesions (Cuvelier *et al*. 1987; Mielants *et al*. 1988). It has been suggested that intestinal *Klebsiella pneumoniae* may be a possible aetiological agent (Ebringer *et al*. 1988; Ebringer 1989; Ebringer 1991), although the cause still remains an open question (Russell 1988; Schwimmbeck and Oldstone 1989; van Kregten *et al*. 1991). It has even been suggested that many inflammatory arthritides may finally turn out to be reactions to microbial agents and thus these arthritides could and will be classified as reactive arthritis (Ford 1986).

Rheumatoid arthritis and intestinal bacteria

The idea that intestinal bacteria might be associated with rheumatoid arthritis is essentially derived from the association between gastrointestinal disease and arthritis (Wands *et al*. 1976; Katz and Hollander 1989; Sartor 1989), and from the fact that several intestinal bacteria cause reactive arthritis (Bennett 1978; Neumann and Wright 1983; Amor 1987; Atkinson and McLeod 1988; Aho 1989). Another reason for associating intestinal bacteria and rheumatoid arthritis is the fact that the genetic component of this disease is unlikely to be greater than 30 per cent, and there are no obvious environmental factors that explain the remaining 70 per cent of cases (Silman 1991). Furthermore, the gut flora can modify induction of arthritis induced experimentally by streptococcal cell wall extracts (van den Broek *et al*. 1992) and pristane (Thompson and Elson 1993). However, there is no definitive proof of a primary role for the gut in the pathogenesis of rheumatoid arthritis. If the link is substantiated, the therapeutic implications would be considerable (Zaphiropoulos 1986).

The main reasons for the difficulties encountered in studies on the association of rheumatoid arthritis and the intestinal microbial flora have been the enormous complexity of the latter (Finegold *et al*. 1974) and the inadequacy of traditional methods to study it (Hentges 1980; Hill 1981; Wayne *et al*. 1987; Woese 1987; Relman 1993). Each gram of wet faeces contains 10^{10} to 10^{12} bacteria recoverable by traditional bacteriological methods (Finegold *et al*. 1974; Moore and Holdeman 1974a; Moore and Holdeman 1974b; Moore and Holdeman 1975). The estimate of the number of bacterial species (more than 400 or 500) present in the intestinal tract of one person at any given time is calculated from the results obtained by culture methods (Moore and Holdeman 1974a; Moore and Holdeman 1975). Based on the studies which have relied on cloned or amplified 16S rRNA molecules using broad-range oligonucleotides (Giovannoni *et al*. 1990; Ward 1990; Schmidt *et al*. 1991; Liesack and Stackebrandt 1992), environmental microbiologists postulate that most of the extant micro-organisms (more than 80 per cent) remain unidentified, because of the insensitivity of bacterial cultures. The same may apply to the commensal microbial flora endogenous to humans (Relman 1993).

Culture-based studies dealing with the relationship between rheumatoid arthritis and bacterial species are presented in Table 7. Most often an initially positive report has been followed by several reports denying an aetiological role of the microbial species (Midvedt 1987), and no definite proof of any association between a change in the intestinal flora and rheumatoid arthritis has been produced. Recently, however, new data have been obtained by a new method—computed gas chromatography of bacterial cellular fatty acids, analysed directly from the stool samples (Peltonen *et al*. 1992). This method is probably the simplest way to establish whether there is a change, and if so the size of that change, in bacterial flora induced by any environmental factor such as diet, medication, or disease. The data show that the intestinal flora is changed in rheumatoid arthritis (Eerola *et al*. 1994), the change is mainly in the anaerobic flora (Eerola *et al*. 1994), and the beneficial effect of dietary treatment in rheumatoid arthritis is linked to changes in the intestinal flora (Peltonen *et al*. 1994).

Table 7 Studies on rheumatoid arthritis (RA) and specific bacteria often found in faeces

Another approach is based not on specific bacterial species or overall changes in the intestinal microflora but on the translocation of bacterial arthritogenic antigens from the gut (Bennett 1978). Non-viable bacterial antigens have been detected in rheumatoid synovial effusions (Bartholomew and Bartholomew 1979). As the translocation of viable bacteria from the gut to extraintestinal sites is not a rare phenomenon (Berg 1992), it is quite feasible that debris (including peptidoglycan) from bacteria normally present in the gastrointestinal tract also traverses the gut–blood barrier and stimulates the immune system. This idea is supported by experimental findings that bacterial cell wall material injected into rats can induce arthritis (Cromartie *et al*. 1977). Bacteria contain peptidoglycan within the capsule. Gram-positive bacteria contain ten times more of this substance than Gram-negative bacteria (Hazenberg *et al*. 1992). Growing evidence indicates that the arthritogenic properties of bacterial adjuvants reside within peptidoglycan dimers. Antibodies against peptidoglycan–polysaccharide polymers have been detected in the sera of patients with juvenile rheumatoid arthritis (Moore *et al*. 1989). Peptidoglycan has been shown to be absorbed from the gut (Lichtman *et al*. 1991). In the light of these observations, it is intriguing to note that past interest has focused mainly on Gram-positive bacteria (Olhagen and Månsson 1968; Svartz 1976), when the role of intestinal bacterial species in the aetiology of rheumatoid arthritis has been discussed (Midvedt 1987). The normal intestinal flora harbours a sufficiently diverse population of bacteria to produce the peptidoglycan–polysaccharide polymers capable of inducing arthropathic responses in the host. In a susceptible person, arthritis might result from the

accumulation of a certain threshold concentration of these arthropathic polymers in the joint. This accumulation might be accelerated greatly when the intestinal epithelium is damaged (Stimpson *et al*. 1986). As in the case of bacterial cell wall material, these peptidoglycan–polysaccharide complexes produced by the human intestinal flora also can induce arthritis when injected subcutaneously into rats (Kool *et al*. 1991). These bacterial cell wall polymers can also reactivate arthritis in rats (Lichtman *et al*. 1993). Several bacterial species of the normal flora contain peptidoglycan which is able to induce arthritis in experimental animals, these bacteria include *Peptostreptococcus productus, Propionibacterium acnes, Methanobacterium formicicum*, group A streptococci (Stimpson *et al*. 1986; Hazenberg *et al*. 1992), *Lactobacillus casei* (Lehman *et al*. 1987), *Eubacterium aerofaciens* (Severijnen *et al*. 1991), other *Eubacterium* species (Severijnen *et al*. 1990; Kool *et al*. 1992), *Bifidobacterium* species (Severijnen *et al*. 1989), and anaerobic coccoid rods (Severijnen *et al*. 1988). Although it is difficult to provide conclusive and persuasive evidence that peptidoglycans or other non-viable microbial components play any major role in rheumatoid arthritis, there is increasing evidence that these arthropathogenic polymers occur in the human intestinal lumen and that they can be released systemically (Schwab 1993).

Conclusions about the aetiological role of intestinal microbial flora in rheumatoid arthritis remain limited. Gullberg (1978) concludes that intestinal absorption of antigens and other noxious substances from intestinal microbes may play an important role in a multifactorial aetiopathogenesis of rheumatoid arthritis. Since arthropathies could be indirect consequences of change in the resident bacterial population in the gut, the same may be applicable to rheumatoid arthritis (Anonymous 1979). Midvedt (1987) states that alterations in the intestinal flora may play a role in the aetiopathogenesis of chronic aseptic inflammation in joints. The suggestion that the aetiology of rheumatoid arthritis lies in the intestine must continue to remain an unproven hypothesis (Doube and Collins 1988), and Phillips (1988) also indicates that at present there is little firm evidence to implicate microbial agents. In reactive arthritis the question is not so much 'do bacteria cause chronic polyarthritis', as 'how do they do it'. In rheumatoid arthritis the first question still remains: do bacteria cause it? (Phillips 1989). Microbially triggered arthritis is usually self-limited, but in some cases it develops into chronic arthritis (Leirisalo *et al*. 1982), and therefore reactive arthritides may serve as a model for studies on the pathogenesis of rheumatoid arthritis (Toivanen and Toivanen 1991). McCulloch *et al*. (1993) favour the hypothesis that rheumatoid arthritis may be a chronic form of reactive arthritis driven by the escape of primed T cells and microbial antigens from an extra-articular site of infection, possibly the gut, lungs, or skin nodules. Schwab (1993) concludes that peptidoglycan–polysaccharides, which are derived from a variety of bacterial species including intestinal bacteria, can be part of the aetiology of rheumatoid arthritis and other chronic inflammatory diseases.

At present there is interest on the subject but little firm evidence that intestinal microbes or their products are aetiologically important in rheumatoid arthritis.

NSAIDs and intestinal microbial flora

There are two reasons why the effects of NSAIDs on the intestinal bacterial flora may be interesting to rheumatologists. One is the pathogenesis of intestinal inflammation observed in patients with rheumatoid arthritis. NSAIDs are able to damage the intestinal mucosal integrity and increase permeability (Bjarnason *et al*. 1987; Mielants *et al*. 1991a). Nearly three-quarters of patients on long-term treatment with NSAIDs (more than 6 months) have small bowel inflammation (Bjarnason *et al*. 1987; Bjarnason 1990), and this may last for up to 16 months after discontinuation of the drug. It is possible that intestinal inflammation is caused by bacterial invasion into the damaged mucosa rather than by NSAIDs directly (Bjarnason *et al*. 1987).

The other reason is the question of the possible bacterial aetiology of rheumatoid arthritis. By altering the intestinal permeability and nutritional environment for bacteria in the gut, NSAIDs may favour the growth of certain bacterial species. It has been suggested that changes observed in the faecal flora of patients with rheumatoid arthritis (especially high counts of *Clostridium perfringens*) (Olhagen and Månsson 1968) might be an effect of NSAIDs (Shinebaum *et al*. 1987). This idea is supported by a finding that *Cl. perfringens* counts are higher both in patients with rheumatoid arthritis (who were all using NSAIDs) and those with osteoarthritis using NSAIDs than in patients with osteoarthritis not taking NSAIDs (Dearlove *et al*. 1992). Also, a similar difference in *Cl. perfringens* counts has been observed between rheumatoid arthritis patients taking NSAIDs and those patients not taking them (Bradley *et al*. 1993). On the other hand, it has been reported that 33 patients with rheumatoid arthritis in the mid-1960s without any medication had an abnormal *Cl. perfringens* flora (Olhagen and Månsson 1968), and that a short-term treatment with NSAIDs does not change faecal flora (Peltonen *et al*. 1993). The latter study is a human experiment where an actual attempt to change faecal flora with NSAIDs was made. The dose of naproxen was fairly high (500 mg twice day), but common in the treatment of rheumatoid arthritis. The treatment period was 2 weeks. Increased intestinal permeability has been induced with half that dose used for half that time (Bjarnason *et al*. 1984). No changes were observed, either in the faecal counts of *Cl. perfringens* or in the bacterial cellular fatty acid profiles of the stool samples. Only a long-term study (more than 6 months) with a similar controlled study design and with more participants will answer the question of whether NSAIDs actually change the faecal flora at the doses used in the treatment of rheumatoid arthritis. Furthermore, increased intestinal permeability and inflammation has been detected in patients with rheumatoid arthritis who were not given NSAIDs (Tagesson and Bengtsson 1983; Smith *et al*. 1985; Segal *et al*. 1986).

The question of the effects of NSAIDs on the intestinal microbial flora remains open as do so many other interesting aspects of associations between rheumatology and gastroenterology.

Chapter References

Abel, G., Zhang, Q.-X., and Agnello, V. (1993). Hepatitis C virus infection in type II mixed cryoglobulinemia. *Arthritis and Rheumatism*, **36**, 1341–9.

Abu-Shakra, M., Guillemin, F., and Lee, P. (1994). Gastrointestinal manifestations of systemic sclerosis. *Seminars in Arthritis and Rheumatism*, **24**, 29–39.

Aho, K. (1989). Bowel infection predisposing to reactive arthritis. *Baillière's Clinical Rheumatology*, **3**, 303–19.

Aho, K., Ahvonen, P., Lassus, A., Sievers, K., and Tiilikainen, A. (1974). HLA-B27 in reactive arthritis. A study of *Yersinia* arthritis and Reiter's disease. *Arthritis and Rheumatism*, **17**, 521–6.

Aho, K., Ahvonen, P., Lassus, A., Sievers, K., and Tiilikainen, A. (1976). *Yersinia* arthritis and related diseases: clinical and immunogenetic implications. In *Infection and immunology in the rheumatic diseases* (ed. D.C. Dumonde), pp. 341–4. Blackwell Scientific Publications, Oxford.

Ahvonen, P., Sievers, K., and Aho, K. (1969). Arthritis associated with *Yersinia enterocolitica* infection. *Acta Rheumatologica Scandinavica*, **15**, 232–55.

Alarcon, C.S., Tracy, I.C., and Blackburn, W.D. (1989). Methotrexate in rheumatoid arthritis. *Arthritis and Rheumatism*, **32**, 671–6.

American College of Rheumatology (1991). American College of Rheumatology position statement: diet and arthritis. *Rheumatic Disease Clinics of North America*, **17**, 443–4.

Amor, B. (1987). Suspected infectious agents and host environment interactions in spondylarthropathies. *Clinical and Experimental Rheumatology*, **5** (Suppl. 1), 19–24.

Amos, R. *et al*. (1986). Sulphasalazine for rheumatoid arthritis: toxicity in 744 patients monitored for one to 11 years. *British Medical Journal*, **293**, 420–3.

Angelin, B. and Palmblad, J. (1988). Fiskolja i stället för farmaka. *Läkartidningen*, **85**, 2831–2.

Anonymous (1979). Rheumatoid arthritis and the gut. (Editorial.) *British Medical Journal*, **i**, 1104.

Atkinson, M.H. and McLeod, B.D. (1988). Case report. Reactive arthritis associated with *Clostridium perfringens* enteritis. *Journal of Rheumatology*, **15**, 520–2.

Augur, N.A. *et al*. (1990). Prospective study of hepatotoxicity in patients receiving methotrexate for rheumatoid arthritis. *Arthritis and Rheumatism*, **33** (Suppl.9), 60.

Bahna, S.I. and Kanuga, J. (1991). Food hypersensitivity. *Rheumatic Disease Clinics of North America*, **17**, 243–9.

Balogh, Z., El-Ghobarey, A.F., Fell, G.S., Brown, D.H., Dunlop, J., and Dick, W.C. (1980). Plasma zinc and its relationship to clinical symptoms and drug treatment in rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **39**, 329–32.

Banford, J.C., Brown, D.H., Hazelton, R.A., McNeil, C.J., Sturrock, R.D., and Smith, W.E. (1982). Serum copper and erythrocyte superoxide dismutase in rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **41**, 458–62.

Bang, H.O. and Dyerberg, J. (1980). Lipid metabolism and ischemic heart disease in Greenland Eskimos. *Advances in Nutrition Research*, **3**, 1–22.

Bang, H.O., Dyerberg, J., and Nielsen, A. (1971). Plasma lipid and lipoprotein pattern in Greenlandic West Coast Eskimos. *Lancet*, **i**, 1143–5.

Bartholomew, L.E. and Bartholomew, F.N. (1979). Antigenic bacterial polysaccharide in rheumatoid synovial effusions. *Arthritis and Rheumatism*, **22**, 969–77.

Belch, J. (1990). Fish oil and rheumatoid arthritis: does a herring a day keep rheumatologists away? *Annals of the Rheumatic Diseases*, **49**, 71–2.

Belch, J.J.F., Ansell, D., Madhok, R., O'Dowd, A., and Sturrock, R.D. (1988). Effects of altering dietary essential fatty acids on requirements for non-steroidal anti-inflammatory drugs in patients with rheumatoid arthritis: a double blind placebo controlled study. *Annals of the Rheumatic Diseases*, **47**, 96–104.

Bennett, J.C. (1978). Future research directions. The infectious etiology of rheumatoid arthritis. New considerations. *Arthritis and Rheumatism*, **21**, 531–8.

Berg, R.D. (1992). Bacterial translocation from the gastrointestinal tract. *Journal of Medicine*, **23**, 217–44.

Beri, D., Malaviya, A.N., Shandilya, R., and Singh, R.R. (1988). Effect of dietary restrictions on disease activity in rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **47**, 69–72.

Bird, H.A., Peacock, M., Storer, J.H., and Wright, V. (1980). Comparison of serum 25-OH vitamin D concentrations in rheumatoid arthritis and osteoarthritis. *British Medical Journal*, **285**, 1416–17.

Bjarnason, I. (1990). Experimental evidence of the benefit of misoprostol beyond the stomach in humans. *Journal of Rheumatology*, **17** (Suppl. 20), 38–41.

Bjarnason, I. *et al*. (1984). Intestinal permeability and inflammation in rheumatoid arthritis: effects of non-steroidal anti-inflammatory drugs. *Lancet*, **ii**, 1171–4.

Bjarnason, I. *et al*. (1987). Blood and protein loss via small-bowel inflammation induced by non-steroidal anti-inflammatory drugs. *Lancet*, **ii**, 711–14.

Bjarnason, I., Hayllar, J., Macpherson, A.J., and Russell, A.S. (1993). Side effects of nonsteroidal anti-inflammatory drugs on the small and large intestine in humans. *Gastroenterology*, **104**, 1842–7.

Blake, D.R. and Bacon, P.A. (1981). Serum ferritin and rheumatoid disease. *British Medical Journal*, **282**, 1273–4.

Bocanegra, T.S., Espinoza, L.R., Bridgeford, P.H., Vasey, F.B., and Germain, B.F. (1981). Reactive arthritis induced by parasitic infestation. *Annals of Internal Medicine*, **94**, 207–9.

Bocanegra, T.S. (1995). Hyperlipidemia, amyloidosis, sarcoidosis, iron storage disease, and Wilson's disease. *Current Opinion in Rheumatology*, **7**, 73–7.

Bollet, A.J. (1981). A report of reactive arthritis following dysentery in 1743. *Arthritis and Rheumatism*, **24**, 860.

Bourne, J.T., Kumar, P., Huskisson, E.C., Mageed, R., Unsworth, D.J., and Wojtulewski, J.A. (1985). Arthritis and coeliac disease. *Annals of the Rheumatic Diseases*, **44**, 592–8.

Bradley, S.M., King, B., Troughton, P.R., Gooi, H.C., Bird, H.A., and Wright, V. (1993). Faecal *Clostridium perfringens* and antibody responses to its antigen in arthritis patients on and off non-steroidal anti-inflammatory drugs. *British Journal of Rheumatology*, **32**, 940–1.

Brazer, S.R. and Medoff, J.R. (1988). Sulfonamide-induced pancreatitis. *Pancreas*, **3**, 538–86.

Brown, F.H., Spitz, P.W., and Fries, J.F. (1980). Unorthodox treatments in rheumatoid arthritis. *Arthritis and Rheumatism*, **23**, 657–8.

Buchanan, H.M., Preston, S.J., Brooks, P.M., and Buchanan, W.W. (1991). Is diet important in rheumatoid arthritis? *British Journal of Rheumatology*, **30**, 125–34.

Callegari, P.E. and Zurier, R.B. (1991). Botanical lipids: potential role in modulation of immunologic response and inflammatory reactions. *Rheumatic Disease Clinics of North America*, **17**, 415–25.

Casewell, M.W. and Warren, R.E. (1980). Faecal klebsiellae in rheumatoid arthritis and spondylitis and in controls—incidence or prevalence in specimens or patients? *British Medical Journal*, **281**, 1495.

Cathcart, E.S. and Gonnerman, W.A. (1991). Fish oil fatty acids and experimental arthritis. *Rheumatic Disease Clinics of North America*, **17**, 235–42.

Cleland, L.G., French, J.K., Betts, W.H., Murphy, G.A., and Elliott, M.J. (1988). Clinical and biochemical effects of dietary fish oil supplements in rheumatoid arthritis. *Journal of Rheumatology*, **15**, 1471–5.

Clemens, P.J. (1991). Ankylating agents. In *Second line agents in treatment of rheumatoid diseases* (ed. J. Dixon and D.E. Furst). Marcel Dekker, New York.

Collin, P., Korpela, M., Hallström, O., Viander, M., Keyriläinen, O., and Mäki, M. (1992). Rheumatic complaints as a presenting symptom in patients with coeliac disease. *Scandinavian Journal of Rheumatology*, **24**, 20–3.

Conn, D.L., McDuffe, F.C., Holley, K.E., and Schroeter, A.L. (1976). Immunological mechanisms in systemic vasculitis. *Mayo Clinic Proceedings*, **51**, 511–18.

Coombs, R.R.A. and Oldham, G. (1981). Early rheumatoid-like joint lesions in rabbits drinking cows' milk. *International Archives of Allergy and Applied Immunology*, **64**, 287–92.

Cooper, C. and Kirwan, J.R. (1990). The risk of local and systemic corticosteroid administration. *Baillière's Clinical Rheumatology*, **4**, 401–17.

Cromartie, W.J., Craddock, J.G., Schwab, J.H., Anderle, S.K., and Yang, C.-H. (1977). Arthritis in rats after systemic injection of streptococcal cells or cell walls. *Journal of Experimental Medicine*, **146**, 1585–602.

Cuéllar, M.L., Carciá, C., and Molina, J.F. (1995). Cryoglobulinemia and other dysproteinemias, familial Mediterranean fever, and POEMS syndrome. *Current Opinion in Rheumatology*, **7**, 58–64.

Culp, K.S. *et al*. (1982). Autoimmune associations in primary biliary cirrhosis. *Mayo Clinic Proceedings*, **57**, 365–70.

Cuvelier, C., Barbatis, C., Mielants, H., De Vos, M., and Roels, H. (1987). Histopathology of intestinal inflammation related to reactive arthritis. *Gut*, **28**, 394–401.

Czaja, A.J. (1981). Current problems in the diagnosis and management of chronic active hepatitis. *Mayo Clinic Proceedings*, **56**, 311–23.

Darlington, L.G. (1985). Does food intolerance have any role in the aetiology and management of rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **44**, 801–4.

Darlington, L.G. (1991). Dietary therapy for arthritis. *Rheumatic Disease Clinics of North America*, **17**, 273–85.

Darlington, L.G. and Ramsey, N.W. (1991 *a*). Diets for rheumatoid arthritis. (Letter.) *Lancet*, **ii**, 1209.

Darlington, L.G. and Ramsey, N.W. (1991 *b*). Is diet important in rheumatoid arthritis? *British Journal of Rheumatology*, **30**, 315–16.

Darlington, L.G. and Ramsey, N.W. (1993). Clinical review. Review of dietary therapy for rheumatoid arthritis. *British Journal of Rheumatology*, **32**, 507–14.

Darlington, L.G., Ramsey, N.W. and Mansfield, J.R. (1986). Placebo-controlled, blind study of dietary manipulation therapy in rheumatoid arthritis. *Lancet*, **i**, 236–8.

Darlington, L.G., Jump, A., and Ramsey, N. (1990). Dietary treatment of rheumatoid arthritis. *Practitioner*, **234**, 456–60.

David, J., Vouyiouka, O., Ansell, B.M., Hall, A., and Woo, P. (1993). Amyloidosis in juvenile chronic arthritis: a morbidity and mortality study. *Clinical and Experimental Rheumatology*, **11**, 85–90.

Dearlove, S.M. *et al*. (1992). The effect of non-steroidal anti-inflammatory drugs on faecal flora and bacterial antibody levels in rheumatoid arthritis. *British Journal of Rheumatology*, **31**, 443–7.

Denman, A.M., Mitchell, B., and Ansell, B.M. (1983). Joint complaints and food allergic disorders. *Annals of Allergy*, **51**, 260–3.

Doube, A. and Collins, A.J. (1988). Is the gut intrinsically abnormal in rheumatoid arthtritis? (Editorial.) *Annals of the Rheumatic Diseases*, **47**, 617–19.

Drasar, B.S. (1974). *Human intestinal flora*. Academic Press, London.

Drücke, T. *et al*. (1980). Hepatitis B antigen-associated periarteritis nodosa in patients undergoing long-term hemodialysis. *American Journal of Medicine*, **68**, 86–90.

Dyerberg, J. and Bang, H.O. (1979). Haemostatic function and platelet polyunsaturated fatty acids in Eskimos. *Lancet*, **i**, 433–5.

Ebringer, A. (1989). The relationship between *Klebsiella* infection and ankylosing spondylitis. *Baillière's Clinical Rheumatology*, **3**, 321–38.

Ebringer, A. (1991). Ankylosing spondylitis and *Klebsiella*—a debate continues. (Editorial.) *Journal of Rheumatology*, **18**, 312–13.

Ebringer, A., Colthorpe, D., Young, A., and Corbett, M. (1980). Increased incidence of faecal *Klebsiella pneumoniae* in patients with HLA B7 CREG antigen and men with rheumatoid arthritis. *British Medical Journal*, **281**, 583–5.

Ebringer, A. *et al*. (1985). Antibodies to *Proteus mirabilis* in rheumatoid arthritis. *Lancet*, **ii**, 305–7.

Ebringer, A. *et al*. (1988). *Klebsiella* antibodies in ankylosing spondylitis and *Proteus* antibodies in rheumatoid arthritis. *British Journal of Rheumatology*, **27** (Suppl. 2), 72–85.

Ebringer, A., Khalafpour, S., and Wilson, C. (1989). Rheumatoid arthritis and *Proteus*: a possible aetiological association. *Rheumatology International*, **9**, 223–8.

Eerola, E. *et al*. (1994). Intestinal flora in early rheumatoid arthritis. *British Journal of Rheumatology*, **33**, 1030–8.

Feller, E. and Schumacher, H.R. (1972). Osteoarticular changes in Wilson's disease. *Arthritis and Rheumatism*, **15**, 259–66.

Finegold, S.M., Attebery, H.R., and Sutter, V.L. (1974). Effect of diet on human fecal flora: comparison of Japanese and American diets. *American Journal of Clinical Nutrition*, **27**, 1456–69.

Ford, D.K. (1986). Reactive arthritis: a viewpoint rather than a review. *Clinics in the Rheumatic Diseases*, **12**, 389–401.

Fries, J.F., Williams, C.A., and Bloch, D.A. (1991). The relative toxicity of nonsteroidal anti-inflammatory drugs. *Arthritis and Rheumatism*, **34**, 1353–60.

Gabriel, S.E. and Bombardier, C. (1990). NSAID induced ulcers. An emerging epidemic? (Editorial.) *Journal of Rheumatology*, **17**, 1–2.

Gabriel, S.E., Jaakkimainen, L., and Bombardier, C. (1991). Risk for serious gastrointestinal complications related to use of nonsteroidal anti-inflammatory drugs: a meta analysis. *Annals of Internal Medicine*, **115**, 787–96.

Geis, G.S., Stead, H., Wallemark, C.-B., and Nicholson, P.A. (1991). Prevalence of mucosal lesions in the stomach and duodenum due to chronic use of NSAIDs in patients with RA and OA, and interim report on prevention by misoprostol of diclofenac-associated lesions. *Journal of Rheumatology*, **18** (Suppl. 28), 11–14.

George, E. and Kirwan, J.R. (1990). Corticosteroid therapy in rheumatoid arthritis. *Baillière's Clinical Rheumatology*, **4**, 621–47.

Giercksky, K.-E., Huseby, G., and Rugstad, H.-E. (1989). Epidemiology of NSAID-related gastrointestinal side effects. *Scandinavian Journal of Gastroenterology*, **24** (Suppl. 163), 3–8.

Giovannoni, S.J., Britschgi, T.B., Moyer, C.L., and Field, K.G. (1990). Genetic diversity in Sargasso sea bacterioplankton. *Nature*, **345**, 60–1.

Golding, P.L., Smith, M., and Williams, R. (1973). Multisystem involvement in chronic liver disease. *American Journal of Medicine*, **55**, 772–82.

Gorbach, S.L. (1986). Bengt E. Gustafsson memorial lecture. Function of the normal human microflora. *Scandinavian Journal of Infectious Diseases*, **49** (Suppl.), 17–30.

Graham, D.Y. and Smith, J.L. (1986). Aspirin and the stomach. *Annals of Internal Medicine*, **104**, 390–8.

Graham, D.Y., Agrawal, N.M., and Roth, S. (1988). Prevention of NSAID-induced gastric ulcer with misoprostol: multicenter, double-blind, placebo-controlled trial. *Lancet*, **ii**, 1277–80.

Granfors, K. *et al*. (1989). *Yersinia* antigens in synovial-fluid cells from patients with reactive arthritis. *New England Journal of Medicine*, **320**, 216–21.

Granfors, K. *et al*. (1990). Salmonella lipopolysaccharide in synovial cells from patients with reactive arthritis. *New England Journal of Medicine*, **335**, 685–8.

Griffin, M.R., Piper, J.M., Daugherty, J.R., Snowden, M., and Ray, W.A. (1988). Nonsteroidal anti-inflammatory drug use and death from peptic ulcer in elderly persons. *Annals of Internal Medicine*, **109**, 359–63.

Gullberg, R. (1978). Possible role of alterations of the intestinal flora in rheumatoid arthritis. *Rheumatology and Rehabilitation*, **17** (Suppl.), 5–10.

Gutierrez, F., Valenzuela, J.E., Ehresmann, G.R., Quismorio, F.P., and Kitridou, R.C. (1982). Esophageal dysfunction in patients with mixed connective tissue disease and systemic lupus erythematosus. *Digestive Diseases and Science*, **27**, 592.

Haddad, J. *et al*. (1992). Lymphocytic sialadenositis of Sjögren's syndrome associated with chronic hepatitis-C virus liver disease. *Lancet*, **339**, 321–3.

Hafström, I., Ringertz, B., Gyllenhammar, H., Palmblad, J., and Harms-Ringdahl, M. (1988). Effects of fasting on disease activity, neutrophil function, fatty acid composition, and leucotriene biosynthesis in patients with rheumatoid arthritis. *Arthritis and Rheumatism*, **31**, 585–92.

Hamberg, J., Lindahl, O., and Lindwall, L. (1981). Fasta och hälsokost vid RA—en kontrollered undersökning. *Swedish Journal of Biological Medicine*, **3**, 8–12.

Hayllar, J. and Bjarnason, I. (1995). NSAIDs, Cox-2 inhibitors, and the gut. (Editorial.) *Lancet*, **346**, 521–2.

Hazenberg, M.P., Klasen, I.S., Kool, J., Ruseler-van-Embden, J.G.H., and Severijnen, A.J. (1992). Are intestinal bacteria involved in the etiology of rheumatoid arthritis? *Apmis*, **100**, 1–9.

Henry, D., Dobson, A., and Turner, C. (1993). Variability in the risk of major gastrointestinal complications from nonaspirin nonsteroidal anti-inflammatory drugs. *Gastroenterology*, **10**, 1078–88.

Hentges, D.J. (1980). Does diet influence human fecal microflora composition? *Nutrition Reviews*, **38**, 329–36.

Hicklin, J.A., McEwen, L.M., and Morgan, J.E. (1980). The effect of diet in rheumatoid arthritis. (Abstract.) *Clinical Allergy*, **10**, 463.

Hickling, P. and Fuller, J. (1979). Penicillamine causing acute colitis. *British Medical Journal*, **ii**, 367.

Hill, M.J. (1981). Diet and the human intestinal bacterial flora. *Cancer Research*, **41**, 3778–80.

Hollinger, F.B. and Glombicki, A.P. (1990). Hepatitis A virus. In *Principle and practice of infectious diseases* (ed. G.L. Mandell, R.G. Douglas, and J.E. Bennett), pp. 911–18. Churchill Livingstone, New York.

Homsy, J., Morrow, W.J.W., and Levy, J.A. (1986). Nutrition and autoimmunity: a review. *Clinical and Experimental Immunology*, **65**, 473–88.

Honkanen, V. E. *et al*. (1991). Plasma zinc and copper concentrations in rheumatoid arthritis: influence of dietary factors and disease activity. *American Journal of Clinical Nutrition*, **54**, 1082–6.

Inman, R.D. (1982). Rheumatic manifestations of hepatitis B infection. *Seminars in Arthritis and Rheumatism*, **11**, 406–20.

Inman, R.D. (1991). Antigens, the gastrointestinal tract and arthritis. *Rheumatic Disease Clinics of North America*, **17**, 309–21.

Inman, R.D. *et al*. (1986). Arthritis, vasculitis, and cryoglobulinemia associated with relapsing hepatitis A virus infection. *Annals of Internal Medicine*, **105**, 700–3.

Isdale, A. and Wright, V. (1989). Seronegative arthritis and the bowel. *Baillière's Clinical Rheumatology*, **3**, 285–301.

Ivey, K.J. and Rooney, P.J. (1989). Non-steroidal anti-inflammatory drugs and the gastrointestinal tract. *Baillière's Clinical Rheumatology*, **3**, 393–409.

Jajic, I. (1987). The significance of the presence of *Cl. perfringens* in the feces and alpha antitoxin in the serum of patients with rheumatoid arthritis. *Clinical Rheumatology*, **6**, 111.

Jajic, I., Udovicic, I., Bilic, I., and Jakovina, M. (1977). O znacenju nalaza *Clostridium perfringensa* u stolici i alfa-antitoksina u serumu bolesnika s reumatoidnim artritisom. (The significance of the presence of *Cl. perfringens* in the feces and of alpha antitoxin in the serum of patients with rheumatoid arthritis). *Lijecnicki Vjesnik*, **99**, 405.

Jäntti, J., Isomäki, H., Laitinen, O., Nikkari, T., Seppälä, E., and Vapaatalo, H. (1985). Linoleic acid treatment in inflammatory arthritis. *Internaional Journal of Clinical Pharmacological Therapy and Toxicology*, **23**, 89–91.

Jäntti, J., Nikkari, T., Solakivi, T., Vapaatalo, H., and Isomäki, H. (1989). Evening primrose oil in rheumatoid arthritis: changes in serum lipids and fatty acids. *Annals of the Rheumatic Diseases*, **48**, 124–7.

Jenner, P.N. (1990). A 12-month postmarketing surveillance study of nabumetone: a preliminary report. *Drugs*, **40** (Suppl. 5), 80–6.

Job, C., Menkes, C.J., and Delbarre, F. (1980). Zinc sulphate in the treatment of rheumatoid arthritis. *Arthritis and Rheumatism*, **23**, 1408–9.

Kahn, M.C. (1924). Intestinal spore-bearing bacteria. *Journal of Infectious Diseases*, **35**, 423.

Katz, K.D. and Hollander, D. (1989). Intestinal mucosal permeability and rheumatological diseases. *Baillière's Clinical Rheumatology*, **3**, 271–84.

Kelley, V.E., Ferretti, A., Izui, S., and Strom, T.B. (1985). A fish diet rich in eicosapentaenoic acid reduces cyclooxygenase metabolites, and suppresses lupus in MRL-1pr-mice. *Journal of Immunology*, **134**, 1914–19.

Kennedy, A.C., Bessent, R.G., and Reynolds, P.M.G. (1980). Effect of oral zinc sulphate and penicillamine on zinc metabolism in patients with rheumatoid arthritis. *Journal of Rheumatology*, **7**, 639–44.

Khalafpour, S., Ebringer, A., and Abuljadayel, I. (1987). Antibodies to proteus species in rheumatoid arthritis. *Clinical Rheumatology*, **6**, 110–11.

Kjeldsen-Kragh, J. *et al*. (1991). Controlled trial of fasting and one-year vegetarian diet in rheumatoid arthritis. *Lancet*, **ii**, 899–902.

Kjeldsen-Kragh, J., Rashid, T., Haugen, M., Förre, Ö., and Ebringer, A. (1993). Anti-proteus antibody titres in rheumatoid arthritis patients during a controlled clinical trial of fasting and one-year vegetarian diet. *Revista Española de Reumatologia*, **20** (Suppl. 1), 471.

Klinenberg, J.R. (1985). ARA presidential address: 1984–2034: the next half-century for American rheumatology. *Arthritis and Rheumatism*, **28**, 1–7.

Kool, J. *et al*. (1991). Induction of arthritis in rats by soluble peptidoglycan–polysaccharide complexes produced by human intestinal flora. *Arthritis and Rheumatism*, **34**, 1611–16.

Kool, J., Severijnen, A.J., and Hazenberg, M.P. (1992). Arthritogenicity of *Eubacterium* species in the adjuvant arthritis model. *Journal of Rheumatology*, **19**, 1000–1.

Kowal, A., Carstens Jr, J.H., and Schnitzer, T.J. (1990). Cyclosporine in rheumatoid arthritis. In *Immunomodulators in rheumatic diseases* (ed. D.E. Furst and M.E. Weinblatt), pp. 61–98. Marcel Dekker, New York.

Kremer, J.M. (1991). Clinical studies of omega-3 fatty acid supplementation in patients who have rheumatoid arthritis. *Rheumatic Disease Clinics of North America*, **17**, 391–402.

Kremer, J.M. *et al*. (1985). Effects of manipulation of dietary fatty acids on clinical manifestations of rheumatoid arthritis. *Lancet*, **i**, 184–7.

Kremer, J.M. *et al*. (1987). Fish-oil fatty acid supplementation in active rheumatoid arthritis. *Annals of Internal Medicine*, **106**, 497–503.

Kremer, J.M., Lee, R.G., and Tolman, K.G. (1989). Liver histology in rheumatoid arthritis patients receiving long-term methotrexate therapy: A propective study with baseline and sequential biopsy samples. *Arthritis and Rheumatism*, **32**, 121–7.

Kremer, J.M. *et al*. (1990). Dietary fish oil and olive oil supplementation in patients with rheumatoid arthritis. *Arthritis and Rheumatism*, **33**, 810–20.

Kroker, G.P. *et al*. (1984). Fasting and rheumatoid arthritis: a multicenter study. In *Clinical ecology* (ed. L.D. Dickley). Charles Thomas, Springfield, Illinois.

Langer, H.E., Hartman, G., Heinemann, G., and Rivhter, K. (1987). Gold enterocolitis induced by auranofin treatment of rheumatoid arthritis: case report and review of the literature. *Annals of the Rheumatic Diseases*, **46**, 787–92.

Langman, M.J.S. (1989). Epidemiological evidence on the association between peptic ulceration and anti-inflammatory drug use. *Gastroenterology*, **96**, 640–6.

Larkai, E.N., Smith, J.L., Lidsky, M.D., and Graham, D.Y. (1987). Gastroduodenal mucosa and dyspeptic symptoms in arthritis patients during chronic nonsteroid anti-inflammatory drug use. *American Journal of Gastroenterology*, **82**, 1153–8.

Larsen, J.H. (1976). *Yersinia enterocolitica* infections and arthritis. In *Infection and immunology in the rheumatic diseases* (ed. D.C. Dumonde), pp. 133–40. Blackwell Scientific Publications, Oxford.

Larsen, J.H. (1980). *Yersinia enterocolitica* infections and rheumatic diseases. *Scandinavian Journal of Rheumatology*, **9**, 129–37.

Lee, T.H. (1988). Benefits from oily fish. (Editorial.) *British Medical Journal*, **297**, 1421–2.

Lee, T.H. and Arm, J.P. (1986). Prospects for modifying the allergic response by fish oil diets. *Clinical Allergy*, **16**, 89–100.

Lee, T.H. *et al*. (1985). Effect of dietary enrichment with eicosapentaenoic and docosahexaenoic acids on *in vitro* neutrophil and monocyte leucotriene generation and neutrophil function. *New England Journal of Medicine*, **312**, 1217–24.

Lehman, T.J.A., Cremer, M.A., Walker, S.M., and Dillon, A.M. (1987). The role of humoral immunity in *Lactobacillus casei* cell wall induced arthritis. *Journal of Rheumatology*, **14**, 415–19.

Leirisalo, M. *et al*. (1982). Followup study of patients with Reiter's disease and reactive arthritis, with special reference to HLA-B27. *Arthritis and Rheumatism*, **25**, 249–59.

Lichtenstein, D.R., Syngal, S., and Wolfe, M.M. (1995). Nonsteroidal anti-inflammatory drugs and the gastrointestinal tract. *Arthritis and Rheumatism*, **38**, 5–18.

Lichtman, S.N., Keku, J., Schwab, J.H., and Sartor, R.B. (1991). Evidence for peptidoglycan absorption in rats with experimental small bowel bacterial overgrowth. *Infectious Immunology*, **59**, 555–62.

Lichtman, S.N. *et al*. (1993). Bacterial cell wall polymers (peptidoglycan–polysaccharide) cause reactivation of arthritis. *Infectious Immunology*, **61**, 4645–53.

Liesack, W. and Stackebrandt, E. (1992). Occurrence of novel groups of the domain bacteria as revealed by analysis of genetic material isolated from an Australian terrestrial environment. *Journal of Bacteriology*, **174**, 5072–8.

Lindberg, E. (1973). Können Ernährungsfaktoren die chronische Polyarthritis beeinflussen? *Zeitschrift für Physiotherapie*, **25**, 119.

Linos, A. *et al*. (1991). The effect of olive oil and fish consumption on rheumatoid arthritis—a case control study. *Scandinavian Journal of Rheumatology*, **20**, 419–26.

Lockshin, M.D. (1981). The unproven remedies committee. *Arthritis and Rheumatism*, **24**, 1188–90.

Lucas, P. and Power, L. (1981). Dietary fat aggravates active rheumatoid arthritis. *Clinical Research*, **29**, 754A.

Luckey, T.D. (1977). Bicentennial overview of intestinal microecology. *American Journal of Clinical Nutrition*, **30**, 1753–61.

Lunardi, C. *et al*. (1988). Food allergy and rheumatoid arthritis. *Clinical and Experimental Rheumatology*, **6**, 423–4.

Maberly, D.J. and Anthony, H.M. (1991). Diets for rheumatoid arthritis. (Letter.) *Lancet*, **ii**, 1210.

Machtey, I. (1991). Vitamin E in arthritis. *Arthritis and Rheumatism*, **34**, 1205.

Mandell, M. and Conte, A. (1980). The role of allergy in arthritis, rheumatism and associated polysymptomatic cerebro-viscero-somatic disorders: a double-blind provocation test study. *Annals of Allergy*, **44**, 51.

Månsson, I. and Olhagen, B. (1966). Intestinal *Clostridium perfringens* in rheumatoid arthritis and other connective tissue disorders. *Acta Rheumatica Scandinavica*, **12**, 167–74.

Månsson, I. and Olhagen, B. (1967). Intestinal *Clostridium perfringens* in arthritis and parakeratosis induced by dietary factors. Experimental studies in pigs. *Bulletin of International Epizoology*, **67**, 1319–28.

Månsson, I. and Olhagen, B. (1968). Intestinal *Clostridium perfringens* in rheumatoid arthritis and other collagen diseases. *Acta Medica Scandinavica*, **184**, 395–402.

Månsson, I., Norberg, R., Olhagen, B., and Björklund, N.-E. (1971). Arthritis in pigs induced by dietary factors. *Clinical and Experimental Immunology*, **9**, 677–93.

Mattingly, P.C. and Mowat, A.G. (1982). Zinc sulphate in rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **41**, 456–7.

McCulloch, J., Lydyard, P.M., and Rook, G.A.W. (1993). Rheumatoid arthritis: how well do theories fit the evidence? *Clinical and Experimental Immunology*, **92**, 1–6.

McKendry, R.J.R. (1991). Purine analogs. In *Second line agents in treatment of rheumatic disease* (ed. J. Dixon and D.E. Furst). Marcel Dekker, New York.

McKenna, P.J., O'Brian, J.T., Scheinman, H.Z., Delaney, W.E., Pellecchia, C., and Lepore, M.J. (1971). Hepatitis and arthritris with hepatitis associated antigen in serum and synovial fluid. *Lancet*, **i**, 214–15.

Mertin, J. and Stackpoole, A. (1981). Anti-PGE antibody inhibit *in vivo* development of cell-mediated immunity. *Nature*, **294**, 456–8.

Midvedt, T. (1987). Intestinal bacteria and rheumatic disease. *Scandinavian Journal of Rheumatology*, **64** (Suppl.), 49–54.

Mielants, H., Veys, E.M., Cuvelier, C., and De Vos, M. (1988). Ileonoscopic findings in seronegative spondylarthropathies. *British Journal of Rheumatology*, **27** (Suppl. II), 95–105.

Mielants, H. *et al*. (1991*a*). Intestinal mucosal permeability in inflammatory rheumatic diseases. I. Role of anti-inflammatory drugs. *Journal of Rheumatology*, **18**, 389–93.

Mielants, H. *et al*. (1991*b*). Intestinal mucosal permeability in inflammatory rheumatic diseases. II. Role of disease. *Journal of Rheumatology*, **18**, 394–400.

Milder, M.S., Cook, J.D., Stray, S., and Finch, C.A. (1980). Idiopathic hemochromatosis, an interim report. *Medicine*, **59**, 34–49.

Minoch, A. and Greenbaum, D.S. (1991). Pill-esophagitis caused by nonsteroidal anti-inflammatory drugs. *American Journal of Gastroenterology*, **86**, 1086–9.

Moore, T.L., El-Najdawi, E., and Dorner, R.W. (1989). Antibody to streptococcal cell wall peptidoglycan–polysaccharide polymers in sera of patients with juvenile rheumatoid arthritis but absent in isolated immune complexes. *Journal of Rheumatology*, **16**, 1069–73.

Moore, W.E.C. and Holdeman, L.V. (1974 *a*). Human fecal flora: the normal flora of 20 Japanese-Hawaiians. *Applied Microbiology*, **27**, 961–79.

Moore, W.E.C. and Holdeman, L.V. (1974 *b*). Special problems associated with isolation and identification of intestinal bacteria in fecal flora studies. *American Journal of Clinical Nutrition*, **27**, 1450–5.

Moore, W.E.C. and Holdeman, L.V. (1975). Discussion of current bacteriological investigations of the relationships between intestinal flora, diet, and colon cancer. *Cancer Research*, **35**, 3418–20.

Morens, D.M., Zweighaft, R.M., and Bryan, J.M. (1979). Non-polio enterovirus disease in the United States. *International Journal of Epidemiology*, **8**, 49–54.

Möttönen, T., Hannonen, P., Seppälä, O., Alftan, G., and Oka, M. (1984). Gluthatione and selenium in rheumatoid arthritis. *Clinical Rheumatology*, **3**, 195–200.

Mørk-Hansen, T., Lerche, A., Kassis, V., Lorenzen, I., and Søndergaard, J. (1983). Treatment of rheumatoid arthritis with prostaglandin E_1 precursors *cis*-linoleic acid and -linolenic acid. *Scandinavian Journal of Rheumatology*, **12**, 85–8.

Neumann, V. and Wright, V. (1983). Arthritis associated with bowel disease. *Clinics in Gastroenterology*, **12**, 767–95.

Neumann, V.C., Shinebaum, R., Cooke, E.M., and Wright, V. (1987). Effects of sulphasalazine on faecal flora in patients with rheumatoid arthritis: a comparison with penicillamine. *British Journal of Rheumatology*, **26**, 334–7.

Numo, R. (1992). Prevention of NSAID-induced ulcers by the coadministration of misopristol: implications in clinical practice. *Scandinavian Journal of Rheumatology*, **92** (Suppl.), 25–9.

O'Driscoll, B.R.C., Milburn, H.J., Kemeny, D.M., Cochrane, G.M., and Panayi, G.S. (1985). Atopy and rheumatoid arthritis. *Clinical Allergy*, **15**, 547–53.

O'Farrelly, C. *et al*. (1988). Association between villous atrophy in rheumatoid arthritis and a rheumatoid factor and gliadin-specific IgG. *Lancet*, **ii**, 810.

Olhagen, B. (1976). Intestinal *Cl. perfringens* in arthritis and allied conditions. In *Infection and immunology in the rheumatic diseases* (ed. D.C. Dumonde), pp. 141–5. Blackwell Scientific Publications, Oxford.

Olhagen, B. (1987). Arthritis in pigs induced by dietary factors. In *Spondylarthropathies: involvement of the gut* (ed. H. Mielants and E.M. Veys), pp. 47–60. Elsevier Science Publishers B.V, (Biochemical Division), Amsterdam.

Olhagen, B. and Månsson, I. (1968). Intestinal *Clostridium perfringens* in rheumatoid arthritis and other collagen diseases. *Acta Medica Scandinavica*, **184**, 395–402.

Paganelli, R., Levinsky, R.J., Brostoff, J., and Wraith, D.G. (1979). Immune complexes containing food proteins in normal and atopic subjects after oral challange and effect of sodium chromoglycate on antigen absorption. *Lancet*, **i**, 1271.

Palmblad, J., Hafström, I., and Ringertz, B. (1991). Anti-rheumatic effects of fasting. *Rheumatic Disease Clinics of North America*, **17**, 351–62.

Panush, R.S. (1984). Controversial arthritis remedies. *Bulletin of the Rheumatic Disease*, **34**, 1–10.

Panush, R.S. (1986). Delayed reactions to foods, food allergy and rheumatic disease. *Annals of Allergy*, **56**, 500–3.

Panush, R.S. (1987). Nutritional therapy for rheumatic diseases. (Editorial.) *Annals of Internal Medicine*, **106**, 619–21.

Panush, R.S. (1990). Food induced ('allergic') arthritis: clinical and serologic studies. *Journal of Rheumatology*, **17**, 291–4.

Panush, R.S. (1991). Does food cause or cure arthritis? *Rheumatic Disease Clinics of North America*, **17**, 259–72.

Panush, R.S., Carter, R.L., Katz, P., Kowsari, B., Longley, S., and Finnie, S. (1983). Diet therapy for rheumatoid arthritis. *Arthritis and Rheumatism*, **26**, 462–71.

Panush, R.S., Stroud, R.M., and Webster, A.M. (1986). Food induced (allergic) arthritis. *Arthritis and Rheumatism*, **29**, 220–6.

Parke, A.L. (1993). Gastrointestinal disorders and rheumatic diseases. *Current Opinion in Rheumatology*, **5**, 79–84.

Parke, A.L. and Hughes, G.R.V. (1981). Rheumatoid arthritis and food: a case study. *British Medical Journal*, **282**, 2027–9.

Parke, A.L. and Parke, D.V. (1994). Hepatic disease, the gastrointestinal tract, and rheumatic disease. *Current Opinion in Rheumatology*, **6**, 85–94.

Parke, A.L., Fagan, E.A., Chadwick, V.S., and Hughes, G.R.V. (1984). Coeliac disease and rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **43**, 478–80.

Parsonnet, J. (1992). Gastrointestinal microbiology. In *Encyclopedia of microbiology* (ed. J. Lederberg), pp. 255–8. Academic Press, San Diego, California.

Pearson, D.J. (1991). Pseudo-food allergy. *Rheumatic Disease Clinics of North America*, **17**, 343–9.

Peltonen, R., Ling, W.H., Hanninen, O., and Eerola, E. (1992). An uncooked vegan diet shifts the profile of human fecal microflora: computerized analysis of direct stool sample gas-liquid chromatography profiles of bacterial cellular fatty acids. *Applied Environmental Microbiology*, **58**, 3660–6.

Peltonen, R., Toivanen, P., and Eerola, E. (1993). Effect of a non-steroidal anti-inflammatory drug, naproxen, on faecal microbial flora. *British Journal of Rheumatology*, **32**, 996–9.

Peltonen, R. *et al*. (1994). Changes of faecal flora in rheumatoid arthritis during fasting and one-year vegetarian diet. *British Journal of Rheumatology*, **33**, 638–43.

Peretz, A., Neve, J., Vertongen, F., Famaey, J.-P., and Molle, L. (1987). Selenium status in relation to clinical variables and corticosteroid treatment in rheumatoid arthritis. *Journal of Rheumatology*, **14**, 1104–7.

Perez-Maceda, B., Lopez-Bote, J.P., Langa, C., and Bernabeu, C. (1991). Antibodies to dietary antigens in rheumatoid arthritis—possible molecular mimicry mechanism. *Clinica Chimica Acta*, **203**, 153–65.

Phillips, P.E. (1988). Evidence implicating infectious agents in rheumatoid arthritis and juvenile rheumatoid arthritis. *Clinical and Experimental Rheumatology*, **6**, 87–94.

Phillips, P.E. (1989). How do bacteria cause chronic arthritis? (Editorial.) *Journal of Rheumatology*, **16**, 1017–19.

Pike, M.C. (1989). Anti-inflammatory effects of dietary lipid modification. (Editorial.) *Journal of Rheumatology*, **16**, 718–19.

Pullman-Mooar, S. *et al*. (1990). Alteration of the cellular fatty acid profile and the production of eicosanoids in human monocytes by gamma-linolenic acid. *Arthritis and Rheumatism*, **33**, 1526–33.

Rabinovitz, M. and Van Thiel, D.H. (1992). Hepatotoxicity of nonsteroidal anti-inflammatory drugs. *American Journal of Gastroenterology*, **87**, 1696–704.

Rangachari, P.K. and Kean, W.F. (1989). Gold and D-penicillamine and the gastrointestinal tract. *Baillière's Clinical Rheumatology*, **3**, 411–23.

Ratner, D., Eshel, E., and Vidger, K. (1985). Juvenile rheumatoid arthritis and milk allergy. *Journal of the Royal Society of Medicine*, **78**, 410–13.

Recht, L., Helin, P., Rasmussen, J.O., Jacobsen, J., Lithman, T., and Schersten, B. (1990). Hand handicap and rheumatoid arthritis in a fish-eating society (the Faroe Islands). *Journal of Internal Medicine*, **227**, 49–55.

Relman, D.A. (1993). The identification of uncultured pathogens. *Journal of Infectious Diseases*, **168**, 1–8.

Relman, D.A., Schmidt, T.M., MacDermott, R.P., and Falkow, S. (1992). Identification of the uncultured bacillus of Whipple's disease. *New England Journal of Medicine*, **327**, 293–301.

Ringertz, B. (1991). Dietary treatment of rheumatoid arthritis. (Editorial.) *Annals of Medicine*, **23**, 1–2.

Rogers, P., Hassan, J., Bresnihan, B., Feighery, C., and Whelan, A. (1988). Antibodies to *Proteus* in rheumatoid arthritis. *British Journal of Rheumatology*, **27** (Suppl. 2), 90–4.

Ross, C.B., Scott, H.W., and Pincus, T. (1989). Jejunoileal bypass arthritis. *Baillière's Clinical Rheumatology*, **3**, 339–54.

Runyon, B.A., LaBreque, D.R., and Anuras, S. (1980). The spectrum of liver disease in systemic lupus erythematosus. *American Journal of Medicine*, **69**, 187–94.

Russell, A.S. (1988). *Klebsiella* and ankylosing spondylitis. (Editorial.) *Clinical and Experimental Rheumatology*, **6**, 1–2.

Ryan, J.C. and Sleisenger, M.H. (1993). Effects of systematic and extraintestinal disease on the gut. In *Gastrointestinal disease* (ed. M.H. Sleisenger and J.S. Fordtran), W.B. Saunders Company, Philadelphia.

Sapico, F.I., Emori, H., Smith, L.D.S., Bluestone, R., and Finegold, S.M. (1973). Absence of relationship of fecal *Clostridium perfringens* to rheumatoid arthritis and rheumatoid variants. *Journal of Infectious Diseases*, **128**, 559–62.

Sartor, B. (1989). Importance of intestinal mucosal immunity and luminal bacterial cell wall polymers in the aetiology of inflammatory joint diseases. *Baillière's Clinical Rheumatology*, **3**, 223–45.

Savolainen, H.A. and Isomäki, H.A. (1993). Decrease in the number of deaths from secondary amyloidosis in patients with juvenile rheumatoid arthritis. *Journal of Rheumatology*, **20**, 1201–3.

Scarpa, R. *et al*. (1992). The arthritis of ulcerative colitis: clinical and genetic aspects. *Journal of Rheumatology*, **19**, 373–7.

Schattenkirchner, M. (1990). An updated safety profile of etodolac in several thousand patients. *European Journal of Rheumatology and Inflammation*, **10**, 56–65.

Scheiman, J.M. (1994). NSAID-induced peptic ulcer disease: a critical review of pathogenesis and management. *Digestive Diseases*, **12**, 210–22.

Scherak, O. and Kolarz, G. (1991). Vitamin E and rheumatoid arthritis. *Arthritis and Rheumatism*, **34**, 1205–6.

Schmidt, T.M., DeLong, E.F., and Pace, N.R. (1991). Analysis of a marine picoplankton community by 16S rRNA gene cloning and sequencing. *Journal of Bacteriology*, **173**, 4371–8.

Schoen, R.T. and Vender, R.J. (1989). Mechanisms of nonsteroidal anti-inflammatory drug-induced gastric damage. *American Journal of Medicine*, **86**, 449–58.

Schwab, J.H. (1993). Phlogistic properties of peptodoglycan–polysaccharide polymers from cell walls of pathogenic and normal-flora bacteria which colonize humans. *Infectious Immunology*, **61**, 4535–9.

Schwimmbeck, P.L. and Oldstone, M.B.A. (1989). *Klebsiella pneumoniae* and HLA-B27-associated diseases of Reiter's syndrome and ankylosing spondylitis. In *Current topics in microbiology and immunology*, pp. 45–56. Springer Verlag, Berlin.

Segal, A.W., Isenberg, D.A., Hajirousou, V., Tolfree, S., Clark, J., and Snaith, M.L. (1986). Preliminary evidence for the gut involvement in the pathogenesis of rheumatoid arthritis. *British Journal of Rheumatology*, **25**, 162–6.

Seignalet, J. (1992). Diet, fasting, and rheumatoid arthritis. *Lancet*, **i**, 68.

Seignalet, J., Pauthe, C., Reynier, J., and Moens, P. (1989). Résultats préliminaires d'un régime sans blé et sans lait dans la polyarthrite rhumatoïde. *La Presse Médicale*, **18**, 1931–2.

Severijnen, A.J., Hazenberg, M.P., and van de Merwe, J.P. (1988). Induction of chronic arthritis in rats by cell wall fragments of anerobic coccoid rods isolated from the faecal flora of patients with Crohn's disease. *Digestion*, **39**, 118–25.

Severijnen, A.J., van Kleef, R., Hazenberg, M.P., and van de Merwe, J.P. (1989). Cell wall fragments from major residents of the human intestinal flora induce chronic arthritis in rats. *Journal of Rheumatology*, **16**, 1061–8.

Severijnen, A.J., van Kleef, R., Hazenberg, M.P., and van de Merwe, J.P. (1990). Chronic arthritis induced in rats by cell wall fragments of *Eubacterium* species from the human intestinal flora. *Infectious Immunology*, **58**, 523–8.

Severijnen, A.J., van Kleef, R., Grandia, A.A., van der Kwast, T.H., and Hazenberg, M.P. (1991). Histology of joint inflammation induced in rats by cell wall fragments of the anaerobic intestinal bacterium *Eubacterium aerofaciens. Rheumatology International*, **11**, 203–8.

Shahani, K.M., and Ayebo, A.D. (1980). Role of dietary lactobacilli in gastrointestinal microecology. *American Journal of Clinical Nutrition*, **33**, 2448–57.

Shaunak, S., Munro, J.M., Weinbrein, K., Walport, M.J., and Cox, T.M. (1988). Cyclophosphamide-induced liver necrosis: a possible interaction with azathioprine. *Quarterly Journal of Medicine*, **67**, 309–17.

Shinebaum, R., Neumann, V.C., Cooke, E.M., and Wright, V. (1987). Comparison of faecal florae in patients with rheumatoid arthritis and controls. *British Journal of Rheumatology*, **26**, 329–33.

Silman, A.J. (1991). Is rheumatoid arthritis an infectious disease? (Editorial.) *British Medical Journal*, **303**, 200–1.

Silverstein, F.E. *et al*. (1995). Misoprostol reduces serious gastrointestinal complications in patients with rheumatoid arthritis receiving nonsteroidal anti-inflammatory drugs. *Annals of Internal Medicine*, **123**, 241–9.

Simkin, P.A. (1976). Oral zinc sulphate in rheumatoid arthritis. *Lancet*, **ii**, 539–42.

Sköldstam, L. (1986). Fasting and vegan diet in rheumatoid arthritis. *Scandinavian Journal of Rheumatology*, **15**, 219–23.

Sköldstam, L. (1989). Vegetarisk kost och reumatoid artrit. Kan hälsokost påverka sjukdomen? *Nordisk Medicin*, **104**, 112–14.

Sköldstam, L. and Magnusson, K.-E. (1991). Fasting, intestinal permeability, and rheumatoid arthritis. *Rheumatic Disease Clinics of North America*, **17**, 363–71.

Sköldstam, L., Larsson, L., and Lindström, F.D. (1979). Effects of fasting and lactovegetarian diet on rheumatoid arthritis. *Scandinavian Journal of Rheumatology*, **8**, 249–55.

Sköldstam, L., Lindström, F.D., and Lindblom, B. (1983). Impaired conA suppressor cell activity in patients with rheumatoid arthritis shows normalization during fasting. *Scandinavian Journal of Rheumatology*, **12**, 369–73.

Sköldstam, L., Börjesson, O., Kjällman, A., Seiving, B., and Åkesson, B. (1992). Effect of six months of fish oil supplementation in stable rheumatoid arthritis: double-blind, controlled study. *Scandinavian Journal of Rheumatology*, **21**, 178–85.

Smith, M.D., Gibson, R.A., and Brooks, P.M. (1985). Abnormal bowel permeability in ankylosing spondylitis and rheumatoid arthritis. *Journal of Rheumatology*, **12**, 299–305.

Smith, W.S. (1990). Infectious arthritis. In *Principle and practice of infectious diseases* (ed. G.L. Mandell, R.G. Douglas, and J.E. Bennett), pp. 911–18. Churchill Livingstone, New York.

Songsiridey, N. and Furst, D.E. (1990). Methotrexate—the rapidly acting drug. *Clinical Rheumatology*, **4**, 575–94.

Sperling, R.I. (1991). Dietary omega-3 fatty acids: effects on lipid mediators of inflammation and rheumatoid arthritis. *Rheumatic Disease Clinics of North America*, **17**, 373–90.

Sperling, R.I. *et al*. (1987). Effects of dietary supplementation with marine fish oil on leucocyte lipid mediator generation and function in rheumatoid arthritis. *Arthritis and Rheumatism*, **30**, 988–97.

Stimpson, S.A. *et al*. (1986). Arthropatic properties of cell wall polymers from normal flora bacteria. *Infectious Immunology*, **51**, 240–9.

Struthers, G.R. (1986). *Clostridium perfringens* and rheumatoid arthritis. *British Journal of Rheumatology*, **25**, 419–20.

Stucki, G., Johanesson, M., and Liang, M.H. (1994). Is misoprostol cost-effective in the prevention of nonsteroidal anti-inflammatory drug-induced gastropathy in patients with chronic arthritis? *Archives of Internal Medicine*, **154**, 2020–5.

Sundqvist, T., Lindström, F., Magnusson, K.-E., Sköldstam, L., Stjernström, I., and Tagesson, C. (1982). Influence of fasting on intestinal permeability and disease activity in patients with rheumatoid arthritis. *Scandinavian Journal of Rheumatology*, **11**, 33–8.

Susuki, A. *et al*. (1993). Cause of death in 81 autopsied patients with rheumatoid arthritis. *Journal of Rheumatology*, **21**, 33–6.

Svartz, N. (1976). Discussions relating to 13–21. Principal evidence associating rheumatic diseases with microbial infection. In *Infection and immunology in the rheumatic diseases* (ed. D.C. Dumonde), pp. 165–9. Blackwell Scientific Publications, Oxford.

Tagesson, C. and Bengtsson, A. (1983). Intestinal permeability to different-sized polyethyleneglycols in patients with rheumatoid arthritis. *Scandinavian Journal of Rheumatology*, **12**, 124–8.

Tarp, U., Overvad, K., Hansen, J.C., and Thorling, E.B. (1985). Low selenium level in severe rheumatoid arthritis. *Scandinavian Journal of Rheumatology*, **14**, 97–101.

Tate, G. *et al*. (1989). Suppression of acute and chronic inflammation by dietary linolenic acid. *Journal of Rheumatology*, **16**, 729–34.

Thompson, S.T. and Elson, C.J. (1993). Susceptibility to pristane-induced arthritis is altered with changes in bowel flora. *Immunology Letters*, **36**, 227–32.

Toivanen, A. and Toivanen, P. (1991). Reactive versus rheumatoid arthritis: what is to be learned? In *Advances in rheumatology and inflammation* (ed. P. Hedqvist, J.R. Kalden, R. Muller-Peddinghaus, and D.R. Robinson), pp. 29–36. EULAR publishers, Basel.

Trang, L.E., Lövgren, O., Bendz, R., and Mjös, O. (1980). The effect of fasting on plasma cyclic adenosine-3_,5_monophosphate in rheumatoid arthritis. *Scandinavian Journal of Rheumatology*, **9**, 229–33.

Trentham, D.E. *et al*. (1993). Effects of oral administration of type II collagen on rheumatoid arthritis. *Science*, **261**, 1727–30.

Trepo, C.G., Zuckerman, A.R., Bird, R.C., and Prince, A.M. (1974). The role of circulating hepatitis B antigen/antobody immune complexes in the pathogenesis of vascular and hepatic manifestations in polyarteritis nodosa. *Journal of Clinical Pathology*, **27**, 863–8.

Tsianos, E.B. *et al*. (1990). Sjögren's syndrome in patients with primary biliary cirrhosis. *Hepatology*, **11**, 730–4.

Tulleken, J.E., Limburg, P.C., and van Rijswijk, M.H. (1988). Fish oil and plasma fibrinogen. *British Medical Journal*, **297**, 615.

Tulleken, J.E., Limburg, P.C., Muskiet, F.A.J., and van Rijswijk, M.H. (1990). Vitamin E status during dietary fish oil supplementation in rheumatoid arhtritis. *Arthritis and Rheumatism*, **33**, 1416–19.

Uden, A.-M., Trang, L., Venizelos, N., and Palmblad, J. (1983). Neutrophil functions and clinical performance after total fasting in patients with rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **42**, 45–51.

van de Merwe, J.P., Schmidt, P.I.M., and Wensick, F. (1981). Antibodies to *Eubacterium* and *Peptostreptococcus* species and the estimated probability of Crohn's disease. *Journal of Hygiene* (*Cambridge*), **87**, 25–33.

van den Broek, M.F., van Bruggen, M.C.J., Koopman, J.P., Hazenberg, M.P., and van den Berg, W.B. (1992). Gut flora induces and maintains resistance against streptococcal cell wall-induced arthritis in F344 rats. *Clinical and Experimental Immunology*, **88**, 313–17.

van der Tempel, H., Tulleken, E.E., Limburg, P.C., Muskiet, F.A.J., and van Rijswijk, M.H. (1990). Effects of fish oil supplementation in rheumatoid arthritis. *Annals of the Rheumatic Disease*, **49**, 76–80.

van Kregten, E., Huber-Bruning, O., Vandenbroucke, J.P., and Willers, J.M.N. (1991). No conclusive evidence of an epidemiological relation between *Klebsiella* and ankylosing spondylitis. *Journal of Rheumatology*, **18**, 384–8.

Veys, E.M., Mielants, H., Joos, R., and De Clercq, L. (1993). Juvenile spondylarthropathies in 1992. *Journal of Rheumatology*, **20** (Suppl. 37), 19–24.

Wands, J.R., LaMont, J.T., Mann, E., Isselbacher, B.S., and Isselbacher, K.J. (1976). Arthritis associated with intestinal-bypass procedure for morbid obesity. *New England Journal of Medicine*, **294**, 121–4.

Ward, D.M. (1990). 16S rRNA sequences reveal numerous uncultured microorganisms in a natural community. *Nature*, **345**, 63–5.

Warren, R.E. and Brewerton, D.A. (1980). Faecal carriage of klebsiella by patients with ankylosing spondylitis and rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **39**, 37–44.

Wasner, C.K., Cassady, J., and Kronenfeld, J. (1980). The use of unproven remedies. *Arthritis and Rheumatism*, **23** (Suppl.), S657–8.

Watson, J., Byars, M.L., McGill, P., and Kelman, A.W. (1993). Cytokine and prostaglandin production by monocytes of volunteers and rheumatoid arthritis patients treated with dietary supplements of blackcurrant seed oil. *British Journal of Rheumatology*, **32**, 1055–8.

Wayne, L.G. *et al*. (1987). Report of the *ad hoc* committee on reconciliation of approaches to bacterial systematics. *International Journal of Systematic Bacteriology*, **37**, 463–4.

Wensick, F., Custers-van-Lieshout, L.M.C., Poppelaars-Kustermans, P.A.J., and Schröder, A.M. (1981). The faecal flora of patients with Crohn's disease. *Journal of Hygiene* (*Cambridge*), **87**, 1–12.

Williams, R. (1981). Rheumatoid arthritis and food: a case study. *British Medical Journa*l,. **283**, 563.

Willoughby, J.M.T., Essigman, W.K., Weber, J.C., and Pinerua, R.F. (1986). Smoking and peptic ulcer in rheumatoid arthritis. *Clinical and Experimental Rheumatology*, **4**, 31–5.

Wilson, G. (1983). Bacterial ecology, normal flora and bacteriocines. The normal bacterial flora of the body. In *Topley and Wilson's principles of bacteriology*. Volume 1. *General microbiology and immunity* (ed. G. Wilson and H.M. Dick), pp. 230–46. Edward Arnold, London.

Woese, C.R. (1987). Bacterial evolution. *Microbiological Reviews*, **51**, 221–71.

Zaphiropoulos, G.C. (1986). Rheumatoid arthritis and the gut. (Editorial.) *British Journal of Rheumatology*, **25**, 138–40.

Zeidler, H. (1992). Epidemiology and economics of NSAID-induced gastropathy. *Scandinavian Journal of Rheumatology*, **92**, 3–8.

Zizic, T.M., Schulman, L.E., and Stevens, M.B. (1975). Colonic perforation in systemic lupus erythematosus. *Medicine*, **54**, 411–26.

1.3.7 The endocrine system

Chad Deal

Many endocrine disorders are associated with well-characterized rheumatic syndromes. Some of the earliest descriptions of classical endocrine diseases, such as acromegaly and hypothyroidism, focused on bone, joint, and muscle manifestations. The principal elements of connective tissue are cells and products of those cells whose growth and metabolism are influenced by hormones. Thus, endocrine glands, through hormone action, directly affect the function of the musculoskeletal system.

Thyroid disease

Hypothyroidism

Rheumatic manifestations of hypothyroidism are listed in Table 1. A characteristic arthropathy has been described by Bland and Frymoyer (1970), Dorwart and Schumacher (1975), and Golding (1970). Of 38 patients with frank hypothyroidism (Bland and Frymoyer 1970), 11 had objective rheumatic findings. Joint involvement was concurrent with the onset of hypothyroidism in 6 of 11, hypothyroidism preceded the arthropathy in 3 of 11, and the arthropathy appeared first in 2 patients. The most characteristic pattern of arthritis resulted in synovial thickening and effusion of large joints, especially the knees. All the patients had stiffness, arthralgia, swelling, and synovial thickening; 10 had synovial effusions; 6 complained of joint pain; and 3 had carpal tunnel syndrome. The effusions had a non-inflammatory appearance (no erythema or warmth), although 1 patient had definite signs of inflammation in the knees.

1. Rheumatoid arthritis-like (large joints, especially prominent in knee)
2. Pseudogout-like
3. Hyperuricaemia/gout; hyperuricaemia common, gout rare
4. Flexor tenosynovitis of the hand
5. Carpal tunnel syndrome
6. Proximal myopathy—non-inflammatory with muscle hypertrophy (Hoffman's syndrome)
7. Skeletal abnormalities in children; slipped capital femoral epiphysis

Table 1 Rheumatic manifestations of hypothyroidism

Dorwart and Schumacher (1975) describe the characteristics of synovial effusions in patients with myxoedema. Of 12 patients with florid myxoedema, 7 had knee effusions, which were most often bilateral. Analysis of synovial fluid revealed volumes of 0.5 to 100 ml, with white blood-cell counts of 50 to 6550/mm ^{3}. There were calcium pyrophosphate dihydrate crystals (both intra- and extracellular) in 6 of 7, but none had acute pseudogout at the time of aspiration. One patient subsequently developed an acute pseudogout attack 5 weeks after thyroid hormone replacement had begun. Synovial biopsies showed mild inflammation with a thickened synovium (3 to 5 layers of synovial lining cells). On radiographs, the knees of 7 of 12 patients had chondrocalcinosis with initial chondral and subchondral erosions and large cyst-like structures. These findings have been reported in the metacarpophalangeal joints also. In summary, the findings in synovial fluid were:

1. usually less than 1000 white cells/mm^3;
2. calcium pyrophosphate dihydrate crystals often present;
3. high viscosity by string test;
4. increase in hyaluronic acid;
5. high total protein in 50 per cent.

The association of hypothyroidism, true pseudogout, and pyrophosphate arthropathy has been reviewed by Alexander *et al*. (1982). Of 105 consecutive patients with pyrophosphate arthropathy, 10.5 per cent had hypothyroidism, compared with 8.1 per cent of osteoarthritic controls. A study of 100 patients with hypothyroidism showed a 17 per cent prevalence of chondrocalcinosis, which was not significantly different from the 10 per cent prevalence in 100 controls (Job-Deslandre *et al*. 1993).

Erickson *et al*. (1994) evaluated 54 consecutive crystal-proven patients with gout, prospectively for hypothyroidism using an ultrasensitive assay for thyroid-stimulating hormone. Fifteen per cent of the patients with gout compared with 4 per cent of the controls ($p < 0.05$) had elevated levels of thyroid-stimulating hormone, an average of 2.5 times higher in women and 6 times higher in men than in controls. Although hyperuricaemia is common in hypothyroidism, gout attacks are rare. In a retrospective group of 137 patients on uric acid lowering agents, 20 per cent had elevated levels of thyroid-stimulating hormone. The occurrence of gouty arthritis in patients with hypothyroidism is unclear, although urate clearance is lower than in the same patients after treatment with thyroid hormone replacement therapy. Screening for occult hypothyroidism in patients with gout is recommended.

Although knees are the most frequently involved joints in hypothyroidism, hand joints may also be affected and this may lead to an initial diagnosis of rheumatoid arthritis. Wrist, metacarpophalangeal, and proximal interphalangeal joints may be affected. Morning stiffness usually lasts less than 30 min and only a minority have signs of acute inflammation. A characteristic feature of hand involvement is a flexor tenosynovitis, present in 4 of 12 patients described by Dorwart and Schumacher (1975). Flexor tenosynovitis and thickening of the transverse carpal ligament may result in carpal tunnel syndrome. While the arthropathy usually responds to thyroid hormone replacement within 2 weeks, the flexor tenosynovitis may persist for weeks after the joint signs have resolved.

In a review of all patients presenting with carpal tunnel syndrome over a 10-year period, 5 of 49 patients had myxoedema (Frymoyer and Bland 1973). Hypothyroidism was the reported cause of carpal tunnel syndrome in up to 7.6 per cent of patients in several other series (Phalen 1966; Phalen 1970). Similarly, approximately 7 per cent of patients with hypothyroidism may have compression of the median nerve. The aetiological background to carpal tunnel syndrome is multifactorial and includes

direct pressure on the nerve from an oedematous transverse carpal ligament, flexor tenosynovitis, infiltration of the perineurium and endoneurium, and neuronal metabolic dysfunction secondary to the hypothyroid state. Acroparaesthesisas are often a prominent component of hypothyroidism, and were present in 5 of 9 patients with arthropathy and hypothyroidism described by Golding (1970), suggesting a metabolic effect of thyroid deficiency on nerve function. Most studies of carpal tunnel syndrome caused by hypothyroidism stress the rapid resolution of symptoms with thyroid hormone replacement therapy.

The diagnosis of hypothyroidism should be considered in patients who have a polymyalgia rheumatica-like picture and a normal or only slightly elevated erythrocyte sedimentation rate. Patients may also present with a fibromyalgia-like picture. These presentations may be prominent in patients who have muscular symptoms as a component of the hypothyroid state.

McLean and Podell (1995) reviewed the skeletal consequences of hypothyroidism in children. These include short stature, retarded bone age, epiphyseal dysgenesis, and delayed dental development. The physeal growth plate is affected, resulting in premature closure and abnormal cartilage and bone, which may result in osteochondritis deformans and slipped capital femoral epiphysis. Crawford *et al*. (1977) reviewed 104 cases of slipped capital femoral epiphysis and found 4 cases of hypothyroidism.

Myopathy

Muscle disease may be an important component of hypothyroidism. In addition to generalized pain and stiffness, a proximal myopathy presenting like polymyositis may be a feature. Approximately 5 per cent of all acquired myopathy is due to hypothyroidism. It has been reported that 25 per cent of patients with myxoedema had proximal muscle weakness. A raised creatine phosphokinase is common, may be marked, and is felt to be secondary to membrane leakage, not membrane disruption as in inflammatory myopathies. A study of 11 patients with hypothyroidism and proximal weakness (Khaleeli *et al*. 1983) found elevations in creatine phosphokinase ranging from 54 to 4640 IU/l (concentrations were normal in 3 of 11; raised by less than twice the normal in 2 of 11; between 440 and 1000 in 4 of 11, and greater than 1000 in 2 of 11). Muscle biopsies were abnormal in 8 of 11. The most common finding was atrophy of type-II fibres; no biopsy showed inflammatory changes. Electromyograms were myopathic in 9 of 11, and showed smaller and shorter individual action potentials than normal, an increased proportion of polyphasic potentials, and increased insertional activity. Fibrillations occurred but were less common. Thyroid hormone replacement in 7 patients with an increased concentration of creatine phosphokinase ranging from 180 to 3000 revealed that the half-life for disappearance of this enzyme was 10 to 12 days (Klein *et al*. 1980). There was often worsening of muscle cramps and stiffness during the first 2 weeks of thyroid hormone replacement. All creatine phosphokinase levels were normal by 9 weeks; this return to normal often preceded the correction of serum levels of thyroid-stimulating hormone.

Weakness, muscular stiffness, and an increase in muscle mass in an adult with myxoedema has been referred to as Hoffman's syndrome (Klein *et al*. 1981). A similar state in children presenting with a striking increase in muscle bulk is referred to as Kocher–Debre–Semelaigne syndrome. The striking increase in muscle bulk may take 6 to 10 months to resolve with thyroid hormone replacement. Myxoedema pseudomyotonia is muscle weakness associated with delayed muscle contraction as well as relaxation (delayed deep tendon reflexes). The electromyogram does not show the high frequency after discharge seen in true myotonic disorders. A prominent feature of hypothyroid myopathy may be muscle cramps, leading to the term pseudotetany.

Thyrotoxicosis

Four types of rheumatic complaints arise in the hyperthyroid state and are listed in Table 2 (Kyle and Hazelman 1981). Myopathy is very common, occurring in 70 per cent of patients, although it is seldom a presenting manifestation. There was periarthritis of the shoulder in 6.7 per cent of 300 patients described by Skillern (1943). Thyroid acropachy (thickening of the extremities) is characterized by clubbing and soft tissue swelling of the hands and feet, often painful, and by radiographs that demonstrate periosteal new bone formation, best seen on the radial aspect of the second and third metacarpals (Fig. 1). A survey of 954 patients with thyrotoxicosis found 19 with acropachy (Gimlette 1960). The syndrome most often develops in treated patients who are hypo- and euthyroid. Osteoporosis is a frequent accompaniment of excess thyroid hormone, which causes accelerated bone turnover. A marked increase in faecal and urinary excretion of both calcium and phosphorus occurs.

1. Myopathy – proximal
2. Periarthritis (shoulder)
3. Thyroid acropachy (thickened skin with periosteal new bone)
4. Osteoporosis

Table 2 Rheumatic manifestations of hyperthyroidism

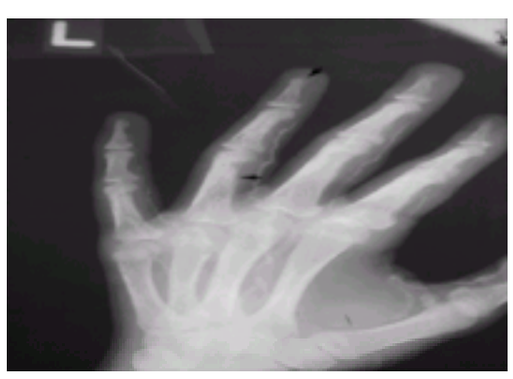


Fig. 1 Rheumatic manifestations of hyperthyroidism

Autoimmune thyroid diseases include Grave's disease and Hashimoto's thyroiditis. These diseases are associated with other disorders (rheumatoid arthritis, Sjögren's syndrome, systemic sclerosis, polymyalgia rheumatica/giant-cell arteritis, polychondritis) that are associated with HLA B8 and DR3 haplotypes. Antibodies to rheumatoid factor and DNA, as well as to thyroglobulin, are frequent in autoimmune thyroiditis. Petri *et al*. (1991) reviewed 26 patients with Grave's disease and 26 patients with Hashimoto's thyroiditis and compared them with 26 control patients for rheumatic disease symptoms and positive antinuclear antibodies (**ANAs**). No differences were noted in rheumatological symptoms when patients with thyroid disease were compared with healthy controls. Positive ANAs using either liver substrate or Hep-2 substrate were more common in patients with Grave's disease than in controls. Although 46.2 per cent of patients with Hashimoto's thyroiditis had positive ANAs using Hep-2 substrate, 26 per cent of controls had positive ANAs and the difference was not significant. Evidence for systemic autoimmune diseases was not found in patients and they did not have antibodies to DNA, SSA, SSB, Sm, or RNP. The development of hypothyroidism in patients with systemic lupus erythematosus has been reported and is often accompanied by antibodies to thyroid tissue. Although controversial, an association between autoimmune thyroid disease, systemic lupus erythematosus, rheumatoid arthritis, and Sjögren's syndrome has been reported.

Becker *et al*. (1963) found a 4 per cent incidence of rheumatoid arthritis in 506 patients with Hashimoto's thyroiditis, compared with 0.4 per cent in a control group with granulomatous thyroiditis. Mulhern *et al*. (1966) found a possible association of Hashimoto's thyroiditis with rheumatoid arthritis but not with systemic lupus

erythematosus or other connective tissue diseases.

LeRiche and Bell (1984) described a group of 15 patients with Hashimoto's thyroiditis and an inflammatory polyarthritis whose arthropathy did not respond to thyroid hormone replacement. Nine of 15 were negative for rheumatoid factor, and the other 6 were positive. These latter patients were positive for HLA DR4, had erosive diseases with nodules, and were felt to have coexistent rheumatoid arthritis and thyroiditis. The 9 patients negative for rheumatoid factor were positive for HLA DR2 and were felt to represent a distinct subset of an inflammatory arthritis associated with thyroiditis. Ninety-one consecutive women with rheumatoid arthritis were evaluated for thyroid dysfunction and compared with 93 controls with osteoarthritis or fibromyalgia. A similar prevalence of Grave's disease and multinodular goitre was seen in the patients and controls, but patients with rheumatoid arthritis had an excess of hypothyroidism and Hashimoto's thyroiditis (30 per cent compared with 11 per cent) (Shiroky *et al*. 1993).

Serological studies have shown an increased incidence of autoimmune thyroiditis in Sjögren's syndrome. Thyroid disease (defined as current or previous clinically overt disease requiring treatment) was found in 10 per cent of patients with primary Sjögren's syndrome and their family members (Foster *et al*. 1993). There are antithyroglobulin and antimicrosomal antibodies in 40 per cent of patients with Sjögren's syndrome (Block and Bunim 1963). Polymyalgia rheumatica or giant-cell arteritis were found in 7 of 250 patients with autoimmune thyroid disease by Dent and Edwards (1978). No association between autoimmune thyroid disease and giant-cell arteritis was discovered in a small prospective study (Barrier *et al*. 1993). In this study, one-quarter of patients with giant-cell arteritis had thyroid dysfunction and/or antibodies compared with one-third of controls. This is in contrast to a larger study by Bowness *et al*. (1991), who reported an increased prevalence of hypothyroidism in patients with polymyalgia rheumatica and giant-cell arteritis.

Hyperparathyroidism

The rheumatic manifestations of hyperparathyroidism are as follows:

1. arthritis/soft tissue:
   a. pseudogout/chondrocalcinosis;
   b. rheumatoid arthritis-like;
   c. azotaemic arthropathy;
   d. shoulder arthropathy;
   e. tendon, capsular, ligament laxity with rupture;
   f. hyperuricaemia (gout is rare);
2. bone/cartilage:
   a. resorption (subperiosteal most diagnostic);
   b. osteoporosis;
   c. chondrocalcinosis;
3. muscle—proximal myopathy.

Musculoskeletal symptoms are reported to be the initial manifestation in up to 16 per cent of patients with primary hyperparathyroidism.

Hyperparathyroidism, chondrocalcinosis, and pseudogout frequently coexist. McCarty *et al*. (1974) reported that 17 (7.2 per cent) of 238 patients with chondrocalcinosis had primary hyperparathyroidism. Conversely, chondrocalcinosis reportedly occurs in 18 to 40 per cent of patients with hyperparathyroidism (less frequently in secondary hyperparathyroidism), and clinical pseudogout occurred in only 1 to 10 per cent of these patients (Bland *et al*. 1979). Pseudogout attacks are frequent within the first few days after parathyroidectomy as the serum calcium falls, chondrocalcinosis persists and attacks of pseudogout continue. A second type of arthropathy has been described by Bywaters *et al*. (1963) in primary hyperparathyroidism and Rubin *et al*. (1984) in secondary hyperparathyroidism, both resemble rheumatoid arthritis. The arthropathy described by Bywaters resembled rheumatoid arthritis in that patients had articular symptoms in the knees, wrists, hand, and shoulders, and radiographic erosions. However, the radiographic changes differ from those of rheumatoid arthritis—as synovial proliferation is absent, the joint space is preserved, pericapsular calcifications are frequently present, erosions have a predilection for the ulnar side of the joint (radial in rheumatoid arthritis), and reactive bone formation with an osteoarthritic appearance eventually develops. The erosions are a result of the direct effect of parathyroid hormone on subchondral bone, and not of synovitis as in rheumatoid arthritis.

Patients with chronic renal failure and secondary hyperparathyroidism on dialysis, in addition to developing the classical changes of renal osteodystrophy, may develop erosive changes resembling rheumatoid arthritis (Fig. 2) Rubin *et al*. (1984) described their findings in 59 patients with chronic renal failure on dialysis; 12 had renal osteodystrophy and erosions, 11 had osteodystrophy without erosions, and 36 had neither. The rheumatoid-like erosions occurred with decreasing frequency in the metacarpophalangeal, distal interphalangeal, proximal interphalangeal, shoulder, wrist, and knee joints. These patients experiences episodes of subacute, often symmetrical, arthritis affecting the metacarpophalangeal, proximal interphalangeal, wrist, and knee joints. Swelling and synovial thickening were seen in 8 of 12 patients with erosive changes. Analysis of synovial fluid from the knees of 5 patients revealed white cell counts of 50 to 1000/mm ^{3}; no crystals were seen. Synovial biopsies showed mild hyperplasia of lining cells (2 to 6 layers).

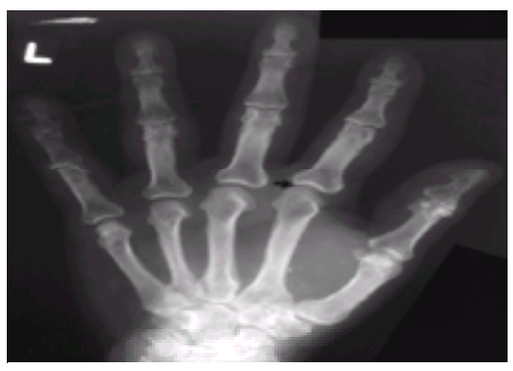


Fig. 2 Thyroid acropachy: note the characteristic lacy subperiosteal new bone in the diaphyses of short tubular bones of the hand.

A third type of arthropathy that may develop in primary or secondary hyperparathyroidism is a shoulder arthropathy with intra-articular and periarticular erosions of the head of the humerus (Nussbaum and Doppman 1982). No calcifications developed and some patients had radiographic findings without clinical symptoms. The effects of parathyroid hormone on soft tissue, capsule, and ligaments may lead to tenderness at the sites. There are numerous reports of hyperuricaemia and gouty arthritis in hyperparathyroidisim (Bland *et al*. 1979).

Neuromuscular disease is well described in primary hyperparathyroidism. Although subjective muscle weakness and fatigue are common, objective weakness is less frequent, and was present in 3 of 76 patients with surgically confirmed hyperparathyroidism reported by Frame *et al*. (1968). Patients have a symmetrical proximal myopathy, mostly in the lower extremities. Fatiguability is a common complaint. Muscle enzymes are normal, and muscle biopsy shows atrophy of type II fibres, often with a neurogenic component. The hallmarks of an inflammatory myopathy (necrosis, regeneration of muscle fibres, with inflammatory cells) are absent. Electromyographic findings are suggestive of both a myopathic (short polyphasic potentials) and neuropathic (long, amplitude polyphasic potentials) process; motor conductions are normal. Other neurological findings include hyperreflexia, tongue fasciculations, gait abnormalities, and paraesthesias (Patten *et al*. 1974). Patients with secondary hyperparathyroidism may have a similar picture. Those with mild, asymptomatic, primary hyperparathyroidism do not develop muscular weakness but do have muscle cramps (52 per cent) and paraesthesias (29 per cent).

Radiographic findings in hyperparathyroidism are as follows:

1. bone resorption:
   a. subperiosteal, intracortical, endoseal, subchondral, trabecular;

   b. subligamentous, localized (brown tumours);
2. bone sclerosis, periostitis;
3. chondrocalcinosis.

Subperiosteal resorption is virtually diagnostic of hyperparathyroidism and most frequently involves the radial aspect of the second and third phalanges. Brown tumours, representing large accumulations of osteoclasts–osteoclastomas and chondrocalcinosis, are most common in primary than secondary hyperparathyroidism, while osteocleosis and periostitis are less frequent.

Hyperparathyroidism is associated with accelerated bone loss, osteoporosis, and an increased fracture risk. However, patients with mild primary hyperparathyroidisim may not develop significant osteoporosis (Rao *et al*. 1988). Bone density measurements are useful in identifying those patients with primary hyperparathyroidism who are losing bone. In a study of 90 patients with primary hyperparathyroidism, fractures prior to the diagnosis of hyperparathyroidism occurred in 30 per cent of patients compared with 18 per cent of controls (Melton *et al*. 1992). Hyperparathyroidism may have differential effects on cortical and trabecular bone, with greater losses in cortical bone (Wishart *et al*. 1990). Surgery only partially reverses bone loss; in 14 patients undergoing parathyroid surgery, 13 showed an increase in bone mass in the lumbar spine (Minisola *et al*. 1993). Martin *et al*. (1990) followed 71 patients with primary hyperparathyroidism after operation. Bone mineral content increased during the first year after surgery, with the greatest gain in those with the lowest preoperative bone mass. However, bone mineral content of the radius remained more than 1 standard deviation below the mean in 61 per cent of patients at 1 year.

Hypoparathyroidism

Spondylitis has been described as a feature of hypoparathyroidism. However, sacroiliac involvement is absent and ankylosis is secondary to extensive calcification of paraspinal ligaments (anterior longitudinal ligament and posterior paraspinal ligament). Calvareal thickening and osteosclerosis have been described (Chaykin *et al*. 1969). Patients with pseudohypoparathyroidism and pseudopseudohypoparathryoidism, in addition to shortening of metacarpals and metatarsals, may have extensive soft tissue calcifications, exostoses, and bowing deformities.

Acromegalic arthropathy

Rheumatological manifestations of acromegaly are as follows:

1. arthropathy:
   a. bursal hyperplasia;
   b. synovial proliferation, prominent in knees;
   c. cartilage hyperplasia, widened joint spaces;
   d. bony proliferation;
   e. an osteoarthritis-like picture;
   f. backache;
   g. hypermobility;
2. carpal tunnel syndrome (50 per cent);
3. proximal muscle weakness (50 per cent);
4. Raynaud's phenomenon.

They are the result of chronic overstimulation of osteoblasts, osteoclasts, fibroblasts, chondrocytes, and muscle cells by growth hormone. Stimulation of chondroitin sulphate and collagen production by chrondrocytes is due to a hepatically derived serum factor called somatomedin (insulin-like growth factor) induced by growth hormone. Lieberman *et al*. (1992) reviewed the rheumatological and skeletal changes in acromegaly, including the effects of insulin-like growth factor 1 on articular cartilage and the effects of chronic somatotropism on the local production of insulin-like growth factor 1 in cartilage. Joint complaints usually arise well after the onset of clinical acromegaly, often 10 years after diagnosis. Large joints (knees, shoulders) are most often affected, and examination reveals synovial proliferation in knees and enlarged hands resulting from bony proliferation. Morning stiffness is not prominent, and joint swelling is present in less than one-half of patients. Back pain is frequent, and carpal tunnel syndrome is present in up to one-half and is usually bilateral. Bursas are often enlarged, especially prepatellar, olecranon, and subacromial.

Three large series of patients with acromegaly have been published by Kellgren *et al*. (1952), Bluestone *et al*. (1971), and Detenbeck *et al*. (1973). Bluestone described 42 patients, 16 with subjective complaints of joint pain (usually in shoulders, knees, hands, or hips) but without objective joint abnormality. While 26 had peripheral joint abnormalities on examination, only 9 of 26 patients had joint swelling, which was most prominent in the knees and consisted of synovial hypertrophy more often than effusions. Twelve had painless bony swelling of distal and proximal interphalangeal, and metacarpophalangeal joints. There was painful limitation of joint motion in 8 patients, often affecting the shoulders. Asymptomatic, coarse crepitance was present in over one-half of patients, usually affecting shoulders and knees. Twenty-two of 42 patients had carpal tunnel syndrome, which was bilateral in 20; 10 patients had mild Raynaud's phenomenon. Axial involvement was present in 20 patients who gave a history of non-traumatic backache, usually in the lumbar spine. Detenbeck evaluated 229 patients with acromegaly; 62 per cent had minor joint complaints; 16 per cent had symptoms that were severe and warranted radiographic evaluation. Kellgren found that 16 of 25 patients with acromegaly had joint complaints. Painless weakness of proximal muscles with normal muscle enzymes may occur in acromegaly. The electromyogram is usually normal; muscle biopsy reveals a myopathic pattern in one-half of patients (Khaleeli *et al*. 1984).

Radiographs are helpful, although the diagnosis is made on clinical grounds (Table 3). Characteristically, cartilage hypertrophy leads to widened joint spaces, most marked in the metacarpophalangeal joints, where the cartilage space is greater than 2.5 mm (Fig. 3). The distal tufts enlarge from the base and tip, and in advanced cases, the bone may meet and form a tunnel. Sesamoids may undergo hypertrophy, and in the thumb, a sesamoid index over 40 (obtained by multiplying the longest diameter by the width) suggests acromegaly. Lateral radiographs of the foot reveal a thickened heel pad (Fig. 4). A thickness of greater than 23 mm in men and 21.5 mm in women is abnormal. The diagnosis of acromegaly should be questioned in the absence of a thick heel pad. Obesity, infection, oedema, and myxoedema may also thicken the heel pad.

Measurement	Men	Women
Heel pad thickness (mm)	>23	>21.5
Sesamoid index (thumb)	>40	>32
Tuft width (mm) (3rd finger)	>12	>10
Joint space thickness (mm) (metacarpophalangeal)	>2.5	>2.5

From Resnick and Niwayama (1988).

Table 3 Bone and soft tissue measurements suggestive of acromegaly

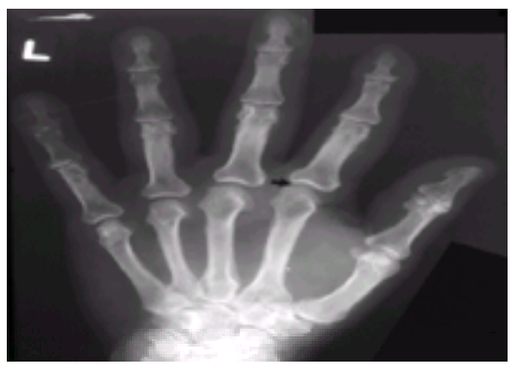


Fig. 3 Bone and soft tissue measurements suggestive of acromegaly.

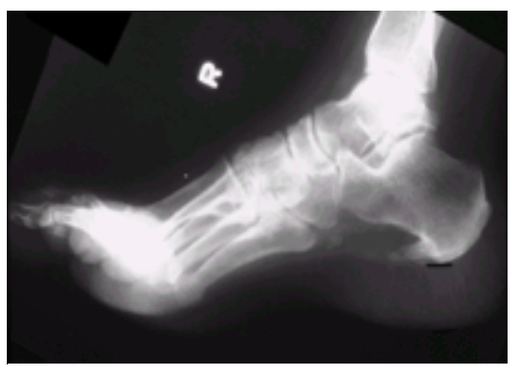


Fig. 4 Bone and soft tissue measurements suggestive of acromegaly.

Correlation of radiographic and pathological findings in acromegaly has been summarized. Stimulation of endochrondral bone formation produces enlargement of costochondral junctions (acromegalic rosary) and thickened intervertebral discs. Stimulation of periosteal bone produces mandibular enlargement, a thickened cranial vault, prominent supraorbital ridges and facial structures, cortical thickening of tubular bones, enlargement of phalangeal tufts, and an increase in the anteroposterior diameter of vertebral bodies. Stimulation of subligamentous bone formation produces calcaneal spurs, and excrescences on the patella, trochanters, and tuberosities. Stimulation of bone resorption produces intracortical striations, widening of medullary cavities, and vertebral scalloping. Proliferation of articular cartilage causes widened articular spaces. Later, cartilage degeneration produces narrowed joint spaces and osteoarthritis with osteophytes. Connective tissue hyperplasia causes increased skin thickness (heel pad sign).

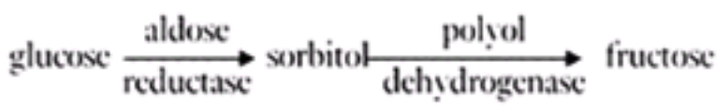
Pituitary

Pituitary and hypothalamic abnormalities may play a role in modulating autoimmune disease. Hyperprolactinaemia has been shown to exacerbate disease activity in the NZB × NZW mouse model of systemic lupus erythematosus. Jara-Quezada *et al*. (1991) reported elevated prolactin levels were associated with disease flares in pregnant patients with lupus. Elevated prolactin levels were found in 10 of 45 patients with lupus compared with 3 per cent of controls (Jara *et al*. 1992). A number of medications that elevate prolactin levels (chlorpromazine, aldomet, dilantin) have been associated with production of autoantibodies. Treatment with bromocriptine, an agent which lowers prolactin levels, has been used in NZB × NZW mice and in isolated case reports in humans with systemic lupus erythematosus, and may have immunosuppressive properties (McMurray *et al*. 1994).

Diabetes

Syndrome of limited joint mobility

The musculoskeletal manifestations of diabetes mellitus are diverse (Table 4), and divided into those that are related to the underlying pathological process of glycosylation of proteins and those that are secondary to the more primary diabetic complications. The syndrome of limited joint mobility, also called diabetic stiff-hand syndrome, was first described by Jung *et al*. (1971). It affects predominantly juveniles with type I diabetes mellitus but also adults with type I and II. Patients with this syndrome experience a slow decrease in mobility of the hands, characterized by a waxy thickening of the skin. The prevalence ranges from 8.4 to 58 per cent; in most series, 30 to 40 per cent of those with insulin-dependent diabetes mellitus develop the syndrome of limited joint mobility. While it occasionally may present before the onset of overt diabetes, the skin changes are more frequent in those with long-standing disease. Rheumatological interest in this syndrome relates to its scleroderma-like appearance, including occasional instances of pulmonary fibrosis and restrictive lung disease. Patients with the syndrome are more likely to have the microvascular complications of diabetes. The skin changes are felt to result from the enzymatic production of excessive sugar alcohols (polyols, such as sorbitol) in connective tissue, causing excessive tissue hydration and increased water content, which seems to parallel the increased stiffness of the skin. The sorbitol dehydrogenase pathway enzymes are saturated, resulting in increased production of sorbitol. Treatment with sorbinil, an inhibitor of the enzyme aldose reductase, has been reported to decrease skin thickening and improve mobility of the hands (Eaton *et al*. 1985):



Intrinsic complications of diabetes mellitus	Increased incidence in diabetes mellitus	Possible associations
Neuropathic joints	Dupuytren's contracture	Gout
Foot disease	Trigger finger	Calcium pyrophosphate deposition disease
Osteolysis	Carpal tunnel syndrome	
Syndrome of limited joint mobility	Shoulder periarthritis	Osteoporosis
Diabetic amyotrophy	Reflex sympathetic dystrophy	Osteoarthritis
Autoimmune insulin resistance		Autoimmune diseases
Diabetic muscular infarction	Diffuse idiopathic skeletal hyperostosis	

From Siebert (1988)

Table 4 Musculoskeletal complications of diabetes mellitus

The changes in skin texture predominate distally with finger involvement, but thickened palmar fascia (which may be misdiagnosed as Dupuytren's contractures, another problem in diabetes) also occurs. The skin changes have been likened to scleroderma and produce what is called the prayer sign; patients cannot closely appose the palms of their hands. They have mild morning stiffness but not synovitis; sensory abnormalities are mild or absent; and the electromyogram shows a mild decrease in conduction times. Radiographs are normal (Kapoor and Sibbitt 1989).

Several large reviews of the syndrome of limited joint mobility have been published. In a review of 137 patients aged between 1 and 24 years of age with

insulin-dependent diabetes mellitus (Seibold 1982), 47 (34 per cent) had palpable induration and thickening of the skin. Of these 47, only 6 had changes proximal to the metacarpophalangeal joints (thus meeting the American College of Rheumatology criteria for systemic sclerosis); 14 had changes at the metacarpophalangeal joints as well as distally, and 27 had changes distal to the proximal interphalangeal joints. Skin changes were graded slight in 25, moderate in 19, and severe in 3 patients. No patient had skin changes proximal to the wrists or on the face. Joint contractures were present in 26 of 47 (23 were symmetrical), were most common in the middle and ring fingers, and were usually angled at 5 to 10°. No patient had flexor tendon rubs as seen in systemic sclerosis. Rosenbloom *et al*. (1981) found the syndrome in 30 per cent of 309 cases of diabetes mellitus. Of those who had had the disease for longer than 4.5 years, 82 of 169 had skin changes. This report emphasized the association of microvascular changes with the syndrome. In the 82 patients with the syndrome and disease for longer than 4.5 years, one-half had the microvascular complications of diabetes (defined as nephropathy, retinopathy, or neuropathy), while only 10 of 87 without the syndrome had such complications. The usefulness of nailfold capillary microscopy was evaluated in a large group of patients with both insulin- and non-insulin-dependent diabetes mellitus, and found mild capillary enlargement and avascular areas in a few patients but no scleroderma-like changes were noted (Gertner *et al*. 1990). Light microscopic and ultrastructural features in skin biopsies were also different in diabetic thick skin and scleroderma. In this report, limited joint mobility was significantly correlated with neuropathy in patients with both insulin- and non-insulin-dependent diabetes and with retinopathy in insulin-dependent diabetes mellitus. No correlation was found between HLA types DR3 and DR4 and limited joint mobility.

Clark *et al*. (1990) reported the syndrome of limited joint mobility in 31 per cent of 70 children with insulin-dependent diabetes mellitus; no correlation was found between the syndrome and diabetic control, or the presence of retinopathy or microalbuminuria. Interestingly, on using HgbA1C as a marker for assessment of diabetes control, patients with the syndrome did not differ from those without it. Stiffening of connective tissue, as assessed by passive extension of the metacarpophalangeal joints, was shown to be associated with diabetic nephropathy in 205 patients with mature-onset diabetes (average age 61 years) (Aoki *et al*. 1993). In a series of 361 adult patients with diabetes mellitus (Starkman *et al*. 1986), 55 per cent of patients with insulin-dependent and 76 per cent with non-insulin-dependent disease had the limited joint mobility syndrome. The syndrome was related to duration of disease only in insulin-dependent diabetes with onset before the age of 35 years. Schulte *et al*. (1993) measured passive range of motion of the upper extremities in 70 adult patients with insulin-dependent diabetes mellitus and 79 non-diabetic controls for age, sex, and activity level. They concluded that the syndrome of limited joint mobility is a generalized phenomenon which occurs throughout the upper extremity including the shoulder and elbow, with the distal interphalangeal and metacarpophalangeal joints affected more than the proximal interphalangeal joints. Motion was significantly correlated ($p < 0.05$) with age, sex, duration of diabetes, and glucose control but not correlated with the presence of diabetic complications such as neuropathy, retinopathy, nephropathy, or peripheral vascular disease.

Articular syndromes and diffuse idiopathic skeletal hyperostosis (DISH)

Diabetes mellitus has been found in a significant number of patients with diffuse idiopathic skeletal hyperostosis (**DISH**). This association has been confirmed in African black patients; one-half of the patients with DISH had diabetes mellitus (Cassim and Rubin 1990). In one series of 510 patients with diabetes mellitus, 13 per cent had DISH (Karava and Viljanen 1966). By the age of 60 to 69 years, 21 per cent of diabetics and 4 per cent of controls had radiographic findings of DISH. One study of patients with DISH showed one-half with abnormal glucose tolerance. The high prevalence of abnormal glucose tolerance tests is partly a result of an association of obesity (present in 30 of 34 patients in one study), with DISH (Harris *et al*. 1974). Many patients have high insulin concentration after glucose challenge, suggesting insulin resistance. Insulin-like growth factors may play a part in stimulation of ligamentous calcifications seen in patients with DISH.

The prevalence of calcium pyrophosphate deposition disease in diabetes is reported to range from 8 to 73 per cent. A controlled trial by McCarty *et al*. (1974), who made glucose tolerance tests in 28 patients with pseudogout and 22 controls, found no difference in the two groups. The consensus is that there is no association between diabetes mellitus and gout, although hyperuricaemia may be more common. Joslin (1952) reported only 1 instance of gout in 1500 with diabetes mellitus.

Neuropathic joints/diabetic foot

Diabetes is the single greatest cause of neuropathic joints. In a review of 101 patients with Charcot joints and diabetes, the estimated prevalence was 1 in 680 patients with diabetes or 0.15 per cent (Sinha *et al*. 1972). Two-thirds of patients were in their fifth and sixth decades, and had had disease for more than 15 years. Diabetic control was often poor and all had peripheral neuropathy. Tarsal and tarsometatarsal joints were involved in 60 per cent, metatarsophalangeal joints in 31 per cent, and ankle in 9 per cent. Trauma and diminished pain perception and position sense appear to play a major part in the development of this arthropathy. The usual presenting complaint was swelling of one foot with vague or no pain. The earliest radiographic findings are joint effusions and subchondral lucency, followed by small fractures and irregular articular surfaces. Bony repair produces periosteal and endochondral ossification. Characteristic radiographic changes include 'sucked candy' appearances, and the balanced pagoda, 'mortar-in-pestle' and 'hour-glass' deformities (Fig. 5). Often patchy demineralization is present. The differential diagnosis of patients with radiographic changes consistent with neuropathic joints include neurosyphilis (common in hips, knees, and vertebral column), syringomyelia (upper extremity), leprosy (hands with calcification in nerves), and congenital indifference to pain (usually in children). Treatment of neuropathic joints usually consists of non-weight bearing, after which the joints may ankylose. Arthrodesis of ankle joints in patients with neuropathic arthropathy has a high incidence of complication (78 per cent) and should be considered with caution (Stuart and Morrey 1990).

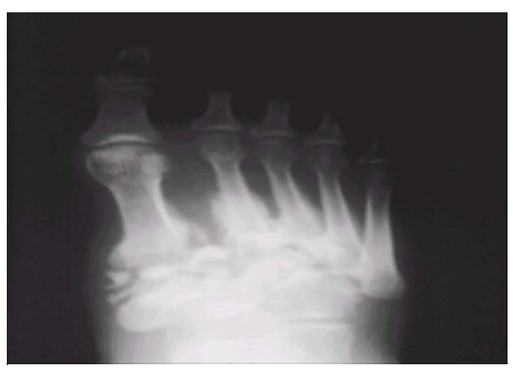


Fig. 5 Diabetic neuroarthropathy. Anteroposterior and lateral views of the right foot will demonstrate extensive destruction of the intertarsal and tarsometatarsal joints. The patient illustrated has a neuropathic Lisfranc dislocation of the tarsometatarsal joint. There is both lateral and dorsal subluxation of the metatarsals in relation to the tarsal bones. The most sensitive place to look for Lisfranc dislocation is at the second tarsometatarsal articulation. Normally the medial border of the second metatarsal will align with the medial border of the intermediate cuneiform. The lack of osteoporosis with this extensive abnormality is suggestive of neuroarthropathy. This patient has been walking on this joint for some time. Diabetic neuroarthropathy predominates in the feet and is most commonly seen at the tarsometatarsal and metatarsal phalangeal joints.

Diabetic osteolysis is characterized by osteopenia, osteolysis, and resorption of the distal metatarsals and proximal phalanges. In contrast to neuroarthropathy, there is relative sparing of the joint (Griffiths 1985). The pathogenesis is unknown but may be secondary to microtrauma in a denervated extremity. Osteomyelitis is not uncommon in diabetes and must be differentiated from cellulitis, neuroarthropathy, and osteolysis. The three-phase bone scan may prove helpful in distinguishing between cellulitis, osteomyelitis, and neuropathic joints. With osteomyelitis, there is usually increased flow in the dynamic phase, with moderate to marked uptake in the immediate and delayed static phases. Cellulitis produces minimal uptake on the delayed images, while neuropathic joints demonstrate minimal activity on dynamic and immediate images and marked uptake on delayed images (Thomasson and Sundaram 1985).

Diabetes and soft tissue rheumatism

A number of soft tissue rheumatological syndromes are more frequent in patients with diabetes, including periarthritis of the shoulder, Dupuytren's contracture of the palmar fascia, carpal tunnel syndrome, and flexor tenosynovitis (Gray and Gottlieb 1976).

A review of 800 outpatients with diabetes mellitus compared with 600 non-diabetic controls showed that 10.8 per cent with diabetes and 2.3 per cent of controls had a history of significant periarthritis of the shoulder (Bridgman 1972). Of those with diabetes and periarthritis, 36 per cent were insulin dependent compared with 23 per cent of the entire group, suggesting an enrichment of periarthritis in type I diabetes. Over 40 per cent of cases were bilateral. In another review of 200 patients with diabetes mellitus compared with 100 controls, 22 per cent with diabetes had calcifications of the shoulder on radiographs compared with 8 per cent of controls. Half of patients with shoulder calcifications had a history of symptomatic periarthritis. Calcifications were often bilateral in multiple sites in patients with diabetes but were

usually unilateral and single in controls. Reflex sympathetic dystrophy of the hand may be a sequel of periarthritis of the shoulder.

Different series have reported varying prevalences of Dupuytren's contractures in diabetes mellitus. Bridgman (1972) found no increased prevalence; other series have reported a prevalence of 3 to 63 per cent (Bland *et al*. 1979). In those series reporting an association, prevalence increases with age and duration of diabetes, but does not appear to be related to the degree of diabetic control.

Phalen (1966) reported that diabetes mellitus was the most common systemic disease associated with carpal tunnel syndrome. Sixty-three of 379 patients (16.6 per cent) with diabetes had this syndrome. Other studies have shown that 5 to 8 per cent of patients with carpal tunnel syndrome have diabetes mellitus. The cause of carpal tunnel syndrome has been suggested to be either compression of the median nerve as a result of tissue infiltration, or ischaemia of the nerve secondary to microvascular disease of the vasa nervorum (carpal tunnel release would not be indicated in these patients). Diabetic hand syndromes secondary to neuropathies result in atrophy of intrinsic muscles of the hand, which may be confused with carpal tunnel syndrome. Electrodiagnostic studies will reveal both median and ulnar nerve involvement in these patients.

Although it is estimated that 10 per cent of adults with trigger fingers are diabetic, this association is poorly documented. MacKenzie (1975) reported on 63 patients with multiple flexor tenosynovitis of the fingers, 11 of whom had, or subsequently developed, diabetes mellitus. A survey of 250 patients with insulin-dependent diabetes mellitus reported that 5 per cent had trigger fingers and this was related to duration of disease. Almost one-half of the patients had the syndrome of limited joint mobility (Yosipovitch 1990).

Diabetic muscle infarction may present as a painful muscle mass, occurs in the setting of poorly controlled diabetes, and is the result of atherosclerotic arterial narrowing. Incisional biopsy may be required to make a definite diagnosis (Rocca *et al*. 1993). Ostrov *et al*. (1995) reviewed three cases of diabetic muscle infarction. Cases usually occurred in patients with poorly controlled diabetes mellitus or significant end-organ complications. This syndrome is often mistaken for thrombophlebitis, myositis, or vasculitis. The authors suggest using magnetic resonance imaging to evaluate diabetic patients with atraumatic, painful, swollen extremities which occur over the course of days to weeks; the images usually show asymmetry of muscle groups with high-intensity 12-weighted images. The painful swelling usually resolves spontaneously. The erythrocyte sedimentation rate is often elevated to over 100 ml/h but creatine phosphokinase and aldolase may be normal. This syndrome must be differentiated from focal myositis and localized nodule myositis; in these cases, biopsy shows normal vasculature without any evidence of atherosclerotic arterial narrowing.

Type I diabetes mellitus is associated with HLA alleles DR3 and DR4 (as are many of the classic connective tissue diseases); 95 per cent of patients with insulin-dependent diabetes mellitus are positive for these alleles. These patients will often have autoantibodies to other organs, and there is an increased incidence of organ-specific autoimmune diseases, usually in those that have HLA-DR3 alleles, such as autoimmune thyroid disease. The association of insulin-dependent diabetes mellitus with other connective tissue diseases has been reported. Thomas *et al*. (1983) found that 13.2 per cent of 295 patients with rheumatoid arthritis had a first- or second-degree relative with insulin-dependent diabetes compared with 3.8 per cent of 307 controls with degenerative joint disease. Other investigators have failed to find an increased prevalence of type I diabetes mellitus in rheumatoid arthritis.

Patients with insulin resistance often develop signs or symptoms of connective tissue diseases. Two types of insulin resistance have been described, A and B. Approximately 80 per cent of patients with type B insulin resistance are women aged 30 to 50 years. The syndrome is characterized by marked insulin resistance, acanthosis nigrans, and features of connective tissue disease including arthralgias, alopecia, enlarged salivary glands, leucopenia, proteinuria, hypergammaglobulinaemia, and high titres of antinuclear and anti-DNA antibodies. Features of connective tissue disease have preceded the appearance of diabetes mellitus in one-third of patients. Individual cases have been diagnosed as Sjögren's syndrome, systemic sclerosis, mixed connective tissue disease, and systemic lupus. A review of 14 patients with type B insulin-resistant diabetes mellitus revealed 12 with autoimmune features; 5 patients had four or more criteria for the diagnosis of systemic lupus erythematosus. Autoimmune features included speckled antinuclear antibodies in all 12 patients, double-stranded DNA in 7, alopecia/nephritis/malar rash in 4, and arthritis in 3 (Tsokos *et al*. 1985).

Polyglandular autoimmune syndromes

Polyglandular autoimmune syndromes are grouped into two distinct disorders. Type I begins at a mean of 12 years of age, and is characterized by hypoparathyroidism, primary adrenal insufficiency, and mucocutaneous candidiasis. Alopecia, chronic active hepatitis, pernicious anaemia, and malabsorption also occur. Type II, also called Schmidt's syndrome, has a mean age of onset of 30 years and is characterized by primary adrenal insufficiency, autoimmune thyroid disease, and insulin-dependent diabetes (Trence *et al*. 1984). These syndromes appear to be autoimmune; 80 per cent of patients have antibodies against one or more endocrine glands. In patients with polyglandular autoimmunity, there is an increased incidence of autoimmune diseases (usually as case reports) involving non-endocrine organs including alopecia, pernicious anaemia, vitiligo, Sjögren's syndrome, rheumatoid arthritis, myasthenia gravis, primary biliary cirrhosis, chronic active hepatitis, systemic lupus, and IgA deficiency (Table 5). Patients with type-II disorders have an increased incidence of the HLA B8, DR3 haplotype. Tucker *et al*. (1987) described 20 patients with type-II polyglandular autoimmune syndrome who had serositis (pleuritis and pericarditis) as a prominent component. Four patients were felt to have systemic lupus, based on the presence of antinuclear antibody titres ranging from 1:80 to 1:640, antibodies to DNA, antibodies to SS-A (Sjögren's syndrome-A), and rheumatoid factor. The 4 patients with systemic lupus had only autoimmune thyroid disease clinically, although 1 had antibodies to adrenal cortical cells.

Endocrinopathies	Non-endocrine disease
Graves' disease	Pernicious anaemia
Hashimoto's thyroiditis	Vitiligo
Addison's disease	Myasthenia gravis
Insulin-dependent diabetes mellitus	Sjögren's syndrome
Autoimmune oophoritis or orchitis	Rheumatoid arthritis
Autoimmune hypoparathyroidism	Idiopathic thrombocytopenic purpura
Autoimmune hypophysitis	Primary biliary cirrhosis

From Volpe (1986).

Table 5 Autoimmune endocrinopathies and non-endocrine organ-specific autoimmune disorders associated with them

Abonen *et al*. (1990) have termed type-I polyglandular endocrinopathy, APECED (autoimmune polyendocrinopathy–candidiasis–ectodermal dystrophy), which emphasizes the three components of disease—endocrine failure, chronic mucocutaneous candidiasis, dystrophy of dental enamel and nails, alopecia, vitiligo, and keratopathy. A report of 68 such patients did not find systemic lupus or other classical connective tissue diseases. The largest series of polyglandular autoimmune syndrome, 71 patients with type I and 224 with type II, does not mention classical connective tissue disease in the patients (Neufeld *et al*. 1981).

POEMS syndrome (a multisystem disorder associated with polyneuropathy, organomegaly, endocrinopathy of various forms, production of a monoclonal antibody, and skin changes) was the subject of a review by Soubrier *et al*. (1994). In this retrospective study of 25 cases, scleroderma-like changes of the hands, mainly sclerodactyly, were noted in 4 of 25 cases. Two patients had Raynaud's phenomenon associated with scleroderma-like skin changes, with oesophageal hypomotility and restrictive pulmonary defect of renal failure, which were evocative of a scleroderma-like syndrome.

Diagnosis and endocrinology

Important points to remember when making a diagnosis are summarized in Box 1.

Box 1 Important points to remember

1. Rheumatic manifestations are common in endocrine disease

2. Polyarthritis, periarthritis tenosynovitis, myopathy, crystal deposition diseases, carpal tunnel syndrome, and osteoporosis are common accompaniments of endocrine disease

3. Differentiated diagnosis of rheumatic complaints should include endocrine disease and appropriate laboratory tests should be ordered (thyroid function tests, serum calcium, glucose)

Chapter References

Abonen, P., Myllarniemi, S., Sipila, I., and Perheentupa, J. (1990). Clinical variations of autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED) in a series of 68 patients. *New England Journal of Medicine*, **332**,829–36.

Alexander, G.M., Dieppe, P.A., Doherty, M., and Scott, D.G.I. (1982). Pyrophosphate arthropathy: a study of metabolic associations and laboratory data. *Annals of the Rheumatic Diseases*, **41**, 337–81.

Aoki, Y. *et al*. (1993). Stiffening of connective tissue in elderly diabetic patients: relevance to diabetic nephropathy and oxidative stress. *Diabetologia*, **36**, 79–83.

Barrier, J.H., Abram, M., Brisseau, J.M., Plancon, B., and Grolleau, J.Y. (1993). Autoimmune thyroid disease, thyroid antibodies and giant cell arteritis: the supposed correlation appears fortuitous. *Journal of Rheumatology*, **20**, 1733–4.

Becker, K.L., Ferguson, R.H., and McConahey, W.M. (1963). The connective tissue diseases and symptoms associated wtih Hashimoto's thyroiditis. *New England Journal of Medicine*, **269**, 277–80.

Bland, J.H. and Frymoyer, J.A. (1970). Rheumatic syndromes of myxedema. *New England Journal of Medicine*, **282**, 1171–4.

Bland, J.H., Frymoyer, J.A., Newberg, A.H., Revers, R., and Norman, J. (1979). Rheumatic syndromes in endocrine diseases. *Seminars in Arthritis and Rheumatism*, **9**, 23–65.

Block, K.J. and Bunim, J.J. (1963). Sjögren's syndrome and its relationship to connective tissue disease. *Journal of Chronic Disease*, **16**, 915–23.

Bluestone, R., Bywaters, E.G.L., Hartog, M., Holt, P.J.L., and Hyde, S. (1971). Acromegalic arthropathy. *Annals of the Rheumatic Diseases*, **30**, 243–58.

Bowness, P., Shotlitt, K., Middlemiss, A., and Myles, A.B. (1991). Prevalence of hypothyroidism in patients with polymyalgia rheumatica and giant cell arteritis. *British Journal of Rheumatology*, **30**, 349–51.

Bridgman, J.F. (1972). Periarthritis of the shoulder and diabetes mellitus. *Annals of the Rheumatic Diseases*, **31**, 69–71.

Bywaters, E.G.L., Dixon, A.S.J., and Scott, J.T. (1963). Joint lesions of hyperparathyroidism. *Annals of the Rheumatic Diseases*, **22**, 171–87.

Cassim, B. and Rubin, D.L. (1990). The prevalence of diffuse idiopathic skeletal hyperostosis in African blacks. *British Journal of Rheumatology*, **29**, 131–2.

Chaykin, L.B., Frame, B., and Sigler, J.W. (1969). Spondylitis: a clue to hypoparathyroidism. *Annals of Internal Medicine*, **70**, 995–1000.

Clark, C.F., Piesowicz, A.T., and Spathis, G.S. (1990). Limited joint mobility in children and adolescents with insulin dependent diabetes mellitus. *Annals of the Rheumatic Diseases*, **49**, 236–7.

Crawford, A.M., MacEwen, G.D., and Fonte, D. (1977). Slipped capital femoral epiphysis co-existent with hypothyroidism. *Clinical Orthopedics*, **122**, 135–40.

Dent, G.R. and Edwards, O.M. (1978). Autoimmune thyroid diseases and polymyalgia rheumatica–giant cell arteritis syndrome. *Clinical Endocrinology*, **4**, 215–19.

Detenbeck, L.C., Tressler, H.A., O'Duffy, J.D., and Randall, R.V. (1973). Peripheral joint manifestations of acromegaly. *Clinical Orthopedics and Related Research*, **91**, 119–27.

Dorwart, B.B. and Schumacher, H.R. (1975). Joint effusions, chondrocalcinosis and other rheumatic manifestations of hypothyroidism. *American Journal of Medicine*, **59**, 780–90.

Easton, P.R., Sibbitt, W.L., and Harsh, A. (1985). The effect of an aldose reductase inhibiting agent on limited joint mobility in diabetic patients. *Journal of the American Medical Association*, **253**, 1437–40.

Erickson, A.R., Erzenauer, R.J., Nordstrom, D.M., and Merenich, J.A. (1994). The prevalence of hypothyroidism in gout. *American Journal of Medicine*, **97**, 231–4.

Foster, H., Fay, A., Charles, P., Walker, D., and Griffiths, I. (1993). Thyroid disease and other autoimmune phenomena in a family study of primary Sjögren's syndrome. *British Journal of Rheumatology*, **32**, 36–40.

Frame, B., Henize, E.G., Block, M.A., and Manson, G.A. (1968). Myopathy in primary hyperparathyroidism. *Annals of Internal Medicine*, **68**, 1022–7.

Frymoyer, J.W. and Bland, J. (1973). Carpal tunnel syndrome in patients with myxedematous arthropathy. *Journal of Bone and Joint Surgery*, **55A**, 78–82.

Gertner, E. *et al*. (1990). HLA antigens and nailfold microscopy studies in patients with insulin dependent and non insulin dependent diabetes mellitus and limited joint mobility. *Journal of Rheumatology*, **17**, 1375–9.

Gimlette, T.M.D. (1960). Thyroid acropachy. *Lancet*, **i**, 22–4.

Golding, D.N. (1970). Hypothyroidism presenting with musculoskeletal symptoms. *Annals of the Rheumatic Diseases*, **29**, 10–14.

Gray, R.G. and Gottlieb, N.L. (1976). Rheumatic disorders associated with diabetes mellitus: literature review. *Seminars in Arthritis and Rheumatism*, **6**, 19–34.

Griffiths, H.J. (1985). Diabetic osteopathy. *Orthopaedics*, **8**, 401–6.

Harris, J., Carter, A.R., Glick, E.N., and Storey, G.O. (1974). Ankylosing hyperostosis: clinical and radiographic features. *Annals of the Rheumatic Diseases*, **33**, 210–15.

Jara-Quezada, L., Graef, A., and Lavalle, C. (1991). Prolactin and gonadal hormones during pregnancy in systemic lupus erythematosus. *Journal of Rheumatology*, **18**, 349–53.

Jara, L.J., Gomez-Sanchez, C., Silveira, L.H., Martinez-Osuna, P., Vassey, F.B., and Espinoza, L.R. (1992). Hyperprolactinemia in systemic lupus erythematosus: association with disease activity. *American Journal of Medical Science*, **303**, 222–6.

Job-Deslandre, C., Menkes, C.J., Guinot, M., and Luton, J.D. (1993). Does hypothyroidism increase the prevalence of chondrocalcinosis? *British Journal of Rheumatology*, **32**, 197–8.

Joslin, E.D. (1952). *Treatment of diabetes mellitus* (9th edn), p. 53. Lea and Febiger, Philadelphia.

Jung, Y., Hohmann, T.C., and Gerneth, J.A. (1971). Diabetic hand syndrome. *Metabolism*, **20**, 1008–15.

Kapoor, A. and Sibbitt, W.L. (1989). Contractures in diabetes mellitus: the syndrome of limited joint mobility. *Seminars in Arthritis and Rheumatism*, **18**, 168–80.

Karava, R. and Viljanen, V. (1966). Hyperostosis of the spine in diabetes mellitus and acromegaly. *Diabetologia* (Berlin), **2**, 123–6.

Kellgren, J.H., Ball, J., and Turton, G.H. (1952). Articular and other limb changes in acromegaly, a clinical and pathological study of 25 cases. *Quarterly Journal of Medicine*, **21**, 405–24.

Khaleeli, A.A. *et al*. (1983). Muscle morphology and metabolism in hypothyroid myopathy: effects of treatment. *Journal of Clinical Pathology*, **36**, 519–26.

Khaleeli, A.A. *et al*. (1984). The neuromuscular features of acromegaly: a clinical and pathological study. *Journal of Neurology, Neurosurgery and Psychiatry*, **47**, 1009–15.

Klein, I., Mantell, P., Parker, M., and Levely, G.S. (1980). Resolution of abnormal muscle enzyme studies in hypothyroidism. *American Journal of Medical Science*, **279**, 159–62.

Klein, I. *et al*. (1981). Hypothyroidism presenting as muscle stiffness and pseudohypertrophy: Hoffman's syndrome. *American Journal of Medicine*, **70**, 891–4.

Kyle, V. and Hazelman, B.L. (1981). The thyroid. *Clinics in the Rheumatic Diseases*, **7**, 711–22.

LaRiche, N.G.H. and Bell, D.A. (1984). Hashimoto's thyroiditis and polyarthritis: a possible subset of seronegative polyarthritis. *Annals of the Rheumatic Diseases*, **43**, 594–8.

Lieberman, S.A., Bjorkengon, A.G., and Hoffman, A.R. (1992). Rheumatologic and skeletal changes in acromegaly. *Endocrinology and Metabolism Clinics of North America*, **21**, 615–31.

MacKenzie, A.H. (1975). Final diagnoses in 63 patients presenting with multiple palmar flexor tenosynovitis. *Arthritis and Rheumatism*, **18**, 415.

Martin, P., Bergmann, P., Gillet, C., Fuss, M., Corvilain, J., and Van Geertruyden, J. (1990). Long term irreversibility of bone loss after surgery for primary hyperparathyroidism. *Archives of Internal Medicine*, **150**, 1495–7.

McCarty, D.J., Silcox, D.C., and Coe, F. (1974). Diseases associated with calcium pyrophosphate dihydrate crystal deposition. *American Journal of Medicine*, **56**, 704–14.

McLean, R.M. and Podell, D.N. (1995). Bone and joint manifestation of hypothyroidism. *Seminars in Arthritis and Rheumatism*, **24**, 282–90.

McMurray, R.W., Allen, S.H., Braun, A.L., Rodriguez, F., and Walker, S.E. (1994). Long-standing hyperprolactonemia associated with systemic lupus erythematosus: possible hormonal stimulation of an autoimmune disease. *Journal of Rheumatology*, **21**, 843–50.

Melton, L.J., Atkinson, E.J., O'Fallen, W.M., and Heath, H. (1992). Risk of age-related fractures in patients with primary hyperparathyroidism. *Archives of Internal Medicine*, **152**, 2269–73.

Minisola, S. *et al*. (1993). Trabecular bone mineral density in primary hyperparathyroidism: relationship to clinical presentation and biochemical markers of bone turnover. *Bone and Mineral*, **20**, 113–23.

Mulhern, L.M., Masi, A.T., and Schulman, L.E. (1966). Hashimoto's disease: a search for associated disorders in 170 clinically detected cases. *Lancet*, **ii**, 508–11.

Neufeld, M., MacLaren, N.K., and Blizzard, R.M. (1981). Two types of autoimmune Addison's disease associated with different polyglandular autoimmune (PGA) syndromes. *Medicine*, **60**, 355–62.

Nussbaum, A.J. and Doppman, J.L. (1982). Shoulder arthropathy in primary hyperparathyroidism. *Skeletal Radiology*, **9**, 98–102.

Ostrov, B.E., Van Slyke, M.A., and Ferriss, J.A. (1995). Diabetic muscle infarction: an underappreciated cause of painful swollen extremity. Evaluation with magnetic resonance imaging. *Journal of Clinical Rheumatology*, **1**, 40–5.

Patten, B.M. *et al*. (1974). Neuromuscular disease in primary hyperparathyroidism. *Annals of Internal Medicine*, **80**, 182–93.

Petri, M., Karlson, E.W., Cooper, D.S., and Lovenson, P.W. (1991). Autoantibody tests in autoimmune thyroid disease. *Journal of Rheumatology*, **18**, 1529–31.

Phalen, G.S. (1966). The carpal tunnel syndrome: seventeen years experience in diagnosis and treatment of 654 hands. *Journal of Bone and Joint Surgery*, **48A**, 211–28.

Phalen, G.S. (1970). Reflections of twenty-one years experience with the carpal tunnel syndrome. *Journal of the American Medical Association*, **212**, 1365–7.

Rao, D.S., Wilson, R.J., Kellerkoper, M., and Parfitt, A.M. (1988). Lack of biochemical progression or continuation of accelerated bone loss in mild asymptomatic primary hyperparathyroidism: evidence for biphasic disease course. *Journal of Clinical Endocrinology and Metabolism*, **67**, 1294–8.

Rocca, P.V., Alloway, J.A., and Nashel, D.J. (1993). Diabetic muscular infarction. *Seminars in Arthritis and Rheumatism*, **22**, 280–7.

Rosenbloom, A.L. *et al*. (1981). Limited joint mobility in childhood diabetes mellitus indicates increased risk for microvascular disease. *New England Journal of Medicine*, **305**, 191–4.

Rubin, L.A. *et al*. (1984). Erosive azotemic arthropathy. *Arthritis and Rheumatism*, **27**, 1086–94.

Schulte, L., Roberts, M.S., Zimmerman, C., Ketler, J., and Simon, L.S. (1993). A quantitative assessment of limited joint mobility in patients with diabetes. *Arthritis and Rheumatism*, **36**, 1429–33.

Seibold, J.R. (1982). Digital sclerosis in children with insulin-dependent diabetes mellitus. *Arthritis and Rheumatism*, **25**, 1357–61.

Shiroky, J.C., Cohen, M., Ballachey, M.L., and Neville, C. (1993). Thyroid dysfunction in rheumatoid arthritis: a controlled prospective survey. *Annals of the Rheumatic Diseases*, **52**, 454–6.

Sibbitt, W.L. (1988). Musculoskeletal complications of diaebtes mellitus. *Mediguide to Inflammatory Diseases*, **7**, 1–6.

Sinha, S., Munichoodappa, C.S., and Kozak, G.P. (1972). Neuro-arthropathy (Charcot joints) in diabetes mellitus. *Medicine*, **51**, 191–210.

Skillern, P.G. (1943). The association of periarthritis of the shoulder with hyperthyroidism. In *Transactions of the American Goiter Association* (ed. C. Thomas), pp. 100–9. Springfield IL.

Soubrier, M.J., Dubert, J.J., and Sauvozie, J.M. (1994). POEMS syndrome: a study of 25 cases and a review of the literature. *American Journal of Medicine*, **97**, 543–53.

Starkman, H.S. *et al*. (1986). Limited joint mobility (LJM) of the hand in patients with diabetes mellitus: relations to chronic complications. *Annals of the Rheumatic Diseases*, **45**, 130–5.

Stuart, M.J. and Morrey, M.D. (1990). Arthrodesis of the diabetic neuropathic ankle joint. *Clinical Orthopaedics*, **253**, 209–11.

Thomas, D.J.B. *et al*. (1983). Evidence for an association between rheumatoid arthritis and autoimmune endocrine disease. *Annals of the Rheumatic Diseases*, **42**, 297–300.

Thomasson, J.L. and Sundaram, M. (1985). The diabetic foot—radiographic appearances. *Orthopaedics*, **8**, 670–7.

Trence, D.L., Morley, J.E., and Handwerger, B.S. (1984). Polyglandular autoimmune syndromes. *American Journal of Medicine*, **77**, 107–16.

Tsokos, G.L. *et al*. (1985). Lupus nephritis and other autoimmune features in patients with diabetes mellitus due to autoantibody to insulin receptors. *Annals of Internal Medicine*, **102**, 176–81.

Tucker, S.W. *et al*. (1987). Serositis with autoimmune endocrinopathy: clinical and immunogenetic features. *Medicine*, **66**, 138–47.

Volpe, R. (1985). *Autoimmunity in endocrine disease*. Marcel Dekker, New York.

Wishart, J., Horowitz, M., Need, A., and Nordin, B.E.C. (1990). Relationship between forearm and vertebral mineral density in post menopausal women with primary hyperparathyroidism. *Archives of Internal Medicine*, **150**, 1329–31.

Yosipovitch, G., Yosipovitch, Z., Karp, M., and Murkamel, M. (1990). Trigger finger in young patients with diabetes mellitus. *Journal of Rheumatology*, **17**, 951–2.

1.3.8 Oncology and haematology

Michael Ehrenfeld and Yehuda Shoenfeld

Rheumatological manifestations of malignant diseases

Rheumatic symptoms and manifestations that may be clues to the existence of cancer may either be caused by direct invasion of the neoplasm or be due to an indirect cause, that is the paraneoplastic syndromes.

A wide range of neoplasia, both haematological and non-haematological, may be associated with rheumatic findings. As these patients may first present to the rheumatologist, he or she should be aware of the extent of these manifestations, thereby enabling earlier recognition and diagnosis. Furthermore, as there is an increased incidence of malignancy in almost all autoimmune rheumatic diseases, early recognition of a predictable malignancy is of great importance (see Chapter 5.19.1).

Direct association between rheumatological symptoms and cancer

Primary and secondary neoplasia may be responsible for this direct association.

Primary neoplastic diseases of bone or joints

Tumours of the joints, such as synovial sarcoma, are extremely rare. They are usually first encountered by orthopaedic surgeons or rheumatologists, as a result of symptoms suggestive of internal derangement of a joint, or in the evaluation of an articular or para-articular mass. Synovial sarcoma can appear at any age, although it is most common in young men. It usually occurs in the extremities, particularly around tendon sheaths and joint bursae, but it can also be found remote from joints. Santavirta *et al*. reviewed 31 cases of synovial sarcoma (Santavirta 1992). The 5-year survival rate was 55 per cent, and local and pulmonary metastases were found to be common. Treatment of choice was radical surgical excision with adjuvant radiotherapy and chemotherapy.

Para-articular involvement by malignant tumours arising from bone, such as osteogenic sarcoma, chondrosarcoma, or fibrosarcoma may give rise to a monoarticular synovitis with a mildly inflammatory fluid (Caldwell and McCallum 1986; Caldwell 1993). Occasionally, the tumour will invade the synovium and malignant cells may be found in the fluid. In a recent study of 47 cases of arthritis secondary to direct tumour involvement of joints, monoarthritis occurred in 45, with knee involvement in the majority (Schwarzer *et al*. 1990). Plain radiographs, tomography, and radionuclide bone scans provide helpful diagnostic information; however, any suspicious, slow-growing mass should be biopsied for diagnosis.

Occasionally, benign tumours such as osteoid osteoma, haemangioma, or lipoma may give rise to arthritis. Haemangiomas are most common in the young and are often associated with cutaneous angiomas. The knee is the joint most frequently involved. Recurrent attacks of acute pain and swelling of the joint, together with repeated aspirations of bloody fluid, are suggestive of the diagnosis. Pigmented villonodular synovitis may also present in a similar way. Radiographs may reveal phlebolites, which are diagnostic. Later findings are similar to those found in haemophiliac arthropathy. Arteriography and arthrography are complementary tests for assessing the size of the lesion before eventual surgical resection.

Secondary neoplasia

Lymphomas and leukaemias may simulate various rheumatic syndromes and cause synovitis (Fig. 1). The skeletal manifestations of leukaemia were reviewed extensively by Rennie and Auchterlonie (Rennie and Auchterlonie 1991). In children, these haematological malignancies may invade the synovium directly and thereby mimic juvenile or adult rheumatoid arthritis. Leukaemia, mainly acute, may present with arthritis in about 13.5 per cent of patients. Both acute and chronic leukaemia may be responsible for either monoarthritis or polyarthritis. In a study of 28 patients with leukaemic arthritis, Spilberg and Meyer found that acute leukaemia was more commonly polyarticular, asymmetrical, and part of the initial presentation (Spilberg and Meyer 1972). In chronic leukaemia the synovitis was usually a later manifestation and tended to be more symmetrical. Synovial biopsy and analysis of synovial fluid may confirm the presence of a leukaemic infiltration into the synovium. Immunocytological examination may facilitate diagnosis of leukaemic cells within the synovial fluid. Aside from migratory polyarthritis or arthralgias, other musculoskeletal manifestation of leukaemia include bone pain and tenderness, intra-articular or periarticular haemorrhage, synovial reaction to adjacent bony periosteal or capsular lesions, and crystal-induced synovitis. Back pain simulating a radiculopathy may be caused by diffuse leukaemic infiltration of the meninges.

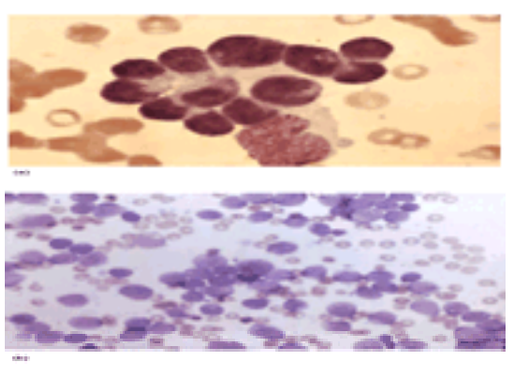


Fig. 1 (a) Neuroblastoma cells in the bone marrow of a patient presenting with musculoskeletal disease symptoms. (b) Acute lymphocytic leukaemia in a patient presenting with a monoarthritis.

Lymphoproliferative diseases may also present with a variety of musculoskeletal manifestations. Bone pain is the most common symptom of bone involvement. Lymphomas, though rare, may directly invade the synovium, with both monoarticular and polyarticular presentations. Radiographs may show bone destruction adjacent to the joints (Isenberg and Shoenfeld 1983).

Non-haematological metastatic disease

Direct invasion of the synovial membrane and the bone by non-haematological malignancy may lead to rheumatic presentations. The most commonly affected joint is the knee, with the hip, ankle, wrist, hand, and foot following. The malignancies are mainly carcinoma of the breast, bronchogenic carcinoma, gastrointestinal tumours, and melanoma (Caldwell and McCallum 1986). The joint disease is usually asymmetrical in distribution. Joint fluid is typically haemorrhagic, culture and crystal analysis are negative. The diagnosis of an underlying malignancy is most often made from a combination of radiographs, bone scan, and needle biopsy of the synovium, which may reveal the histological features of the tumour. Aspiration of synovial fluid may allow identification of malignant cells in some cases (Fam *et al*. 1980; Newton *et al*. 1984).

When radiographic films suggest a particularly destructive process accompanied by constitutional symptoms, a previous history of malignancy, and a protracted type of synovitis, metastatic carcinoma should be suspected.

Indirect associations between rheumatological syndromes and cancer

These paraneoplastic rheumatic syndromes can be subdivided into myopathies (particularly dermatomyositis and polymyositis) and arthropathies such as hypertrophic osteoarthropathy, amyloidosis in chronic haemodialysis, multiple myeloma and the POEMS syndrome (polyneuropathy, organomegaly, endocrinopathy, M protein, and skin change syndrome), secondary gout, carcinomatous polyarthritis, and the palmar fasciitis–polyarthritis syndrome. Miscellaneous indirect associations include: necrotizing vasculitis and polyarteritis nodosa-like syndrome; polymyalgia rheumatica; lupus-like syndrome; pyrophosphate arthropathy; cryoglobulinaemia; scleroderma; panniculitis, subcutaneous fat necrosis, and arthritis; relapsing polychondritis; and erythema nodosum with arthritis (see also Chapter 5.19.1). The syndrome of hypertrophic osteoarthropathy should be described in more detail. It consists of clubbing, periostitis of tubular bones, and an arthropathy ranging from mild arthralgia to a diffuse polyarthritis. If frank arthritis occurs, it usually involves the knees, ankles, wrists, and metacarpophalangeal joints. Diagnostic measures include a technetium bone scan, which may reveal an increased uptake around the shafts of the affected bones. Radiographs will usually reveal later periosteal elevation as a sign of new bone formation. Hypertrophic pulmonary osteoarthropathy (HPO) occurs as a presenting symptom in up to 20 per cent of primary lung tumours, though association with other malignancies is also known (Segal and Mackenzie 1982; Martinez-Lavin *et al*. 1993).

Malignancies in autoimmune rheumatic disease

Autoimmune rheumatic diseases and cancer are diagnosed with increasing frequency and there are many reports of an association between them, usually with lymphoreticular malignancies (Sela and Shoenfeld 1988) (see also Chapter 5.19.1). Rheumatoid arthritis, Sjögren's syndrome, systemic lupus, dermato-/polymyositis, systemic sclerosis, eosinophilic fasciitis, mixed connective tissue disease, lymphomatoid granulomatosis, and Paget's disease of bone have all been specifically associated with cancers (Haga *et al*. 1993; Gridley *et al*. 1993; Morel *et al*. 1993; Menon *et al*. 1993; Rosenthal *et al*. 1993).

Various autoantibodies have been identified in both haematological and epithelial malignancies. There are various methodological limitations in many of the studies that have tried to establish an association between autoimmune rheumatic disease and malignancy, leading to potential errors. These problems include the diagnostic criteria, the reliability of information on deaths, problems with follow-up data, inadequate confirmation of the histological type of the tumour, referral bias, imperfect selection of the control group in case–control studies, inadequate methods of analysis, therapy-related cancers, host susceptibility in both diseases, and environmental factors such as radiation, chemicals, or viral agents (Sela and Shoenfeld 1988). Even given these limitations, it is clear that some specific malignancies are still found more frequently than expected among most patients with autoimmune connective tissue diseases.

Malignancy as a complication of therapy

Immunosuppression

Immunosuppressive therapy has been used effectively in the treatment of autoimmune and rheumatic diseases because immune mechanisms are thought to be fundamental to the pathogenesis of these diseases. The drugs commonly used include the alkylating agents cyclophosphamide and chlorambucil, a purine analogue azathioprine, and the folic acid antagonist methotrexate.

The potential for the development of malignancy after the use of immunosuppressive drugs in non-malignant conditions has been a concern since 1970 when reticulum-cell sarcoma was documented to have developed in an unusual number of renal transplant recipients. Similarly, cancers have been described after the use of immunosuppressive drugs in rheumatic diseases; however, the actual frequency of these complications is difficult to determine because of methodological problems listed above.

In a long-term, retrospective, case–control study, Baker *et al*. have shown that prolonged daily treatment of rheumatoid arthritis with cyclophosphamide was associated with an increased risk of cancer of the urinary bladder, skin, or haematopoietic system (Baker *et al*. 1987). This risk was greatest in patients receiving the highest total dose and persisted for many years after the drug was discontinued. In a 7-year study of 61 cyclophosphamide-treated patients with rheumatoid arthritis and an equal number of controls, Baltus *et al*. reported malignancies in 25 per cent of treated patients and 7 per cent of the controls (Baltus *et al*. 1983). If data from Baker's and Baltus' studies are combined, there were haematological malignancies in 8 of 200 cyclophosphamide-treated patients versus 1 of 200 controls. It was later shown that regimens which involve the intensive use of cyclophosphamide followed by a drug-free period result in a reduced risk of untoward effects, including neoplasia and haemorrhagic cystitis. To date, there are only a few reported instances of malignancies in patients receiving pulse doses of cyclophosphamide.

Chlorambucil has become a popular alternative alkylating agent to cyclophosphamide, principally in Europe, in treating rheumatoid arthritis, systemic lupus, or vasculitis. This is primarily because of a lower incidence of bladder toxicity and alopecia, and its preferential lymphopoietic inhibition at low dosage. The potential for oncogenesis with chlorambucil is similar to that of cyclophosphamide, and of 1853 patients with rheumatoid arthritis treated mainly with chlorambucil or cyclophosphamide, 15 (0.85 per cent) developed acute leukaemia (Kahn *et al*. 1979).

Azathioprine appears to be the favoured alternative immunosuppressive agent to chlorambucil in the United States. The relatively high incidence of malignancies associated with azathioprine in renal transplant patients does not appear to apply to patients with rheumatoid arthritis, though there are occasional reports of lymphoreticular malignancies. Malignancies associated with azathioprine include non-Hodgkin's lymphoma, pancreatic sarcoma, reticulum-cell sarcoma and adenocarcinoma of the lung, squamous-cell carcinoma of the skin, acute myeloid leukaemia, transitional-cell tumour of the bladder, and carcinoma of the breast and the cervix.

Interestingly, Hazleman showed that the prevalence of neoplasia in patients with rheumatoid arthritis treated with azathioprine was less than in those who were treated with other drugs (Hazleman 1982). There have been similar results from other studies. It seems therefore that though the risk of azathioprine-induced malignancies exists, these are unlikely to achieve great importance for rheumatologists treating patients with rheumatoid arthritis (Lewis *et al*. 1980). Recently, Matteson *et al*. examined prospectively the safety of azathioprine and other disease modifying antirheumatic drugs (DMARDs) in the treatment of rheumatoid arthritis (Matteson *et al*. 1991). In a yearly follow-up over a 7-year period, 20 malignant conditions were diagnosed in 530 DMARD-treated adult patients with rheumatoid arthritis. Based on their results, the authors concluded that patients with rheumatiod arthritis requiring DMARD therapy may be at an increased risk of malignancy (particularly lymphoproliferative disorders) as compared with the general population.

The growing enthusiasm for the use of intermittent, low-dose methotrexate in rheumatoid arthritis has led to concern that the drug might carry an increased risk of associated neoplastic disease. There have been reports of malignancy occurring in psoriatic patients treated with methotrexate; to date, however, there is no convincing evidence implicating methotrexate as a cancer-causing agent in rheumatoid arthritis, though several cases of thymomas, lymphomas, and carcinomas have been reported in patients with rheumatoid arthritis treated with methotrexate. Though a relationship between methotrexate intake and various lymphoproliferative disorders has been suggested in several case reports, large studies in cancer and psoriasis, using both high- and low-dose methotrexate respectively, have not established a true relationship between methotrexate and cancer as has been shown for cyclophosphamide (Nyfors and Jensen 1983).

Arellano and Krupp recently calculated the relative risk of developing malignancies following treatment with cyclosporin (Arellano and Krupp 1993). Seventeen such patients were observed in over 1000 rheumatoid arthritis patients involved in clinical trials. The authors conclude that the use of cyclosporin, increases the already existing risk for malignancies by approximately the same degree as other DMARDs.

Because of potential toxicity and carcinogenicity, the decision to use immunosuppressive drugs in rheumatic diseases should be reserved for patients with active disease that is expected to be at least partly reversible.

Irradiation

X-ray therapy has been used in the past to treat ankylosing spondylitis, and this treatment may be implicated in the development of malignancy (Darby *et al*. 1987). Leukaemia, aplastic anaemia, and basal-cell carcinoma in the irradiated area may be late complications and this treatment has been abandoned.

More recently, total-lymphoid or total-body irradiation has been used in a small number of patients with rheumatoid arthritis refractory to all conventional therapy. This procedure has also been used as an immunosuppressive in patients with other autoimmune diseases. Though a few patients have been reported to develop a myeloproliferative or lymphoproliferative disorder, the overall incidence of malignancies in treated patients (with both rheumatoid arthritis and systemic lupus) was not found to be increased (Urowitz and Rider 1985).

Cancer therapy and musculoskeletal disorders

Loprinzi *et al*. recently described eight patients with a hitherto unrecognized new syndrome of myalgia–arthralgia developing within 1 to 3 months after completion of cyclophosphamide–5-fluorouracil combination chemotherapy for breast cancer (Loprinzi *et al*. 1993). Further cases of this 'postchemotherapy rheumatism' syndrome have been reported, also in association with other chemotherapeutic regimens and other malignancies, such as haematological malignancies. The symptoms usually clear within 4 to 9 months from their onset.

Haematological disorders in rheumatic diseases (seeChapter 4.2)

Anaemias

The rheumatic disease are commonly complicated by haematological abnormalities including those of all cell lines, coagulation abnormalities, and the haematological malignancies referred to above (Mowat 1971; Bennett 1977; Hamilton 1983). The most common finding is the anaemia of chronic disease, a mild anaemia that is generally asymptomatic. Its severity correlates with the activity of the disease. The two red-cell indices—mean corpuscular volume and mean corpuscular haemoglobin concentration—tend to be decreased, often to the range of values characteristic of microcytic and hypochromic anaemias. The serum ferritin rises while the transferrin and iron fall. Typically there is decreased iron absorption and impaired release of iron from the reticuloendothelial system (Baynes *et al*. 1987). Haemosiderin is usually found in the bone marrow. The cellularity of the marrow is usually abnormal, with plasma cells being increased in the majority of cases and associated with lymphoid aggregates. Marrow findings are, however, unpredictable, usually reflecting the diverse causes of cytopenias in patients with rheumatic diseases (Fig. 2).

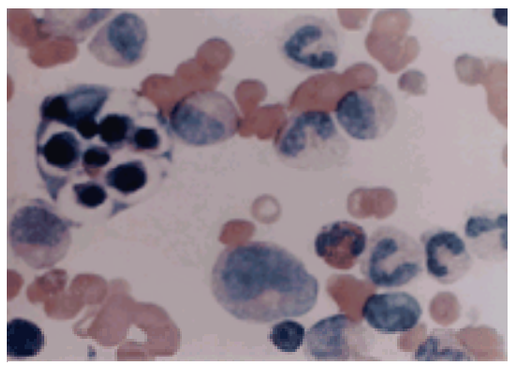


Fig. 2 Haemocytophagocytosis in a patient with juvenile rheumatoid arthritis.

The pathogenesis of the anaemia of chronic disease is still unclear. Inflammatory mediators, particularly the cytokines interleukin 1 and tumour necrosis factor, appear to be important in the impairment of erythropoiesis (Maury 1989).

One of the factors that could be related to the anaemia of rheumatoid arthritis appears to be a reduction in marrow response to erythropoietin as well as reduced erythropoietin levels. Patients in whom the hormone was measured by radioimmunoassay seemed to have a suppressed serum erythropoietin response to anaemia where the iron concentration was normal. Treatment with the human recombinant hormone resulted in improvement of the anaemia in a few of the patients, though no apparent changes in overall rheumatological status were noticed (Baer *et al*. 1990; Pincus *et al*. 1990).

In rheumatoid arthritis, more than one cause of anaemia is usually found and patients may simultaneously have true iron deficiency anaemia, which is found in up to 50 to 75 per cent of patients with chronic active disease, as well as anaemia of chronic disease. When these two conditions occur together, the haemoglobin drops, usually to below 9.5 g/dl, and the mean corpuscular volume is less than 80 fl. Iron deficiency is most often caused by chronic blood loss from gastritis induced by non-steroidal anti-inflammatory drugs, peptic ulcer, or from the diaphragmatic hernia frequently encountered in rheumatoid patients. An anaemia out of proportion to the clinical activity of the arthritis should therefore alert the physician to an unrelated cause.

Less frequently, a macrocytic megaloblastic anaemia may be found in rheumatoid arthritis with vitamin B_{12} deficiency. Folic acid deficiency has also been encountered, most often in iron-deficient patients with rheumatoid arthritis. The characteristic morphological changes in red blood cells may not always be present as they may be masked by the other common causes of anaemia.

Haemolysis with haemolytic anaemia is not a feature of rheumatoid arthritis, although antibody-mediated, Coombs'-positive, haemolytic anaemia has been described in rheumatoid arthritis and particularly in Felty's syndrome (see next subsection). Drug-induced haemolysis may also occur.

Pure red-cell aplasia is a rare but treatable cause of anaemia in rheumatoid arthritis and should be considered when there is severe normochromic anaemia in the absence of obvious blood loss or haemolysis. Autoimmune suppression of erythroid stem cells or antirheumatic drugs have been implicated in this complication, although single case reports suggest that pure red-cell aplasia could be an extra-articular manifestation of rheumatoid arthritis. Bone marrow hypoplasia may also complicate Felty's syndrome or renal failure.

As in rheumatoid arthritis, anaemia is the most common haematological abnormality in systemic lupus and occurs in over 50 per cent of the patients with active disease (Shoenfeld and Schwartz 1983). The most frequent cause is decreased production of red blood cells as a result of the chronic inflammation. The lupus

anaemia is usually normocytic, normochromic, with a reticulocyte count that is disproportionally low for the degree of anaemia. The bone marrow has a normal cellularity and normal or even increased iron. Iron studies are similar to those described in the anaemia of chronic disease in rheumatoid arthritis.

An unusual cause of impaired production of red blood cells in systemic lupus depends on a serum antibody that inhibits erythropoiesis. In these instances, the patients' serum suppresses *in-vitro* growth of autologous and allogeneic bone marrow cells. This type of immune-mediated anaemia can respond to either corticosteroids or plasma exchange.

Pure red-cell aplasia can be an unusual complication of systemic lupus. Antibodies directed against erythroblasts are present, and the condition may also respond to corticosteroid therapy. Although systemic lupus is often associated with anaemia, leucopenia, and thrombocytopenia, pancytopenia is present in only 5 per cent of these patients. Bone marrow findings in these patients show a variety of morphological abnormalities, that is hypoplasia, dyserythropoiesis, hyperplasia, gelatinous transformation, or lymphocytosis with or without plasmacytosis (Feng *et al*. 1991).

Autoimmune haemolytic anaemia may be the presenting sign of systemic lupus or one of its major manifestations. About 15 per cent of lupus patients will develop haemolytic anaemia at some time during the course of their disease. It is accompanied by a normal marrow response, reticulocytosis, a positive Coombs' test, hyperbilirubinaemia, and splenomegaly. Positive Coombs' tests are found in up to 30 per cent of patients, in half of whom there is no evidence of haemolysis. A positive Coombs' test may indicate the presence of circulating immune complexes but not necessarily of erythrocyte antibodies or autoimmune haemolytic anaemia. Systemic lupus patients with autoimmune haemolytic anaemia or immune thrombocytopenia probably make up two related subsets of lupus patients, who also have a high frequency of antiphospholipid antibodies. Corticosteroids, immunosuppressives, or splenectomy are the usual therapeutic measures employed, though newer therapeutic methods, such as plasmapheresis and high-dose intravenous gammaglobulin, have been proved to be beneficial in corticosteroid-resistant cases of autoimmune haemolytic anaemia (Miller 1990). When using prednisone or its equivalent, the initial dose should be 1.0 to 1.5 mg/kg, continued at that level for at least 4 to 6 weeks. Following a satisfactory response, the dose should be reduced gradually, generally at a rate of about 10 mg/week. The reticulocyte count is a reliable indicator of both responsiveness to therapy and relapse. A fall will indicate a response, whereas a rise signals a relapse. In cases of severe fulminant haemolysis, pulse methylprednisolone should be instituted at 1 g intravenously for 3 consecutive days.

As with rheumatoid arthritis, iron deficiency anaemia is not unusual in systemic lupus and is usually the result of either gastrointestinal drug-induced bleeding or menorrhagia. The anaemia is microcytic, hypochromic, with hypoferraemia and an increased serum iron-binding capacity with reduced iron stores.

Disorders of leucocytes

Leucopenia occurs in many rheumatic diseases (Bennett 1977; Hamilton 1983). In patients with rheumatoid arthritis several types of leucopenia have been found. The triad of neutropenia splenomegaly, and deforming rheumatoid arthritis (Felty's syndrome) occurs in about 1 per cent of patients with rheumatoid arthritis. Serious infections are frequent in these patients, most commonly skin infections and leg ulcers. Other factors associated with the increased incidence of infections are corticosteroid therapy and its dose, hypocomplementaemia, and high levels of immune complexes. These complexes coat granulocytes, leading to their sequestration and a reduced survival; furthermore, granulocyte-specific antinuclear antibodies are directed against white blood cell surface antigens, both contributing to the leucopenia. The neutrophil counts are usually between 500 and 2500/mm^3. In most patients, this may be accompanied by a mild anaemia of chronic disease and thrombocytopenia. The most common bone-marrow abnormality in Felty's syndrome is a maturation arrest of the granulocyte cell line. The neutropenia of Felty's syndrome as well as other rheumatic diseases is believed to be due to several mechanisms, including inadequate marrow granulopoiesis, accelerated disappearance of mature blood neutrophils often associated with antineutrophil antibodies, and splenic sequestration.

Treatment of Felty's syndrome is controversial and includes splenectomy, corticosteroids or immunosuppressives, and, to a limited extent, lithium salts. Leucopenia is not a contraindication to gold therapy and, when the leucopenia is mild, immunosuppressive agents may also be used when indicated. Gold may also improve the cutaneous vasculitis and may decrease the susceptibility to infection (Dillon *et al*. 1986).

Although haematological improvement is frequent after splenectomy, it does not always correlate well with reduction in the risk of recurrent infections or with ulcer healing. Plasmapheresis has also been used in these patients with some success.

Leucocytosis with proliferation of polymorphonuclear leucocytes can occur during an inflammatory flare of rheumatoid arthritis. In the differential diagnosis it is important to exclude a bacterial infection. In systemic lupus, leucocytosis is rare in the absence of infection or corticosteroid therapy; a depression of the total white blood-cell count to below 4500/mm^3 is characteristic, occurs in as many as two-thirds of lupus patients, and is mediated by immune complexes or complement-mediated aggregation. Other causes of granulocytopenia include drugs, decreased marrow production, and increased marginal or splenic pooling. Several mechanisms may exist at the same time in any given patient. Lymphopenia is possibly one of the most common manifestation of systemic lupus and is usually mediated by anti-T-cell autoantibodies. Other causes of granulocytopenia or lymphopenia in systemic lupus may be hypersplenism, or corticosteroids and immunosuppressive agents.

Eosinophilia

It has long been recognized that certain patients with rheumatoid arthritis show a significant eosinophilia; this usually correlates with the presence of vasculitis, pleuropericarditis, pulmonary fibrosis, and subcutaneous nodules. Eosinophilia also accompanies gold-induced skin rashes.

Platelet abnormalities

Thrombocytosis is common in rheumatoid arthritis and a positive correlation has been found between the platelet count and disease activity. Extreme thrombocytosis has been noticed with extra-articular manifestations of the disease—in particular, pulmonary involvement, peripheral neuropathy, and vasculitis. Thrombocytopenia is rare in rheumatoid arthritis, except when related to drug treatment or Felty's syndrome.

In systemic lupus, thrombocytopenia is the most common and important clinical disorder of haemostasis. Between 25 and 50 per cent of lupus patients will have mild, clinically negligible thrombocytopenia during the course of the disease, but only about 10 per cent will develop a marked reduction in platelets (less than 50 000/µl) and bleeding manifestations (Gladman *et al*. 1983). The bone marrow usually shows normal or increased numbers of megakaryocytes. Abnormalities of platelet function are commonly seen in patients with immune thrombocytopenia, and in patients with systemic lupus, whether or not they are thrombocytopenic.

At times, autoimmune thrombocytopenia develops simultaneously with autoimmune haemolytic anaemia. These instances, referred to as Evan's syndrome, almost always occur in patients with systemic lupus. In some patients autoimmune thrombocytopenia may precede other signs of lupus by months or years.

The finding of a positive antinuclear antibody test in a patient with an apparently idiopathic autoimmune thrombocytopenia thus warrants continued observation. An association between the presence of anticardiolipin and other antiphospholipid antibodies and thrombocytopenia in systemic lupus, as well as in chronic immune thrombocytopenic purpura, has been described. Antiplatelet antibodies are probably responsible for the observed thrombocytopenia, as well as for hypersplenism. Although complement-mediated lysis may occur, the major effector is the splenic macrophage, which opsonizes the antibody-coated platelets.

Corticosteroids are the mainstay of treatment. The initial dose in prednisone equivalents is usually 1 mg/kg, but 2 mg/kg may be required in more severe instances. An increase in the platelet count should occur within 1 to 3 weeks in responsive cases. The dose can be gradually reduced after 4 weeks, with careful monitoring of the platelet count. Clinical findings must guide the management. Petechiae and ecchymoses in the skin ('dry' purpura) are harmless, whereas 'wet' purpura with gastrointestinal bleeding or bleeding in the genitourinary system are much more dangerous. Bleeding in the central nervous system is rare. Steroid failures have been treated with splenectomy, immunosuppressive drugs, vinca alkaloids, cyclosporin, and high-dose intravenous gammaglobulin—all with varying degrees of success (Boumpas *et al*. 1990; Miller 1990).

Immune thrombocytopenic purpura has also been reported in mixed connective tissue disease, dermatomyositis, and systemic sclerosis.

Haemostatic changes in rheumatic diseases

The lupus anticoagulant and antiphospholipid syndrome (see also Chapter 5.7.3)

Various alterations in measures of coagulation have been described in association with most rheumatic diseases (Hamilton 1983). Most naturally occurring, circulating anticoagulants are antibodies to specific clotting factors. In systemic lupus specific antibodies against factors VIII, IX, XI, XII, and XIII have been reported. The discovery of the lupus anticoagulant has made this disorder the most studied in relation to haemostatic changes. Lupus anticoagulant was first described by Conley and Hartmann, who noticed that the addition of plasma from a patient with systemic lupus to normal blood prolonged the whole-blood clotting time but not the thrombin clotting time (Conley and Hartmann 1952). It was later shown that these *in-vitro* abnormalities can occur despite normal activity of the individual clotting factors. Furthermore, an association between a biological false-positive test for syphilis and the circulating anticoagulant became apparent. This anticoagulant, called the lupus inhibitor, is now known to be an antibody directed against the negatively charged phospholipid portion of the prothrombin activator complex. It is an immunoglobulin of the IgG or M class, which interferes with the phospholipid-dependent coagulation tests. It has recently been shown that a cofactor, b $_2$ glycoprotein I (apolipoprotein H), which is a normal serum protein, binds to negatively charged phospholipids and enhances autoimmune antiphospholipid binding (McNeil *et al*. 1990).

This lupus anticoagulant is one of a family of antiphospholipid antibodies. Others include anticardiolipin antibodies and an antibody that causes the false-positive result in reaginic tests for syphilis (VDRL, RPR, or Wasserman test). Whether antiphospholipid antibodies and the lupus anticoagulant are the same or two closely related but different antibodies is not yet known. Antiphospholipid antibodies probably cross-react with anti-DNA antibodies (Harris *et al*. 1985). Acidic phospholipids have now been recognized to bind to DNA-polymerase, topoisomerase I, and Z-DNA binding proteins. It has also been shown that low-avidity serum anti-DNA antibodies cross-react extensively with cardiolipin, whereas high-avidity antibodies do not (Smeenk *et al*. 1987). The lupus anticoagulant and anticardiolipin antibodies are highly correlated, although there are patients who have positive tests for one without the other. Passive transfer of anticardiolipin to naïve mice induces an experimental model of antiphospholipid syndrome, including thrombocytopenia and fetal loss (Blank *et al*. 1991). It seems, in general, that anticardiolipin antibodies are a more sensitive measure of antiphospholipid antibodies than the lupus anticoagulant. They are both associated with the clinical features listed below (Hughes 1993; Bick and Baker 1994; Kampe 1994).

1. Thromboembolic phenomena:
    a. venous; thrombosis, recurrent (deep vein, axillary, inferior vena cava, cerebral and retinal, Budd–Chiari syndrome, renal vein)
    b. arterial;
        i. cerebrovascular accidents, including transient ischaemic attacks and amaurosis
        ii. coronary thrombosis
        iii. iliofemoral artery-related peripheral arterial gangrene
        iv. mesenteric thrombosis
        v. retinal artery thrombosis
        vi. renal microangiopathy and renal failure
        vii. adrenal haemorrhage/failure
        viii. avascular necrosis.
2. Fetal loss: placental thrombosis and infarction (spontaneous recurrent abortion or miscarriage).
3. Thromboctyopenia: Coombs'-positive haemolytic anaemia.
4. Neurological disease:
    a. chorea and seizures;
    b. migraine headaches;
    c. progressive multifocal-infarct dementia;
    d. transverse myelitis;
    e. 'pseudomultiple sclerosis';
    f. pseudotumour syndrome.
5. Heart valve vegetations and endocarditis.
6. Pulmonary hypertension.
7. Livedo reticularis.
8. Chronic leg ulcers.

It has been suggested that many of these clinical associations are consequences of recurrent thrombosis. The incidence of lupus anticoagulant, the most common anticoagulant found in patients with systemic lupus, is not known but has been estimated at 5 to 10 per cent, though some series quote figures ranging up to 21 to 65 per cent. It has also been shown that lupus anticoagulant occurs in conditions other than systemic lupus and may be associated with drug administration, AIDS, or neoplasia. Diseases associated with antiphospholipid antibodies (anticardiolipin or lupus anticoagulant) are as follows:

1. Systemic diseases

    Autoimmune:

    a. systemic lupus erythematosus (frequent);
    b. rheumatoid arthritis (rare);
    c. vasculitis, including Takayasu's arteritis;
    d. systemic sclerosis and polymyositis;
    e. juvenile chronic arthritis;
    f. mixed connective tissue disease and 'overlap' syndromes;
    g. Sjögren's syndrome;
    h. autoimmune haemolytic anaemia;
    i. autoimmune thrombocytopenic purpura;
    j. Behçet's disease;
    k. hypothyroidism and Addison's disease.
2. Haematological disorders:
    a. Hodgkin's and non-Hodgkin's lymphomas;
    b. leukaemias, including hairy-cell leukaemia;
    c. Waldenström's macroglobulinaemia;
    d. bone marrow aplasia.
3. Liver disorders:
    a. cirrhosis;
    b. chronic active hepatitis;
    c. Budd–Chiari syndrome.
4. Drug associated:
    a. oral contraceptives;
    b. drugs associated with drug-induced lupus (procainamide, hydralazine, chlorpromazine, phenytoin, antibiotics).
5. Infections: bacterial, viral, fungal, and protozoa (e.g. AIDS).
6. Vascular thrombosis: recurrent venous or arterial.
7. Spontaneous abortions.
8. Neurological disorders:
    a. focal cerebral ischaemia;
    b. ocular ischaemia;
    c. chorea;
    d. migraines;
    e. dementia, psychiatric illness.
9. Others:
    a. epithelial malignancies;

b. psoriatic arthritis and osteoarthritis;
c. Raynaud's phenomenon;
d. ulcerative colitis;
e. rheumatic fever (acute);
f. normal population.

Paradoxically, patients with this antibody do not have a bleeding tendency unless they have another haematological abnormality such as thrombocytopenia or vascular deficiency. In contrast they are, as noted above, at increased risk of developing recurrent venous or arterial thromboembolic events. The explanation of this tendency to thrombosis in 5 to 30 per cent of patients with systemic lupus is still unclear. Thrombotic events have been considered to be related to the titre of the antibody, the isotype present (IgG), and perhaps the length of time the antibodies have been present. Several possible mechanisms have been proposed. One suggests that the reaction of lupus anticoagulant with platelet phospholipids may increase platelet adhesiveness, thus lowering the threshold for aggregation. Another possibility is the interaction of the anticoagulant with prostacycline, which is a potent vasodilator and inhibitor of platelet aggregation. In addition, the lupus anticoagulant can inhibit prekallikrein, thereby promoting clotting. Other options include interference with activation of protein C and antithrombin III (Triplett 1990).

The antiphospholipid syndrome is the disorder characterized by production of high levels of antiphospholipid antibodies in association with recurrent of venous or arterial thromboses, fetal wastage, and thrombocytopenia (Alarcon-Segovia *et al*. 1989). A variety of clinical features have been added to the initial definition of this condition. Though many of these patients are found to have systemic lupus or other connective tissue diseases, patients with the syndrome who do not have classical systemic lupus or other connective tissue disease may be classified as having 'primary' antiphospholipid syndrome (Asherson *et al*. 1989; Mackworth-Young *et al*. 1989). It appears that these patients do not have major clinical or serological features of systemic lupus. Otherwise there are no overt differences between primary antiphospholipid syndrome patients and those with antiphospholipid syndrome during the course of systemic lupus. In primary antiphospholipid syndrome there is, however, a lower incidence of heart-valve lesions, livedo reticularis, and chorea, and there is absence of acute illness with systemic symptoms and the typical lupus flares. These patients present with an idiopathic deep-vein thrombosis, pulmonary thromboembolism or thromboembolic pulmonary hypertension, transient ischaemic attacks or strokes, and myocardial infarctions or other arterial occlusions in the absence of systemic lupus or any other defined connective tissue disease. They are often antinuclear-antibody negative, and are always negative for typical serological markers of systemic lupus, such as antibodies to double-stranded DNA or to extractable nuclear antigen.

Recently, Vianna *et al*. published a European multicentre study, involving 114 patients with either primary antiphospholipid syndrome or secondary antiphospholipid syndrome (Vianna *et al*. 1994). They were able to show that patients with secondary antiphospholipid syndrome had a significant incidence of autoimmune haemolytic anaemia, endocardial valve disease, neutropenia, and abnormal C4 levels.

The 'catastrophic' antiphospholipid antibody syndrome

A few patients with antiphospholipid antibodies develop a sudden widespread organ failure with thrombocytopenia, wide spread thrombosis, and 'adult respiratory distress syndrome'. Prognosis is poor and the causes of the sudden catastrophe are unclear (Asherson 1992).

Obstetric complications (see Chapter 1.3.1.1)

The presence of lupus anticoagulant in pregnancy has been associated with recurrent early abortions and intrauterine death (Branch *et al*. 1985; Triplett 1989). It is most commonly associated with a first-trimester abortion or second-trimester fetal death. It has been suggested that placental infarctions are the cause of the fetal loss. However, fetal loss is now not thought to be as frequent as the published record suggests. Some women with high levels of lupus anticoagulant activity do not have thrombotic events or suffer fetal death. Prospective studies are needed to determine the true risk of a positive lupus anticoagulant test for the outcome of the pregnancy. An incidence of 5 to 48 per cent of lupus anticoagulant in women with histories of recurrent abortions has been recorded, compared with an incidence of less than 5 per cent in normal, uncomplicated pregnancies. However, prospective studies of normal pregnant women (using a high titre cut-off for antiphospholipid antibody) found that antiphospholipid antibody was rare (0–2.2 per cent). It seems, therefore, that screening for antiphospholipid antibody in unselected women is probably not indicated (Love and Santoro 1990).

Treatment of lupus anticoagulant (see Chapter 5.7.3)

The presence of lupus anticoagulant or antiphospholipid antibody in asymptomatic patients is not an indication for treatment. Some evidence suggested that patients with IgG antiphospholipid antibody of high titre are a high-risk group but other studies have not confirmed this. Patients with antiphospholipid antibodies who have recurrent venous or arterial thrombosis are best treated initially with heparin and then with oral anticoagulants such as warfarin or drugs that inhibit platelet aggregation (e.g. aspirin, dipyridamole). It is probably best to treat these patients as long as the antibody is present at high titre.

Although corticosteroids and immunosuppressive agents may lead to the normalization of the activated partial thromboplastin time, they are not recommended for long-term use as a preventive measure for recurrent thrombosis in patients with antiphospholipid antibody or lupus anticoagulant. On the other hand, for women with antiphospholipid antibody-associated fetal loss, relatively high doses of prednisone and low-dose aspirin have been used as treatment during pregnancy, with improved outcome (Parke 1989; Harris 1990). However, there are conflicting data that pregnant women receiving aspirin alone did not do worse than those on prednisone with low-dose aspirin (Lockshin *et al*. 1989). Though the optimal therapy for the fetal wastage syndrome has not yet been defined, the common practice in many institutions is to use low-dose aspirin (about 80 mg/day) in combination with low-dose heparin (2500 to 5000 u subcutaneously, twice daily).

The combination of low molecular weight heparin and low-dose aspirin has also been shown recently to be effective in the treatment of antiphospholipid antibody syndrome in pregnancy. Other forms of therapy, mainly for women with high-titre antiphospholipid antibodies and prior fetal death, include heparin, plasma exchange, intermittent intravenous high doses of immunoglobulin, and immunosuppressive agents.

The presence of antiphospholipid antibody or lupus anticoagulant in patients who are undergoing major surgery probably warrants the use of low-dose subcutaneous heparin prophylaxis.

Haematological effects of drugs

Non-steroidal anti-inflammatory drugs

A wide variety of adverse reactions in most body systems has been described with the use of non-steroidal anti-inflammatory drugs. Idiosyncratic blood dyscrasias are extremely rare among the adverse haematological effects of these drugs, even though they are used extensively and long term. The mechanism of action is either by exerting bone marrow toxicity or by direct toxic effects on circulating blood cells. Haemolytic anaemia (usually autoimmune in nature), aplastic anaemia, agranulocytosis, isolated neutropenia, and thrombocytopenia are the haematological disorders most frequently reported (O'Brien and Bagby 1985; Henry 1988).

Anaemia and agranulocytosis

Though it is difficult to be certain of the extent of association because of multiple drug exposures, cases of aplastic anaemia and agranulocytosis associated with the ingestion of non-steroidal anti-inflammatory drugs have been described with most of those available. In single cases, aspirin, indomethacin, sulindac, diclofenac, ibuprofen, fenoprofen, naproxen, and piroxicam have all been associated with aplastic anaemia.

Phenylbutazone and oxyphenylbutazone have long been known to be associated with a higher risk of aplastic anaemia and have therefore been abandoned in some countries and limited to hospital use only in others. In one large study of toxicity over a 17-year period in Denmark, 7 per cent of all reports were of haematological adverse reactions. About two-thirds of them dealt with bone marrow depression. Granulocytopenia and agranulocytosis were described in one-quarter of the reports on haematological disorders. Fatal agranulocytosis was particularly found in women over the age of 60 years (Kroman-Andersen and Pedersen 1988).

Pure red-cell aplasia, though extremely rare, has been described with phenylbutazone, indomethacin, and fenoprofen.

Effects on the blood-clotting mechanism

Non-steroidal anti-inflammatory drugs may prolong the bleeding time by inhibiting synthesis of prostaglandin endoperoxidase and thromboxane A_2 and consequently inhibiting platelet aggregation. This effect is usually also manifest at the lowest dosages used, i.e. antipyretic and analgesic doses. This effect has led to the use of aspirin in doses as low as 80 mg daily to prevent platelet aggregation and embolization in patients prone to transient ischaemic attacks. As a result of the effect on platelet function, most non-steroidals may rarely cause clinical bleeding of varying severity. However, blood loss from gastric lesions may be partially influenced by effects on platelet aggregation. A second mechanism by which protein-bound non-steroidals may effect coagulation is by displacing warfarin from plasma protein-binding sites. This in turn may increase warfarin's anticoagulant effect.

Thrombocytopenia

This has been described with most non-steroidal anti-inflammatory drugs, some of the cases being fatal. Fatal thrombotic thrombocytopenic purpura with acute renal failure has also been described. Thrombocytopenia was seen in one-quarter of the haematological adverse reactions in the large-scale Danish study (Kroman-Andersen and Pedersen 1988).

All non-steroidal anti-inflammatory drugs will interfere with platelet aggregation.

Autoimmune haemolytic anaemia

This anaemia induced by non-steroidals is extremely rare but may be severe. The condition tends to be reversible when the offending drug is withdrawn, although most patients require steroid therapy. In some cases the haemolytic anaemia is accompanied by signs of hypersensitivity.

Leukaemia

The possible induction of leukaemia by phenylbutazone was suggested in the early 1960s. Later, in the long-term Danish study, four male patients were reported to have died of leukaemia related to oxyphenylbutazone, phenylbutazone, mophebutazone, indomethacin, and ketoprofen, respectively. All were elderly, between 62 and 90 years of age.

Antimalarials

Various rare blood dyscrasias have been associated with antimalarials, including agranulocytosis, aplastic anaemia, leucopenia, and reversible toxic granulation. The agranulocytosis appears to be reversible (Maksymowych and Russel 1987).

A few of the antimalarials, particularly primaquine, have striking haemolytic effects when given to patients with glucose 6-phosphate dehydrogenase deficiencies. However, chloroquine and hydroxychloroquine have only a weak haemolytic activity.

Hydroxychloroquine has been suggested and used in some institutions as prophylaxis against thromboembolism after total hip replacement (Loudon 1988). In man, hydroxychloroquine can reduce red blood-cell aggregates without prolonging the bleeding time. A variable reduction in platelet aggregation and blood viscosity have been shown as well. Animal studies have shown a reduction in the size of the thrombus with hydroxychloroquine.

Gold compounds

The greatest concern when using these agents is bone marrow suppression (1–3 per cent), which may present as leucopenia, thrombocytopenia, agranulocytosis, or aplastic anaemia (Kean 1990). Blood dyscrasias are especially likely to develop after a total dose of approximately 300 mg of gold has been reached and exceeded, although agranulocytosis and aplastic anaemia may develop at any time during the course of treatment with gold. The untoward effects seen with intramuscular and oral gold are the same but their incidence differs. Because of these risks, full blood-cell and platelet counts must be made at least monthly.

The most common haematological side-effect of gold is thrombocytopenia, which occurs in up to 10 per cent of patients at some stage during gold therapy (Adachi *et al*. 1987). It is immune mediated with the production of platelet-associated IgG leading to peripheral platelet destruction. A second mechanism that has been proposed is complete marrow aplasia, which is usually associated with a gradual fall in platelet counts, as compared with the precipitous fall in the immune-mediated type. An association with HLA-DR3 has been demonstrated. The condition is usually responsive to corticosteroid therapy, although other modes of therapy may be as efficacious (i.e. *N*-acetylcysteine, penicillamine, vincristine, high-dose intravenous gammaglobulin, and splenectomy). Isolated thrombocytopenia usually has a good prognosis. It should be noted that oral gold preparations are safer and less toxic than injectable gold—thrombocytopenia being a rare side-effect with this medication.

Gold-induced neutropenia and marrow aplasia with pancytopenia are the most serious toxic effects of gold compounds, but these are rare. The mortality in this group is estimated at 60 to 75 per cent; there is a little evidence that corticosteroids, anabolic steroids, or D-penicillamine as a chelating agent, or plasmapheresis may be helpful.

Antithymocyte globulin and bone marrow transplantation have been used successfully in treating gold-induced aplastic anaemia. The condition has been associated in some patients with HLA-DR4, although earlier reports found an association with HLA-DR3. The findings are not conclusive.

Eosinophilia has long been recognized as an accompanying feature of gold therapy. If persistent and increasing, it may herald a more serious adverse drug reaction and is often associated with gold-induced dermatitis. The incidence of eosinophilia at the time of adverse reaction varies between 24 and 78 per cent.

D-*Penicillamine*

This drug has a rather high incidence (50 per cent) of adverse effects. Toxic side-effects are believed to be dose related, except for hypersensitivity reactions and autoimmune disorders. Haematological toxicity is the most serious of the untoward effects (Jaffe 1986; Kay 1986). Fatal reactions are associated predominantly with bone marrow aplasia. Other observed toxic effects include thrombocytopenia and neutropenia. These may be of two types. One is idiosyncratic and is most often seen within the first 12 months of treatment. It is independent of dosage and is acute in nature, similar to the aplastic anaemia. The second type, which is more common, is dose related and more gradual in its onset. Other rare side-effects include haemolytic anaemia, sideroblastic anaemia, transient eosinophilia, thrombocytosis, and polycythaemia. Frequent haematological monitoring is mandatory because recovery from bone marrow suppression is possible after prompt discontinuation of the drug (Howard-Lock *et al*. 1986).

Sulphasalazine

The adverse effects of sulphasalazine are well known because it has been used for over 40 years in the treatment of inflammatory bowel disease. A wide range of blood dyscrasias have been described over the years. These include agranulocytosis, leucopenia, aplastic anaemia, thrombocytopenia, haemolytic anaemia, megaloblastic anaemia with macrocytosis, and methaemoglobinaemia (Pinals 1988; Hopkinson *et al*. 1989). Agranulocytosis has most commonly occurred after a mean duration of treatment of 4 weeks, and appears to be sudden in onset and is a very serious complication. Dose-related adverse reactions appear to be related to acetylator phenotype in some studies.

Though fatalities have been described, withdrawal of treatment in cases of early neutropenia leads to restoration of the white-cell count to normal.

A rise in mean corpuscular volume is frequently seen during therapy, although true macrocytosis is less common, and macrocytic anaemia is even more rare, and is usually associated with low red-cell folate levels. However, when macrocytic anaemia does arise, it responds to folate supplements and treatment may be continued.

Methotrexate

This drug is an antimetabolite that acts as a folic acid antagonist. Its specificity toward rapidly dividing cells accounts for its adverse effects on the bone marrow and the gastrointestinal mucosa. Reported blood dyscrasias include leucopenia and thrombocytopenia (2–5 per cent and 1–2 per cent of treated patients, respectively) (Mackinnon *et al*. 1985; Weinblatt 1985; Furst and Kremer 1988). Although bone marrow toxicity, including pancytopenia, usually has a benign outcome and toxicity may be overcome by dose adjustment or temporary discontinuation of the drug, several instances of sepsis and death have been reported. Factors that increase the risk of blood dyscrasias include renal insufficiency, folate deficiency, and coadministration with trimethoprim sulphamethoxazole.

Other, less likely, risk factors for haematological toxicity include concomitant use of non-steroidal anti-inflammatory drugs or probenecid, competition for protein binding, or hypoalbuminaemia. The concomitant use of sulphasalazine, which is also a folate antagonist, require great caution. These haematological adverse reactions are usually dose related.

The use of weekly or daily oral folic or folinic acid in avoiding the cytopenia is controversial. In a double-blind, placebo-controlled trial, daily supplementation of low-dose methotrexate with 1 mg of folic acid significantly reduced the toxic manifestations without altering the efficacy (Morgan *et al*. 1990). The effect of weekly folic or folinic acid may simply be to correct subclinical folate deficiency, which has been shown to be associated with a higher incidence of severity of toxicity. Furthermore, if haematological toxicity is suspected, folinic acid should be given within 24 to 48 h of the methotrexate dose. However, the optimum dosing interval for the combination of folic acid and methotrexate has still not been determined.

Isolated case reports linking low-dose methotrexate therapy in rheumatoid arthritis and the development of cancer have been published over the last few years. Malignancies described include lymphoma, acute leukaemia, adenocarcinoma of the lung, and bladder cancer. Although still anecdotal, patients should be carefully followed for the possibility of the development of a malignancy.

Azathioprine

The main haematological side-effect of azathioprine is marrow suppression. Leucopenia (11–15 per cent) is more common than thrombocytopenia or anaemia, though pancytopenia may occur (Huskisson 1984; Singh *et al*. 1989). Leucopenia is generally mild and reversible. Haematological side-effects of azathioprine are dose dependent. When the usual dosage is given, the interval before maximal bone marrow suppression is about 2 weeks. If rapid agranulocytosis occurs, it is possibly an idiosyncratic reaction. Severe leucopenia may lead to serious infections. Thrombocytopenia rarely occurs without leucopenia.

Other haematological changes include macrocytosis and pure red-cell aplasia. When erythroid toxicity occurs, megaloblastoid changes are not related to the serum folate or vitamin B_{12} concentrations (McGrath *et al*. 1975).

The issue of azathioprine-related lymphoproliferative malignancies has been dealt with elsewhere in this chapter.

Chapter References

Adachi, J.D. *et al*. (1987). Gold induced thrombocytopenia: 12 cases and a review of the literature. *Seminars in Arthritis and Rheumatism*, **16**, 287–93.

Alarcon-Segovia, D. *et al*. (1989). Antiphospholipid antibodies and the anti-phospholipid syndrome in systemic lupus erythematosus: a prospective analysis of 500 consecutive patients. *Medicine*, **68**, 353–65.

Arellano, F. and Krupp, P. (1993). Malignancies in rheumatoid arthritis patients treated with cyclosporin A. *British Journal of Rheumatology*, **32** (suppl. 1), 72–5.

Asherson, R.A. (1992). The catastrophic antiphospholipid syndrome. *Journal of Rheumatology*, **19**, 508–13.

Asherson, R.A. *et al*. (1989). The 'primary' antiphospholipid syndrome: major clinical and serological features. *Medicine*, **68**, 366–74.

Baer, A.N., Dessypris, E.N., and Krantz, S.B. (1990). The pathogenesis of anemia in rheumatoid arthritis: a clinical and laboratory analysis. *Seminars in Arthritis and Rheumatism*, **19**, 209–23.

Baker, G.L., Kahl, L.E., Lee, B.C., Stolzer, B.L., Agarwal, A.L., and Medsger, T.A. (1987). Malignancy following treatment with cyclophosphamide: long term case-control follow up study. *American Journal of Medicine*, **83**, 1–9.

Baltus, J.A.M., Boersma, J.W., Hartman, A.A., and Vandenbroucke, J.P. (1983). The occurrence of malignancies in patients with rheumatoid arthritis treated with cyclophosphamide: a controlled retrospective followup. *Annals of the Rheumatic Diseases*, **42**, 368–73.

Baynes, R.D., Bothwell, T.H., Bezwoda, W.R., Gear, A.J., and Atkinson, P. (1987). Hematologic and iron-related measurements in rheumatoid arthritis. *American Journal of Clinical Pathology*, **87**, 196–200.

Bennett, R.B. (1977). Hematological changes in rheumatoid disease. *Clinics in Rheumatic Diseases*, **3**, 433–65.

Bick, R.L. and Baker, W.F. (1994). Antiphospholipid and thrombosis syndromes. *Seminars in Thrombosis and Hemostasis*, **20**, 3–15.

Blank, M., Cohen, J., Toder, V., and Shoenfeld, Y. (1991). Induction of anti-phospholipid syndrome in naive mice with mouse lupus monoclonal and human polyclonal anti-cardiolipin antibodies. *Proceedings of the National Academy of Sciences* (USA), **88**, 3069–73.

Boumpas, D.T., Barez, S., Klippel, J.H., and Balow, J.E. (1990). Intermittent cyclophosphamide for the treatment of autoimmune thrombocytopenia in systemic lupus erythematosus. *Annals of Internal Medicine*, **112**, 674–7.

Branch, W.B., Scott, J.R., Kochenour, N.K., and Hershgold, E. (1985). Obstetric complications associated with the lupus anticoagulant. *New England Journal of Medicine*, **313**, 1322–6.

Caldwell, D.S. (1993). Musculoskeletal syndromes associated with malignancy. In *Textbook of rheumatology* (eds N.K. Kelley *et al*.), pp. 1552–66. W.B. Saunders, Philadelphia.

Caldwell, D.S. and McCallum, R.M. (1986). Rheumatologic manifestations of cancer. *Medical Clinics of North America*, **70**, 385–417.

Conley, C.L. and Hartmann, R.C. (1952). A hemorrhagic disorder caused by circulating anticoagulant in patients with disseminated lupus erythematosus. *Journal of Clinical Investigation*, **31**, 621–2.

Darby, S.C., Doll, R., Gill, S.K., and Smith, P.G. (1987). Long term mortality after a single treatment course with X-rays in patients treated for ankylosing spondylitis. *British Journal of Cancer*, **55**, 179–90.

Dillon, A.M., Luthra, H.S., and Conn, D.L. (1986). Parenteral gold therapy in the Felty syndrome. Experience with 20 patients. *Medicine*, **65**, 107–19.

Fam, A.G., Kolin, A., and Lewis, A.J. (1980). Metastatic carcinomatous arthritis and carcinoma of the lung. A report of two cases diagnosed by synovial fluid cytology. *Journal of Rheumatology*, **7**, 98–104.

Feng, C.S., Ng, M.H., Szeto, R.S., and Li, E.K. (1991). Bone marrow findings in lupus patients with pancytopenia. *Pathology*, **23**, 5–7.

Furst, D.E. and Kremer, J.M. (1988). Methotrexate in rheumatoid arthritis. *Arthritis and Rheumatism*, **31**, 305–14.

Gladman, D.D., Urowitz, M.B., Tozman, E.C., and Glynn, M.F.X. (1983). Haemostatic abnormalities in systemic lupus erythematosus. *Quarterly Journal of Medicine*, **52**, 424–33.

Gridley, G., McLaughlin, J.K., Ekbom, A., Klareskog, L., Adanic, H.O., Hacker, D.G., Hoover, R., and Franmeni, J.F. Jr (1993). Incidence of cancer among patients with rheumatoid arthritis. *Journal of the National Cancer Institute*, **85**, 307–11.

Haga, H.J., Eide, G.E., Brun, J., Johansen, A., and Langmark, F. (1993). Cancer in association with polymyalgia rheumatica and temporal arteritis. *Journal of Rheumatology*, **20**, 1335–9.

Hamilton, P.J. (1983). The haematology laboratory and the rheumatologist. *Clinics in Rheumatic Diseases*, **9**, 69–92.

Harris, E.N. (1990). Maternal autoantibodies and pregnancy—I: The anti-phospholipid antibody syndrome. *Baillière's Clinical Rheumatology*, **4**, 53–68.

Harris, E.N. *et al*. (1985). Crossreactivity of antiphospholipid antibodies. *Journal of Clinical and Laboratory Immunology*, **16**, 1–6.

Hazleman, B.L. (1982). The comparative incidence of malignant disease in rheumatoid arthritis exposed to different treatment regimens. *Annals of the Rheumatic Diseases*, **41** (suppl.1), 12–17.

Henry, D.A. (1988). Side-effects of non-steroidal anti-inflammatory drugs. *Baillière's Clinical Rheumatology*, **2**, 425–54.

Hopkinson, N.D., Saiz Garcia, F., and Gumpel, J.M. (1989). Haematological side-effects of sulphasalazine in inflammatory arthritis. *British Journal of Rheumatology*, **28**, 414–17.

Howard-Lock, H.E., Lock, C.J.L., Mewa, A., and Kean, W.F. (1986). D-Penicillamine: chemistry and clinical use in rheumatic disease. *Seminars in Arthritis and Rheumatism*, **15**, 261–81.

Hughes, G.R.V. (1993). The antiphospholipid syndrome: ten years on. *Lancet*, **342**, 341–4.

Huskisson, E.C. (1984). Azathioprine. *Clinics in Rheumatic Diseases*, **10**, 325–32.

Isenberg, D.A. and Shoenfeld, Y. (1983). The rheumatologic complications of hematologic diseases. *Seminars in Arthritis and Rheumatism*, **12**, 348–58.

Jaffe, I.A. (1986). Adverse effects profile of sulfhydryl compounds in man. *American Journal of Medicine*, **80**, 471–6.

Kahn, M.F., Arlet, J., Bloch-Michel, H., Caroit, M., Chaouat, Y., and Renier, J.C. (1979). Laucemies aigues apres traitement par agents cytotoxiques en rheumatologie: 19 observations chez 2,006 patients. *Nouvelle Presse Médicale*, **8**, 1393–7.

Kampe, K.E. (1994). Clinical syndromes associated with lupus anticoagulants. *Seminars in Thrombosis and Hemostasis*, **20**, 16–26.

Kay, A. (1986). European League against Rheumatism study of adverse reactions to D-penicillamine. *British Journal of Rheumatology*, **25**, 193–8.

Kean, W.F. (1990). Intramuscular versus oral gold therapy. *Baillière's Clinical Rheumatology*, **4**, 219–46.

Kroman-Andersen, H. and Pedersen, A. (1988). Reported adverse reactions to and consumption of nonsteroidal anti-inflammatory drugs in Denmark over a 17-year period. *Danish Medical Bulletin*, **35**, 187–92.

Lewis, P., Hazleman, B.L., Hanks, R., and Roberts, S. (1980). Cause of death in patients with rheumatoid arthritis with particular reference to azathioprine. *Annals of the Rheumatic Diseases*, **39**, 457–61.

Lockshin, M.D., Druzin, M.L., and Qamar, T. (1989). Prednison does not prevent recurrent fetal death in women with antiphospholipid antibody. *American Journal of Obstetrics and Gynecology*, **160**, 439–43.

Loprinzi, C.L., Duffy, J., and Ingle, J.N. (1993). Postchemotherapy rheumatism. *Journal of Clinical Oncology*, **11**, 768–9.

Loudon, J.R. (1988). Hydroxychloroquine and postoperative thromboembolism after total hip replacement. *American Journal of Medicine*, **85** (suppl. 4A), 57–61.

Love, P.E. and Santoro, S.A. (1990). Antiphospholipid antibodies: Anticardiolipin and the lupus anticoagulant in systemic lupus erythematosus (SLE) and in non-SLE disorders: Prevalence and clinical significance. *Annals of Internal Medicine*, **112**, 682–98.

Mackinnon, S.K., Starkebaum, G., and Willkens, R.F. (1985). Pancytopenia associated with low dose pulse methotrexate in the treatment of rheumatoid arthritis. *Seminars in Arthritis and Rheumatism*, **15**, 119–26.

Mackworth-Young, C.G., Louizou, S., and Walport, M.J. (1989). Primary antiphospholipid syndrome: features of patients with raised anticardiolipin antibodies and no other disorder. *Annals of the Rheumatic Diseases*, **48**, 362–7.

Maksymowych, W. and Russel, A.S. (1987). Antimalarials in rheumatology efficacy and safety. *Seminars in Arthritis and Rheumatism*, **16**, 206–21.

Martinez-Lavin, M., Matucci-Cerinic, M., Jajic, I., and Pineda, C. (1993). Hypertrophic osteoarthropathy: Conseusus on its definition, classification assessment and diagnostic criteria. *Journal of Rheumatology*, **20**, 1386–7.

Matteson, E.L., Hickey, A.R., Maguire, L., Tilson, H.H., and Urowitz, M.B. (1991). Occurance of neoplasia in patients with rheumatoid arthritis enrolled in a DMARD registry. Rheumatoid arthritis azathioprine registry steering committee. *Journal of Rheumatology*, **18**, 809–14.

Maury, C.O.J. (1989). Anemia in rheumatoid arthritis: role of cytokines. *Scandinavian Journal of Rheumatology*, **18**, 3–5.

McGrath, B.P., Ibels, L.S., Raik, E., Hargrave, M., Mahony, J.F., and Stewart, J.H. (1975). Erythroid toxicity of azathioprine: macrocytosis and selective marrow hypoplasia. *Quarterly Journal of Medicine*, **XLIV** (No. 173), 57–63.

McNeil, H.P., Simpson, R.J., Chesterman, C.N., and Krilis, S.A. (1990). Anti-phospholipid antibodies are directed against a complex antigen that includes a lipid-binding inhibitor of coagulation: Beta2-glycoprotein I (apoliprotein H). *Procedings of the National Academy of Sciences* (USA), **87**, 4120–4.

Menon, S., Snaith, M.L., and Isenberg, D.A. (1993). The association of malignancy with SLE: an analysis of 150 patients under long term review. *Lupus*, **2**, 177–81.

Miller, M.L. (1990). Treatment of systemic lupus erythematosus in adults and children. *Current Opinion in Rheumatology*, **2**, 712–16.

Morel, P., Quiquandon, I., Janin, A., Dupriez, B., Desableus, B., Gosselin, B., and Bauters F. (1993). High incidence of lymphoid infiltration on labial salivary gland biopsy in non-Hodgkin's lymphomas: clinical implications. *British Journal of Haematology*, **85**, 93–8.

Morgan, S.L. *et al*. (1990). The effect of folic acid supplementation on the toxicity of low-dose methotrexate in patients with rheumatoid arthritis. *Arthritis and Rheumatism*, **33**, 9–18.

Mowat, A.G. (1971). Hematologic abnormalities in rheumatoid arthritis. *Seminars in Arthritis and Rheumatism*, **1**, 195–219.

Newton, P., Freemont, A.T., Noble, J., Scott, N., Grennan, D.M., and Hilton, R.C. (1984). Secondary malignancy synovitis: report of three cases and review of the literature. *Quarterly Journal of Medicine*, **53**, 135–43.

Nyfors, A. and Jensen, H. (1983). Frequency of malignant neoplasms in 248 long-term methotrexate treated psoriatics. *Dermatologica*, **167**, 260–1.

O'Brien, W.M. and Bagby, G.F. (1985). Rare adverse reactions to nonsteroidal antiinflammatory drugs. *Journal of Rheumatology*, **12**, 13–20; 347–53; 562–7; 785–90.

Parke, A.L. (1989). Antiphospholipid antibody syndromes. *Rheumatic Disease Clinics of North America*, **15**, 275–86.

Pinals, R.S. (1988). Sulfasalazine in the rheumatic disease. *Seminars in Arthritis and Rheumatism*, **17**, 246–59.

Pincus, T. *et al*. (1990). Multicenter study of recombinant human erythropoietin in correction of anemia in rheumatoid arthritis. *American Journal of Medicine*, **89**, 161–8.

Rennie, J.A.N. and Auchterlonie, I.A. (1991). Rheumatological manifestations of the leukemias and graft vs. host disease. *Baillière's Clinical Rheumatology*, **52**, 231–51.

Rosenthal, A.K., McLaughlin, J.K., Linet, M.S., and Persson, I. (1993). Sclerodrma and malignancy: an epidemiologic study. *Annals of the Rheumatic Diseases*, **52**, 531–3.

Santavirta, S. (1992). Synovial sarcoma: a clinicopathological study of 31 cases. *Archives of Orthopaedic Trauma Surgery*, **111**, 155–9.

Schwarzer, A.C., Fryer, J., Preston, S.J., Schreiber, L., and McNaught, P.J. (1990). Metastatic adenosquamous carcinoma presenting as an acute monoarthritis, with review of the literature. *Journal of Orthopaedic Rheumatology*, **3**, 175–85.

Segal, A.M. and Mackenzie, A.H. (1982). Hypertrophic osteoarthropathy: a 10-year retrospective analysis. *Seminars in Arthritis and Rheumatism*, **12**, 220–32.

Sela, O. and Shoenfeld, Y. (1988). Cancer in autoimmune diseases. *Seminars in Arthritis and Rheumatism*, **18**, 77–87.

Shoenfeld, Y. and Schwartz, R.S. (1983). Hematologic manifestations. In *The clinical management of systemic lupus erythematosus* (ed. P.H. Schur), pp. 123–36. Grune and Stratton, New York.

Singh, G., Fries, J.F., Spitz, P., and Williams, C.A. (1989). Toxic effects of azathioprine in rheumatoid arthritis. *Arthritis and Rheumatism*, **32**, 837–43.

Smeenk, R.J.T., Lucassen, W.A.M., and Swank, T.J.G. (1987). Is anticardiolipin activity a cross-reaction of anti-DNA or a separate entity? *Arthritis and Rheumatism*, **30**, 607–17.

Spilberg, I. and Meyer, G.J. (1972). The artritis of leukemia. *Arthritis and Rheumatism*, **15**, 630–5.

Triplett, D.A. (1989). Antiphospholipid antibodies and recurrent pregnancy loss. *American Journal of Reproductive Immunology*, **20**, 52–67.

Triplett, D.A. (1990). Laboratory diagnosis of lupus anticoagulant. *Seminars in Thrombosis and Hemostasis*, **16**, 182–92.

Urowitz, M.B. and Rider, W.D. (1985). Myeloproliferative disorders in patients with rheumatoid arthritis treated with total body irradiation. *American Journal of Medicine*, **78** (suppl. 1A), 60–4.

Vianna, J.L., Khamashta, M.A., Ordi-Ros, J., Cervera, R., Lopez-Soto, A., Tolosa, C., Franz, J., Selva, A., Ingelmo, M. *et al*. (1994). Comparison of the primary and secondary antiphospholipid antibody syndromes: A European multi center study of 114 patients. *American Journal of Medicine*, **96**, 3–9.

Weinblatt, M.E. (1985). Toxicity of low dose methotrexate in rheumatoid arthritis. *Journal of Rheumatology*, **12** (Suppl. 12), 35–9.

1.3.9 An anaesthetic perspective

Allan I. Binder and Frances Dormon

Introduction

Rheumatic diseases have important implications in anaesthetic practice (Binder and Isenberg 1989; Skues and Welchew 1993). Arthritis affecting the cervical spine, temporomandibular and cricoarytenoid joints, and involvement of the mouth, lips, nasopharynx, and oesophagus may result in increased risk and difficulty during intubation, maintenance, and the recovery from general anaesthesia. Constitutional illness, anaemia, and the involvement of specific organs are common and may complicate anaesthetic management. Medication, in particular systemic steroids, may also be important. The sedentary life-style of many patients with rheumatic disease accelerates atherosclerotic vascular disease and increases the risk of infection postoperatively.

This chapter will concentrate on the problems from the anaesthetist's point of view, so the rheumatologist can help by emphasizing the problems of an individual patient in relation to the particular procedure. We include discussion of alternatives to general anaesthesia, such as regional and local blocks, with their relative advantages and disadvantages. Newer techniques to increase the safety of intubation, including fibreoptic laryngoscopy and 'minitracheotomy' devices that can be used with jet ventilators, and anaesthetic drugs with specific advantages are mentioned. Finally, the role of the pain clinic, usually run by anaesthetists, is examined.

Systemic factors in rheumatic disease and their relevance to routine preanaesthetic assessment and anaesthesia

Autoimmune diseases such as systemic lupus erythematosus, juvenile chronic arthritis, scleroderma, and rheumatoid arthritis are multisystem diseases. As the emphasis of clinical features varies greatly from patient to patient, with many patients showing features of more than one autoimmune disease, preanaesthetic investigation needs to be tailored specifically to the individual patient. In some circumstances, such as the pregnant patient with systemic lupus erythematosus, the anaesthetist needs to consider the risks of anaesthesia to both the mother and fetus (Davies 1991). In order that the possible dangers of anaesthesia in patients with autoimmune disease may be anticipated and hence reduced, consideration will be given to the impact of the disease on major organs, with appropriate preanaesthetic investigation.

Cardiac disease

Patients with systemic lupus erythematosus, even if drug-induced (Goldberg *et al*. 1980), juvenile chronic arthritis (Goldenberg *et al*. 1992), scleroderma, and sometimes other autoimmune diseases can have serious cardiovascular pathology. The presence of valvular vegetations is associated with anticardiolipin antibodies (Leung *et al*. 1990; Metz *et al*. 1994) but also occurs in lupus patients without these antibodies (Gleason *et al*. 1993). Many other patients with autoimmune disease have subclinical cardiac abnormalities (Giunta *et al*. 1993; Corrao *et al*. 1995) which are unlikely to affect anaesthetic management, provided that the procedure is uncomplicated. However, they can prove important if the anaesthetist experiences difficulties with intubation and the delay leads to inadvertent hypoxia. For this reason, careful preoperative evaluation of the cardiovascular system should always be carried out.

Clinical assessment, electrocardiography, and chest radiographs will exclude major cardiac disease. Echocardiography and echo-Doppler cardiography (Cujec *et al*. 1991; Giunta *et al*. 1993) provide non-invasive methods to define the presence and severity of cardiac lesions, and so avoid unexpected difficulty during anaesthesia. If abnormalities are found, further investigation of cardiac reserve may be required. With a significant conduction defect, as occurs in some patients with spondylarthropathies (Tucker *et al*. 1982), the preoperative insertion of a temporary pacing wire is advisable or, as a minimum requirement, facilities should be available for pacing during surgery. Patients with reduced cardiac reserve need more detailed monitoring of their central venous pressure, arterial pressure, and sometimes cardiac output. In all cases, the cardiac status should be improved as far as possible preoperatively.

In patients with cardiac disease, general anaesthesia is often preferable. The modern technique of 'balanced anaesthesia' (Bailey and Stanley 1994) depends on the use of several drugs for induction. The smaller dose of each drug in the combination is associated with fewer undesirable side-effects. However, particular care still needs to be taken with the dosage of drugs during induction and reversal of anaesthesia. If the patient awakes too rapidly or has excessive pain, an endogenous surge of adrenaline may occur, which can destabilize the cardiac status.

If general anaesthesia is considered undesirable, local anaesthetic techniques can be considered. Limiting factors are the ability of patients to lie still for sufficiently long and the positioning required by the surgeon (Skues and Welchew 1993). A spinal block can cause a precipitous fall in the blood pressure in the compromised patient (Malmqvist *et al*. 1987). Although ephedrine and fluid will reverse this fall, the cardiac status may be too critical for these to be used. For these reasons, epidural blockade which develops more slowly, is safer.

Patients with suspected valvular abnormality, septal defect, or patent ductus arteriosus should be prescribed antibiotic prophylaxis (White 1985). The antibiotic should be started 1 h preoperatively if the oral or intramuscular route is used, or just before intubation if given intravenously.

Respiratory disease

Pulmonary disease may be of several types. Pulmonary fibrosis is the most characteristic and important for anaesthesia as it causes a progressive reduction in lung compliance and vital capacity. Impairment of gas transfer can lead to pulmonary hypertension (Winslow *et al*. 1995) and respiratory failure.

In systemic lupus erythematosus, pleuropulmonary complications are common and often severe, especially when associated with vascular abnormalities as a result of antiphospholipid antibodies (Asherson and Cervera 1995). Scleroderma also frequently has respiratory complications, resulting in complex abnormalities which

include a restrictive ventilatory defect, airflow obstruction, and a reduced diffusing capacity for carbon monoxide. Pulmonary vascular disease and restriction of chest expansion may also coexist, increasing the risk of respiratory failure and pulmonary hypertension (Arroliga *et al*. 1992). The severity of lung involvement in scleroderma does not always correlate with the extent of extrapulmonary disease (Tashkin *et al*. 1994), and can dominate the clinical picture.

Ankylosis of the thoracic spine and involvement of costovertebral joints typical of ankylosing spondylitis, as well as the seronegative spondylarthropathies, can seriously impair chest expansion. These patients usually maintain adequate ventilation, unless other factors such as bronchitis, pneumonia, or pneumothorax intervene (Radford *et al*. 1977).

Reduction in chest expansion can also be found in some patients with severe kyphosis due to advanced osteoporotic collapse of thoracic vertebrae or weakness of the respiratory muscle as a result of dermatomyositis (Schwarz 1992). In dermatomyositis, general anaesthesia is preferable, as respiration can be controlled with greater certainty by mechanical ventilation. After anaesthesia, pulmonary reserves may decline in all the above types of patients, leading to respiratory failure.

Preoperative assessment should identify significant respiratory disease. As a general rule, the ability to climb two flights of stairs without undue breathlessness suggests adequate respiratory reserve for general anaesthesia. If a problem is suspected, pulmonary function tests and baseline blood–gas analyses are necessary. The function tests should include transfer factor and a vitalograph, so the severity of restrictive and obstructive components of lung disease can be defined, and reversible elements treated. The baseline blood–gas analyses are especially helpful if respiratory failure develops in the postoperative period, where the preoperative carbon dioxide can indicate the level at which weaning off the ventilator can be attempted.

The choice of anaesthetic technique usually depends upon the surgical requirements, with some consideration to the likely postoperative course. For major thoracic and abdominal surgery, the postoperative period can be complicated by respiratory insufficiency, owing to pain and the residual effect of the anaesthetic (Catley *et al*. 1985). Regional techniques are often preferred to general anaesthesia in patients with serious respiratory disease, although long operations and the supine position may be particularly unpleasant for breathless and disabled patients. If sedation is added to a regional technique, great care is necessary to avoid oversedation and dangerous respiratory depression. A good method of analgesia is just as important following regional as general anaesthesia, as respiratory failure is as likely with either technique if pain control is inadequate.

Renal disease

Serious renal failure is a common finding in patients with systemic lupus erythematosus, but can occur in scleroderma and other autoimmune diseases. Subclinical renal dysfunction can also occur in lupus (Cottiero *et al*. 1995), rheumatoid (Pedersen *et al*. 1995), and other connective tissue diseases, and can be exacerbated by non-steroidal anti-inflammatory drugs (Segasothy *et al*. 1995), amyloidosis, or unrelated causes.

A mild reduction in renal function is not a problem to the anaesthetist, but if present should be noted so as to avoid prolonged perioperative fluid restriction. A drip will maintain hydration but fluid overload should be avoided if the patient also has cardiac disease.

With more serious renal failure, creatinine clearance and isotopic renograms (O'Malley and Ziessman 1993) help define the severity of the renal disease and residual reserve of renal function, although neither is accurate in advanced disease. In these patients, expert nephrological advice should be sought.

Monitoring of the central venous pressure is usually needed to help maintain the fluid balance in patients with renal disease, especially during and after major surgery, where there are large shifts of fluid complicating the fluid balance. Whichever technique of anaesthesia is used in patients with autoimmune diseases, the plasma concentration of anaesthetic drugs may be affected by the degree to which they are protein-bound (Wood 1986) or by the presence of renal dysfunction.

Haemopoietic disease

Anaemia is a common feature of rheumatic diseases and can result from many varied causes. The 'anaemia of chronic disease', typical of inflammatory diseases such as rheumatoid arthritis, results from ineffective erythropoiesis (Vreugdenhil *et al*. 1990), and rarely drops the haemoglobin to below 9 g/100 ml. Ideally, anaemia should be assessed, and where possible treated, before elective surgery. A haemoglobin of 10 g/100 ml is considered adequate for general anaesthesia, although lower levels are acceptable for simple surgical procedures which are unlikely to cause appreciable blood loss, provided the anaemia is chronic and physiological adaptation has been achieved. Preoperative oximetry is very helpful in the assessment of anaemic patients as it demonstrates the presence of desaturation without cyanosis.

If preoperative transfusion is necessary, it should be completed at least 24 h before surgery to allow the transfused red cells to replenish the levels of 2,3-diphosphoglycerate and function in oxygen transfer. If a delay in surgery is not possible, the transfusion should be delayed until the time of the surgery, to avoid inadvertent fluid overload developing preoperatively.

Complex haematological abnormalities can develop as a result of systemic lupus erythematosus, Felty's syndrome, and occasionally other rheumatic diseases. The drugs used in treatment can also cause abnormalities. Full haematological assessment in these patients should include white blood count, platelet count, and a coagulation screen. Anticardiolipin antibodies and cryoglobulins should also be measured in some cases. Anticardiolipin antibodies increase the risk of intravascular coagulation during surgery (Menon and Allt-Graham 1993) and paradoxically also occasionally cause excessive bleeding, especially when hypoprothrombinaemia and/or thrombocytopenia are also present (Shaulian *et al*. 1981). In these patients, fresh frozen plasma and platelet packs should be available for the surgery. Expert haematological advice should be sought if complex abnormalities are found.

Low white-cell counts as a result of systemic lupus, Felty's syndrome, or some rheumatic drugs increase the risks of postoperative sepsis, and decrease the white cell response and hence identification of infection. Prophylactic antibiotics should be prescribed in susceptible patients.

Prophylaxis against deep venous thrombosis is increasingly important, especially in patients undergoing elective surgery for total hip and knee arthroplasty (Lieberman and Geerts 1994) and other conditions with a high risk of venous thrombosis and pulmonary embolism (White 1985). Adjusted-dose heparin, low-dose warfarin, or the low molecular weight heparins all significantly reduce the risks of thromboembolism following hip arthroplasty. Although less thoroughly assessed, low-dose warfarin, the low molecular weight heparins, and pneumatic compression boots have similar value in the prevention of thromboembolism after knee arthroplasty. The optimum duration of thromboembolic prophylaxis has yet to be determined, with recommendations varying from discontinuing therapy when the patient begins to mobilize, to maintenance for as long as 3 months after the arthroplasty. This issue is especially relevant in view of the current trends toward earlier discharge from hospital and the concern about thromboembolic events following discharge (Trowbridge *et al*. 1994).

Gastrointestinal disease

Oesophageal fibrosis and hypomotility are common in scleroderma and, even when not clinically apparent, markedly increase the risk of reflux of gastric contents after anaesthesia (Orringer *et al*. 1976; Smith and Shribman 1984). Extubation should be delayed in these cases until the patient is able to maintain an independent airway. Positioning and careful observation should be continued well into the postoperative period to avoid aspiration and identify cricoarytenoid inflammation (Funk and Raymon 1975), both of which could result in respiratory embarrassment after extubation.

Skin and joints

Particular care is necessary in the handling, positioning, and movement of patients with generalized arthropathy who undergo surgery. Corticosteroid therapy and rheumatoid nodules increase the risks of pressure-induced skin damage. Unusual parts of the body may become weight bearing during surgery and will require additional padding and support. All patients with significant arthritis should be sent to theatre wearing a soft collar, as this serves to remind the anaesthetist and other theatre staff to avoid excessive movement of the neck and to handle the patient with extreme care. Dry eyes need to be protected with regular methylcellulose eyedrops during surgery. The application of petroleum jelly to the lips and nose will provide similar protection to oral and nasal mucosa.

Antirheumatic drugs

Prolonged corticosteroid therapy is associated with osteoporosis, deficient wound healing, reduced resistance to infection, and accelerated atherosclerosis. However,

the suppressive effect on the hypothalamic–pituitary–adrenal axis is the most important complication to the anaesthetist, as it can persist for up to a year after corticosteroid therapy has ended. The levels of cortisol do not show whether the axis has recovered and other functional tests of the axis are complicated and unreliable. Steroid cover should therefore be given to these patients. Intravenous hydrocortisone is started with surgery, the doses being tapered over the next 2 to 4 days and returned to the maintenance levels within a week. For major surgery, 200 mg of intravenous hydrocortisone is given with anaesthetic induction. The dosage of hydrocortisone infused is then reduced by 25 per cent each day untill the maintenance dose is reached. For lesser procedures, half the initial dose, that is 100 mg of hydrocortisone, is adequate (Roizen 1994). The trend is to shorten the period of steroid cover, because of its adverse effect on wound healing and increased risk of infection (White 1985).

The anaesthetist must minimize the risks of infection by maintaining scrupulous cleanliness of all equipment used. Bacterial filters should be incorporated into the breathing circuits to reduce the risk of pulmonary infection. Preoperative dental hygiene is particularly important in scleroderma and Sjögren's syndrome and prophylactic antibiotics are sometimes needed.

Azathioprine, methotrexate, and other immunosuppressive drugs are also associated with an increased risk of infection and similar precautions to those for steroids should be adopted. A reduction in the white blood count and platelet counts, and possible adverse effects on wound healing, may also be associated with immunosuppressive drugs. Bridges *et al*. suggested some reduction in early infection rates if methotrexate was stopped 4 weeks before elective orthopaedic surgery, but Kasdan and June and Sany *et al*. failed to confirm this, and advocated continuation of methotrexate through the surgical period, until larger prospective studies have clarified the relative infection risks of continuing immunosuppression (Bridges *et al*. 1991; Kasdan and June 1993; Sany *et al*. 1993).

Emergency work-up for anaesthesia

Ideally there should be plenty of time for a full anaesthetic work-up to assess potential risks in patients with rheumatological disease (Binder and Isenberg 1989; Skues and Welchew 1993) and this has been considered above. However, these patients often require emergency surgery, precluding an ideal work-up. Under these circumstances, general assessment of mouth opening, neck movement, and laryngeal function at the bedside may help predict a difficult intubation. Two simple tests which may be helpful in predicting a difficult direct laryngoscopy are:

1. an inability to visualize the soft palate or uvula when the mouth is opened as widely as possible and the tongue is protruded (Mallampati *et al*. 1985);
2. an inability of the patient to protrude the lower teeth beyond the upper incisors suggesting reduced temporomandibular joint movement (Calder 1992).

If the patient has serious peripheral arthritis, instability of the cervical spine can be assumed and radiological assessment of the neck should be made. If there is serious concern about the neck, awake intubation (Sinclair and Mason 1984), especially via a fibreoptic laryngoscope (Sidhu *et al*. 1993), should also be considered. A soft collar to protect the vulnerable neck and careful handling have been mentioned earlier.

To avoid serious danger to the patients, the anaesthetist must be alerted to the presence and severity of systemic ill-health, anaemia, and the cardiac, respiratory, and renal complications of their rheumatic disease. Details of drug therapy such as corticosteroids and immunosuppressive agents should also be provided. All these factors may influence the approach to anaesthesia and the choice and dosage of anaesthetic agents.

Intubation difficulties

Rheumatoid arthritis

Rheumatoid arthritis is by far the most common autoimmune rheumatic disease presenting to the anaesthetist. The advent of joint replacement has markedly increased planned surgical intervention, even in patients with advanced disease. Emergency operations are also frequently needed, especially for upper gastrointestinal bleeding induced by non-steroidal anti-inflammatory drugs. Rheumatoid arthritis can affect any synovial joint and often affects the cervical spine (Crosby and Lui 1990), temporomandibular (Ericson and Lundberg 1967; Redlund-Johnell 1988), and cricoarytenoid joints (Funk and Raymon 1975; Lawry *et al*. 1984).

Adequate extension at the cervical spine is important for direct laryngoscopy and endotracheal intubation, and can be compromized by rheumatoid neck involvement. Cervical subluxation, which usually occurs at the atlantoaxial joint, can affect any level. While subluxation can cause cervical myelopathy or radiculopathy, it is often asymptomatic. Atlantoaxial subluxation usually occurs anteriorly but can also occur posteriorly, as a result of odontoid peg erosion (Castro *et al*. 1994). Progressive subaxial cervical subluxation affecting multiple levels, the so-called 'stair-case spine', typically occurs in seropositive rheumatoid patients with severe deforming peripheral arthropathy (Yonezawa *et al*. 1995). With progressive cervical disease, the neck becomes short and rigid, with superior migration of the dens, further complicating intubation. While neurological changes in rheumatoid patients usually results from subluxation, soft-tissue pannus formation can contribute to spinal cord compression (Castro *et al*. 1994). Temporomandibular joint involvement which limits mouth opening can also complicate direct laryngoscopy and intubation (Aiello and Metcalf 1992).

A hoarse voice, a sense of fullness in the throat, dysphagia, exertional dyspnoea, and rarely stridor may indicate disease of the cricoarytenoid joints which, if suspected, may require further investigation. Indirect laryngoscopy or fibreoptic examination of the cords may reveal rheumatoid nodules on the cords (Bridger *et al*. 1980) or cricoarytenoid joint involvement, with decreased cricoarytenoid movement, bowing of the cords during inspiration, fixed adduction of the cords, or glottic stenosis. In extreme cases, these findings indicate the need for elective tracheostomy or consideration of alternative methods of anaesthesia. The presence of such findings necessitates particular vigilance for signs of postextubation airway obstruction (Funk and Raymon 1975).

The anaesthetic work-up should identify potential problems with the airway and those patients likely to present difficulties during intubation (Macarthur and Kleiman 1993), although this assessment is notoriously unreliable (Wilson 1993). Bedside examination of neck and jaw movement, as described above, should be routinely carried out on all patients. Even under ideal circumstances, unexpected problems may arise (Keenan *et al*. 1983) and the anaesthetist needs to anticipate the possibility of a failed intubation.

In rheumatoid arthritis, radiographs of the cervical spine in flexion/extension with a through-mouth view of the odontoid peg are helpful in defining the presence and severity of any spinal instability, although the routine use of this investigation in all asymptomatic rheumatoid arthritis patients has been questioned (Campbell *et al*. 1995). Magnetic resonance imaging (**MRI**) (Castro *et al*. 1994; Yonezawa *et al*. 1995) is the investigation of choice in defining the presence and severity of subluxation and soft-tissue pannus affecting the cervical spine (Fig. 1).

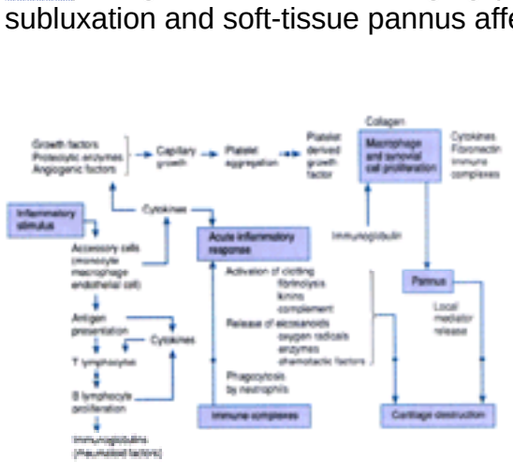


Fig. 1 Magnetic resonance scan of a rheumatoid patient showing atlantoaxial subluxation and compression of the spinal cord.

Having made a full assessment, the anaesthetist will choose to anaesthetize the patient in one of the ways listed in Table 1. The choice of anaesthetic approach is

often based on individual preference and, with some of the techniques, the level of individual expertise (Vaughan 1991).

No relaxant needed
Intravenous induction
Inhalational induction: with/without Guedel airway with/without laryngeal mask airway
Relaxant needed
Intravenous induction: delay relaxant until positive-pressure ventilation (PPV) is established with mask; intubate under (i) direct vision, or (ii) with introducer, or (iii) blind oral/nasal, or (iv) with a fibreoptic 'scope
Inhalational induction: delay relaxant until PPV established with mask intubate as for intravenous induction
Awake intubation: using fibreoptic laryngoscope
Awake insertion of 'minitracheotomy': ventilate with jet ventilator through minitracheotomy for induction and intubation
Elective tracheostomy in the awake patient

Table 1 Methods of airway maintenance during anaesthetic induction

Juvenile chronic arthritis

Juvenile chronic arthritis, particularly of polyarticular or systemic onset, often results in serious and early involvement of the cervical spine (Hensinger *et al*. 1986). In addition to instability, growth abnormalities of the vertebral bodies, bony fusions, and sometimes torticollis add to the intubational difficulties. Development of the temporomandibular joint is similarly affected, with micrognathia and reduced mouth opening giving clues to the severity of temporomandibular involvement (Stabrun *et al*. 1988). MRI scanning of the temporomandibular joints is also valuable in their assessment (Taylor *et al*. 1993). Serious disease of the cricoarytenoid joints can also occur (Jacobs and Hui 1977).

The younger patient (below the age of 14 years) with potential for a difficult intubation presents the anaesthetist with a far greater problem (Smith 1990). There are fewer anaesthetic options in the small child because techniques such as awake intubation are unsuitable and venous access may be difficult to achieve before induction of anaesthesia. The anatomy of the child is different from the adult, making visualization of the larynx more difficult. Sizes of the tracheal tube need to be scaled down according to age, and adult equipment cannot be used, increasing the hazards for anaesthetists unfamiliar with young patients. The physiological differences include a reduced respiratory reserve which leads to more rapid onset of cyanosis if there is a delay in intubation or difficulty in maintaining the airway.

For all these reasons, very young children with significant arthritis should be treated in specialist paediatric centres where expertise with difficult intubation techniques is greater. However, the options listed in Table 1 still generally apply. The laryngeal mask (see Techniques for difficult intubation below), can be used in all age groups and there is a paediatric version of the 'minitracheotomy' for awake intubation or airways maintenance during induction and intubation in children aged 6 years or older. Furthermore, intubation can be achieved using a smaller fibreoptic laryngoscope or a guide wire technique with the adult fibreoptic laryngoscope.

For certain procedures where intubation is unnecessary, ketamine can be used for anaesthesia (D'Arcy *et al*. 1976). This drug usually results in the maintenance of adequate respiration and a clear airway. However, it does not guarantee a safe airway and should not be used unless full intubation facilities and equipment for emergency tracheostomy are available. Ketamine is also unsuitable for older children, who are more liable to suffer from troublesome hallucinations during the recovery phase.

Ankylosing spondylitis and seronegative spondylarthropathies

Significant ankylosis of the cervical spine is common in ankylosing spondylitis, and can occur in psoriatic arthritis (Salvarani *et al*. 1992) and any of the seronegative spondylarthropathies (Dekker-Saeys *et al*. 1978; Olivieri 1988). Disease of the cricoarytenoid joints can also be present. Ankylosis is rarely missed during the preanaesthetic assessment in patients with severe involvement, although these patients often generate more anxiety than is warranted. While difficult to intubate because of the rigidity of the neck, the airway is easy to maintain without intubation or manipulation. The laryngeal mask airway (Brain 1983) is easy to insert and, provided that the intercostal joints are not severely ankylosed, it is possible to ventilate the patients with the mask. If intubation is considered essential, it should be achieved with minimal manipulation in the awake patient (Sinclair and Mason 1984) or using a 'blind' technique. Care must be taken in these patients as atlantoaxial subluxation or vertebral fracture can occur from relatively minor trauma (Murray and Persellin 1981; Suarez-Almazor and Russell 1988).

Sjögren's syndrome and scleroderma

These conditions, alone or in conjunction with other autoimmune states, can cause problems with intubation because of their specific effects on the mouth and nose.

In Sjögren's syndrome, xerostomia may result in poor mouth opening, and dryness and friability of the oral and nasal mucosa, making it difficult to insert the laryngoscope. The poor dental state may also result in loose teeth, which could dislodge during intubation, or single-standing teeth, which could impede laryngoscopy.

In scleroderma, the reduced mouth opening can be even more severe and further complicate an intubation made potentially difficult by abnormalities of cervical and mandibular joints. Poor dental hygiene is almost invariable in patients with severe scleroderma and patients need to be warned that if teeth are electively removed for anaesthesia, dentures may be impossible to fit. Whether the oral or nasal route is used for intubation, severe lacerations or bleeding from telangectasia may follow (Davidson-Lamb and Finlayson 1977).

Pulmonary fibrosis is commonly found in patients with scleroderma, Sjögren's syndrome, and other autoimmune diseases and progressive reduction in lung compliance and vital capacity can also result from sclerodermatous involvement of the chest wall.

Venous access is often a problem in scleroderma, limiting the choice for induction of anaesthesia (Smith and Shribman 1984). Vasospastic abnormalities are also common and vasodilatation must be anticipated after induction of anaesthesia (Eisele 1990), with plasma expanders or blood being available to maintain the blood pressure after induction. Infused fluids must be warmed to prevent vasospastic irritability in these patients. Intravenous induction may provoke a painful cyanotic reaction, similar to that experienced by patients with Raynaud's phenomenon (Davidson-Lamb and Finlayson 1977). Intra-arterial cannulation should be avoided, as gangrene of the limb can result from thrombosis. The maintenance of the ambient temperature in theatre and throughout the postoperative period is another important consideration in these patients.

If local anaesthetic techniques are to be used, vasoconstrictors should be avoided, as the blood supply may be further compromized (Eisele 1990). Regional anaesthesia can also result in prolonged but reversible sensory blockade, which is useful as analgesia, but can be associated with prolonged vasospastic change (Neill 1980; Smith and Shribman 1984).

Techniques for difficult intubation

Table 1 lists the range of techniques available to the anaesthetist. Not all patients require intubation, particularly for short procedures on the limbs, and local anaesthetic techniques may be more suitable in these cases. Repeated attempts at endotracheal intubation can increase the risk of neurological damage (Hastings and Kelley 1993) and incidence of airway obstruction due to glottic trauma in the postoperative period.

Induction of general anaesthesia can be achieved smoothly with an intravenous or inhalational agent. If the airway is difficult to maintain, an oral or nasal airway may help. The laryngeal mask airway (Brain 1983; Calder *et al*. 1990) passed orally, fits snugly over the larynx (Fig. 2), and is of proven value in the safer management of anaesthesia in arthritic patients. However, the laryngeal mask airway does not protect the airway from soiling by gastric contents and is therefore not a replacement for intubation in all cases. Intubation is essential if the surgery requires muscle relaxation, the operation is longer, the patient needs to lie prone, or there is pulmonary

disease.

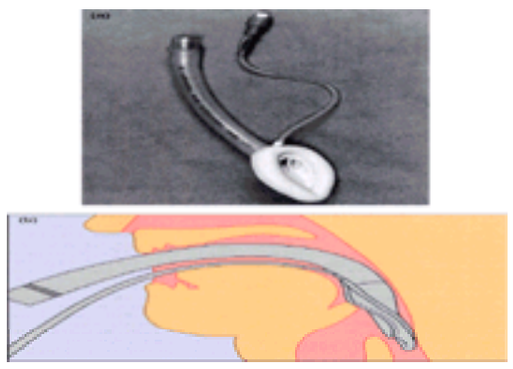


Fig. 2 (a) Laryngeal mask airway. (b) Diagram of a laryngeal mask airway in place.

Despite careful assessment, it is not always possible to predict what will happen to the airway after induction (Wilson 1993). Sometimes, even with mild disease, the airway becomes obstructed soon after induction and the use of airway devices does not help. On other occasions, despite apparently adequate mouth opening preoperatively, problems arise after induction. Therefore, with all these patients, induction of anaesthesia is undertaken slowly, allowing repeated reassessment of the adequacy of the airway. Inhalational anaesthesia is the safer option in these cases, as induction can be halted if the patient obstructs. After induction, and before any muscle relaxants are given, the anaesthetist must ensure that it is possible to ventilate the patient manually with a bag, mask, and airway. Provided this is possible, there is then enough time to try the several available techniques. If laryngoscopy is only moderately difficult and a grade 3 view of the larynx is obtained (Cormack and Lehane 1984) (Fig. 3), it is possible to intubate directly or by use of a soft bougie as an introducer over which the tracheal tube is passed. If this proves impossible, some anaesthetists are sufficiently experienced with blind-nasal intubation. The use of a small, fibreoptic intubating laryngoscope (Fig. 4) has radically altered the management of rheumatoid patients with cervical spine disease (King and Adams 1990), and may also reduce the risk of life-threatening postoperative complications, especially in patients undergoing posterior cervical spine surgery (Wattenmaker *et al* 1994). Several other methods of intubation have been described. A 'minitracheotomy' (Matthews *et al*. 1986) can be made and a catheter passed through the larynx from below. This can then be used as a guide through the oral route. The laryngeal mask airway can also be used as a guide, either for the blind passage of a soft-bougie introducer or to guide the fibreoptic intubating laryngoscope towards the larynx. Once the larynx has been identified, a tracheal tube can be 'railroaded' over the laryngoscope.

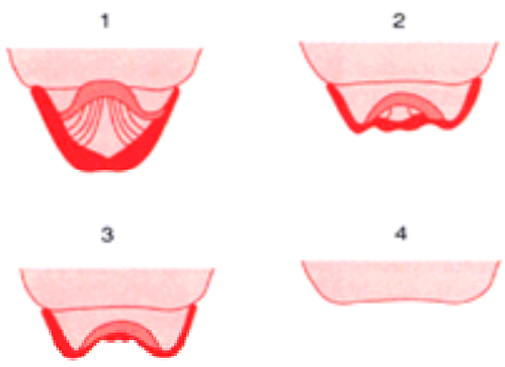


Fig. 3 Diagrammatic classification of the four grades of view of the larynx obtained at laryngoscopy in a difficult intubation: grade 1, most of the glottis visible; grade 2, only epiglottis and posterior commissure seen; grade 3, only epiglottis seen; grade 4, not even the epiglottis can be exposed.

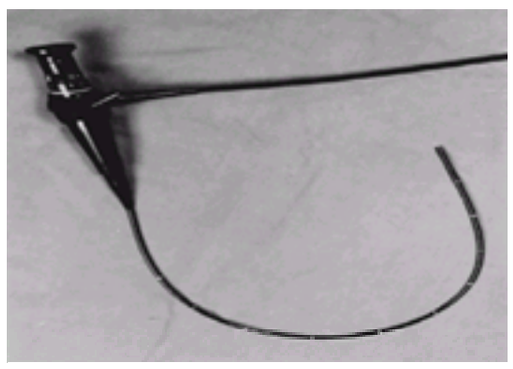


Fig. 4 Fibreoptic intubating laryngoscope.

If intubation is considered too hazardous following induction of anaesthesia, the patient should be intubated before induction, that is 'awake' (Sinclair and Mason 1984; Sidhu *et al*. 1993). This technique can be learnt with relative ease but should not be attempted unless the operator is familiar with the fibreoptic intubating laryngoscope in normal patients (Vaughan 1991).

Preoperative explanation of the procedure will reassure the patient and anticholinergic premedication can help by reducing secretions. Anxiolytic agents should be used with care because respiratory depression may occur or the airway may become obstructed. Several methods have been described to achieve analgesia of the pharynx and upper trachea but care must be taken to avoid toxicity if a 4 per cent lignocaine spray is used.

A 'minitracheotomy' device inserted under local anaesthesia can be used with a jet ventilator to ensure adequate ventilation (Matthews *et al*. 1986) but should only be used by those familiar with the jet, because a failure to maintain a clear airway for expiration through the larynx would result in pneumothorax.

Finally, if intubation is considered impossible by any other route, and surgery can only be carried out under general anaesthesia, an elective tracheostomy should be performed under local anaesthesia. New kits are now available for percutaneous dilational tracheostomy, which is quicker to perform with no increase in the complication rate (Bodenham 1993)

Postoperative care

Immediately following extubation prolonged observation is necessary to ensure that airway obstruction does not develop, especially in patients with serious cricoarytenoid (Funk and Raymon 1975), temporomandibular, or cervical spine pathology. In some cases, postoperative ventilation may be necessary for 24 to 48 h, to ensure adequacy of the airway and oxygenation. Aspiration of gastric contents, which can lead to respiratory embarrassment, is a particular risk in scleroderma patients where gastro-oesophageal reflux is common (Orringer *et al*. 1976; Smith and Shribman 1984).

The main challenge facing the anaesthetist immediately following major surgery is the provision of effective analgesia without respiratory depression. Ventilatory insufficiency in the postoperative period can be precipitated by a number of causes such as restrictive lung disease, infection, obesity, diaphragmatic splinting due to

pain, and the residual effect of the anaesthetic (Catley *et al*. 1985) superimposed on pre-existing borderline respiratory function. If the carbon dioxide is significantly raised preoperatively, then postoperative ventilation is usually necessary, although weaning off the ventilator should be attempted with a higher carbon dioxide than normal.

Early and regular chest physiotherapy, bronchodilators, and antibiotics are important in postoperative management. If the retention of sputum becomes a problem, a 'minitracheotomy' (Matthews *et al*. 1986) should be inserted to assist in the removal of secretions before the patient tires and develops respiratory failure.

Adjuvant epidural analgesia, cryotherapy to the intercostal nerves during thoracotomy, or patient-controlled intravenous opioid analgesia are all methods that optimize postoperative pain control and hence reduce the risk of respiratory failure. Patient-controlled analgesia systems may not be appropriate in rheumatoid patients with severe hand involvement (Jani and Trujillo 1994) and catheter techniques of regional blockade using local anaesthetic agents (see below) may be more effective.

Opioid drugs need to be used with particular caution in rheumatoid patients as obstructive sleep apnoea (Mahowald *et al*. 1989) may occur, especially in patients with temporomandibular joint involvement (Redlund-Johnell 1988). Non-steroid anti-inflammatory drugs given by the suppository, intramuscular, or buccal routes have particular value in postoperative pain control, by limiting the need for opioids which often cause vomiting and respiratory depression.

After major abdominal or thoracic surgery there are large shifts of fluid into the interstitium, thoracic cavity, and bowel, complicating the fluid balance. The high incidence of renal and cardiac abnormalities may result in early cardiac or renal failure with only minor variations in fluid balance and therefore continuous monitoring of the central venous pressure may need to be extended well into the recovery period.

Careful positioning with additional padding and support to weight bearing parts of the body should be maintained, with pressure relieving mattresses used soon after surgery if indicated. However, passive exercises and mobilization also need to be started as early as possible to help prevent joint contracture, pressure sores, and other postoperative complications which are common in severely disabled patients. Rheumatoid therapy should also be reinstituted as soon as feasible and prophylaxis to prevent deep vein thrombosis, infection, and peptic ulceration should be continued in appropriate cases.

Alternatives to general anaesthesia

Local anaesthetic techniques, often combined with sedation, have gained increasing popularity in recent years. Improvement in education of the general public has contributed more to the increased usage than the introduction of new drugs or techniques. There are numerous techniques (Table 2) ranging in complexity from simple infiltration to blockade at the spinal level (Wildsmith and Armitage 1993).

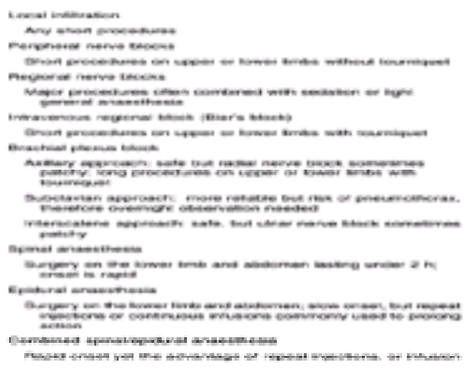


Table 2 Alternatives to general anaesthesia

Local infiltration

Careful explanation, with or without anxiolytic premedication, is all that is required before local infiltration. The appropriate use of sedative drugs enables fairly complex operations to be undertaken. The combination of 1 to 2 mg of midazolam and 10 to 30 mg of propofol (see Table 3) provides sedation and amnesia with relative safety. Full monitoring and anaesthetic facilities should however be available, because of the rare (1–2 per cent) occurrence of respiratory depression in especially sensitive patients.

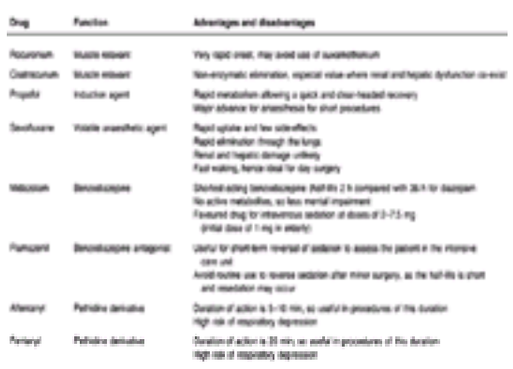


Table 3 Newer drugs of value in anaesthesia

Peripheral nerve blocks

These techniques are suitable for many surgical procedures on the limbs, especially where a tourniquet is unnecessary. Several texts list the individual blocks (Wildsmith and Armitage 1993).

Regional nerve blocks

These are useful for many major procedures, sometimes in combination with sedation or even light general anaesthesia without intubation.

Intravenous regional anaesthesia (Bier's block)

This form of anaesthesia is safe, simple, and useful for a wide range of procedures, given the necessary expertise. The local anaesthetic is injected intravenously into an arm or leg that has been isolated from the general circulation by a tourniquet. Care is necessary to ensure that the cuff is not inadvertently deflated for at least 20 min to avoid risks associated with the release of a bolus of local anaesthetic. Furthermore, full resuscitation facilities should be available. This type of anaesthesia is suitable for operations lasting about 40 min, as tourniquet pain occurs after this time. However, a major advantage is that patients can be discharged within an hour of the procedure.

Brachial plexus blocks

The brachial plexus can be approached by three distinct routes, each having some advantages and disadvantages.

The axillary approach

The axillary approach is the safest and therefore the most commonly used route. However, the axillary block has the disadvantages that the success rate is variable and supplementary peripheral nerve blocks are often required. The block of the radial nerve is often patchy and, if a tourniquet is used, nerve blocks may be needed for the tourniquet area.

The subclavian approach

The subclavian approach is more reliable, particularly in the upper arm and tourniquet areas, but is associated with a greater risk of pneumothorax, making overnight observation in hospital desirable. A bilateral block should not be attempted because slow development of pneumothoraces may occur.

The interscalene approach

The interscalene approach is also reliable but the ulnar block may be patchy. It should be reserved for patients with emphysema or chronic obstructive airways disease, where the risk of pneumothorax is high, and for operations on the upper arm.

Spinal anaesthesia

Increased use of this form of anaesthesia has followed the introduction of narrow-gauge needles, which have reduced the incidence of spinal headache to under 5 per cent, and obviated the need for 24 h of bed-rest after the procedure. The block, which is established within 10 min, is more reliable and complete than an epidural block and lasts for up to 2 h. The time limit can be extended by the insertion of a fine-bore catheter for continuous spinal anaesthesia, although the introduction of the combined spinal/epidural technique has reduced the use of spinal catheters. This type of block should be used with great care in patients with poor myocardial reserve, as its rapid onset can lead to sudden hypotension due to peripheral vasodilatation (Malmqvist *et al*. 1987).

The use of a hyperbaric mixture of anaesthetic, such as bupivacaine in 4 per cent dextrose, allows the establishment of a unilateral block for surgery on one leg. For this type of block, the spinal needle is inserted with the patient lying on the side requiring surgery for at least 10 min while the block is established. The drop in blood pressure is far less with a unilateral block.

Epidural anaesthesia

This has gained wide usage even in the presence of extensive spinal disease. The wider-bore epidural needle allows a catheter to be inserted easily for continuous infusion. The block takes longer to establish and there is a small chance of 'unblocked' segments, where nerve roots are left unaffected by the anaesthetic. If the needle is advanced too far, the dura may be pierced producing a dural leak which can cause severe headache and very rarely complications such as arachnoiditis (Sklar *et al*. 1991) have been reported.

Combined spinal and epidural anaesthesia

Using a special needle pack, the epidural space is identified and a long spinal needle is used to pierce the dura and perform a spinal block. The spinal needle is then removed and a catheter inserted into the epidural space. This combined technique offers the advantages of a rapid onset spinal block with the facility to administer 'top-ups' or a continuous infusion for longer operations or to provide postoperative analgesia.

Newer drugs of value in anaesthesia

A few agents recently introduced into anaesthetic practice have improved the safety and efficiency of anaesthesia for patients with multisystem disease. These drugs are shown in Table 3 with their relative advantages and disadvantages.

The pain clinic

Pain clinics are being established in large numbers because of their value in the management of patients with intractable pain (Butler and Murphy 1989). Treatment is tailored to the individual need of each patient and includes physical and psychological measures to reduce pain, and, where this is impossible, to help patients to cope with their pain and modify their life-styles accordingly. Pain clinics are usually run by anaesthetists, assisted by a multidisciplinary team. Detailed evaluation of these clinics is outside the scope of this chapter but the major treatment options are shown in Table 4. Although a large variety of conditions are referred to pain clinics, chronic mechanical backache, neck and shoulder pain, and reflex sympathetic dystrophy are the most commonly referred by rheumatologists.

Multidisciplinary team	Anaesthetist, other physicians, social worker, physiotherapist, occupational therapist, psychologist
Drug therapy	Conventional analgesics, e.g. non-narcotic, analgesics
	Narcotic analgesics as a last resort
	Centrally-acting, pain-modifying drugs: tricyclic antidepressants, antiepileptic agents
Nerve blocks	Local, regional, spinal blocks
	Autonomic blocks: stellate ganglion, coeliac plexus and lumbar sympathectomy
	Facet joint injection or cryotherapy
Other	Guanethidine blocks
	Calcitonin injection therapy
	Clonidine injection
	Percutaneous cordotomy and other rare neurosurgical techniques

Table 4 The approach of a 'pain clinic' and its therapeutic modalities

Intractable back pain, having failed to respond to conventional treatment, may benefit from the approach of the pain clinic. Transcutaneous nerve stimulation, acupuncture, nerve blocks, epidural and facet-joint injections are some of the mechanical means of achieving pain control. Where the pain continues despite these measures, psychological assessment and treatment of depression may be valuable. Behaviour modification and other techniques to improve coping can also be taught.

Reflex sympathetic dystrophy (algodystrophy) is often intractable and when unresponsive to physiotherapy may benefit from guanethidine blocks, continuous epidural infusion, calcitonin injections, and chemical sympathectomy.

Patients with rheumatoid arthritis and inflammatory arthropathies at present form a very small proportion of referrals to pain clinics. However, there is increasing recognition of the value of pain specialists in the amelioration of pain from some sources. An example of this is the treatment of pain arising from the cervical spine, whether due to inflammatory or mechanical causes. Cervical epidural injection should be considered where the cervical spine pain is severe and unresponsive to physiotherapy but without any serious neurological abnormality. This treatment is particularly effective where muscle spasm and radicular pain are prominent features (Ferrante *et al*. 1993; Castagnera *et al* 1994). Suprascapular blocks for intractable shoulder pain, and local injection around the hip, can have similar value. Moreover, the multidisciplinary approach of the pain clinic may help in the management of pain and development of coping mechanisms in patients with generalized inflammatory arthropathies whose pain is inadequately controlled despite optimal medical and surgical treatment.

The true value of the pain clinics to rheumatology patients has not yet been fully explored. The marked improvement in postoperative pain control achieved by pain specialists over recent years suggests that there could be similar advances in the control of some types of pain in patients with inflammatory arthropathies.

Conclusions

This chapter has considered the systemic and joint abnormalities in patients with rheumatic disease from the perspective of the anaesthetist. The preanaesthetic assessment and techniques for general and regional anaesthesia have been described, with an insight into the difficulties often encountered and methods used for solving these problems. Finally, the role of the pain clinic has been considered.

Acknowledgement

The authors wish to thank Mr Alan Crockard for kindly providing the MRI of a rheumatoid neck, shown in Fig. 1.

Chapter References

Aiello, G. and Metcalf, I. (1992). Anaesthetic implications of temporomandibular disease. *Canadian Journal of Anaesthesia*, **39**, 610–7.

Arroliga, A.C., Podell, D.N., and Matthay, R.A. (1992). Pulmonary manifestations of scleroderma. *Journal of Thoracic Imaging*, **7**, 30–45

Asherson, R.A. and Cervera, R. (1995). Review: antiphospholipid antibodies and the lung. *Journal of Rheumatology*, **22**, 62–6.

Bailey, P.L. and Stanley, T.H. (1994). Intravenous opioid anaesthetics. In *Anesthesia*, (4th edn) (ed. R.D. Miller), pp 332–6. Churchill Livingstone, New York.

Binder, A.I. and Isenberg, D.A. (1989). Rheumatic diseases. In *Medicine in the practice of anaesthesia* (eds L. Kaufman and D.J. Betteridge), pp. 338–54. Edward Arnold, London.

Bodenham, A.R. (1993). Percutaneous dilational tracheostomy. Completing the anaesthetists range of airway techniques. (Editorial.) *Anaesthesia*, **48**, 101–2.

Brain, A.I. (1983). The laryngeal mask: a new concept in airway management. *British Journal of Anaesthesia*, **55**, 801–5.

Bridger, M.W., Jahn, A.F., and Van Nostrand, A.W. (1980). Laryngeal rheumatoid arthritis. *Laryngoscope*, **90**, 296–303.

Bridges, S.L. Jr., Lopez-Mendez, A., Han, K.H., Tracy, I.C., and Alarcon, G.S. (1991). Should methotrexate be discontinued before elective orthopedic surgery in patients with rheumatoid arthritis? *Journal of Rheumatology*, **18**, 984–8.

Butler, S.H. and Murphy, T.M. (1989). The management and functions of pain centres. In *Relief of intractable pain*, (4th revised edn) (eds M. Swerdlow and J.E. Charlton,), pp. 63–85. Elsevier, New York.

Calder, I. (1992). Predicting difficult intubation. *Anaesthesia*, **47**, 528–30.

Calder, I., Ordman A.J., Jackowski, A., and Crockard, H.A. (1990). The brain laryngeal mask airway: an alternative to emergency tracheal intubation. *Anaesthesia*, **45**, 137–9.

Campbell, R.S., Wou, P., and Watt, I. (1995). A continuing role for pre-operative cervical spine radiography in rheumatoid arthritis? *Clinical Radiology*, **50**, 157–9.

Castagnera, L., Maurette, P., Pointillart, V., Vital, J.M., Erny, P., and Senegas, J. (1994). Long-term results of cervical epidural steroid injection with and without morphine in chronic cervical radicular pain. *Pain*, **58**, 239–43.

Castro, S., Verstraete, K., Mielants, H., Vanderstraeten, G., de-Reuck, J., and Veys, E.M. (1994). Cervical spine involvement in rheumatoid arthritis: a clinical, neurological and radiological evaluation. *Clinical and Experimental Rheumatology*, **12**, 369–74.

Catley, D.M., Thornton, C., Jordan, C., Lehane, J.R., Royston, D., and Jones, J.G. (1985). Pronounced episodic oxygen desaturation in the postoperative period: its association with ventilatory pattern and analgesic regimen. *Anesthesiology*, **63**, 20–8.

Cormack, R.S. and Lehane, J. (1984). Difficult tracheal intubation in obstetrics. *Anaesthesia*, **39**, 1105–11.

Corrao, S. *et al*. (1995). Cardiac involvement in rheumatoid arthritis: evidence of silent heart disease. *European Heart Journal*, **16**, 253–6.

Cottiero, R.A., Madaio, M.P., and Levey, A.S. (1995). Glomerular filtration rate and urinary albumin excretion rate in systemic lupus erythematosus. *Nephron*, **69**, 140–6.

Crosby, E.T. and Lui, A. (1990). The adult cervical spine: implications for airway management. *Canadian Journal of Anaesthesia*, **37**, 77–93.

Cujec, B., Sibley, J., and Haga, M. (1991). Cardiac abnormalities in patients with systemic lupus erythematosus. *Canadian Journal of Cardiology*, **7**, 343–9.

D'Arcy, E.J., Fell, R.H., Ansell, B.M., and Arden, G.P. (1976). Ketamine and juvenile chronic polyarthritis (Still's disease). Anaesthetic problems in Still's disease and allied disorders. *Anaesthesia*, **31**, 624–32.

Davidson-Lamb, R.W. and Finlayson, M.C. (1977). Scleroderma. Complications encountered during dental anaesthesia. *Anaesthesia*, **32**, 893–5.

Davies, S.R. (1991). Systemic lupus erythematosus and the obstetrical patient—implications for the anaesthetist. *Canadian Journal of Anaesthesia*, **38**, 790–5.

Dekker-Saeys, B.J., Meuwissen, S.G., Van Den Berg-Loonen, E.M., De Haas, W.H., Agenant, D., and Tytgat, G.N. (1978). Ankylosing spondylitis and inflammatory bowel disease. II Prevalance of peripheral arthritis, sacroiliitis and ankylosing spondylitis in patients suffering from inflammatory bowel disease. *Annals of the Rheumatic Diseases*, **37**, 33–5.

Eisele, J.H. (1990). Connective tissue disease. In *Anaesthesia and uncommon diseases* (3rd edn) (eds J. Katz, J.L. Benumof, and L.B. Kadis), pp. 645–67. W.B. Saunders, Philadelphia.

Ericson, S. and Lundberg, M. (1967). Alterations in the temporomandibular joint at various stages of rheumatoid arthritis. *Acta Rheumologica Scandinavica*, **13**, 257–74.

Ferrante, F.M., Wilson, S.P., Iacobo, C., Orav, E.J., Rocco, A.G., and Lipson, S. (1993) Clinical classification as a predictor of therapeutic outcome after cervical epidural steroid injection. *Spine*, **18**, 730–6.

Funk, D. and Raymon, F. (1975). Rheumatoid arthritis of the cricoarytenoid joints: an airway hazard. *Anesthesia and Analgesia*, **54**, 742–5.

Giunta, A. *et al*. (1993). Spectrum of cardiac involvement in systemic lupus erythematosus:echocardiographic, echo-Doppler observations and immunological investigation. *Acta Cardiologica*, **48**, 183–97.

Gleason, C.B., Stoddard, M.F., Wagner, S.G., Longaker, R.A., Pierangelli, S., and Harris, E.N. (1993). A comparison of cardiac valvular involvement in the primary antiphospholipid syndrome versus anti-cardiolipin-negative systemic lupus erythematosus. *American Heart Journal*, **125**, 1123–9.

Goldberg, M.J., Husain, M., Wajszczuk, W.J., and Rubenfire, M.(1980). Procainamide-induced lupus erythematosus pericarditis encountered during coronary bypass surgery. *American Journal of Medicine*, **69**, 159–62.

Goldenberg, J. *et al*. (1992). Symptomatic cardiac involvement in juvenile rheumatoid arthritis. *International Journal of Cardiology*, **34**, 57–62.

Hastings, R.H. and Kelley, S.D. (1993). Neurological deterioration associated with airway management in a cervical spine-injured patient. *Anesthesiology*, **78**, 580–3.

Hensinger, R.N., Devito, P.D., and Ragsdale, C.G. (1986). Changes in the cervical spine in juvenile rheumatoid arthritis. *Journal of Bone and Joint Surgery*, **68**, 189–98.

Jacobs, J.C. and Hui, R.M. (1977). Cricoarytenoid arthritis and airway obstruction in juvenile rheumatoid arthritis. *Paediatrics*, **59**, 292–4.

Jani, K. and Trujillo, M. (1994) Getting to grips with patient controlled analgesia—an evaluation of handgrip disability. *British Journal of Anaesthesia*, **72** Suppl. 1, 118.

Kasdan, M.L. and June, L. (1993). Postoperative results of rheumatoid arthritis patients on methotrexate at the time of reconstructive surgery of the hand. *Orthopedics*, **16**, 1233–5.

Keenan, M.A., Stiles, C.M., and Kaufman, R.L. (1983). Acquired laryngeal deviation associated with cervical spine disease in erosive polyarticular arthritis: use of the fibreoptic bronchoscope in rheumatoid disease. *Anesthesiology*, **58**, 441–9.

King, T.A. and Adams, A.P. (1990). Failed tracheal intubation. *British Journal of Anaesthesia*, **65**, 400–14.

Lawry, G.V., Finerman, M.L., Hanafee, W.N., Mancuso, A.A., Fan, P.T., and Bluestone, R.(1984). Laryngeal involvement in rheumatoid arthritis: a clinical, laryngoscopic and computerized tomographic study. *Arthritis and Rheumatism*, **27**, 873–82.

Leung, W.H., Wong, K.L., Lau, C.P, Wong, C.K., and Liu, H.W. (1990). Association between antiphospholipid antibodies and cardiac abnormalities in patients with systemic lupus erythematosus. *American Journal of Medicine*, **89**, 411–9.

Lieberman, J.R. and Geerts, W.H. (1994). Prevention of venous thromboembolism after total hip and knee arthroplasty. *Journal of Bone and Joint Surgery*, **76**, 1239–50.

Macarthur, A. and Kleiman, S. (1993). Rheumatoid cervical joint disease: a challenge to the anaesthetist. *Canadian Journal of Anaesthesia*, **40**, 154–9.

Mahowald, M.W., Mahowald, M.L., Bundlie, S.R., and Ytterberg, S.R. (1989). Sleep fragmentation in rheumatoid arthritis. *Arthritis and Rheumatism*, **32**, 974–83.

Mallampati, S.R., *et al*. (1985). A clinical sign to predict difficult tracheal intubation: a prospective study. *Canadian Anaesthetic Society Journal*, **32**, 429–34.

Malmqvist, L.A., Bengtsson, M., Bjornsson, G., Jorfeldt, L., and Lofstrom, J.B. (1987). Sympathetic activity and haemodynamic variables during spinal analgesia in man. *Acta Anaesthetica Scandinavica*, **31**, 467–73.

Matthews, H.R., Fischer, B.J., Smith, B.E., and Hopkinson, R.B. (1986). Minitracheotomy: a new delivery system for jet ventilation. *Journal of Thoracic and Cardiovascular Surgery*, **92**, 673–5.

Menon, G. and Allt-Graham, J. (1993). Anaesthetic implications of the anti-cardiolipin antibody syndrome. *British Journal of Anaesthesia*, **70**, 587–90.

Metz, D. *et al*. (1994). Prevalence of valvular involvement in systemic lupus erythematosus and association with antiphospholipid syndrome: a matched echocardiographic study. *Cardiology*, **85**, 129–36.

Murray, G.C. and Persellin, R.H. (1981). Cervical fracture complicating ankylosing spondylitis: a report of eight cases and review of the literature. *American Journal of Medicine*, **70**, 1033–41.

Neill, R.S. (1980). Progressive systemic sclerosis. Prolonged sensory blockade following regional anaesthesia in association with a reduced response to systemic analgesics. *British Journal of Anaesthesia*, **52**, 623–5.

Olivieri, I. *et al*. (1988). Ankylosing spondylitis with predominant involvement of the cervical spine in a woman with Behçet's syndrome. *Annals of the Rheumatic Diseases*, **47**, 780–3.

O'Malley, J.P. and Ziessman, H.A. (1993). Quantitation of renal function using radioisotopic techniques. *Clinics in Laboratory Medicine*, **13**, 53–68.

Orringer, M.B., Dabich, L., Zarafonetis, C.J., and Sloan, H. (1976). Gastro-oesaphogeal reflux in oesophogeal scleroderma: diagnosis and implications. *Annals of Thoracic Surgery*, **295**, 120–30.

Pedersen, L.M., Nordin, H., Svensson, B., and Bliddal, H. (1995). Microalbuminuria in patients with rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **54**, 189–92.

Radford, E.P., Doll, R., and Smith, P.G. (1977). Mortality amongst patients with ankylosing spondylitis not given X-Ray therapy. *New England Journal of Medicine*, **297**, 572–6.

Redlund-Johnell, I. (1988). Upper airway obstruction in patients with rheumatoid arthritis and temporomandibular joint destruction. *Scandinavian Journal of Rheumatolology*, **17**, 273–9.

Roizen, M.F. (1994). Anaesthetic implications of concurrent disease. In *Anesthesia* (4th edn) (ed R.D. Miller), pp. 919–22. Churchill Livingstone, New York.

Salvarani, C. *et al*. (1992). The cervical spine in patients with psoriatic arthritis: a clinical, radiological and immunogenetic study. *Annals of the Rheumatic Disease*, **51**, 73–7.

Sany, J. *et al*. (1993). Influence of methotrexate on the frequency of postoperative infectious complications in patients with rheumatoid arthritis. *Journal of Rheumatology*, **20**, 1129–32.

Schwarz, M.I. (1992). Pulmonary and cardiac manifestations of polymyositis-dermatomyositis. *Journal of Thoracic Imaging*, **7**, 46–54.

Segasothy, M., Chin, G.L., Sia, K.K., Zulfiqar, A., and Samad, S.A. (1995). Chronic nephrotoxicity of anti-inflammatory drugs used in the treatment of arthritis. *British Journal of Rheumatology*, **34**, 162–5.

Shaulian, E., Shoenfeld, Y., Berliner, S., Shaklai, M., and Pinkhas, J. (1981). Surgery in patients with circulating lupus anticoagulant. *International Surgery*, **66**, 157–9.

Sidhu, V.S., Whitehead, E.M., Ainsworth, Q.P., Smith, M., and Calder, I. (1993). A technique of awake fibreoptic intubation. Experience in patients with cervical spine disease. *Anaesthesia*, **48**, 910–3.

Sinclair, J.R. and Mason, R.A. (1984). Ankylosing spondylitis. The case for awake intubation. *Anaesthesia*, **39**, 3–11.

Sklar, E.M., Quencer, R.M., Green, B.A., Montalvo, B.M., and Post, M.J. (1991). Complications of epidural anaesthesia: MR appearance of abnormalities. *Radiology*, **181**, 549–54.

Skues, M.A. and Welchew, E.A. (1993). Anaesthesia and rheumatoid arthritis. *Anaesthesia*, **48**, 989–97.

Smith B. (1990). Anaesthesia in paediatric rheumatology. In *Paediatric rheumatology update* (ed P. Woo, P.H. White, and B.M. Ansell), pp. 124–30. Oxford University Press.

Smith, G.B. and Shribman, A.J. (1984). Anaesthesia and severe skin disease. *Anaesthesia*, **39**, 443–55.

Stabrun, A.E., Larheim, T.A., Hoyeraal, H.M., and Rosler, M. (1988). Reduced mandibular dimensions and asymmetry in juvenile rheumatoid arthritis. Pathogenetic factors. *Arthritis and Rheumatism*, **31**, 602–11.

Suarez-Almazor, M.E. and Russell, A.S. (1988). Anterior atlantoaxial subluxation in patients with spondyloarthropathies: association with peripheral disease. *Journal of Rheumatology*, **15**, 973–5.

Tashkin, D.P. *et al*. (1994). Interrelationships between pulmonary and extrapulmonary involvement in systemic sclerosis: a longitudinal analysis. *Chest*, **105**, 489–95.

Taylor, D.B. *et al*. (1993). MR evaluation of the temporomandibular joint in juvenile rheumatoid arthritis. *Journal of Computer Assisted Tomography*, **17**, 449–54.

Trowbridge, A., Boese, C.K., Woodruff, B., Brindley, H.H., Lowry, W.E., and Spiro, T.E. (1994). Incidence of posthospitalization proximal deep venous thrombosis after total hip arthroplasty: a pilot study. *Clinical Orthopaedics*, **299**, 203–8.

Tucker, C.R., Fowles, R.E., Calin, A., and Popp, R.L. (1982). Aortitis in ankylosing spondylitis: early detection of aortic root abnormalities with two-dimensional echocardiography. *American Journal of Cardiology*, **49**, 680–6.

Vaughan, R.S. (1991). Training in fibreoptic laryngoscopy. *British Journal of Anaesthesia*, **66**, 538–40.

Vreugdenhil, G., Kroos, M.J., van Eijk, H.G., Lowenberg, B., and Swaak, A.J. (1990). Impaired iron uptake and transferrin binding by erythroblasts in the anaemia of rheumatoid arthritis. *British Journal of Rheumatology*, **29**, 335–9.

Wattenmaker, I., Concepcion, M., Hibberd, P., and Lipson, S. (1994). Upper airway obstruction and perioperative management of the airway in patients managed with posterior operations on the cervical spine for rheumatoid arthritis. *Journal of Bone and Joint Surgery*, **76**, 360–5.

White, R.H. (1985). Preoperative evaluation of patients with rheumatoid arthritis. *Semininars of Arthritis and Rheumatism*, **14**, 287–99.

Wildsmith, J.A.W. and Armitage, E.N. (eds) (1993). *Principles and practice of regional anaesthesia*, (2nd edn). Churchill Livingstone, Edinburgh.

Wilson, M.E. (1993). Predicting difficult intubation. *British Journal of Anaesthesia*, **71**, 333–4.

Winslow, T.M., Ossipov, M.A., Fazio, G.P., Simonson, J.S., Redberg, R.F., and Schiller, N.B. (1995). Five-year follow-up study of the prevalence and progression of pulmonary hypertension in systemic lupus erythematosus. *American Heart Journal*, **129**, 510–5.

Wood, M. (1986). Plasma drug binding: implications for anesthesiologists. *Anesthesia and Analgesia*, **65**, 786–804.

Yonezawa, T., Tsuji, H., Matsui, H., and Hirano, N. (1995). Subaxial lesions in rheumatoid arthritis. Radiographic factors suggestive of lower cervical myelopathy. *Spine*, **20**, 208–15.

1.3.10 The eye

Elizabeth M. Graham and Alison M. Leak

Introduction

The ophthalmologist has many roles in the management of patients with rheumatological problems. Most rheumatological diseases have ocular complications, many of which are specific to the disease, and consequently the ophthalmologist has an important diagnostic role. Many ophthalmic complications require specific therapy and in some cases may actually dictate the systemic therapy for the disease, giving the ophthalmologist a therapeutic role. Rheumatological diseases require a wide range of drugs and several of these have important ocular side-effects; the ophthalmologist therefore has a guardian's role in monitoring the eye and preventing complications.

Most rheumatological diseases affect either the uvea, sclera, retina, or optic nerve, and produce symptoms and signs that confirm both their nature and aetiology. Involvement of the uvea or sclera usually produces painful red eyes with preserved vision, whereas involvement of the retina or optic nerve causes profound visual loss without pain or redness. Fortunately, with the particular exception of juvenile chronic arthritis, most of the ocular complications that can affect the sight are symptomatic, so that the rheumatologist can at once alert the patient to the possibility of future eye disease and then await events.

The ocular examination

All patients referred to the ophthalmologist undergo the same routine examination, whether or not they are symptomatic. Visual function is assessed by testing visual acuity via a pinhole to obtain the effects of refractive error. The visual field is assessed using a red pin, as red is the most sensitive colour for elucidating disease of the optic nerve. Colour vision, changes in which may indicate early disease of the optic nerve or severe macular disease, is tested with Ishihara colour plates.

The eyes are examined in detail for evidence of redness or squint, and then with a slit lamp to observe any medial opacity (e.g. uveitis, cataract) and to measure the intraocular pressure. Pupillary reactions and eye movements are tested before the pupils are dilated for examination of the fundi.

Eye symptoms and signs

The more common eye symptoms encountered in patients with rheumatological diseases are dry, gritty eyes, photophobia, watering, redness, pain, floaters, and, most importantly, blurring or actual loss of vision. The differential diagnosis of the painful red eye and of sudden loss of vision is outlined in Table 1 and Table 2.

Ocular pathology	Vision	Pain	Distribution of redness	Extra signs	Likely systemic disease
Conjunctivitis	Good	Mild	Diffuse	Purulent sticky exudate	Reiter's syndrome
Dry eyes	Good	Gritty	Mild diffuse	Schirmer test Reduced tears Rose Bengal stain	Rheumatoid arthritis Systemic vasculitis Sjögren's syndrome (primary and secondary)
Episcleritis	Good	Irritation	Diffuse or nodular	Mobile nodules	Rheumatoid arthritis Systemic vasculitis
Scleritis	May be reduced	Very severe	Diffuse or nodular	Keratitis	Rheumatoid arthritis
Anterior uveitis	Reduced	Mild to severe	Circumcorneal	Keratic precipitates Small pupil Abnormal intraocular pressure Swollen optic disc	Seronegative spondylo-arthropathy Sarcoidosis Behçet's disease Vogt–Koyanagi disease

Table 1 The differential diagnosis of painful red eye

Cause of visual loss	History	Field	Pupil reaction	Media	Fundus	Likely systemic disease
Vitreous haemorrhage	Sudden	Generalised constriction	Normal	Hazy (blood)	Not seen or new vessels Retinal vascular occlusion	Systemic vasculitis Behçet's disease
Macular oedema	Gradual, with distortion	Small, relatively central scotoma	Normal	Hazy Cells in anterior chamber and vitreous	Swollen disc	Behçet's disease Sarcoidosis Seronegative arthropathy
Central retinal artery occlusion	Sudden	Dense, large, central scotoma	Afferent pupillary defect	Clear	Normal or pale optic disc Cherry-red spot at macula Dark blood in arterioles	Systemic vasculitis, especially giant-cell arteritis Polyarteritis nodosa
Ischaemic optic neuropathy	Sudden	Dense altitudinal defect	Afferent pupillary defect	Clear	Pale, swollen, optic disc	Giant-cell arteritis Systemic vasculitis
Inflammatory optic neuropathy	Progressive	Central scotoma	Afferent pupillary defect	Clear or hazy	Optic disc oedema Optic atrophy	Sarcoidosis Whipple's granulomatosis

All these are painless and occur in white eyes, except occasionally inflammatory optic neuropathy.

Table 2 The differential diagnosis of sudden loss of central vision

The combination of symptoms with ocular signs indicates the site of the eye disease, and this may reflect both the nature and severity of the joint disease.

Common eye disorders encountered in rheumatological disease

Conjunctivitis

Bacterial conjunctivitis commonly presents as sore, red, sticky eyes with a purulent discharge. Patients complain of being unable to open their eyes in the morning and of blurred vision. Viral conjunctivitis is very common. It presents with red, watering, irritable eyes and only minimal discharge, but often excessive photophobia as the cornea may become secondarily infected.

Dry eyes

Patients with dry eyes may complain of a foreign-body sensation, dryness, redness, grittiness, and excessive secretion of mucus. Their symptoms are aggravated by hot, dry, polluted atmospheres. Dry eyes may be due to local eye disease, particularly blepharitis and allergic eye disease, or keratoconjunctivitis sicca, which is associated with systemic disease. Keratoconjunctivitis sicca occurs when secretion of the lacrimal gland is of insufficient quality or quantity to maintain the tear film and ocular surface. It may be caused by reduction of tear secretion, destruction of the lacrimal glands or scarring of the ducts. The systemic diseases associated with dry eyes are primarily autoimmune diseases, for example primary Sjögren's syndrome, systemic lupus erythematosus, rheumatoid arthritis and systemic sclerosis, but also include infiltrative disorders including lymphoma, sarcoidosis or amyloidosis.

Slit-lamp examination is the essential investigation for patients with dry eyes. Offending local ocular disease can be diagnosed. The tear film is examined, particularly the break-up time of a fluorescent dye and the integrity of the cornea, which indicates adequacy of lubrication. In the absence of a slit lamp, Schirmer's test will identify patients with bone dry eyes but is otherwise not helpful. Rose Bengal stain denotes areas with an incomplete tear film. In early cases of keratoconjunctivitis sicca the stain is limited to the conjunctiva and in severe cases can be seen all over the cornea.

Treatment of dry eyes requires patience from the sufferer and the physician. Frequent regular tear supplements are the mainstay and these range from viscous preparations to normal saline. Temporary or permanent canalicular ablation is rarely required. Modern treatments which are under investigation are epidermal growth factor, aldose reductase inhibitors or cyclosporin eye drops (Tsubota 1994).

Episcleritis

The episclera is the thin layer of transparent tissue lying between the conjunctiva and the white sclera. It has its own blood supply and is incompletely fixed to the underlying sclera. This is in contrast to the conjunctiva, which is freely mobile over the episclera. Episcleritis is a benign, self-limiting condition that is usually unilateral and often nodular, involving only a small segment of the eye's surface. Patients complain of irritation and redness rather than pain, although the nodules may be tender to the touch. The affected episcleral vessels are dilated, within a swollen episclera. The condition subsides spontaneously within 4 to 6 weeks.

Scleritis

Scleritis is a serious, potentially sight-threatening condition, which in contrast to episcleritis is painful and does not subside spontaneously. Pain is the most constant feature and may be excruciating. The inflamed area is red and may produce a nodular, diffuse, or necrotizing pattern. If adjacent tissues are involved in the inflammatory process, additional signs and symptoms develop: spread to the cornea produces peripheral ulceration, keratitis and photophobia, whereas involvement of the posterior sclera causes proptosis, serous retinal detachment, a swollen optic disc, and, rarely, angle-closure glaucoma.

An unusual variant of scleritis is scleromalacia perforans, which occurs in elderly women with severe rheumatoid arthritis (see also below). This is characteristically painless and the affected sclera becomes dead white because the underlying lesion is arteritic infarction (Watson and Hazleman 1976). Eventually the sclera becomes atrophic and the blue, underlying choroid can be seen bulging through it. Actual perforation of the globe is very rare, despite the name. All forms of scleritis are very important to recognize for both the ophthalmologist and rheumatologist because they often herald activity of the systemic disease and may require systemic immunosuppression to prevent permanent ocular damage and blindness.

Acute anterior uveitis

Anterior uveitis is inflammation within the anterior chamber of the eye. The patient complains of an acutely painful, red eye with photophobia, sometimes associated with blurred vision. The pain is exacerbated by close work because the inflamed iris is stretched when the pupil constricts. The condition is generally unilateral and frequently recurrent. The ophthalmological signs, easily seen with the slit lamp, consist of circumcorneal redness and keratic precipitates on the endothelial surface of the cornea with flare and cells in the anterior chamber, looking like a cinema projection beam with specks of dust. In severe cases the cells gravitate to form a fluid level called a hypopyon. The pupil is small and poorly reactive; the intraocular pressure is often low in the acute stage and high in the convalescent stage.

Retinal vasculitis and posterior uveitis

The pattern of retinal disease in systemic inflammatory disease frequently reflects the nature of the disease and consequently may have important diagnostic value (Sanders 1987; Graham *et al*. 1989). In patients with seronegative arthropathies, Behçet's disease, and sarcoidosis, there may be inflammation in the posterior chamber of the eye (cells in the vitreous); the patient complains of floaters and of blurred or distorted central vision if the macula is oedematous. In mild cases, fundoscopy reveals focal or diffuse white sheathing of the retinal veins; in severe cases there are also scattered haemorrhages. Actual occlusion of the retinal veins and new vessels on the optic disc or peripheral retina may develop. The new vessels, similar to those in diabetes mellitus, generally develop as a consequence of retinal ischaemia secondary to occlusion of retinal veins. Their walls are fragile and vitreous haemorrhage is an important cause of visual morbidity in these patients.

By contrast, patients with systemic vasculitis do not develop inflammation within the eye but the vasculitic process affects the retinal capillaries and arterioles to produce retinal ischaemia, sometimes resulting in new vessels and vitreous haemorrhage. The ophthalmoscopic features range from asymptomatic, cotton-wool spots (due to occlusion of the retinal capillaries) to complete occlusion of the retinal artery causing blindness.

Optic neuropathy

Disease of the optic nerve invariably causes profound visual loss. Symptoms range from vague, blurred central vision and difficulty in differentiating colours to complete blindness. The cardinal ophthalmological signs of optic nerve disease are reduced visual acuity, impaired or absent colour vision, a central scotoma, and an afferent pupillary defect. The optic disc may look swollen, pale, or even normal. The pattern of visual loss varies from sudden, complete blindness, as experienced by

patients with giant-cell arteritis in whom the optic nerve undergoes infarction, to a slowly progressive loss in patients with granulomatous disorders when the optic nerve becomes compressed by abnormal tissue. All patients who develop disease of the optic nerve require urgent investigation and treatment.

Specific ocular manifestations of systemic inflammatory diseases

The ocular manifestations of systemic inflammatory disease vary according to the size and type of vessel predominantly affected by the disease process. As a general rule, diseases that affect arterioles involve the cornea, episclera, sclera, and retinal arterioles, whereas those affecting the venules produce uveitis (intraocular inflammation), macular oedema, and retinal venous disease.

Giant-cell arteritis

Giant-cell arteritis is an ophthalmic emergency and an important cause of preventable blindness in elderly people. Twenty-five per cent of patients develop ocular disease and the majority experience sudden loss of vision in one eye, with a dense, inferior, altitudinal field defect (Ross Russell 1959; Ross Russell 1986). The cause of the blindness is usually infarction of the optic nerve (ischaemic optic neuropathy) (Fig. 1), and occlusion of the central retinal artery in a minority of cases. Treatment is immediate, high-dose systemic steroids (e.g. prednisolone 1–2 mg/kg daily) to produce symptomatic relief of headache and, most importantly, to prevent blindness in the other eye, as recovery of vision is very unusual. If the patient has experienced symptoms of fluctuating vision in the good eye, treatment with intravenous steroids, heparin, and plasma expanders is indicated.

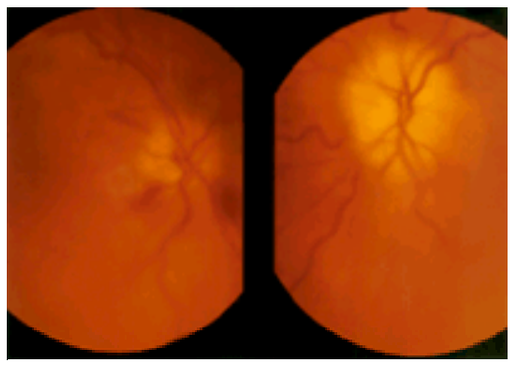


Fig. 1 Giant-cell arteritis. An 81-year-old lady with severe scalp tenderness woke up blind in her left eye, followed 3 days later by blindness in her right eye. The right fundus (left) has a swollen optic disc with peripapillary haemorrhages. The left optic disc (right) is swollen and pale, typical of ischaemic optic neuropathy.

Unusual neurophthalmological complications of giant-cell arteritis include oculomotor palsies, Horner's syndrome, cortical blindness (due to embolization from affected vertebral arteries), internuclear ophthalmoplegia, and visual hallucinations (Cullen and Coleiro 1976).

Systemic lupus erythematosus

Eye symptoms are very common in patients with systemic lupus erythematosus and range from minor, external problems (Frith *et al*. 1990) to severe retinopathy. Five per cent of patients will develop scleritis some time during the course of their disease, and this is occasionally seen at presentation (Fig. 2). As usual, active scleritis indicates active systemic disease and may require adjustment of systemic therapy. Uveitis is very rare in systemic lupus and only occurs in association with severe scleritis.

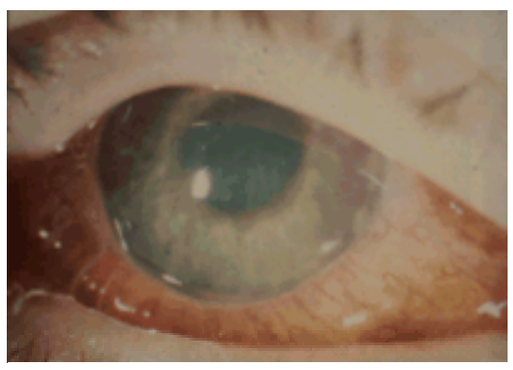


Fig. 2 Scleritis. A 32-year-old Caucasian woman presents with a 3-week history of general malaise and fever and a 1-week history of loss of vision associated with severe pain. The eye is red, the sclera is inflamed, the conjunctiva is chemosed. The cornea appears lazy because of corneal folds caused by hypotension in the eye secondary to severe ischaemia. Investigations confirmed acute, fulminant systemic lupus erythematosus.

The most important ophthalmic manifestation is the retinopathy described by Maumenee (1940). The hallmark of this is the cytoid body, which is the descriptive term used in systemic lupus for cotton-wool spot. Retinal haemorrhages and occlusions of central and branch retinal arteries may also occur; these are due to deposition of immune complexes rather than vasculitis (Graham *et al*. 1985). Occlusions of branch retinal veins are usually secondary to hypertension. Serous detachment of the retina and choroidal infarcts (Fig. 3) are seen in patients with active disease due to vasculitis of the choroidal vessels (Eckstein *et al*. 1993). The presence of anticardiolipin antibodies may be a marker for the development of retinal vascular occlusion (Levine *et al*. 1988) (Fig. 4).

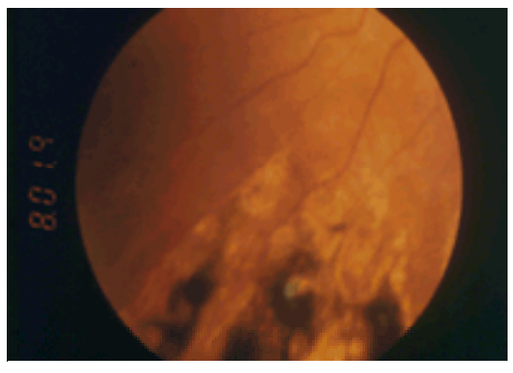


Fig. 3 Choroidal infarcts in systemic lupus erythematosus. A 22-year-old female presents with headaches and confusional state. Systemic lupus is diagnosed. 10 years later she has further thrombotic events when antiphospholipid antibodies were identified and these infarcts observed. They are asymptomatic.

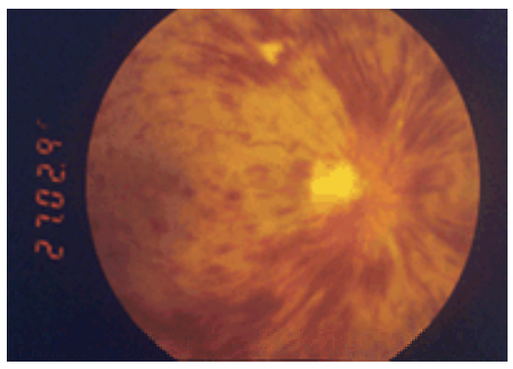


Fig. 4 Retinal vein occlusion in a woman who suffers with systemic lupus erythematosus and has antiphospholipid antibodies.

Polyarteritis nodosa

This disease has similar ophthalmic manifestations to those of systemic lupus. However, necrotizing scleritis with peripheral corneal ulceration is relatively common. The retinopathy is usually secondary to accelerated hypertension and therefore characterized by arteriovenous nipping, cotton-wool spots, and flame-shaped haemorrhages. The appearance of isolated cotton-wool spots or branch-artery occlusion is unusual; sudden visual loss, due to either ischaemic optic neuropathy or occlusion of the central retinal artery, rarely occurs. The development of scleritis or retinal vascular disease in the absence of hypertension warrants investigation of the systemic disease.

Wegener's granulomatosis

Eye problems develop in about half of patients with Wegener's granulomatosis, particularly those with limited disease (Spalton *et al*. 1981). This is the only systemic inflammatory disease that commonly presents with orbital disease, which is due to infiltration of the orbit with granulomatous tissue (Fig. 5). The patient complains of proptosis, which is caused by the orbital mass, red eyes from the scleritis, and loss of vision from compression of the optic nerve by the granulomatous tissue, often accompanied by oedema of the optic disc and choroidal folds.

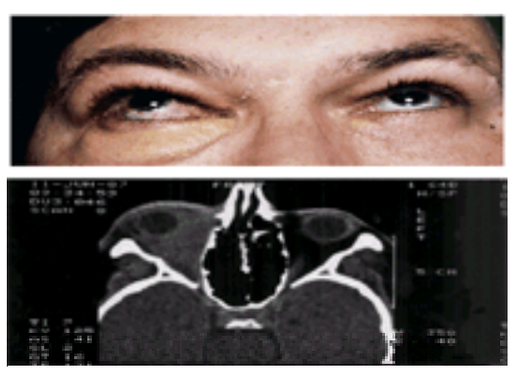


Fig. 5 Wegener's granulomatosis. A 32-year-old man presented with fever, malaise, sore red right eye, and blurred vision. The right eye was proptosed with dilated scleral vessels and an incidental finding of extensive xanthelasmas (top). A CT scan revealed a 'dirty' right orbit where the muscles and optic nerve could not be identified due to extensive granulomatous tissue.

Systemic sclerosis

The most common eye complaint is dryness. A minority develop retinopathy similar to that seen in the other systemic vasculitides, with cotton-wool spots, haemorrhages, and occlusion of retinal arteries and also with choroidal infarction (Grennan and Forrester 1977).

Rheumatoid arthritis

One-third of patients suffer with dry eyes and in 10 per cent this is combined with a dry mouth (Sjögren's syndrome). Five per cent of patients presenting with episcleritis and 30 per cent presenting with isolated scleritis have rheumatoid arthritis (Watson and Hazleman 1976). Patients with scleritis associated with rheumatoid arthritis more commonly have bilateral scleritis than patients with scleritis due to other systemic immune disease (Sainz de la Meza *et al*. 1994). The scleritis is associated with an occlusive vasculitis; this may be very severe and cause corneal melt and occasionally perforation requiring surgical repair. Scleromalacia perforans affects elderly women predominantly (Fig. 6). Advancing scleral disease in these patients almost always heralds a flare-up of the systemic disease, particularly the vasculitic complications. Uveitis is not a feature of rheumatoid arthritis and involvement of the retinal vessels is very unusual.

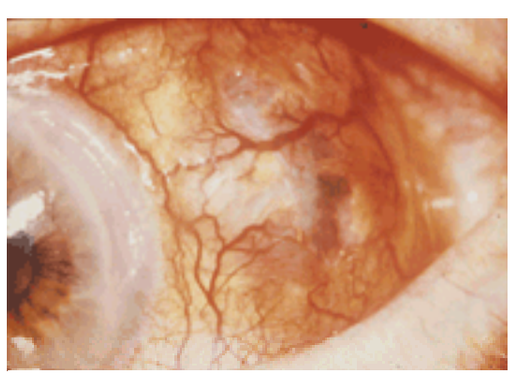


Fig. 6 Scleromalacia perforans. A 75-year-old man has suffered with severe rheumatoid arthritis and scleritis for many years. He now has an area of very thin sclera due to ischaemia and ulceration of the adjacent peripheral cornea.

Ankylosing spondylitis and the seronegative arthropathies

Acute anterior uveitis is the most important ocular disease in this group of patients. It occurs in 25 per cent of patients with ankylosing spondylitis and in 10 to 15 per cent of patients with other causes of seronegative arthropathy. The clinical symptoms and signs are identical, regardless of the accompanying disease. Attacks are usually unilateral but both eyes become affected during the course of the disease (Rothova *et al*. 1987). In severe cases, slit-lamp examination reveals a sticky, fibrinoid aqueous in the anterior chamber; posterior synechiae (adhesions between the iris and the lens) are common and raised intraocular pressure occurs. Patients who are HLA-B27 positive tend to be younger at the time of the first attack, have more severe ocular inflammation, and a longer duration of the acute attack than those who are HLA-B27 negative, in whom uveitis may be milder but sometimes more chronic (Linssen 1990). Conjunctivitis occurs in one-third of patients with

Reiter's disease but there is no increased incidence in the other groups (Fig. 7).

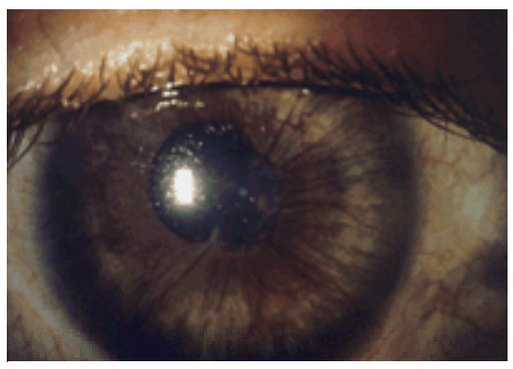


Fig. 7 Uveitis associated with ankylosing spondylitis. A 24-year-old man is blind from chronic anterior uveitis. The eye is white but the pupil is irregular due to posterior synechiae and abnormal new vessels are present on the iris.

Interestingly, the posterior segment of the eye is rarely affected in seronegative arthropathies: only 10 per cent of patients develop cells in the vitreous from retinal vasculitis affecting predominantly the capillaries and the postcapillary venules, but the consequence of this may be cystoid macular oedema and oedema of the optic disc. In very severe cases, serous retinal detachment ensues (Rodriguez *et al*. 1994).

Sarcoidosis

Eye involvement occurs in 30 to 40 per cent of patients with systemic sarcoidosis (Jabs and Johns 1986; Karma *et al*. 1988). The most common ocular manifestation is acute or chronic, relapsing uveitis, often in the absence of active pulmonary disease. The anterior uveitis is characterized by granulomatous, 'mutton fat', keratic precipitates and nodules on the iris (Fig. 8). Patients may complain of dry eyes from involvement of the lacrimal glands, which can be very dramatic, particularly in patients of African ethnicity. Asymptomatic conjunctival and scleral nodules may develop and are of diagnostic use.

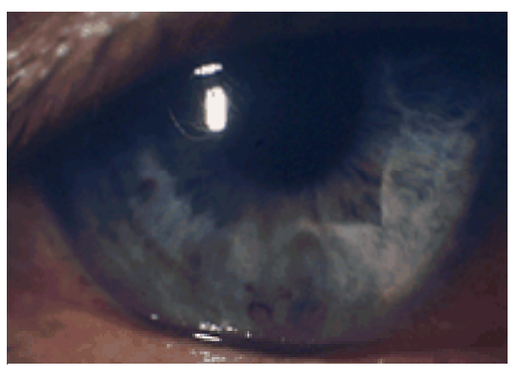


Fig. 8 Granulomatous anterior uveitis. A 58-year-old man presents with episodes of breathlessness and red eyes, due to sarcoidosis. The conjunctiva is injected; large, grey, keratic precipitates are seen inferiorly on the corneal endothelium.

Twenty-five per cent of patients with ocular sarcoidosis develop posterior uveitis with cells in the vitreous, and retinal vasculitis. The diagnostic retinal features of sarcoidosis are periphlebitis, focal cuffing of the retinal veins, new vessels on the optic disc or in the peripheral retina (often in the absence of retinal ischaemia), choroidal nodules (Fig. 9), and pigment epithelial lesions (Spalton and Sanders 1981).

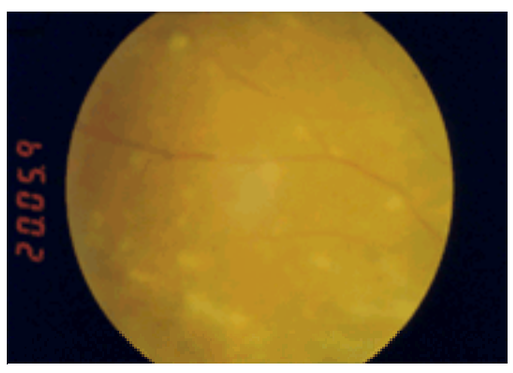


Fig. 9 Sarcoidosis. A 40-year-old Asian man has suffered with uveitis and sarcoidosis for 15 years. He has sheathing of his retinal veins and multiple pale areas in the pigment epithelium and choroid, which probable represent granulomas.

Five per cent of patients develop granulomatous optic neuropathy with profound visual loss (Graham *et al*. 1986). This may mimic acute demyelinating optic neuritis, with pain and rapid visual loss, or it may produce progressive visual loss owing to a granuloma compressing the optic nerve (Fig. 10). Both types may be acutely steroid-sensitive and therefore it is important to make the diagnosis rapidly.

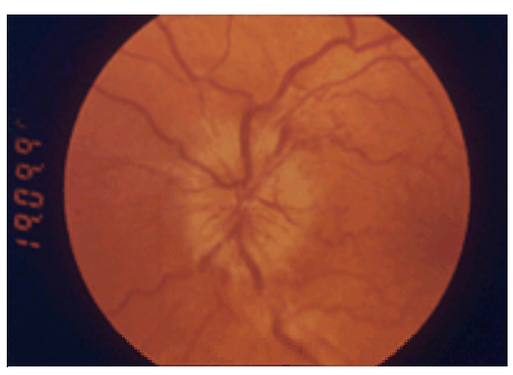


Fig. 10 Optic nerve infiltration by sarcoidosis. A 24-year-old Caucasian man presents with sudden loss of vision associated with pain in the left eye. Investigations reveal a lymphopenia, a high angiotensin-converting enzyme, raised protein, and 11.0 million lymphocytes in the cerebrospinal fluid.

Behçet's disease

Ocular manifestations in Behçet's disease occur in 70 per cent of patients and are very important as 25 per cent of patients will go blind (Michelson and Chisari 1982). Together with orogenital ulceration and thrombophlebitis, uveitis forms the major triad described in 1937 by Hulusi Behçet.

Severe anterior uveitis, often in a white eye, is common, with hypopyon formation in extreme cases. The anterior uveitis is typically recurrent but can resolve spontaneously. The high incidence of blindness in patients with Behçet's disease is due to the severity of the retinal changes. The pathognomonic features are sequential occlusions of branch retinal veins (Fig. 11) and retinal infiltrates. The recurrent nature of the venous occlusions results in progressive retinal ischaemia and eventually the entire retinal vasculature becomes obliterated and the optic disc atrophic (Atmaca 1989). Retinal infiltrates, which are due to collections of polymorphs in the superficial retinal layers, are pathologically similar to a hypopyon and will resolve spontaneously over the course of a week.

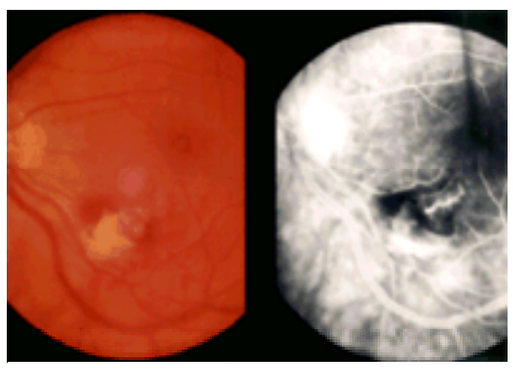


Fig. 11 Behçet's disease. A 24-year-old woman with a history of severe orogenital ulceration and thrombophlebitis suddenly developed blurred central vision in her left eye. An area of haemorrhage and infiltrate was present inferior to the macula (left) and a fluorescein angiogram (right) revealed occlusion of a small retinal vein.

Ocular involvement in rheumatic conditions of childhood

Introduction

Uveitis occurring before the age of 16 years is approximately four times less common than uveitis in the general population (Perkins 1966). In two reviews of uveitis in childhood, systemic illnesses such as sarcoidosis, Behçet's disease, and vasculitis accounted for less than 8 per cent of patients (Perkins 1966; Kanski and Shun-Shin 1984). The most common associated disorder in Kanski's series was seronegative juvenile chronic arthritis. Ocular complications of rheumatic diseases in childhood were reviewed by Petty (1990).

Juvenile chronic arthritis

The chronic anterior uveitis that occurs in association with juvenile chronic arthritis deserves special mention, particularly as it does not occur in adults. Affected children are mostly girls, and are usually young at onset of arthritis (mean age, 3 years). Joint involvement is usually pauciarticular, although it may spread to become polyarticular. Most patients have associated antinuclear antibodies in their serum.

Arthritis generally predates the eye involvement by 1 to 3 years or sometimes much longer, but 5 to 10 per cent of cases present with uveitis and the ophthalmologist should therefore ask for a rheumatological opinion in any young child with chronic anterior uveitis. Similarly, because the eye disease is usually asymptomatic until complications occur, the rheumatologist must arrange for an early slit-lamp examination on all children. The ophthalmologist will then arrange suitable follow-up screening at regular intervals (Leak 1992).

The non-granulomatous iridocyclitis runs an uncomplicated course in only one-third of affected children (Petty 1987). Although the majority of cases develop the complications of band keratopathy, posterior synechiae, cataracts, or glaucoma, only 10 to 20 per cent of affected eyes become blind (Rosenberg 1987; Wolf *et al*. 1987). Early onset of uveitis, either symptomatic or when picked up in routine screening within the first year of arthritis, and particularly if associated with posterior synechiae at this first examination, is associated with a significantly worse prognosis for vision (Wolf *et al*. 1987)). Most eye involvement develops for the first time within 5 years of the onset of arthritis, and after this the risk is much less, unlike the acute symptomatic anterior uveitis in ankylosing spondylitis, which may occur for the first time very many years later than the original musculoskeletal symptoms.

Spondylarthropathies

As many as 8.6 per cent of patients with ankylosing spondylitis have their onset before the age of 16 years (Bennet and Wood 1968), frequently in an adolescent boy with pauciarticular arthritis of the legs, or enthesitis. Acute anterior uveitis probably occurs as frequently in juvenile-onset spondylitis as in adult ankylosing spondylitis, and over long follow-up up to 27 per cent of such children develop episodes of uveitis.

Bilateral, painless conjunctivitis is common in Reiter's syndrome but, in addition, acute anterior uveitis and keratitis with corneal ulcers have been described in children. Chronic anterior uveitis is only described in occasional children with juvenile spondylitis or Reiter's disease and always requires reconsideration of the diagnosis.

Although juvenile psoriatic arthritis is often grouped with the spondarthropathies, acute uveitis is uncommon and ocular involvement is usually chronic and clinically indistinguishable from that which occurs in association with early-onset juvenile chronic arthritis.

Multisystem systemic illness

Many systemic rheumatic diseases with ocular involvement in adults have counterparts in childhood, including systemic lupus erythematosus, Sjögren's syndrome, and systemic sclerosis (Leak and Isenberg 1989). Behçet's disease is rarely diagnosed before adolescence, although a history of oral ulcers may predate the appearance of other major manifestations by more than 8 years (Kim *et al* 1994).

Sarcoidosis beginning under the age of six years with rash and arthropathy may be associated with chronic anterior uveitis (Hoover *et al* 1986, Sahn *et al* 1990). Features distinguishing this from juvenile chronic arthritis include the iris nodules, posterior eye involvement, and rather painless, non-evasive, boggy tenosynovitis. Blau syndrome is a familial form of granulomatous arthritis, uveitis, and rash with similar features to sarcoidosis in childhood (Blau 1985).

In childhood, classical polyarteritis nodosa and Kawasaki disease are the most common forms of systemic vasculitis. Bilateral, non-suppurative, bulbar, conjunctival injection is a recognized diagnostic criterion in Kawasaki disease, occurring in virtually all cases in the early stages of the illness; acute, bilateral uveitis is also very common (Ohno *et al*. 1982; Smith *et al*. 1989).

Differential diagnosis of coexistent ocular and arthritic problems in childhood includes infections such as Lyme disease, infantile-onset, multisystem inflammatory disease (Yarom *et al*. 1985), and lysosomal storage diseases (Cassidy and Petty 1990).

Infective arthropathies and eye disease

Secondary syphilis may be accompanied by an acute panuveitis with an active anterior uveitis, cells in the vitreous, and retinal changes involving predominantly the capillaries, which produce oedema of the macula and optic disc, and also subretinal fluid in the inferotemporal quadrant, leading to a characteristic pale appearance.

Other infectious diseases that must be considered in undiagnosed patients presenting with arthritis and ocular inflammation include those involving Epstein–Barr virus, Lyme disease, Whipple's disease, brucellosis, and human immunodeficiency virus (Nussenblatt and Palestine 1989).

Temporal relationship of ocular involvement with other disease manifestations in specific conditions

The chronic anterior uveitis associated with juvenile chronic arthritis usually runs an intermittent course that is not associated with the current activity of the arthritis, although it is common for their onset to be within 1 to 2 years of each other. The severity of ocular and joint involvement is also independent such that blindness may occur in a child with minimal arthritis and no residual deformity. By contrast, in adult rheumatoid arthritis, scleritis indicates generalized, active disease.

The acute attacks of unilateral anterior uveitis that characterize the seronegative spondarthropathies occur randomly, whereas ocular involvement in Reiter's syndrome occurs as part of a symptom complex, usually within the first 2 to 3 years (Lee *et al*. 1986).

In Behçet's disease and sarcoidosis, uveitis is frequently found with other systemic manifestations at presentation, but subsequent flare-ups frequently occur in isolation such that eye and joint disease are not significantly associated.

In Kawasaki disease, eye involvement occurs as part of the acute inflammatory process, and Ohno *et al*. (1982) found a strong statistical correlation between the activity of uveitis, as measured by the numbers of inflammatory cells in the anterior chamber, and the erythrocyte sedimentation rate (**ESR**) and C-reactive protein.

Diagnostic and laboratory tests

The initial history and examination will provide the diagnosis in the majority of patients. The use of radiological or laboratory tests in the assessment of a patient with eye disease is limited. A useful screen for a patient with uveitis would be a full blood count, ESR, chest radiographs, and serology for syphilis. Further tests such as HLA-B27, angiotensin-converting enzyme, and antinuclear antibodies would be indicated by clinical evidence of disease elsewhere. In a prospective study of 865 patients, intensive investigation established a specific diagnosis in 628 patients (73 per cent) but an associated systemic disease was found in only 220 patients (26 per cent) (Rothova *et al*. 1992).

Evidence of an acute-phase response, for example a raised ESR or C-reactive protein, suggests a generalized systemic illness with active ocular involvement. However, in diseases in which the activity of eye and joint disease do not run in parallel, active uveitis alone is usually not associated with a raised ESR.

Autoantibodies

Rheumatic diseases such as rheumatoid arthritis and systemic lupus erythematosus are often associated with serum autoantibodies but in this situation the antibodies are not of diagnostic use as the systemic disease would usually be obvious clinically.

Amongst a mixed group of patients with uveitis, antibodies to tissue antigens do not occur more frequently than in controls, except in children with juvenile chronic arthritis in which 80 per cent or more have antinuclear antibodies (Murray 1986). Patients with retinal vasculitis may have circulating immune complexes and/or antiretinal antibodies. The presence of antibodies alone may be associated with more severe disease (Kasp *et al*. 1986). Antineutrophil cytoplasmic antibodies have been found in 'fit' patients with severe necrotizing scleritis and may indicate an underlying systemic vasculitis that requires persistent vigilance by the ophthalmologist (Talks *et al*. 1994).

Immunogenetics

Determination of HLA-B27 in patients with acute anterior uveitis is valuable because B27-positive patients are more likely to have an associated spondarthropathy and should be referred for a rheumatological opinion (Feltkamp 1990).

Treatment of the ocular complications of systemic inflammatory disease

Mild, anterior-segment complications such as conjunctivitis, dry eyes, and episcleritis will invariably respond to local treatment.

Treatment of scleritis, anterior uveitis, retinal vasculitis, and optic neuropathy depends firstly on the severity of the eye disease and its effect on vision, and secondly on the activity of the systemic disease. In general, systemic therapy is dictated by the eye disease if the vision is reduced or potentially threatened by inflammation.

Scleritis

Scleritis is an indication for systemic therapy. Adequate doses of non-steroidal, anti-inflammatory drugs may suffice (e.g. flurbiprofen, 50 mg, three times a day), but if corneal or retinal involvement develops or the inflammation is uncontrolled, systemic steroids are indicated. In rare instances it is necessary to supplement the steroids by pulse cyclophosphamide (Meyer *et al*. 1987) or cyclosporin (Wakefield and McCluskey 1989). If the scleritis coincides with activity of the systemic disease, adjusting the treatment for the latter is usually sufficient.

Uveitis

The treatment of uveitis depends primarily on the visual acuity and the site of the inflammatory process. The rationale of treatment of uveitis is set out in Table 3. In the majority of cases, anterior uveitis can be treated topically with steroid eye drops (prednisolone betamethasone/dexamethasone) and a dilating agent (tropicamide/cyclopentolate/atropine), occasionally supplemented by a subconjunctival injection. A trial of systemic steroids may be necessary in selected patients with frequent persistent relapses of anterior uveitis. If the inflammation involves the posterior segment and the visual acuity drops below 6/12 (0.5), due either to cystoid macular oedema or vitreous cells, a subtenons injection of steroid or systemic steroids are indicated. If the vision drops because of occlusion of the retinal vessels or involvement of the optic nerve, systemic immunosuppression is immediately indicated. Corticosteroids and cyclosporin are effective (Towler *et al*. 1990; Howe *et al*. 1994), but these may be used in conjunction with azathioprine (Lightman 1991) or methotrexate (Shah *et al*. 1992).

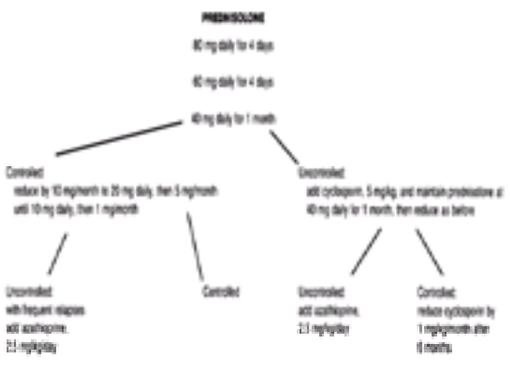


Table 3 Changing course of systemic steroids for patients with posterior uveitis

The most difficult management problem is Behçet's disease because of the severity and frequency of the attacks of uveitis, which almost invariably affect the retinal vessels. A combination of prednisolone, cyclosporin, azathioprine, and colchicine appears to be most effective in refractory patients. However, there are no controlled trials of these drugs in Behçet's uveitis and therefore the perplexing problem of benefits of treatment versus side-effects must be continuously evaluated. The management of Behçet's disease is discussed in Chapter 5.11.7.

Antirheumatic therapy with ocular side-effects

Most antirheumatic drug therapy causes no significant ocular side-effects but attention is drawn to toxicity of antimalarial drugs prescribed for rheumatoid arthritis, systemic lupus erythematosus, and related disorders, and to corticosteroids, widely used in many rheumatic conditions.

Antimalarial drugs

Use of these compounds in patients with rheumatic diseases has been hindered by fears about ophthalmological toxicity and the need for frequent monitoring by an ophthalmologist. Corneal deposits and poor contraction of the ciliary body are minor, reversible problems and the real worry is the development of retinopathy.

Chloroquine has a high affinity for the melanin-containing tissue of the eye, including the retinal pigment epithelium (Maksymowych and Russell 1987). The pathophysiology of the retinal damage is related to malfunction of the phagolysosomes, resulting in faulty clearance of ageing photoreceptor membranes. The build-up of lamellar myelin bodies disturbs the metabolism of the retinal pigment epithelium.

In the advanced stage with an absolute scotoma and loss of visual acuity, there is granular pigmentation of the macula surrounded by a 'bull's eye' of rings of depigmentation and pigmentation. However, in the early stages, fundoscopy and other standard optic screening tests are normal despite the presence of a scotoma on a red Amsler grid (Fig. 12). Identification at this stage is the goal of screening programmes, as cessation of treatment will usually prevent progression.

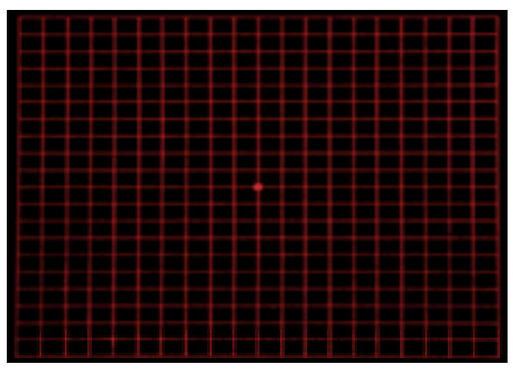


Fig. 12 The red Amsler grid.

Evaluation of published reports on retinal toxicity is difficult because of the differences in dosage over variable periods of time and the exact criteria used to define retinopathy, as well as lack of baseline data in some patients.

The total dose of drug used before development of retinopathy is very variable whereas the daily dosage in relation to the patient's weight now seems to be the important factor. At daily doses of 250 mg for chloroquine or 400 mg for hydroxychloroquine, there are few cases of mild, reversible retinopathy. Reports of progressive retinopathy at 250 mg/day of chloroquine contain insufficient information on patients' weight.

Several reviewers are agreed that the risk of toxicity is very small at the recommended daily dosages: chloroquine 3.5 mg/kg; hydroxychloroquine 6.5 mg/kg (Maksymowych and Russell 1987; Rynes 1987; Easterbrook 1988; Lozier and Friedlaender 1989). In practical terms, for a patient of 65 kg, 400 mg of hydroxychloroquine are safe, or 225 mg of chloroquine (not 250 mg). However, smaller patients must take a reduced dose, as should those with renal or hepatic impairment. Current opinion favours hydroxychloroquine as having the best safety record (Berstein 1991; Easterbrook 1993; Spalton *et al*. 1993).

An adequate, baseline ophthalmological examination is the ideal and should include visual acuity (corrected for distance and near), colour vision, red Amsler chart, and fundoscopy. Detection of coexistent eye disease such as age-related macular pigmentation will necessitate regular follow-up at 6-monthly intervals. In the absence of ocular abnormalities, opinion is divided as to the necessity and frequency of repeat tests. Some suggest no other ophthalmological review is necessary (Morsman *et al*. 1990), although the prescribing physician may wish to recheck the Amsler grid test every 6 months (Benstein 1991). By contrast, Easterbrook (1993) still recommends annual ophthalmological assessment, whereas Spalton *et al*. (1993) feel this is only necessary after 5 years on antimalarial therapy. Self-testing by the patient with an Amsler grid at 2-weekly intervals is an inexpensive and reproducible test that is both simple and sensitive (Easterbrook 1988; Lozier and Friedlaender 1989). Detection of faded or distorted areas should prompt early evaluation with field testing by the ophthalmologist.

Corticosteroids

Both oral and topical corticosteroids may cause a variety of ophthalmic side-effects:

1. Corneal and scleral thinning can occur with topical steroids and lead to perforation.
2. Cataracts are common, especially after higher doses of systemic steroids. Their development in chronic uveitis is partly related to the treatment but the advantages of topical therapy in particular far outweigh the side-effect.
3. Raised intraocular pressure is another common side-effect of steroids. Five per cent of the population are termed 'steroid reactors' because of a significant but asymptomatic increase in pressure. The frequency, duration, and type of topical or oral steroid all influence this effect, and use of an alternative steroid should be considered, for example from dexamethasone to fluoromethalone.
4. Infectious problems: herpes simplex can cause an acute, unilateral anterior uveitis and topical steroids risk exacerbating such eye infections; topical steroids should not be used indiscriminately on an undiagnosed red eye.

Immunosuppressives

Patients on long-term systemic immunosuppression may develop viral retinitis (herpes simplex, herpes zoster, cytomegalovirus) causing retinal necrosis and eventual blindness (Fig. 13). Early symptoms include red eyes, floaters, and field defect, and require an urgent ophthalmological opinion.

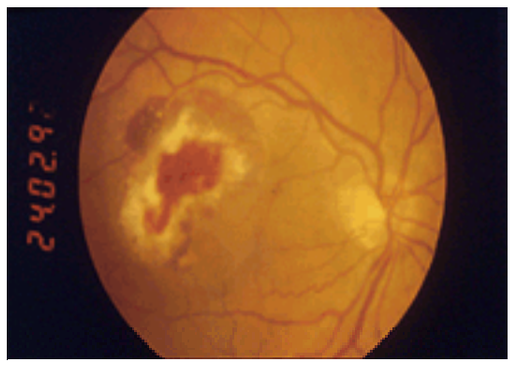


Fig. 13 Cytomegalovirus retinitis. A middle-aged man has Wegener's granulomatosis controlled with prednisolone and azathioprine. After 12 months' treatment he develops blurred vision. The area of retinal necrosis with haemorrhage and satellite lesions is typical of cytomegalovirus retinitis. The retinitis resolved with ganciclovir and reduction of immunosuppressive therapy.

Box 1 Important key points for ophthalmologists to remember

1. Examine all children suspected of juvenile rheumatoid arthritis as soon as possible. It takes seconds!
2. Visual loss due to ischaemic optic neuropathy with cotton wool spots and/or minor retinal arteriolar occlusion is due to systemic vasculitis.
3. Uveitis associated with choroidal or pigment epithelial disease suggests sarcoidosis..
4. Uveitis and retinal infiltrates or retinal vein occulsions suggests Behçet's disease.
5. Scleritis associated with peripheral corneal ulceration indicates systemic vasculitis, particularly Wegener's granulomatosis or polyarteritis nodosa.

Chapter References

Atmaca, L. S. (1989). Fundus changes in Behçet's disease. *Graefe's Archives of Clinical and Experimental Ophthalmology*, **227**, 340–4.

Bennett, P. A. and Wood, P. H. N. (ed.) (1968). *Population studies of the rheumatic diseases*, Excerpta Medica International Congress series, No. 148. Excerpta Medica, Amsterdam.

Bernstein, H. N. (1991). Ocular safety of hydroxychloroquine. *Annals of Ophthalmology*, **23**, 292–6.

Blau, E. B. (1985). Familial granulomatous arthritis, iritis and rash. *Journal of Paediatrics*, **5**, 689–93.

Cassidy, J. T. and Petty, R. E. (ed.) (1990). Storage diseases. In *Textbook of pediatric rheumatology*, (2nd edn), pp. 555–7. Churchill Livingstone, New York.

Cullen, J. E. and Coleiro, J. A. (1976). Ophthalmic complications of giant cell arteritis. *Surveys of Ophthalmology*, **20**, 247–60.

Easterbrook, M. E. (1988). Ocular effects and safety of antimalarial agents. *American Journal of Medicine*, **85** (Suppl. 4A), 23–9.

Easterbrook, M. E. (1993). The ocular safety of hydroxychloroquine. *Seminars in Arthritis and Rheumatism*, **23** (Suppl. 1), 62–7.

Eckstein, M. B., Spalton, D. J., and Holder G. (1993). Visual loss from central serous retinopathy. *British Journal of Ophthalmology*, **77**, 263–4.

Feltkamp, T. E. W. (1990). Ophthalmological significance of HLA associated uveitis. *Eye*, **4**, 839–44.

Frith, P., Burge, S. M., Millard, P. R., and Wojnarowski, F. (1990). External ocular findings in lupus erythematosus; a clinical and immunopathological study. *British Journal of Ophthalmology*, **74**, 163–7.

Graham, E. M., Spalton, D. J., Barnard, R. O., Garner, A., and Ross Russell, R. W. (1985). Cerebral and retinal changes in systemic lupus erythematosus. *Ophthalmology*, **92**, 444–8.

Graham, E. M., Ellis, C. J. K., Sanders, M. D., and Macdonald, W. I. (1986). Optic neuropathy in sarcoidosis. *Journal of Neurology, Neurosurgery, and Psychiatry*, **49**, 756–63.

Graham, E. M., Stanford, M. R., Sanders, M. D., Kasp, E., and Dumonde, D. C. (1989). A point prevalence study of 150 patients with idiopathic retinal vasculitis. 1. Diagnostic value of ophthalmological features. *British Journal of Ophthalmology*, **73**, 714–21.

Grennan, D. M. and Forrester, J. (1977). Involvement of the eye in SLE and scleroderma. *Annals of the Rheumatic Diseases*, **36**, 152–6.

Hoover, D. L., Khan, J. A., and Giangiacoma, J. (1986). Pediatric ocular sarcoidosis. *Surveys of Ophthalmology*, **30**, 215–28.

Howe, L. J., Stanford M. R., Edelsten C., and Graham, E. M. (1994). The efficacy of systemic corticosteroids in sight threatening retinal vasculitis. *Eye*, **8**, 443–7.

Jabs, D. A. and Johns, C. J. (1986). Ocular involvement in chronic sarcoidosis. *American Journal of Ophthalmology*, **102**, 297–301.

Kanski, J. J. and Shun-Shin, A. (1984). Systemic uveitis syndromes in childhood: an analysis of 340 cases. *Ophthalmology*, **91**, 1247–52.

Karma, A., Huhti, E., and Poukkla, A. (1988). Course and outcome of ocular sarcoidosis. *American Journal of Ophthalmology*, **106**, 467–72.

Kasp, E., Graham, E. M., Stanford, M. R., Sanders, M. D., and Dumonde D. C. (1986). A point prevalence study of 150 patients with idiopathic retinal vasculitis. 2. Clinical relevance of antiretinal autoimmunity and circulating immune complexes. *British Journal of Ophthalmology*, **73**, 722–30.

Kim, D-K., Chang, S. N. Bang D., Lee E-S., and Lee S. (1994). Clinical analysis of 40 cases of childhood-onset Behçets disease. *Paediatric Dermatology*, **11**, 95–101.

Leak, A. M. (1992). Ophthalmological screening in seronegative juvenile chronic arthritis: a personal view. *British Journal of Rheumatology*, **31**, 631–2.

Leak, A. M. and Isenberg, D. A. (1989). Autoimmune rheumatic disorders in childhood—a comparison with adult onset disease. *Quarterly Journal of Medicine*, **73**, 875–93.

Lee, D. A., Barker, S. M., Su, W. P. D., Allen, G. L., Liesegant, T. J., and Illustrup, D. M. (1986). The clinical diagnosis of Reiter's syndrome, ophthalmic and non-ophthalmic aspects. *Journal of Ophthalmology*, **93**, 350–6.

Levine, S. R., Crofts, J. W., Lesser, G. R., Floberg, J., and Welch, K. M. A. (1988). Visual symptoms associated with the presence of lupus anticoagulant. *Ophthalmology*, **95**, 686–92.

Lightman, S. (1991). Use of steroids and immunosuppressive drugs in the management of posterior uveitis. *Eye*, **5**, 294–8.

Linssen, A. (1990). B27+ disease versus B27– disease. *Scandinavian Journal of Rheumatology*, Suppl. 87, 111–19.

Lozier, J. R. and Friedlaender, M. H. (1989). Complications of antimalarial therapy. *International Ophthalmology Clinics*, **29**, 172–8.

Maksymowych, W. and Russell, A. S. (1987). Antimalarials in rheumatology: efficacy and safety. *Seminars in Arthritis and Rheumatism*, **16**, 206–21.

Maumenee, A. E. (1940). Retinal lesions in lupus erythematosus. *American Journal of Ophthalmology*, **23**, 971–81.

Meyer, A. R., Watson, P. G., and Franks, W. (1987). Pulsed immunosuppressive therapy in the treatment of immunologically advanced corneal and scleral disease. *Eye*, **1**, 487–95.

Michelson, J. B. and Chisari, F. V. (1982). Behçet's disease. *Surveys in Ophthalmology*, **26**, 190–203.

Morsman, C. D. G., Livesey, S. J., Richards, I. M., Jessop, J. D., and Mills, P. V. (1990). Screening for hydroxychloroquine retinal toxicity: is it necessary? *Eye*, **4**, 572–6.

Murray P. (1986). Serum autoantibodies and uveitis. *British Journal of Ophthalmology*, **70**, 266–8.

Nussenblatt, R. B. and Palestine, A. G. (1989). *Uveitis: fundamentals and clinical practice*. Year Book Medical Publishers, Chicago.

Ohno, S. *et al*. (1982). Ocular manifestations of Kawasaki's disease (mucocutaneous lymph node syndrome). *American Journal of Ophthalmology*, **93**, 713–17.

Perkins, E. S. (1966). Pattern of uveitis in childhood. *British Journal of Ophthalmology*, **50**, 169–85.

Petty, R. E. (1987). Current knowledge of the etiology and pathogenesis of chronic uveitis accompanying juvenile rheumatoid arthritis. *Rheumatic Disease Clinics of North America*, **13**, 19–36.

Petty, R. E. (1990). Ocular complications of rheumatic diseases of childhood. *Clinical Orthopaedics and Related Research*, **259**, 51–9.

Rodriguez, A., Akova, Y. A., Pedroza-Seres, M., and Foster, C. S. (1994). Posterior segment ocular manifestations in patients with HLA-B27-associated uveitis. *Ophthalmology*, 101, 1267–74.

Rosenberg, A. M. (1987). Uveitis associated with juvenile rheumatoid arthritis. *Seminars in Arthritis and Rheumatism*, **16**, 158–73.

Ross Russell, R. V. (1959). Giant cell arteritis: a review of 35 cases. *Quarterly Journal of Medicine*, **112**, 471–89.

Ross Russell, R. W. (1986). Giant cell (cranial) arteritis. In *Handbook of clinical neurology*, Vol. 4, (ed. C. Rose), pp. 309–28. Elsevier, Amsterdam.

Rothova, A., van Veenendaal, W. G., Linssen, A., Glasius, E., Kjilstra, A., and de Jong, P. T. V. H. (1987). Clinical features of acute anterior uveitis. *American Journal of Ophthalmology*, **103**, 137–45.

Rothova, A. *et al*. (1992). Uveitis and systemic disease. *British Journal of Ophthalmology*, **76**, 137–41.

Rynes, R. I. (1987). Side effects of antimalarial therapy. *British Journal of Clinical Practice*, Suppl. 52, 42–5.

Sahn, E. E., Hampton, M. T., Garen, P. D., Warnick J., Smith, D., and Silver R. M. (1990). Preschool sarcoidosis masquerading as juvenile rheumatoid arthritis: two case reports and a review of the literature. *Paediatric Dermatology*, **7**, 208–13.

Sainz de la Maza, M., Foster, C-S., and Jabbur, N. S. (1994). Scleritis associated with rheumatoid arthritis and with other systemic immune diseases. *Ophthalmology*, **101**, 1281–8.

Sanders, M. D. (1987). Retinal arteritis, retinal vasculitis and autoimmune retinal vasculitis. *Eye*, **1**, 441–65.

Shah, S. S. *et al*. (1992). Low dose methotrexate therapy for ocular inflammatory disease. *Ophthalmology*, **99**, 1419–23.

Smith, L. B. H., Newburger, J. W., and Burns J. C. (1989). Kawasaki syndrome and the eye. *Pediatric Infectious Disease Journal*, **8**, 116–18.

Spalton, D. J. and Sanders M. D. (1981). Fundus changes in histologically confirmed sarcoidosis. *British Journal of Ophthalmology*, **65**, 348–58.

Spalton, D. J., Graham, E. M., Page, N. G. R., and Sanders, M. D. (1981). Ocular changes in limited forms of Wegener's granulomatosis. *British Journal of Ophthalmology*, **65**, 553–63.

Spalton, D. J., Verdon Roe, G. M. and Hughes, G. R. V. (1993). Hydroxychloroquine, dosage, parameters and retinopathy. *Lupus*, **2**, 355–8.

Talks, S. J., Luqmani, R. A., and McDonell, P. J. (1994). A severe antineutrophil cytoplasmic antibody associated anterior segment vasculitis. (Letter). *Eye*, **8**, 698–700.

Towler, H. M. A., Whiting, P. H., and Forrester, J. V. (1990). Combination low dose cyclosporin A and steroid therapy in chronic intraocular inflammation. *Eye*, **4**, 514–20.

Tsubota, K. (1994). New approaches to dry eye therapy. In *Dry eye. International ophthalmic clinics*, Vol. 34, (ed. G. Smolin and M. Friedlaender), pp. 115–28. Little, Brown, Boston.

Wakefield, D. and McCluskey, P. (1989). Cyclosporin therapy for severe scleritis. *British Journal of Ophthalmology*, **73**, 743–6.

Watson, P. G. and Hazleman, B. L. (1976). *The sclera and systemic disorders*. Saunders, London.

Wolf, M. D., Lichter, P. R., and Ragsdale, C. G. (1987). Prognostic factors in the uveitis of juvenile rheumatoid arthritis. *Ophthalmology*, **94**, 1242–8.

Yarom, A., Rennebohm, R. M., and Levinson, J. E. (1985). Infantile multisystem inflammatory disease: a specific syndrome? *Journal of Pediatrics*, **106**, 390–6.

1.3.11 The kidney

Carolyn P. Cacho, Jay B. Wish, Frances B. Wheeler and Gary M. Kammer

Introduction

Renal involvement is often the most catastrophic and potentially lethal complication of systemic rheumatic diseases. As such, it is vital for the rheumatologist to be familiar with the pathophysiology and clinical manifestations of renal disease. Many renal diseases are rapidly progressive and, therefore, the rheumatologist should recognize the early signs and symptoms, and have a framework for determining when consultation with a nephrologist is indicated. In many cases, urinary abnormalities such as haematuria and proteinuria will predate the onset of renal insufficiency, giving the rheumatologist an opportunity to determine the nature and severity of the renal lesion in consultation with the nephrologist. Similarly, the acute onset or worsening of hypertension may indicate renal involvement, warranting collaboration with the nephrologist for establishing both diagnostic and therapeutic strategies.

Several rheumatic disorders have a relatively high incidence of renal involvement (Table 1). It is appropriate for patients with these illnesses to undergo a thorough initial assessment of renal function as well as frequent reassessments of renal parameters, even when the initial nephrological examination is negative. Renal involvement in the systemic rheumatic diseases is variable and unpredictable. For example, systemic lupus erythematosus may present with proteinuria without hypertension or abnormalities of renal function; the proteinuria may or may not progress to renal failure over time. At the other extreme, lupus nephritis may present as a rapidly progressive glomerulonephritis, characterized by haematuria, proteinuria, severe hypertension, and advanced renal insufficiency. Other rheumatic diseases with a relatively high incidence of renal involvement also show marked variability of signs and symptoms of kidney disease (Table 1). Furthermore, it is often difficult to differentiate rheumatic diseases either on the clinical or pathological basis of renal involvement. In many instances, the presentation of the renal disease and even the findings from renal biopsy may be identical. Thus, diagnoses of rheumatic disorders are defined on the basis of the clinical range of extrarenal findings and on serological analyses. The overlap of rheumatic and renal diseases further underscores the advantages of close collaboration between the nephrologist and the rheumatologist in their evaluation, diagnosis, and treatment. The following sections seek to provide a broad perspective of the interaction between nephrologist and rheumatologist in the evaluation and management of renal complications in systemic rheumatic diseases by focusing on proteinuria, haematuria, acute renal failure, and hypertension.

Systemic lupus erythematosus	60–70
Mixed connective tissue disease	6–47
Henoch–Schönlein purpura	30–70
Sjögren's syndrome	40–50
Systemic sclerosis	45
Wegener's granulomatosis	50
Classic polyarteritis	30–60
Microscopic polyarteritis	60–70

Numbers are percentages.

Table 1 Rheumatic disorders with a relatively high incidence of renal involvement

Proteinuria

Definitions

The upper limit of protein excretion in healthy adults is approximately 150 mg/day. Functional proteinuria describes proteinuria without intrinsic renal disease. Such protein excretion is usually less than 500 mg/day and may be associated with high fever, strenuous exercise, or exposure to cold. Although proteinuria up to 1 to 2 g daily has been recorded during severe congestive heart failure without intrinsic renal disease, proteinuria greater than 500 mg daily is strongly suggestive of intrinsic renal disease. Abnormal protein excretion of less than 3 to 5 g/day can occur in both glomerular and tubular interstitial diseases; however, more than 2 g of protein daily implicates glomerular disease. When proteinuria greater than 3.5 g daily is sufficient to lead to hypoalbuminaemia, hypercholesterolaemia, and peripheral oedema, the patient is said to have the nephrotic syndrome. Protein excretion of greater than 3.5 g daily can be found in the absence of hypoalbuminaemia, lipid abnormalities, and/or peripheral oedema. This is called nephrotic-range proteinuria, and protein excretion less than 3.5 g/day is termed non-nephrotic range proteinuria.

The quantity of proteinuria as a result of glomerular disease does not correlate with the nature or severity of the glomerular lesion. Proteinuria of greater than 20 g daily may be associated with minimal histological changes on renal biopsy and may be exquisitely sensitive to treatment with corticosteroids. By contrast, 3 to 5 g of proteinuria daily can be found in the presence of active lesions that are poorly responsive to therapy.

The prolonged presence of nephrotic-range proteinuria often leads to progressive renal insufficiency. While persistent proteinuria does not invariably progress to chronic renal failure, patients with a partial (less than 1 g daily) or complete (less than 150 mg daily) resolution of proteinuria in response to therapy have a lower rate of progression to endstage renal disease (Appel *et al*. 1987). Although remission of proteinuria is a favourable prognostic sign, rheumatic diseases do have periodic exacerbations and remissions. Thus, there is a risk for recurrence of proteinuria, which may become persistent and ultimately progress to chronic insufficiency in disorders such as systemic lupus erythematosus or systemic sclerosis.

The monitoring of proteinuria and renal function over time can be done by the nephrologist, but the rheumatologist must be kept informed in order to adjust therapy.

Patients receiving therapy aimed at reducing proteinuria should undergo serial determination of protein excretion in order to assess the response to therapy and to determine a prognosis for renal survival. Patients with complete resolution of proteinuria also should have close follow-up by urinalysis approximately 3 to 4 times per year. In addition, screening for proteinuria, examination of the urinary sediment, and measurement of renal function should be undertaken whenever patients present with an extrarenal exacerbation of a rheumatic disorder. Monitoring of certain serological factors may also yield useful information that suggests the exacerbation of renal disease. An increase in the concentration of antinative-DNA antibody and a decrease in C3 and/or C4 complement levels often predate the onset or exacerbation of renal disease in systemic lupus (Garin *et al*. 1979; Pillemer *et al*. 1988). However, as the serum C3 and C4 may or may not be chronically decreased, an isolated low complement level is not necessarily predictive of active disease. In this circumstance, both the blood pressure and renal function must be closely monitored.

Diagnostic studies in the patient with proteinuria

The onset of proteinuria in a rheumatic disease does not necessarily implicate renal involvement from the systemic illness. None the less, when it complicates a systemic rheumatic illness that has a high incidence of renal involvement (Table 1), the onset of proteinuria is certainly of concern. It is in this circumstance that it is appropriate to consult a nephrologist in order to begin a well-conceived and thorough diagnostic evaluation. Moreover, there are many patients under the care of the rheumatologist in whom the diagnosis of a systemic illness has not yet been made. With this in mind, the onset of proteinuria should trigger a more complete diagnostic evaluation. This should include quantification of urinary protein excretion and a careful microscopic examination of the urine sediment and an estimation of renal function. Additional diagnostic studies for the new onset of proteinuria should include antinuclear antibodies, serum complement (C4 and C3), hepatitis B surface antigen, and anti-HCV antibodies. As long-standing diabetes remains one of the most common causes of secondary nephrotic syndrome, an evaluation of proteinuria should include measurement of serum glucose and/or glycohaemoglobin. The increasing prevalence of human immunodeficiency virus (**HIV**) infection in the general population and the association of this infection with focal segmental glomerulosclerosis and heavy proteinuria warrants an anti-HIV antibody determination. Finally, paraproteinaemia is a significant cause of nephrotic syndrome due to glomerular amyloidosis and 'light chain nephropathy', a glomerular and tubular interstitial disorder. Therefore, a serum protein electrophoresis as well as the urine immunoelectrophoresis for the measurement of Bence-Jones proteins should be done in patients with unexplained proteinuria.

Differential diagnosis of proteinuria in systemic rheumatic disease

The main causes of heavy proteinuria in rheumatic illnesses are divisible into two groups. The first group presents with isolated nephrotic range proteinuria with or without renal insufficiency. Systemic lupus erythematosus commonly presents in this manner and, indeed, presentation with the nephrotic syndrome may precede the diagnosis of lupus. The renal biopsy generally shows a non-inflammatory form of lupus nephritis, characterized either by (a) membranous nephropathy (WHO classification type V) with thickening of the glomerular basement membrane and subepithelial deposition of immune complexes; or (b) mesangial proliferative disease (WHO classification type II) with increased mesangial cellularity and matrix but without significant proliferation of glomerular endothelial or epithelial cells and inflammatory infiltration. Patients with mesangial proliferation will generally have extensive deposition of immune complexes within the mesangial matrix.

Renal amyloidosis occurs in up to 15 per cent of patients with rheumatoid arthritis and may also complicate Sjögren's syndrome, scleroderma, and long-standing systemic lupus. This entity typically presents with heavy proteinuria in the absence of haematuria, hypertension, or renal insufficiency. The diagnosis is often inferred by the demonstration of amyloid in a biopsy from an extrarenal site, such as the rectum, gingiva, or abdominal fat pad. However, in some cases, renal biopsy is required for the diagnosis. Renal amyloidosis may be associated with massive proteinuria of greater than 20 g daily and generally leads to progressive renal insufficiency that is unresponsive to treatment.

Several therapeutic agents frequently used for connective tissue diseases have been associated with the development of heavy proteinuria. Non-steroidal anti-inflammatory drugs (**NSAIDs**) can induce heavy proteinuria, acute renal failure, and electrolyte disorders. Renal dysfunction associated with these agents is most frequently the result of haemodynamic changes within the glomerulus secondary to inhibition of prostaglandin synthesis, but these drugs may also produce acute interstitial nephritis and significant renal failure. Unlike other forms of drug-induced interstitial nephritis, peripheral eosinophilia, fever, and eosinophiluria are generally absent with NSAIDs. Proteinuria is variable but may be as high as 20 g/day. Renal biopsy, when done, has usually revealed minimal-change glomerular disease and acute interstitial nephritis. Immunofluorescence staining is negative. Recent studies have indicated that 75 to 80 per cent of the mononuclear interstitial infiltrate is composed of T lymphocytes. It has been speculated that inhibition of cyclo-oxygenase by NSAIDs shunts arachidonic acid to the lipoxygenase pathway, with subsequent production of mediators of inflammation, such as the leukotrienes. The development of heavy proteinuria with or without renal insufficiency in a patient treated with these drugs should prompt discontinuation. In most cases, the proteinuria and renal insufficiency have resolved after elimination of the offending drug (Levin 1988).

Approximately 3 to 4 per cent of patients treated with gold for rheumatoid arthritis develop proteinuria; however, only one-quarter of these show the nephrotic syndrome. Renal biopsy of these individuals has generally revealed a classic membranous glomerulopathy, a non-inflammatory lesion characterized by thickening of the glomerular basement membrane resulting from presumptive *in situ* immune-complex deposits. The pathogenesis of this lesion is poorly understood. The development of increasing proteinuria or nephrotic syndrome during gold therapy should prompt discontinuation of the drug. In over 80 per cent of patients, the proteinuria resolves within 1 year, and steroid therapy does not appear to be warranted (Katz *et al*. 1984).

Administration of penicillamine is associated with proteinuria in about 10 per cent of cases. There is no clear relationship between the development of proteinuria and either the daily or cumulative dose of the drug, suggesting an idiosyncratic reaction. As with gold therapy, renal biopsy in patients with penicillamine-associated proteinuria generally reveals a membranous glomerulopathy. Although the pathogenesis is unknown, the renal lesion usually resolves after discontinuation of the drug, but the proteinuria may persist for as long as 2 years. Renal insufficiency rarely develops, even when complicated by heavy proteinuria (Stein *et al*. 1980).

The second type of presentation is that of proteinuria of variable quantity accompanied by haematuria and possibly renal insufficiency. Systemic rheumatic illnesses which may present in this manner include polyarteritis nodosa, Wegener's granulomatosis, Henoch–Schönlein purpura, and cryoglobulinaemia. Systemic lupus erythematosus may also fall into this category, especially when it is associated with a focal or diffuse proliferative glomerular lesion (WHO classification type III to IV). Cryoglobulinaemia, also an immune-complex disease, is characterized by deposition of circulating immune complexes within the glomerulus and a proliferative glomerular lesion not unlike that seen in the more severe cases of lupus nephritis. Henoch–Schönlein purpura more often presents with haematuria (which may be gross) than with proteinuria; impairment of renal function may or may not complicate this type of purpura. Renal insufficiency, if it does occur, may be self-limiting or progressive. Polyarteritis nodosa is associated with two relatively distinct renal syndromes. In classical polyarteritis nodosa, there is a vasculitis involving the large and medium-sized vessels of the kidney and this results in hypertension as the primary presenting symptom. When present, glomerular lesions are the result of ischaemia, not inflammation, and manifest primarily as haematuria, not proteinuria. In microscopic polyarteritis, a rapidly progressive glomerulonephritis is characterized by red blood cell casts, proteinuria, acute renal failure, and hypertension. Wegener's granulomatosis, also associated with inflammation of the small vessels of the kidney including the glomeruli, presents with a similar clinical picture. Although the **c-ANCA** (antineutrophil cytoplasmic autoantibody; see below) is present in probably 90 per cent of patients with active, untreated systemic Wegener's granulomatosis (Jennette and Falk 1995), the only definitive way to distinguish these various forms of acute necrotizing glomerulonephritis is to perform a renal biopsy (Madaio 1990).

Haematuria

Definitions

Haematuria is defined as the presence of red blood cells in the urine. The normal urine should be negative for haem pigment on dipstick examination and should reveal less than three red blood cells per high-power field on examination of the sediment. Haematuria is said to be microscopic if the urine colour is normal and blood can be detected only by microscopic examination or a urine dipstick method. The haematuria is termed gross if red blood cell excretion is large enough to cause a visible reddish or brownish discoloration of the urine. Significant bleeding from the lower portions of the urinary tract such as the bladder or urethra usually produces a red discoloration of the urine. Bleeding from the kidney, on the other hand, and especially from glomeruli, results in denaturation of haem proteins as the red blood cells traverse the urinary tract; this produces a 'tea-coloured' or 'cola-coloured' urine. The dipstick haem test is not specific for intact red blood cells and will register positive if any haem pigment, including haemoglobin or myoglobin, is present in the urine. Free haemoglobin may be present in the urine as a result of massive intravascular haemolysis due to an immune mechanism in patients with severe systemic lupus or a non-immune mechanism in severe microangiopathies. Myoglobin may be present in the urine as the result of lysis of skeletal muscle. This is usually seen in traumatic or non-traumatic rhabdomyolysis but rarely can be seen with chronic inflammatory muscle diseases in exacerbation, such as dermatomyositis and polymyositis. When the haematuria is accompanied by red blood cell casts, hypertension, and oedema, the patient is said to be nephritic. Nephritic patients have a variable amount of proteinuria and decrement in renal function.

Anatomical sources of haematuria in a patient with systemic rheumatic disease

Subjects with rheumatic disease may develop bleeding at various sites along the urinary tract either as a result of the underlying systemic disease or as a result of therapy for the rheumatic disorder. Lower urinary tract (bladder and urethra) bleeding most commonly results from urinary tract infection. Chronic immunosuppressive therapy and underlying renal insufficiency predispose to urinary tract infections, which can also present with either painless or painful microscopic or gross haematuria. The presence of white blood cells, leucocyte esterase, or nitrates on urinalysis is suggestive but not diagnostic. Therefore, a urine culture should be included in the evaluation of unexplained haematuria. In addition, cyclophosphamide, a pharmacological agent used in the therapy of several rheumatic diseases, can induce haemorrhagic cystitis and has been associated with transitional cell tumours of the uroepithelium and occasionally squamous cell carcinoma of the bladder. Haemorrhagic cystitis can present as painless microscopic or painful gross haematuria. This side-effect may be prevented by hydration and the use of mesna (2-mercaptoethane sulphonate).

Bleeding from the upper urinary tract (ureters and urinary collecting system) is most often associated with long-term ingestion of non-narcotic analgesic agents. Renal papillary necrosis occurs with the accumulation of large doses of analgesic agents containing a combination of aspirin and phenacetin or with sustained use NSAIDs. The process may present with either painless microscopic or gross haematuria. Papillary necrosis can also induce renal colic, owing to ureteral spasm as the sloughed papilla passes down the ureter to the bladder. Occasionally the diagnosis can be made by detecting the passage of gross tissue into the urine or from the presence of fragments of the sloughed papilla in the microscopic examination of the urine sediment. More commonly, the diagnosis is made from radiographic studies. The complications of renal papillary necrosis are pain and the potential for urinary tract obstruction. Papillary necrosis itself is generally not associated with a decline in renal function, although chronic interstitial nephritis is often associated with this process and may lead to progressive renal failure. Recently, chronic heavy use of over-the-counter analgesics including NSAIDs and acetaminofen, but not aspirin, has been associated with the development of endstage renal disease (Sandler *et al*. 1991; Perneger *et al*. 1994). Whether this association is due to analgesic nephropathy or to the increased prevalence of analgesic ingestion among patients with diseases which predispose them to endstage renal disease remains unclear. NSAIDs can also produce an acute interstitial nephritis, which usually presents with proteinuria and renal dysfunction. Associated inflammatory changes within the renal interstitium usually resolve, but interstitial fibrosis, which may also occur, is irreversible. The key to treatment in either case is to recognize the diagnosis and to eliminate the offending agent (Levin 1988). Occasionally, microscopic haematuria can complicate interstitial nephritis. A special case of this phenomenon may occur in the interstitial nephropathy which is associated with Sjögren's syndrome. In this entity, the haematuria may be secondary to nephrocalcinosis resulting from the distal renal tubular acidosis not infrequently encountered in patients with this disease (Moutsopoulos *et al*. 1991).

Aside from the presence of a malignancy, the greatest concern in the patient with haematuria associated with systemic rheumatic illness is the development of glomerulonephritis. Glomerular haematuria, as opposed to isolated proteinuria, generally reflects an inflammatory process of the glomerulus, often resulting from the deposition of immune complexes or antiglomerular basement membrane (**anti-GBM**) antibody. Rheumatic disorders associated with this entity include systemic lupus erythematosus, Henoch–Schönlein purpura, Wegener's granulomatosis, and cryoglobulinaemia. Non-nephrotic or nephrotic-range proteinuria, cellular casts in the urinary sediment, and a deterioration of renal function may be present. The immediate danger from acute glomerulonephritis is deterioration of renal function.

Vascular diseases of the kidney can lead to haematuria by induction of ischaemic glomerular damage. Such injury causes destruction of the glomerular filtration barrier and leakage of red blood cells into Bowman's space. This mechanism occurs in both classical and microscopic polyarteritis as well as the microangiopathic syndromes such as systemic sclerosis, malignant hypertension, thrombotic thrombocytopenic purpura, and haemolytic uraemic syndrome. Finally, renal vein thrombosis, which may occur in the nephrotic syndrome, may also lead to gross or microscopic haematuria. It should be pointed out that renal vein thrombosis is not a cause of heavy proteinuria but, rather, a consequence of it. Nephrotic syndrome, especially that due to membranous glomerulonephropathy, predisposes to the development of renal vein thrombosis because of a hypercoagulable state induced by the urinary loss of certain natural anticoagulants, including antithrombin III, and increased plasma fibrinogen resulting from accelerated liver synthesis.

Diagnostic studies

It is important to point out that the presence of a known rheumatic illness does not exclude a malignant cause of haematuria. One of the greatest concerns with the new onset of haematuria in the presence or absence of rheumatic disease is the possibility of a tumour of the urinary tract. Therefore, in contrast to the evaluation of isolated proteinuria, the work up of haematuria should prompt both a urological and a nephrological evaluation. Normal urinary tract imaging with intravenous pyelography or renal ultrasonography and cystoscopy essentially rule out the bladder, ureters, and collecting system as the source of bleeding, while the coexistence of haematuria and significant proteinuria or the presence of red blood cell casts in the urine sediment offer reasonable security that the haematuria is of glomerular origin and not due to neoplasm. Protein excretion should be quantified and a baseline measure of renal function obtained. Once non-glomerular sources of the haematuria have been excluded, the exact nature of the glomerular disorder may be determined from the results of the serological tests and laboratory examination as previously described for the proteinuric patient. In addition, the ANCA test may be helpful in diagnosing a vasculitis or, in the patient with an established vasculitis, may distinguish between vasculitic disease activity and complications of cytotoxic therapy. A renal biopsy may be indicated to establish the diagnosis or to ascertain the nature, activity, or chronicity of the glomerular lesion. These may be important in determining prognosis as well as potential response to therapy.

Role of renal biopsy in the evaluation of proteinuria and haematuria

The role of renal biopsy in the evaluation of proteinuria and haematuria has been controversial. The decision to perform a renal biopsy in a specific clinical setting varies with the practice style of the nephrologist involved. Inferences can be made about the nature and activity of the renal lesion on the basis of renal functional abnormalities, the activity of the urine sediment, and the presence or severity of hypertension. Often, a diagnosis can be established from the extrarenal manifestations of the disease and its serological abnormalities. A therapeutic decision can then be made from this clinical information. The response of the renal lesion to therapeutic intervention can be assessed by changes in renal function, reduced cast excretion in the urine sediment, diminished quantity of proteinuria, and improvement in serological abnormalities. Such an approach can certainly be defended in the patient who may be at higher than usual risk for renal biopsy because of obesity, coagulation abnormalities, or the presence of a solitary kidney.

On the other hand, there are clinical circumstances in which a renal biopsy may be essential to establish a diagnosis. Neither serological studies nor the pattern of extrarenal involvement are specific for a particular renal pathology and, therefore, renal biopsy is a fundamental diagnostic procedure. Although the presence of the C-ANCA strongly supports the diagnosis of Wegener's granulomatosis, the identification of immune deposits on a renal biopsy specimen in a patient with rapidly progressive glomerulonephritis is still the 'gold standard' and an important determinant of whether corticosteroids, immunosuppressive agents, or plasmapheresis are the most appropriate course of therapy.

An intermediate circumstance occurs when the diagnosis of systemic disease is established and a renal biopsy is being considered to ascertain the nature and severity of the glomerular lesion. The most typical case would be that of lupus nephritis, in which some nephrologists might recommend renal biopsy because of the prognostic implications of certain renal lesions. In addition to classifying patients with lupus nephritis according to the pattern of glomerular involvement (minimal, mesangial, focal proliferative, diffuse proliferative, and membranous), renal pathologists make note of the extent of activity (inflammatory, potentially reversible changes) and chronicity (sclerosing, irreversible changes) within the kidney (Balow and Austin 1988; Esdaile *et al*. 1989). Thus, given a patient with moderate renal failure, the presence of a high chronicity index may be considered a relative contraindication for intensive corticosteroid or immunosuppressive therapy because the likelihood of a therapeutic response would be outweighed by the potential toxicity. By contrast, a high activity index may prompt the nephrologist to institute aggressive treatment because of the potential for reversal of the underlying pathology (Boumpas *et al*. 1995). The more conservative nephrologist might defer renal biopsy, attempt a limited trial of corticosteroid or immunosuppressive therapy, and then, if the renal disease fails to respond, consider renal biopsy at a later date to investigate a pathological explanation for the lack of response. Both approaches can be justified, and in fact, decision-analysis studies of this issue have failed to reveal any clearly favourable approach. The rheumatologist may, therefore, wish to choose the nephrology consultant on the basis of compatibility of their practice styles in this context.

Acute renal failure

Definitions

Acute renal failure is defined as a sudden decline in renal function, usually measured by an increase in serum creatinine and blood urea nitrogen (**BUN**) or a decrease in glomerular filtration rate. The serum creatinine is a reliable marker of changes in the glomerular filtration rate if its limitations are understood. The BUN is a less reliable marker of changes in renal function, primarily because it may be elevated in several conditions where renal function is stable. These include

gastrointestinal bleeding and steroid administration. As a result, rises in the BUN are best interpreted in relation to changes in the serum creatinine. Approximately 20 per cent of urinary creatinine excretion under physiological conditions is due to tubular secretion. Therefore, as glomerular filtration declines and tubular secretion of creatinine contributes a greater fraction of creatinine excretion, increases in the serum creatinine and a fall in the creatinine clearance tend to underestimate the true decrease in glomerular filtration rate that has occurred. Furthermore, therapeutic agents such as cimetidine and trimethoprim may competitively inhibit tubular creatinine secretion, leading to increases in the serum creatinine in the absence of any changes in glomerular filtration rate. Finally, because the serum creatinine is roughly inversely proportional to creatinine clearance or glomerular filtration rate, at relatively normal levels of renal function, changes in the serum creatinine tend to be relatively insensitive. For instance, an increase in the serum creatinine from 1 to 2 mg/dl would represent a halving (50 per cent decline) of the glomerular filtration rate. Once such a change has occurred, substantial renal function has been lost, despite what appears to be only a modest increase in the serum creatinine. An increase in the serum creatinine from 1.0 to 1.1 mg/dl, if real, represents a 9 per cent decline in glomerular filtration rate, also rather significant. However, the standard laboratory assay for serum creatinine has a variation of plus 10 per cent, meaning that any change of 0.1 mg/dl in the serum creatinine could correctly or incorrectly be attributed to laboratory variation. Therefore, it is important that even seemingly small changes in the serum creatinine should not be discounted and, if a more reliable indicator of renal function is required, creatinine clearance or radionuclide studies (such as iothalamate clearance) should be considered.

Patients with acute renal failure are generally described as being oliguric or non-oliguric. Oliguria is defined as a daily urine output of less than 400 to 500 ml. This quantity represents the smallest volume of urine that can serve as a vehicle for the excretion of a typical daily osmotic load of waste products, assuming that the urine is maximally concentrated throughout the 24-h period. Therefore, a urine volume of less than 400 to 500 ml/day will necessarily imply an incomplete excretion of waste products, and will be accompanied by an increase in the concentration of some of those waste products in the blood, such as the blood urea nitrogen and serum creatinine. Oliguric renal failure has traditionally been divided into prerenal, renal, and postrenal or obstructive causes. Prerenal acute renal failure is a physiological decrease in the volume of urine in response to a true decrease in intravascular volume or an effective decrease in circulating volume. This is an appropriate response by the kidney because it preserves the extracellular fluid volume while sacrificing solute excretion. It does not represent intrinsic renal disease and, by definition, should be reversed upon restoration of an adequate effective circulating volume. Prerenal azotaemia is characterized by a blood urea nitrogen-to-creatinine ratio greater than 20, a low urinary sodium, and high urine osmolarity.

In the nephrotic syndrome, hypoalbuminaemia results in transudation of fluid from the vasculature into the interstitial space. Baroreceptors in the renal vasculature then 'sense' a decreased circulating volume. Sodium and water reabsorption increase and the volume of urine declines. Similarly in congestive heart failure, as may complicate systemic lupus erythematosus, low cardiac output is 'sensed' by the kidneys and a low urine output state develops. NSAIDs, by inhibiting renal vasodilatory prostaglandins, may also produce adverse haemodynamic changes and a decline in renal function independent of their potential to produce interstitial nephritis or proteinuria. It has been reported that sulindac has a 'renal sparing' effect because the enzyme that converts sulindac to its active form is not present in the kidney. However, the sparing effect of sulindac appears to be lost at daily dosages of greater than 300 mg. Volume depletion, hepatic disease, and lupus nephritis predispose to the adverse renal haemodynamic changes associated with the NSAIDs. The alterations in renal function are reversed upon discontinuation of the offending drug. The angiotensin-converting enzyme (**ACE**) inhibitors, including captopril, enalapril, and lisinopril, can induce decreased renal perfusion and an acute decline in renal function. Such alterations in renal function are enhanced when there is underlying volume depletion, renal vascular disease, or pre-existing glomerular disease. Although this group of drugs may be the treatment of choice for the malignant hypertension associated with scleroderma renal crisis and the microangiopathies, renal function must be monitored closely because the ACE inhibitors can dilate intrarenal efferent arterioles, leading to a decline in intraglomerular pressure and glomerular filtration rate.

Cyclosporin, a potent immunosuppressive agent used extensively in solid-organ transplantation, has been shown to be effective in selected cases of rheumatoid arthritis. The drug has also been used in a number of other autoimmune diseases including Sjögren's syndrome and psoriasis. Cyclosporin's major liability is its nephrotoxicity, which may manifest as an acute, subacute, or chronic decline in renal function. Cyclosporin produces a dose-related decline in renal blood flow and glomerular filtration rate, which usually is reversible upon decreasing or discontinuing the drug. The mechanism for this effect is not well understood but is felt by some investigators to be related to the stimulation of vasoconstrictor prostaglandins within the kidney. The NSAIDs, which inhibit the vasodilator prostaglandins within the kidney, may have an additive effect on cyclosporin nephrotoxicity.

This deleterious dose-dependent effect of cyclosporin was confirmed in another study in which renal biopsy specimens from 192 patients with autoimmune diseases treated with cyclosporin were analysed to determine the effect of cyclosporine therapy. The investigators found that cyclosporin toxicity could be minimized by limiting the cyclosporin dose to no more than 5 mg/kg body weight per day and by adjusting the medication dose to avoid increases in the serum creatinine of more then 30 per cent of baseline (Feutren *et al*. 1992). In a study comparing high- and low-dose cyclosporin therapy in rheumatoid arthritis, glomerular filtration rate declined by an average of 40 per cent in patients receiving 10 mg/kg body weight per day and by an average of 17 per cent in patients receiving 1 mg/kg body weight per day (Yocum *et al*. 1988). Other, less common, forms of cyclosporin nephrotoxicity include a renal microangiopathy and chronic interstitial nephritis, both of which are irreversible. Close monitoring of renal function is vital in all patients receiving cyclosporin so that dosage modifications can be made as soon as nephrotoxicity is detected.

Postrenal failure is oliguria caused by an intrarenal or extrarenal obstruction to urine flow. Intrarenal obstruction may be due to crystallization of a substance within the renal tubules. This might be endogenous, such as uric acid or a calcium salt, or exogenous, such as crystals of methotrexate or a sulphonamide. Extrarenal obstruction to urinary flow most commonly involves the lower urinary tract because the obstruction of only one of two functioning kidneys will allow the non-affected kidney to continue to make urine, so oliguria and renal failure will not result. Severe haemorrhagic cystitis complicating cyclophosphamide therapy may produce enough blood clot within the bladder to obstruct the flow of urine and acute oliguric renal failure develops. Bilateral papillary necrosis could lead to obstruction of the upper urinary tract through occlusion of both ureters, but this would be a most unusual event. Postrenal failure can be reversed if the site of obstruction is identified, and the obstruction is relieved before onset of renal parenchymal damage. Postrenal failure is characterized by low-to-absent urine output. Urinary sodium concentration and urine osmolarity are variable. In the acute phase, the urine electrolytes and osmolarity may mimic a prerenal picture with a low urine sodium concentration and high urine osmolarity, however, as the obstruction becomes chronic, increasing tubular dysfunction is reflected by a high urine sodium concentration and low urine osmolarity. Obstruction is best diagnosed by renal ultrasound, which may reveal hydronephrosis, and cystoscopy, which could identify the presence of clots or tissue in the bladder.

Of most concern in the patient with acute renal failure is renal parenchymal disease, because this is most likely to lead to irreversible renal damage if appropriate diagnostic and therapeutic procedures are not undertaken. Acute renal failure due to renal parenchymal disease may be oliguric or non-oliguric, the latter term meaning that, despite an apparently normal daily volume of urine, solute clearance continues to decline as measured by serum creatinine or clearance techniques. The presence of normal urinary volumes may make patient management somewhat easier. Fluid and sodium restrictions may not have to be as severe, and not as much fluid may have to be removed on dialysis. However, non-oliguric renal failure may be just as ominous as oliguric renal failure with regard to prognosis and potential response to therapy. Therefore, attention only to urinary volume without consideration to changes in the serum creatinine or creatinine clearance is misguided because severe, progressive renal insufficiency may develop without an oliguric phase. Serial tests of the serum creatinine must be done in the unstable patient to assess adequately any changes in renal function. An estimation of creatinine clearance using the Cockroft–Gault formula can then be made in order to adjust medication dosing. Urine collections of 24-h duration for creatinine clearance are less helpful in the assessment of acute renal failure because, aside from being cumbersome, the large changes in the serum creatinine over a relatively short period of time render the test inaccurate. The Cockroft–Gault formula for the estimation of endogenous creatinine clearance is calculated as follows:

Creatinine clearance = (140 – age) (ideal body weight, kg) for men
(× 0.85 for women)
(72) (serum creatinine, mg/dl)

Anatomical sources of acute renal failure in the patient with systemic rheumatic disease

Acute renal failure has several pathophysiological mechanisms in systemic rheumatic diseases.

A. Prerenal:
    1. True decreased circulating volume
        a. overdiuresis
        b. dehydration
    2. 'Effective' decreased circulating volume
        a. nephrotic syndrome
        b. congestive heart failure
        c. non-steroidal anti-inflammatory drugs

        d. cyclosporin
        e. angiotensin-converting enzyme inhibitors
B. Intrinsic renal:
    1. Acute interstitial nephritis
        a. non-steroidal anti-inflammatory agents
        b. Sjögren's syndrome
    2. Acute tubular necrosis
        a. sepsis
        b. pigment
            i. haemoglobin
            ii. myoglobin
        c. toxins/drugs
            i. antibiotics
            ii. radiocontrast media
    3. Rapid progressive glomerulonephritis
        a. immune complex disease
            i. systemic lupus erythematosus
            ii. Henoch–Schönlein purpura
            iii. cryoglobulinaemia
        b. anti-GBM disease
        c. pauci-immune
            i. polyarteritis
            ii. Wegener's granulomatosis
    4. Microangiopathic
        a. scleroderma
        b. thrombotic thrombocytopenic purpura/haemolytic uraemic syndrome
        c. cyclosporin
C. Postrenal/obstructive:
    1. Extrarenal
        a. cyclophosphamide-induced haemorrhagic cystitis
        b. bilateral papillary necrosis
    2. Intrarenal
        a. Methotrexate-induced intratubular crystallization

The NSAIDs, used very often in patients with rheumatic disease, may be associated with the development of heavy proteinuria and acute renal failure. Renal biopsy reveals an inflammatory infiltrate in the interstitial or tubular portion of the kidney parenchyma which is called an interstitial nephritis. Interstitial nephritis induced by NSAIDs appears to be unique among the drug-induced tubular interstitial nephritides in its association with heavy proteinuria. Tubular interstitial nephritis generally occurs after prolonged exposure (mean, 5 to 6 months) to the offending agent. The propionic acid derivatives—fenoprofen, ibuprofen, and naproxen—have accounted for three-quarters of the cases reported, with fenoprofen alone accounting for over one-half of the incidence of tubular interstitial disease induced by NSAIDs. The lesions have also been found after the self-administration of over-the-counter ibuprofen and may recur after exposure to different NSAIDs of the same group or upon re-exposure to the same agent. The onset of nephrotic syndrome generally precedes that of renal failure. Approximately 10 per cent of cases have only nephrotic syndrome, about 15 per cent have only renal failure, and the remainder develop both (Abraham and Keane 1984). Wright's or Hansen's stain of the urinary sediment may rarely reveal eosinophiluria. Supportive dialysis becomes necessary in about one-third of patients. An improvement in renal function occurs within days of discontinuing the offending agent, although proteinuria may persist for weeks to months. There appear to be no definite predisposing factors that identify subjects prone to developing nephritis induced by these drugs. Alternative agents infrequently associated with nephrotoxicity include meclofenamate and sulindac.

Sjögren's syndrome is associated with a chronic inflammatory infiltrate of the renal interstitium that may lead to a progressive deterioration in renal function. More common renal manifestations are those related to tubular dysfunction, including distal renal tubular acidosis and nephrogenic diabetes insipidus. Glomerular obsolescence with sclerosis may occur in association with severely damaged tubules; heavy proteinuria is uncommon. Renal involvement is generally inferred from a rising creatinine and tubular dysfunction, which may present with metabolic acidosis, hypokalaemia, hypophosphataemia, hypouricaemia, aminoaciduria, or polyuria (El-Mallakh *et al*. 1985). However, renal involvement is generally not a limiting factor in Sjögren's syndrome. A specific medical therapy for the renal lesion is not available.

Acute tubular necrosis has several potential pathophysiological mechanisms in systemic rheumatic diseases. Perhaps the most common is sepsis complicating chronic immunosuppression. Sepsis with endotoxaemia may lead to substantial renal dysfunction, even in the absence of overt septic shock, because local release of vasoactive substances in the kidney produces vasoconstriction and shunts blood flow away from the kidneys. While early acute renal failure of sepsis may resemble prerenal azotaemia and may be characterized by a low urinary sodium and a high urinary osmolarity, it is more often associated with high urinary sodium and a low urinary osmolarity. Aggressive hydration often fails to reverse the oliguria and rising serum creatinine, which indicates that the defect is not merely a function of decreased renal perfusion. Sepsis-associated renal failure may respond to the renal vasodilator dopamine. None the less, the treatment of choice is eradication of the infection with appropriate antibiotic therapy.

Large loads of haem pigment can result in renal tubular toxicity and acute renal failure with acute tubular necrosis. Severe intravascular haemolysis, which can occur in systemic lupus and the microangiopathic diseases, can result in the formation of haemoglobin casts within the tubules, intratubular obstruction, and direct tubular toxicity. There is considerable experimental evidence that red cell stroma contributes to the nephrotoxicity of acute haemolysis by inducing vasoconstrictive changes within the kidney. Such alterations in vascular tone promote tubular reabsorption of fluid, thereby further concentrating and enhancing the precipitation of haemoglobin within the tubular lumens. Polymyositis and dermatomyositis are rarely associated with myoglobinuric acute renal failure, because these diseases tend to be chronic rather than acute diseases of muscle.

The clinical course of pigment-induced nephropathy is similar to that of other forms of acute tubular necrosis with oliguria and (because of cell breakdown) significant hyperkalaemia. Mannitol may be useful in the early stages of the renal failure to increase urine flow, thereby washing out obstructing tubular casts, producing direct renal vasodilatation, and possibly preventing swelling of glomerular cells. Once renal failure due to pigment nephropathy becomes established, management is similar to that of other forms of acute renal failure. The electrolyte disorders, such as hyperkalaemia, that may be associated with cell lysis may necessitate early dialysis.

Rapidly progressive glomerulonephritis is a catastrophic form of inflammatory renal disease that sometimes complicates systemic rheumatic disease. It is a syndrome characterized by hypertension, proteinuria, haematuria with red blood cell casts, and a rapid decline in renal function. The pathological correlate of rapidly progressive glomerulonephritis is the formation of epithelial cell crescents in Bowman's space, often accompanied by a necrotizing vasculitis of the arterioles and glomerular capillaries. There are a number of systemic rheumatic diseases that can be associated with this type of glomerulonephritis, including systemic lupus erythematosus, polyarteritis nodosa, Wegener's granulomatosis, and cryoglobulinaemia. If the diagnosis can be established on the basis of the extrarenal manifestations of these diseases, then the cause of a rapidly progressive glomerulonephritis should be apparent. However, the glomerulonephritis can be the initial manifestation of disease. At times, its extrarenal manifestations are not sufficiently specific to allow for a diagnosis. In such circumstances, the immunopathology of the renal lesion has traditionally been the basis for classifying the various forms of the rapidly progressive disease. The detection of deposits within the glomerular capillaries by immunofluorescence microscopy, or of subepithelial or subendothelial deposits by electron microscopy, suggests an immune-complex pathogenesis that is characteristic of lupus nephritis, postinfectious glomerulonephritis, or cryoglobulinaemic nephritis. The presence of linear immunofluorescence along the glomerular basement membrane is diagnostic for glomerulonephritis mediated by anti-GBM antibody which, in the presence of pulmonary haemorrhage, is termed Goodpasture's syndrome.

Much less well defined, however, are the forms of rapidly progressive glomerulonephritis that are not associated with immune-complex deposits or anti-GBM binding within the glomerulus, which have been termed pauci-immune glomerulonephritis. The diagnosis usually depends upon the pattern of extrarenal manifestations and includes polyarteritis nodosa in the presence of systemic necrotizing arteritis, Wegener's granulomatosis in the presence of pulmonary necrotizing granulomas, and idiopathic crescentic glomerulonephritis when no extrarenal disease is present. A scheme for the differential diagnosis of rapidly progressive glomerulonephritis is summarized in Table 2.

Group	Glomerular immunofluorescence	Serological markers	Extrarenal manifestations	Diagnosis
Immune complex	Granular	Antinuclear antibodies	Skin, joints, haematological	Systemic lupus
		Cryoglobulins	Skin, joints	Cryoglobulinaemic glomerulonephritis
		Antistreptococcal antibodies	None	Poststreptococcal glomerulonephritis
		HBsAg	Skin, joints	Hepatitis-associated glomerulonephritis
		Increased IgA	Skin, joints, abdominal pain	Henoch–Schönlein purpura
Anti-GBM[a]	Linear	Anti-GBM antibodies	Lung haemorrhage	Goodpasture's syndrome
			None	Anti-GBM glomerulonephritis
Pauci-immune	None	Mostly c-ANCA[b]	Lungs, sinuses, joints	Wegener's granulomatosis
		c-ANCA or p-ANCA[c]	Skin, joints, lungs, nerves	Polyarteritis nodosa
		Mostly p-ANCA	None	Idiopathic crescentic glomerulonephritis

[a]Antiglomerular basement membrane. [b]Cytoplasmic, and [c]perinuclear antineutrophil cytoplasmic autoantibodies.

Table 2 Differential diagnosis of rapidly progressive glomerulonephritis

It has now been demonstrated that idiopathic, rapidly progressive glomerulonephritis and vasculitis-associated glomerulonephritis may share a common serological marker, antineutrophil cytoplasmic autoantibodies (**ANCA**). However, the ANCA test has also been associated with other systemic disorders in which renal disease may be absent. The ANCA test produces three patterns of indirect immunofluorescent staining: fine granular cytoplasmic (**c-ANCA**), perinuclear (**p-ANCA**), and a diffuse, atypical ANCA (**a-ANCA**) (Gross and Csernok 1995). The c-ANCA and p-ANCA immunofluorescent patterns are shown in Fig. 1.

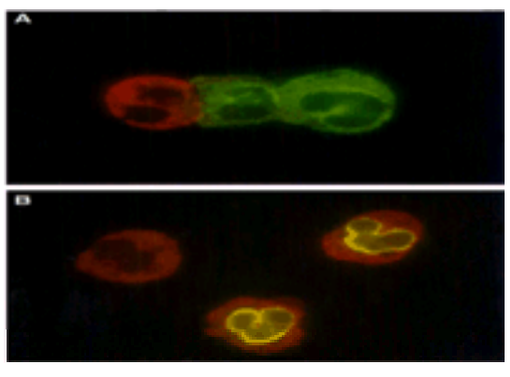


Fig. 1 Antineutrophil cytoplasmic autoantibody (ANCA) indirect immunofluorescence patterns on neutrophils (× 1000). (a) Cytoplasmic (c-ANCA) pattern with antigen specificity for proteinase 3 (PR3-ANCA), and a negative eosinophil (as seen on both ethanol- and formalin-fixed neutrophils). (b) Artefactual perinuclear (p-ANCA) staining on ethanol-fixed neutrophils, with a negative eosinophil. p-ANCA immunofluorescence due to myeloperoxidase antibodies (MPO-ANCA) appears cytoplasmic on formalin-fixed neutrophils.

c-ANCA reflects an autoantibody response against a 29-kDa serine protease, known as proteinase 3 (**PR3**), which is localized within the azurophilic granules of quiescent neutrophils (Niles *et al*. 1989) as well as on the cell surface of activated neutrophils (Csernok *et al*. 1990). Localized in lysosomes, PR3 is a bifunctional protein with both enzymatic and antimicrobial functions. Acting as a proteinase, it hydrolyses naphthol-ASD-chloracetate.

Although several azurophilic granule enzymes can drive the p-ANCA response, the principal autoantigen is myeloperoxidase (**MPO**) (Falk and Jennette 1988; Wiik *et al*. 1995). MPO is a 140-kDa cationic protein of azurophilic granules of neutrophils and certain granules of monocytes that generates chlorinated oxygen species upon activation of the cell by a microbe.

The autoantigen(s) stimulating the a-ANCA response has not been identified. PR3 and MPO are the most commonly-identified autoantigens triggering the c-ANCA and p-ANCA responses, respectively, and therefore the terms PR3-ANCA and MPO-ANCA are coming into more common usage.

PR3-ANCA (c-ANCA) is a sensitive and specific autoantibody for active Wegener's granulomatosis. Although the frequency of this autoantibody in all patients with Wegener's granulomatosis approximates 90 to 95 per cent (Kallenberg *et al*. 1994), the overall sensitivity is estimated to be 66 per cent and the overall specificity is 98 per cent (Rao *et al*. 1995). During active disease, the pooled sensitivity increases to 91 per cent and the pooled specificity rises to 99 per cent. It remains uncertain, however, whether or not disease activity correlates directly with autoantibody titres. Based on an estimated 5 per cent prevalence of Wegener's granulomatosis, 37 per cent of patients with a positive c-ANCA may have a false-positive test result. These data underscore the importance of employing the c-ANCA test as corroborating evidence for the diagnosis of Wegener's granulomatosis in patients with the appropriate clinical stigmas and disease activity.

The p-ANCA is most commonly associated with microscopic polyangiitis, Churg–Strauss syndrome, and pauci-immune glomerulonephritis. These disorders tend to be associated with MPO-ANCA. The p-ANCA pattern occurs in less than 5 per cent of subjects with Wegener's granulomatosis and classic polyarteritis nodosa.

At a recent consensus conference, the idiopathic vasculitides were reclassified into large, medium, and small vessel vasculitis (Jennette *et al*. 1994). Although Wegener's granulomatosis, microscopic polyangiitis, and Churg–Strauss syndrome were the small vessel vasculitic disorders predominantly associated with a positive ANCA test, it is also known that other non-renal and non-vasculitic processes, including ulcerative colitis and primary sclerosing cholangitis, are often associated with a positive p-ANCA.

Perhaps the most malignant form of renal involvement in the systemic rheumatic diseases is the microangiopathy that may be seen in systemic sclerosis, the thrombotic thrombocytopenic purpura/haemolytic uraemia syndrome, and with the use of cyclosporin. The presence of microscopic renal lesions in patients with systemic sclerosis may be as high as 80 per cent, although only 40 to 50 per cent will present with clinical renal involvement, including hypertension, proteinuria, or azotaemia. Hypertension is very ominous in patients with systemic sclerosis, with the death rate among hypertensive patients being about 2.5 times greater than among those with normal blood pressure. There are two distinct forms of hypertension seen in patients with systemic sclerosis. Mild to moderate hypertension associated with some degree of proteinuria may lead to the slow development of renal failure, and the extrarenal manifestations of systemic sclerosis usually dominate the clinical picture. However, in 15 to 20 per cent of patients who develop hypertension in the setting of systemic sclerosis, the onset is abrupt with severe elevation of blood pressure to greater than 130 mmHg diastolic, usually associated with cardiac failure and rapid deterioration of renal function (Steen *et al*. 1984). In some patients, severe hypertension may be the presenting feature of systemic sclerosis, with little involvement of skin, peripheral vessels, or joints. Classically, the small arterioles of the kidney in patients with the malignant form of hypertension in systemic sclerosis show extensive fibrinoid necrosis and microthrombus formation. Glomerular changes are variable and show varying degrees of thrombosis and necrosis secondary to ischaemia. Obviously, the accelerated hypertension and rapidly progressive renal failure of systemic sclerosis are life-threatening and demand prompt, intensive intervention. The initial priority is to bring the blood pressure under control, and the ACE inhibitors have been the mainstay of therapy in this regard (Asher and Murray 1991). Even if oliguria and progressive renal failure supervene despite control of blood pressure, antihypertensive therapy should be continued because it will protect other organs from the damaging effects of severe hypertension and renal function may return after months of dialysis support (Chapman *et al*. 1986).

Thrombotic thrombocytopenic purpura and haemolytic uraemic syndrome are classical examples of microangiopathic haemolytic anaemia and have many features in common including anaemia, thrombocytopenia, and renal involvement. The two syndromes are similar in most of their clinical and laboratory findings, differing primarily in age of onset and extent of renal involvement. An abnormal urinalysis characterized by haematuria, haemoglobinuria, proteinuria, hyaline, and granular casts is seen in the vast majority of patients with either disease, but those with haemolytic uraemic syndrome have a 90 per cent incidence of associated renal function abnormalities compared with a 50 per cent incidence in patients with thrombotic thrombocytopenic purpura. The classical pathological finding in the kidney in both syndromes is occlusive lesions or arterioles and glomerular capillaries with fibrin, platelets, and red blood cells. Glomerular capillaries also show hyaline thrombosis with areas of infarction. The pathogenesis of these diseases is not completely understood, but it appears that some form of endothelial injury may be the

underlying insult. Patients with systemic lupus erythematosus have been known to progress to thrombotic thrombocytopenic purpura (Gelfand *et al*. 1985), and it is also seen in the setting of severe disseminated intravascular coagulation, sepsis, and other vasculitides including polyarteritis nodosa and Wegener's granulomatosis. Plasma exchange has become the mainstay of therapy in thrombotic thrombocytopenic purpura and haemolytic uraemic syndrome. Renal failure with its associated complications remains the most common cause of death (Eknoyan and Riggs 1986).

Principles of treatment

Several general principles of evaluation and management of acute renal dysfunction in systemic rheumatic diseases should be emphasized. Deterioration of renal function is usually associated with moderate to severe hypertension. If incompletely treated, hypertension will further compound the renal injury by producing more intrarenal vascular damage. Therefore, intensive control of blood pressure is imperative. Malignant hypertension associated with renal vasculitis, systemic sclerosis, or microangiopathic disease results from extremely high concentrations of renin and angiotensin, released in response to focal areas of renal ischaemia. The ACE inhibitors, therefore, may be the only oral agents effective in the sustained reduction of blood pressure. Severe hypertension complicated by stroke, hypertensive encephalopathy, coronary ischaemia, pulmonary oedema, or aortic dissection requires concentrated intervention with confinement to an intensive care unit and use of a rapidly acting, parenteral, antihypertensive agent. There may be a further decline in renal function after hypertension has come under control, because of decreased renal perfusion pressure in the setting of high renal vascular resistance. However, once renal autoregulatory processes have stabilized, gradual return of renal function to physiological levels takes place after the successful control of blood pressure.

The chief cause of complications in the setting of acute renal failure is not the azotaemia but the associated fluid and electrolyte abnormalities. The rheumatologist should manage these abnormalities with the nephrologist. In the acutely ill, highly catabolic patient with systemic illness, early and frequent dialysis may be required. Fluid overload may compound hypertension, leading to further renal injury, and should be managed intensively with high-dose diuretics. The rheumatologist and nephrologist should review the prescribed medications critically as soon as acute renal insufficiency is recognized. The dosage schedule for potentially nephrotoxic agents such as NSAIDs, aminoglycoside antibiotics, and amphotericin B should be adjusted for the level of renal failure or discontinued if possible. Intravascular radiocontrast procedures should be avoided during acute renal insufficiency. If it is necessary to use a diagnostic procedure requiring contrast, a minimal dose should be employed and the patient intensively hydrated before, during, and after the procedure to promote excretion of the radiocontrast. The presence of a highly catabolic state caused by corticosteroid therapy, gastrointestinal bleeding, parenteral nutrition, or septicaemia can induce a rise in the blood urea nitrogen by as much as 50 mg/dl per day in the setting of acute renal failure. Such levels of azotaemia are clearly detrimental to platelet and neutrophil function and may potentiate a bleeding diathesis or further compromise host defenses. Therefore, intensive daily or continuous dialysis may be warranted.

Course and prognosis

The prognosis of acute renal failure is determined both by the underlying disease and by the quality of medical management. Subjects with renal dysfunction complicating a rheumatic disorder should be followed-up jointly by a rheumatologist and nephrologist.

Hypertension

Introduction

Hypertension is a very common finding in the adult population, affecting as many as 20 to 30 per cent of individuals over the age of 40. Raised blood pressure may be an incidental finding in the patient with systemic rheumatic illness and does not necessarily represent renal involvement from their disease. However, the onset of acute hypertension in a patient who was previously normotensive or whose blood pressure was well controlled may suggest renal involvement and may predate the onset of haematuria, proteinuria, or renal insufficiency. Anyone with a rheumatic disease who develops severe hypertension should be intensively investigated for the cause, and referral to a nephrologist should be considered. Intensive control of the blood pressure is important to minimize the renal and extrarenal complications of the severe hypertension, and a nephrologist may be best equipped to choose appropriate antihypertensive therapy in this setting.

Specific diseases

Classical polyarteritis may produce hypertension through involvement of medium-sized arteries within the kidney, leading to distal areas of decreased perfusion and the release of potent vasoconstrictor substances. The glomeruli may not be directly affected in this polyarteritis and, therefore, the presence of minimal haematuria, proteinuria, or other abnormal urinary sediment findings may presage abnormal results from renal function tests. Severe hypertension is, in fact, the most common renal manifestation of classical polyarteritis, affecting as many as 35 per cent of subjects with no other clinical renal manifestations of the disease. The diagnosis of classical polyarteritis in this setting is generally made from the extrarenal manifestations, most commonly mesenteric vasculitis with abdominal pain, abnormal liver function, and mononeuritis multiplex. Up to one-third will have a positive serology for hepatitis B surface antigen. In the absence of proteinuria or haematuria suggesting active glomerular involvement, renal biopsy is not helpful in the confirmation of the diagnosis. However, even without hypertension, polyarteritis nodosa is often found to involve medium-sized arteries on renal arteriography, characterized by segmental narrowing and aneurysmal dilatation.

Rapidly progressive glomerulonephritis can be associated with severe hypertension in the setting of acute glomerular injury. Intensive control of the blood pressure should be the objective, even before a specific diagnosis of the glomerular disease has been established. Inflammation is often compounded by ischaemia and, therefore, the hypertension will often prove refractory. High doses of multiple antihypertensive agents may be required for its control. A further decline in renal function may be anticipated as blood pressure control is achieved, because profound abnormalities in autoregulation of renal blood flow can occur, impairing the ability of the glomeruli to adapt to a fall in renal perfusion pressure. None the less, a rise in the serum creatinine does not obviate the necessity to reduce the diastolic pressure to between 90 and 100 mmHg.

Summary and conclusions

The purpose of this chapter has been to provide the rheumatologist with a survey of the most important renal manifestations of systemic rheumatic illnesses, grouped according to the presentation most likely to be encountered in clinical practice. Table 3 is a list of important points to remember. Proteinuria, haematuria, acute renal failure, and hypertension can be ominous signs of direct renal involvement in systemic rheumatic diseases. However, they may also represent indirect involvement by complications of the primary disorder or its treatment, in which case the prognosis tends to be more favourable. The role of the rheumatologist here is to be familiar with the diseases most likely to affect the kidney and their renal manifestations, and to obtain the appropriate screening tests so that renal involvement can be identified promptly and appropriate intervention initiated without delay. When the patient does develop urinary abnormalities, decreasing renal function, or severe hypertension, the nephrologist should be involved early to assist the rheumatologist in the diagnosis and treatment of the renal disorder. Renal biopsy is a useful tool in many of these settings, but it should be understood that the information gained from it is more likely to be helpful in establishing the pathophysiological background, extent of activity, and chronicity of the renal lesion than it is to provide a specific diagnosis. The diagnosis of most of these systemic diseases is more likely to be established from the nature and pattern of extrarenal manifestations, and from serological abnormalities. Findings on renal biopsy can be helpful in choosing therapy in certain individuals, based on the confirmation of a suspected diagnosis, the activity compared with the chronicity of the lesion, the likelihood of responsiveness to therapy, and occasionally, an unexpected finding such as drug nephrotoxicity or amyloidosis.

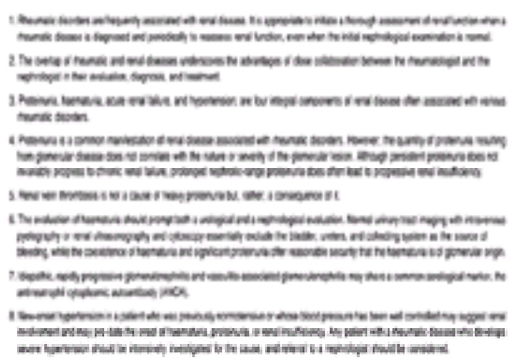


Table 3 Important points to remember

Many deaths associated with systemic rheumatic disease are due to renal involvement. The improvement in diagnosis and treatment as well as the widespread availability of dialysis for chronic conditions have significantly reduced fatality for many of these illnesses. None the less, the presence of renal involvement remains an ominous prognostic sign in many of these diseases and should be treated seriously. Relatively recent discoveries concerning the immunopathogenesis of the renal involvement in some of these illnesses are promising but also reminds much that yet remains unknown about these disorders.

Chapter References

Abraham, P.A. and Keane, W.F. (1984). Glomerular and interstitial disease induced by nonsteroidal anti-inflammatory drugs. *American Journal of Nephrology*, **4**, 1–6.

Appel, G.B. *et al*. (1987). Long term follow-up of patients with lupus nephritis. *American Journal of Medicine*, **83**, 877–85.

Asher, J.P and Murray, K.M. (1991). Use of ngiotensin-converting-enzyme inhibitors in the management of renal disease. *Clinical Pharmacy*, **10**, 25–31.

Balow, J.E. and Austin, H.A. (1988). Renal disease in systemic lupus erythematosus. *Rheumatic Disease Clinics of North America*, **14**, 117–33.

Boumpas, D.T. *et al*. (1995). Systemic lupus erythematosus: emerging concepts Part 1: renal, neuropsychiatric, cardiovascular, pulmonary and hematologic disease. *Annals of Internal Medicine*, **122**, 940–50.

Chapman, P.J., Pascoe, M.D., and van Zyl-Smit, R. (1986). Successful use of captopril in the treatment of 'scleroderma renal crisis'. *Clinical Nephrology*, **26**, 106–8.

Csernok, E. *et al*. (1990). Ultrastructural localization of proteinase 3, the target antigen of anti-cytoplasmic antibodies circulating in Wegener's granulomatosis. *American Journal of Pathology*, **137**, 1113–20.

Eknoyan, G. and Riggs, S.A. (1986). Renal involvement in patients with thrombotic thrombocytopenic purpura. *American Journal of Nephrology*, **6**, 117–31.

El-Mallakh, R.S. *et al*. (1985). Long term low dose glucocorticoid therapy associated with remission of overt renal tubular acidosis in Sjögren's syndrome. *American Journal of Medicine*, **79**, 501–14.

Esdaile, J.M. *et al*. (1989). The clinical and renal biopsy predictors of long-term outcome in lupus nephritis: a study of 87 patients and review of the literature. *Quarterly Journal of Medicine*, **72**, 779–833.

Falk, R.J. and Jennette, J.C. (1988). Anti-neutrophil cytoplasmic autoantibodies with specificity for myeloperoxidase in patients with systemic vasculitis and idiopathic necrotizing and crescentic glomerulonephritis. *New England Journal of Medicine*, **318**, 1651–7.

Feutren, G. *et al*. (1992). Risk factors for cyclosporine induced nephropathy in patients with autoimmune diseases. *New England Journal of Medicine*, **326**, 1654–60.

Garin, E.H. *et al*. (1979). The significance of serial measurements of serum complement C3 and C4 components and DNA binding capacity in patients with lupus nephritis. *Clinical Nephrology*, **12**, 148–54.

Gelfand, J. *et al*. (1985). Thrombotic thrombocytopenic purpura syndrome in systemic lupus erythematosus: treatment with plasma infusion. *American Journal of Kidney Diseases*, **6**, 154–60.

Gross, W.L. and Csernok, E. (1995). Immunodiagnostic and pathophysiologic aspects of antineutrophil cytoplasmic antibodies in vasculitis. *Current Opinion in Rheumatology*, **7**, 11–19.

Jennette, J.C. and Falk, R.J. (1995). Clinical and pathological classification of ANCA-associated vasculitis: what are the controversies? *Clinical and Experimental Immunology*, **101**, 18–22 (Suppl., The Proceedings of the 6th International ANCA Workshop).

Jennette, J.C. *et al*. (1994). Nomenclature of systemic vasculitides: proposal of an international conference. *Arthritis and Rheumatism*, **37**, 187–92.

Kallenberg, C.G.M. *et al*. (1994). Anti-neutrophil cytoplasmic antibodies: current diagnostic and pathophysiologic potential. *Kidney International*, **46**, 1–15.

Katz, W.A., Blodgett, R.C., and Pietrusko, R.G. (1984). Proteinuria in gold-treated rheumatoid arthritis. *Annals of Internal Medicine*, **101**, 176–9.

Levin, M.L. (1988). Patterns of tubulo-interstitial damage associated with nonsteroidal antiinflammatory drugs. *Seminars in Nephrology*, **8**, 55–61.

Madaio, M.P. (1990). Renal biopsy. *Kidney International*, **38**, 529–43.

Moutsopoulos, H.M. *et al*. (1991). Nephrocalcinosis in Sjögren's syndrome: a late sequela of renal tubular acidosis. *Journal of Internal Medicine*, **230**, 187–91.

Niles, J.L. *et al*. (1989). Wegener's granulomatosis autoantigen is a novel neutrophil serine proteinase. *Blood*, **74**, 1888–93.

Perneger, T.V. *et al*. (1994). Risk of kidney failure associated with the use of acetominophen, aspirin, and nonsteroidal anti-inflammatory drugs. *New England Journal of Medicine*, **331**, 1675–9.

Pillemer, S.R. *et al*. (1988). Lupus nephritis: association between serology and renal biopsy measures. *Journal of Rheumatology*, **15**, 284–8.

Rao, J.K. *et al*. (1995). The role of antineutrophil cytoplasmic antibody (c-ANCA) testing in the diagnosis of Wegener granulomatosis. *Annals of Internal Medicine*, **123**, 925–32.

Sandler, D.P. *et al*. (1991). Nonsteroidal anti-inflammatory drugs and the risk for chronic renal disease. *Annals of Internal Medicine*, **115**, 165–72.

Steen, V.D. *et al*. (1984). Factors predicting the development of renal involvement in progressive systemic sclerosis. *American Journal of Medicine*, **76**, 779–86.

Stein, H.B. *et al*. (1980). Adverse effects of D-penicillamine in rheumatoid arthritis. *Annals of Internal Medicine*, **92**, 24–9.

Wiik, A. *et al*. (1995). The diversity of perinuclear antineutrophil cytoplasmic antibodies (pANCA) antigens. *Clinical and Experimental Immunology*, **101**, 15–17 (Supplement, The Proceedings of the 6th International ANCA Workshop).

Yocum, D.E. *et al*. (1988). Cyclosporin A in severe, treated-refractory rheumatoid arthritis. *Annals of Internal Medicine*, **109**, 863–9.

1.3.12 Psychiatric issues in rheumatology

Malcolm P. Rogers and Simon Helfgott

Introduction

The practice of rheumatology is not unlike the practice of psychiatry. Relationships with patients tend to be long term, most of the diseases are chronic and unpredictable in course, and the treatments available seldom cure but often are effective in increasing function and controlling the underlying disease. Function is an important focus of both specialties. The two disciplines have avoided a high-technology, disease-orientated approach in favour of a more comprehensive view of the patient's illness, open to relevant psychological and social data. Interestingly, a study by the Rand Corporation on the effects of six chronic disorders, including major depression and rheumatoid arthritis, showed that both had a similarly adverse affect on quality of life (Stewart *et al*. 1989). The number of days of missed work and time spent in bed at home were roughly equivalent for both disorders.

Stress and psychoneuroimmunology

While most of this chapter will concentrate on the recognition and treatment of the psychiatric complaints and disorders that occur frequently among patients with rheumatic disease, it is worthwhile first briefly discussing the topic of psychoneuroimmunology. There is an extensive literature on the possible effects of psychological stress and other emotional states on the onset and the course of rheumatic disease, particularly rheumatoid arthritis and systemic lupus erythematosus (Rimon 1989). For years, experienced clinicians and patients frequently connected flare-ups of disease with stressful life events. It is fair to say that the issue is not fully settled from a scientific point of view, despite many years of effort. The principal difficulty has been the practical necessity of relying on retrospective analyses. Some of the more imaginative approaches have included studies of identical twins discordant for rheumatoid arthritis (Meyerowitz *et al*. 1968), and attempted quantitation of the amount of life stress during the 6 months immediately before a flare-up of disease (Heisel 1972). These and other studies have provided some suggestive evidence for a link between stress and illness but there have been negative results as well. These questions helped to stimulate what has now become the new and evolving field of psychoneuroimmunology (Kiecolt-Glaser and Glaser 1989). George Solomon was one of the first to ask whether psychological stress might influence rheumatoid arthritis by its effect on the immune system (Solomon 1969). Using an animal model of inflammatory arthritis he demonstrated that stress could indeed alter parts of the immune response. Subsequent studies with the model of arthritis induced by type II collagen in rodents have shown that various kinds of psychological stress have important but often conflicting effects on inflammation in joints, apparently without directly affecting antibody or cellular immune responses (Rogers *et al*. 1980).

Psychoneuroimmunology has also explored the effects of bereavement, depression, and relaxation on immunity in a variety of clinical states (Ader 1991). In simple terms, experimental and naturally occurring stress has been associated with altered measures of both cellular and humoral immunity. However, there is no clear demonstration as yet that these altered measures are biologically significant in the sense that they cause or adversely affect human disease. Certain forms of psychological intervention in cancer, for example, have been shown positively to affect the incidence of illness and death. Whether the psychological mechanism responsible for this clinical benefit acts through an immunological pathway remains unanswered.

Psychiatric manifestations of rheumatic diseases

The psychiatric manifestations of rheumatic diseases fall into three basic categories: (i) organic diseases caused by central nervous involvement of the underlying condition (primary disorders); (ii) psychiatric disturbances that are psychological 'reactions' to the impairments of rheumatic disease (secondary disorders); (iii) drug-induced psychiatric disorders (tertiary disorders). Each group is outlined below in Table 1, Table 3, and Table 4, respectively. The difficult task for the rheumatologist or psychiatrist is to define which of the three groups best describes the underlying cause of the psychiatric impairment or whether there is an independent, primary psychiatric disorder. This chapter will attempt to outline an approach to the evaluation of psychiatric illness in the patient with rheumatic disease.

Systemic lupus erythematosus	Affective disorders Psychosis Organic mental disorder
Primary Sjögren's syndrome	Affective disorders Paranoid ideation Personality changes
Rheumatoid arthritis	Organic mental disorder (rare)
Giant-cell arteritis/polymyalgia rheumatica	Organic mental disorder Depression
Granulomatous angiitis of the central nervous system	Organic mental disorder
Systemic necrotizing vasculitis–polyarteritis nodosa/ Wegener's granulomatosis	Organic mental disorder Seizures
Cogan's syndrome	Seizures
Wegener's disease	Organic mental disorder Hypersomnia
Behçet's syndrome	Organic mental disorder Seizures
Sarcoidosis	Organic mental disorder Psychosis Affective disorder, hyper-controllable Confusion
Lyme disease	Organic mental disorder Cognitive impairment Sleep disturbance
Amyloidosis	Organic mental disorder

Table 1 Rheumatic diseases with primary psychiatric involvement

Anti-double-stranded DNA antibodies
Depressed complement levels; C3, C4
Lymphopenia
Antineurofilament antibody
Anticardiolipin antibody
Antiribosomal-P antibodies
Antineuronal antibodies

Table 3 Proposed laboratory correlates of neuropsychiatric systemic lupus erythematosus

Disease	Psychiatric manifestations
Rheumatoid arthritis Juvenile chronic arthritis	Mood disturbance (from endogenous depression to grief), anxiety
Spondyarthropathies Osteoarthritis Polymyalgia rheumatica Systemic sclerosis Systemic lupus erythematosus Primary Sjögren's syndrome Myositis	

Table 4 Rheumatological disease with secondary psychiatric involvement

Rheumatic disease with primary psychiatric involvement (Table 1)

Organic mental syndromes occur frequently in this category. They include dementia, delirium, cognitive impairment, and sometimes more subtle changes of personality and affect. A disturbance of orientation, memory, language, or attention (and obviously specific motor and sensory symptoms) all strongly suggest organic disease of the central nervous system, and therefore help to identify primary psychiatric involvement.

Systemic lupus erythematosus

Systemic lupus can have major effects on the central nervous system, producing a variety of neuropsychiatric symptoms (see Chapter 5.7.1). Nearly a century ago Sir William Osler described these central nervous symptoms, including delirium, aphasia, and hemiplegia, and speculated that they may relate to vascular changes in the brain (Osler 1903). Daly (1942) was the first to emphasize the varied psychiatric presentations seen with systemic lupus, reporting 'toxic delirium' with confusion and disorientation as well as psychoses with paranoid features. Other psychiatric features include affective disorders, dementia, phobias, and autistic behaviour (Adelman *et al*. 1986). The true incidence of psychiatric symptoms is difficult to ascertain; most large series report an incidence of 15 to 20 per cent (Hay *et al*. 1992).

Clinical features

The psychiatric manifestations can sometimes predate the other clinical manifestations of systemic lupus. Most patients show some 'organic' symptoms suggestive of an organic mental disorder. However, a schizophrenic-like picture with hallucinations and paranoia and a grossly clear sensorium is sometimes found. Autistic behaviour has been seen, but infrequently. About half of the psychiatric episodes in the Johns Hopkins' study were accompanied by neurological signs or symptoms, with seizures being significantly associated (Feinglass *et al*. 1976). One-third of patients had no neurological features at any point in the course of the disease.

Using standardized neuropsychological tests, a number of investigators noted significantly more cognitive impairment in patients with systemic lupus than in those with rheumatoid arthritis and in controls (Carbotte *et al*. 1986; Hanly *et al*. 1992a; Denburg *et al*. 1993). Neuropsychological testing has the advantage of being more directly relevant to the functional and behavioural capacity of the patient. Easily administered tests include the Bender Gestalt (with immediate and delayed recall), trail making, verbal fluency, and the Wechsler Adult Intelligence Scale–Revised.

Psychiatric assessment

Psychiatric assessment consists of a careful history, including the history of pre-existing psychiatric disorders or substance abuse (Table 2). It is very useful to ascertain the patient's usual baseline of psychological functioning, including educational background, coping mechanisms, and view of their current illness. Particular investigation should focus on possible somatoform symptoms, such as conversion reactions, hypochondriasis, and frequent procedures or surgical interventions without evidence for organic disease. In addition, mental status should be carefully examined, with particular attention to orientation, memory (both remote and short-term), attention, language, reasoning, and visuospatial disturbance. Tests should include memory for remote events, such as names of political figures and the detail of historical events, as well as for repetition and retention of new information, such as four random words. Attention can be tested by having the patient repeat digits backward and forward, by naming the days of the week in reverse, or by subtracting serial 3s or 7s. Brief tests of naming objects in the room, reading, repeating a phrase, and writing a sentence should be given. Having the person copy a simple geometric figure or drawing a clock and putting in the time are effective screening techniques for visuospatial deficits. The best way to elicit symptoms of psychosis is to inquire about any unusual occurrences, unusual perceptions, the hearing or seeing of things that seem strange or unreal, or of possible fears that others are trying to hurt them.

History
Prior psychiatric illness
Prior substance abuse
Educational level
Family support
Mental status
Appearance and behaviour
Mood and affect
Speech
Thought content and process: rationality, delusions, hallucinations
Attentiveness (digit span, serial 7 subtraction, days of the week)
Orientation (time, person, place)
Memory (remote: political, historical, personal dates, etc.); new: remember four random words after 5 min)
Judgement
Intellect (abstraction, comprehension)

Table 2 The bedside psychiatric examination

Patients who show clear difficulties with such 'bedside' screening tests are most probably experiencing some organic brain disturbance. Others who do well on such tests yet still report prominent subjective difficulties may be suffering from more subtle organic disturbances. Both should have more detailed neuropsychological tests from experienced neuropsychologists.

The psychiatric disturbances may be due to failure of other organs. For example, uraemic encephalopathy may ensue, and renal failure or vasculitis with secondary severe hypertension can also result in hypertensive encephalopathy with the development of headaches, disorientation, coma, and seizures. Occult central nervous infections, especially in the immunocompromised patient, can lead to an organic mental syndrome. As mentioned later, high doses of corticosteroids, the mainstay of therapy for systemic lupus, can directly cause psychiatric disturbances.

Establishing the diagnosis of neuropsychiatric systemic lupus erythematosus

This remains a difficult task, even for the most astute clinician. The diagnosis should be considered in a young woman presenting with psychotic features, even if no other clinical or laboratory features of systemic lupus are present, as these psychiatric signs and symptoms may well be the initial presentation. In a study of 296 patients admitted to an acute psychiatric ward in Nottingham, United Kingdom, approx. 1 per cent of the patients met criteria for the diagnosis of systemic lupus (Hopkinson and Powell 1990). There is no consensus on the association between systemic and psychiatric lupus; most series have failed to document any relation

between the two.

A number of immunological tests have been proposed as useful in the evaluation of neuropsychiatric systemic lupus (Table 3). In a series of 70 female patients (Hanly *et al*. 1993), antineuronal antibodies were detected in 34 per cent, lymphocytotoxic antibodies in 47 per cent, anti-P antibodies in 17 per cent, and anticardiolipin antibodies in 24 per cent. However, there was no significant difference in the prevalence of any of these antibodies in patients with and without cognitive impairment. On the other hand, work by Denburg and colleagues (Denburg *et al*. 1994) has suggested a relation between specific cognitive deficits, namely visuospatial dysfunction, and the presence of lymphocytotoxic antibodies.

Anticardiolipin antibodies can be found in patients with systemic lupus erythematosus or with the primary antiphospholipid syndrome (Love and Santoro 1990). Cerebrovascular thrombotic events, particularly stroke, can develop. An organic mental syndrome may arise, although isolated psychiatric disease without overt neurological signs is probably rare.

Concentrations of interferon-a were increased in the cerebrospinal fluid in 5 of 6 patients with lupus psychosis (Shiozawa *et al*. 1992); these decreased when the psychosis subsided and they did not rise following seizures alone. These limited findings suggest that interferon-a, possibly synthesized in the brain, may play some part in the pathogenesis of psychosis in systemic lupus.

One study measured a variety of cytokines, prostaglandins, and autoantibodies in the cerebrospinal fluid of patients with systemic lupus and infection (Tsai *et al*. 1994). The results suggested that high levels of interleukin (**IL**)-6 and prostaglandin E_2 in the cerebrospinal fluid favoured a diagnosis of infection of the central nervous system while modestly elevated IL-6, high IgG, and autoantibodies against calf thymus antigens in the cerebrospinal fluid suggested a diagnosis of neuropsychiatric systemic lupus.

The autopsy findings in lupus of the central nervous system have recently been reviewed. Ellison *et al*. (1993) described small-vessel hyalinization and thickening, with fragments of platelet membrane found in the wails of small cortical and meningeal vessels. Concurrent thrombus formation, possibly related to antiphospholipid antibody formation, may facilitate the incorporation of platelet fragments into small-vessel walls, with resultant thickening and irregularity of the vessels. In another series of seven patients with confirmed neuropsychiatric systemic lupus (Hanly *et al*. 1992b), four were found at autopsy to have multifocal cerebral cortical microinfarcts associated with microvascular injury. These studies suggest a vascular basis for neuropsychiatric systemic lupus in many, but not all patients.

Electroencephalography

The electroencephalogram may help to confirm the presence of an abnormal seizure focus that manifests as episodic changes in behaviour and mentation (e.g., temporal lobe epilepsy). It may help to distinguish delirium from a primary mental disorder with cognitive dysfunction such as acute mania. Clinically, delirium is characterized by diminished attention and sometimes diminished level of consciousness, associated with disorientation, agitation, reversal of the sleep-wakefulness cycle, perceptual distortions and misinterpretations, and impairment of short-term memory. Most patients with delirium have diffuse slowing of the background electroencephalogram, unlike in most primary mental disorders (McNamara 1991). Ritchlin *et al*. (1992) found that using neurometric quantitative electroencephalography as an indicator of cerebral dysfunction in patients with systemic lupus produced a diagnostic sensitivity of 87 per cent and specificity of 75 per cent.

Brain imaging

Advances in technology have resulted in four imaging techniques that may be helpful in evaluating neuropsychiatric systemic lupus; these four are outlined next.

Computed tomography (CT)

This technique initially described the presence of steroid-induced cortical atrophy (Bentson 1978) and calcification of the paraventricular areas (Daud and Norudin 1988) seen in systemic lupus. For the most part, CT has been supplanted by MRI. However, CT scans continue to serve an important diagnostic role in identifying emergent haemorrhage or cerebral infarction. They are also considerably less expensive than MRI.

Magnetic resonance imaging (MRI)

MRI is far more sensitive than CT for detecting disease of the central nervous system, and can demonstrate the reversal of some of the brain lesions in lupus (McCune *et al*. 1988). Three different patterns of disease can be demonstrated by MRI: cerebral infarction, multiple small areas of increased signal intensity secondary to microinfarctions, and focal areas of increased intensity in the cerebral grey matter (Bell *et al*. 1991).

A recent study (Baum *et al*. 1993) highlights the possibilities and difficulties in interpreting the MRI in neuropsychiatric systemic lupus. Twenty-one consecutive outpatients with systemic lupus were studied: 12 had focal lesions, primarily in the frontal lobes. Patients with focal neurological signs were more likely to have lesions on MRI. Ten patients had paraventricular hyperintensities, a finding noted in several other series as well. Yet seven patients with abnormal MRI had no neuropsychiatric symptoms, and eleven had neuropsychiatric symptoms and signs that did not correlate with MRI findings.

To summarize, MRI may fail to detect abnormalities in diffuse, non-focal, or mild lupus of the central nervous system. It is sensitive to focal lesions that are large enough to produce clear-cut localizing signs on neurological or cognitive examination.

Single photon-emission CT (SPECT)

SPECT appears to have limited value in the evaluation of neuropsychiatric systemic lupus. Rogers *et al*. (1992) noted abnormal findings in 8/18 patients with subtle cognitive and affective changes, including a diffuse bilateral temporal-parietal pattern previously noted only in Alzheimer's disease. Another study (Rubbert *et al*. 1993) found normal SPECT scans in 90 per cent of those with overt neurological signs. SPECT may have a useful role in patients with diffuse symptoms, such as cognitive dysfunction and psychosis, in which MRI may be insensitive.

Positron-emission tomography (PET)

PET assesses cerebral blood flow and glucose metabolism. Carbotte *et al*. (1992) described an intensive longitudinal study of three women with neuropsychiatric systemic lupus. Fluorodeoxyglucose uptake indicated abnormalities in all three that had not been identified on plain CT, yet corresponded well with localizable cognitive deficits. Changes in each patient's cognitive profile on reassessment paralleled changes on PET. The high cost and limited availability of the necessary equipment has precluded more widespread use of PET.

Treatment of the psychiatric symptoms

The usual treatments, ranging from antidepressants to antipsychotic medications and psychotherapy, have an important role. However, one must treat the underlying biological cause, the systemic lupus, as well. If the patient develops a frank psychosis on corticosteroids, the first task would be to define the aetiology of the change in mental status; is this drug-induced or a manifestation of the lupus or, less commonly, due to an occult infection of the central nervous system in an immunocompromised host? Generally, doses of prednisone in the range of 1 to 1.5 mg/kg per day have been given orally or systemically with variable results. Although steroid-induced psychosis can occur, the neuropsychiatric manifestations in patients taking corticosteroids are much more likely to be due to the disease than to the drug.

Patients who fail to respond to corticosteroids have been treated with cyclophosphamide: one study from the National Institutes of Health suggests the use of monthly pulse administrations of 0.75 to 1.0 g/m^2 of cyclophosphamide (Boumpas *et al*. 1991). The sample in most of these studies of treatment is too small for any meaningful extrapolation. In a retrospective study of 31 patients (Neuwelt *et al*. 1995) with severe neuropsychiatric systemic lupus refractory to corticosteroids and other oral immunosuppressive agents, intravenous cyclophosphamide was associated with substantial improvement in 18 patients, and stabilization of the condition in nine others. Patients with an organic brain syndrome or a large number of neuropsychiatric manifestations had a poorer outcome. The treatment of anticardiolipin-related thrombotic events remains controversial; recent studies suggest that warfarin appears to be superior to low dose aspirin in the management of this condition

(Khamashta *et al*. 1995)

Sjögren's syndrome

One study found a high incidence of psychopathology in a group of patients with primary Sjögren's syndrome, in particular, high levels of hostility and paranoid ideation (Angelopoulos *et al*. 1988). Other studies have identified an increased incidence of affective disorders, cognitive impairment (in attention and concentration), and a correlation between neurological and psychiatric symptoms suggestive of an organic mental disorder (Malinow *et al*. 1985). An increased incidence of neuropsychiatric symptoms may also accompany secondary Sjögren's syndrome.

Rheumatoid arthritis

The only primary neuropsychiatric syndrome in rheumatoid arthritis, described in two cases, is an organic mental disorder caused by rheumatoid nodules in the choroid plexus (Kim *et al*. 1982). Rarely, rheumatoid cranial vasculitis also appears as an organic mental disorder.

The vasculitides

The vasculitides associated with neuropsychiatric factors are listed in Table 1. Granulomatous angiitis of the central nervous system is a rare condition that presents with non-specific signs and symptoms including acute or subacute onset of confusion, headache, personality change, paresis, cranial neuropathy, or loss of consciousness (Sigal 1987). An elevated erythrocyte sedimentation rate may be the only abnormal laboratory finding. Cerebral angiography may reveal changes including vascular beading or aneurysms, but a definite diagnosis can only be made by brain or leptomeningeal biopsy. Pathological changes include an inflammatory process involving small arteries, with intimal proliferation, fibrosis, and multinucleated giant cells (Sigal 1987). A few cases have followed herpes zoster ophthalmicus.

Systemic vasculitides such as polyarteritis nodosa or Wegener's granulomatosis can involve the central nervous system. Psychiatric disturbances are rarely seen without neurological signs or symptoms. Depression and changes in mental status are not uncommon presentations for giant-cell arteritis. Helpful clinical clues for establishing this diagnosis include the presence of jaw claudication, new onset of headaches, visual changes, and proximal myalgias. An elevated erythrocyte sedimentation rate is generally found. Biopsy of the temporal artery is recommended for the patient with a new onset of depression and any of the aforementioned features. Improvement is noted after treatment with corticosteroids.

Behçet's syndrome

The central nervous manifestations of Behçet's syndrome consist of subacute, haemorrhagic, and necrotizing meningoencephalitis, most typically affecting the brainstem and hypothalamus. In addition to the focal neurological symptoms associated with these lesions, organic mental disorders, disorders of consciousness, and seizures have all been described (Bousser *et al*. 1988).

Whipple's disease

Behavioural changes, hypersomnia, memory disturbance, and dementia, which may antedate its usual gastrointestinal or joint manifestations, characterize some of the central nervous manifestations of Whipple's disease (Fleming *et al*. 1988). One case report described dementia partially reversed by antibiotic therapy (Tarter *et al*. 1990).

Sarcoidosis

Involvement of the central nervous system occurs in 5 per cent of patients with sarcoidosis (Stoudemire *et al*. 1983). Delirium, dementia, personality change, hypersomnolence, depression, and psychosis have all been described. Steroids have reversed at least some of these neuropsychiatric manifestations.

Lyme disease

Lyme disease, caused by the tick-borne spirochaete *Borrelia burgdorferi*, is associated with a wide variety of neurological and some psychiatric manifestations. Logigian *et al*. (1990) studied 27 patients with signs of earlier Lyme disease, current evidence of immunity to *B. burgdorferi*, and chronic neurological symptoms with no other identifiable cause. Twenty-four of them had a mild encephalopathy, which had begun 1 month to 14 years after the onset of the disease and was characterized by memory loss, mood changes, or sleep disturbance. Fourteen had memory impairment on neuropsychological testing and 18 had increased protein in cerebrospinal fluid, evidence of intrathecal production of antibody to *B. burgdorferi*, or both. Associated symptoms included fatigue, headaches, arthritis, and hearing loss. After a 2-week course of intravenous ceftriaxone, two-thirds of the patients had improved, one other had improved and then relapsed, and the remainder had no change in their condition.

A long-term follow-up (mean of 6 years) of a cohort of patients revealed significantly more verbal memory deficits in patients with higher IgG Lyme antibody titres who received treatment later (mean of 3 years) (Shadick 992; Shadick *et al*. 1994**)**. Similarly, Krupp *et al*. (1991) noted more memory impairment on formal testing of Lyme patients than healthy controls.

Sigal (1995) has described the medical challenges associated with 'pseudo-Lyme' disease. Subjective complaints such as chronic fatigue, headaches, or memory deficits may be incorrectly attributed to Lyme disease. It is of utmost importance that the diagnosis of Lyme disease be made using accepted serological techniques. The entity of 'seronegative' neuropsychiatric Lyme disease is probably quite rare.

A recent review of experience with Lyme disease in children (Adams *et al*. 1994) found no differences between disease and control groups for a number of neuropsychological measures. These data contrast with a number of case reports in the literature suggesting the existence of Lyme-related psychiatric disease, including paranoia, dementia, schizophrenia, bipolar disorder, panic attacks, major depression, anorexia nervosa, and obsessive–compulsive disorder (Fallon and Nields 1994).

Amyloidosis

Some of the types of primary amyloidosis may involve the blood vessels of the central nervous system. Cerebral amyloid angiopathy can produce a variety of organic mental disorders including dementia, a fact of increasing interest because of the isolation of b-amyloid protein in both plaques and skin in patients with Alzheimer's disease. It is unclear whether secondary amyloidosis may also be associated with brain pathology.

Rheumatological disease with secondary psychiatric involvement (Table 4)

Depression in rheumatoid arthritis

Estimates of the incidence of depression in rheumatoid arthritis have varied widely, from 30 per cent (more typically) to as high as 70 per cent. The variability has resulted from differences in the definition of depression as well as differences among the populations studied. Some studies have found that the prevalence of depression is no higher in rheumatoid arthritis than in other chronically ill medical patients (Wolfe and Howley 1993).

Frank *et al*. (1988a) made one of the few studies using current operational criteria for depression in the context of rheumatoid arthritis. By using a structured diagnostic interview of predominantly male patients, they found a prevalence of major depressive disorder of 17 per cent. A much larger number, 41 per cent, met the criteria for dysthymic disorder, a less severe but more chronic form of depression in which symptoms must be present for at least 2 years and for more than half of the time. These diagnostic categories are not mutually exclusive: almost all of the patients with major depressive disorder met criteria for dysthymic disorder, so-called double depression. Overall, 41 per cent of this sample of 137 outpatients with rheumatoid arthritis were found to be depressed, with either of the two major depressive disorders or both. No correlation was found between the presence of a history of depression and higher levels of reported pain, as has been found in other studies.

Contrary to intuition, most studies have not found a significant correlation between depression and measures of disease activity, but rather with socioeconomic

factors, such as decreased social support and economic deprivation. One of the most comprehensive investigations of the origins of depression in rheumatoid arthritis found that a wide range of factors, including disability measures, duration of disease, social isolation, and economic deprivation, all made significant contributions to the explanation of depressed mood, together accounting for 44 per cent of the variation in depressed mood. Others have confirmed the complexity of determinants for depression in rheumatoid arthritis (Blalock and DeVellis 1992), indicating that loss of valued activities (Katz and Velin 1995) and social relationships (Crotty *et al*. 1994) may be as important as disease activity or functional status.

While most studies have found an increased incidence of depression in patients with rheumatoid arthritis, a few reports have not borne this out (Cassileth *et al*. 1984; Wolfe and Howley 1993). Cassileth measured, on psychological rating scales, the psychological health of patients with a variety of chronic medical disease including rheumatoid arthritis. They showed that these persons were overall indistinguishable from the general population and significantly healthier mentally than a population of psychiatric outpatients (Cassileth *et al*. 1984). However, they did not use the standard, operationalized definitions of clinical depression.

Systemic lupus erythematosus

Because of the potential life-threatening nature of systemic lupus and the alteration in appearance, both from the underlying disease and from use of corticosteroids, the psychological toll can be enormous (Omdal *et al*. 1995). Given the younger age of onset and the female predominance, issues related to family planning and pregnancy are of foremost concern to both patient and physician. Most observers have found an increase in clinical depression and anxiety (Liang *et al*. 1984). Providing patients with information relating to their disease and its management may increase their knowledge base but does not affect their psychological response (Konttinen *et al*. 1991). Peer support groups for systemic lupus may have a beneficial role to play (Peterson *et al*. 1993).

Juvenile chronic arthritis

Measurement of pain behaviours may be especially useful in studies of treatment outcome because these behaviours are relatively independent of depression (Jaworski *et al*. 1995). One recent survey found no correlation between scores on psychological testing and measures of functional status (Baildam *et al*. 1995), whereas another found that the proportion of patients demonstrating an anxious preoccupation with their disease increased with the degree of disability (David *et al*. 1994).

Miscellaneous disorders including systemic sclerosis, inflammatory myositis, spondylitis, severe osteoarthritis

The main connecting link in these conditions is the impairment in function due to joint and muscle problems and changes in appearance. These changes produce sadness, grief reactions, social isolation, and economic hardships. Some patients emphasize the positive, character-building aspects of these diseases and their potential for strengthening marital and family relationships.

Drug-induced psychiatric manifestations (Table 5)

Medication	Psychiatric manifestations
Corticosteroids	Mania, depression, organic mental disorders, insomnia
Hydroxychloroquine	Psychosis, organic mental disorders, personality change, insomnia
Cyclosporin	Mania, depression, lethargy, organic mental disorders
Non-steroidal anti-inflammatory drugs	Depression, organic mental disorders
Methotrexate Azathioprine Gold salts D-Penicillamine	Mood changes, headaches
Monoclonal antibodies (e.g., anti-TNF, anti-IL-2)	Fatigue, headaches

IL, interleukin; TNF, tumour necrosis factor.

Table 5 Drug-induced psychiatric manifestations

Corticosteroids

Major psychiatric side-effects have been well described (Table 6), occurring in approx. 6 per cent of all patients on steroids (Kershner and Wang-Cheng 1989). Female sex appears to be an important risk factor (Boston Collaborative Drug Surveillance Program 1972). Starting treatment with corticosteroids often results in euphoria. This may be due to a direct central nervous effect of the drug, or may relate to the psychological sense of well being from the beneficial effects of treatment on the disease process. Occasionally, steroids can induce a depressive state, especially with doses exceeding 0.5 mg/day or when rapidly tapered off after prolonged use. Wolkowitz *et al*. (1990) found that healthy volunteers given either a single, 1-mg dose of dexamethasone or 80 mg/day of prednisone for 5 days made significantly more errors of commission in verbal memory tasks, suggesting the possibility of specific, corticosteroid-related, cognitive impairments. Psychiatric symptoms are generally reversible when the drug is reduced or stopped. One prospective study found a good response to prophylactic treatment with lithium carbonate (Falk *et al*. 1979).

Depression	40
Mania	27
Psychosis	14
Delirium	10
Mixed depression and mania	8

Table 6 Steroid-induced psychiatric manifestations (percentages)

Cyclosporin

Cyclosporin produces its anti-inflammatory activity by binding to the interleukin-2 receptor on T cells. The observed psychiatric side-effects have included mania (Wamboldt *et al*. 1984), depression, and lethargy (Adams *et al*. 1987). With the higher doses (5-15 mg/day) given after organ transplantation, irreversible dementia, encephalopathy, and fatal progressive neurological deterioration have been seen (Bertoli *et al*. 1988). Cyclosporin can attenuate the opiate withdrawal syndrome precipitated by naloxone in morphine-dependent animals and this effect can be adoptively transferred by splenic mononuclear cells (Dougherty and Dafny 1987). The drug can also alter the electrophysiological properties of discrete brain nuclei (Dougherty *et al*. 1987).

Antimalarials

Chloroquine and to a lesser extent hydroxychloroquine have associated toxic psychiatric reactions including personality change, depression, depersonalization,

neurotic symptoms, confusional states, psychosis, and suicide (Mohan *et al*. 1981; Good and Shader 1982). Chloroquine is concentrated in brain and spinal cord. The psychiatric effects may be related to acetylcholinesterase inhibitory activity (Mustakillio *et al*. 1962).

Others

Gold salts, D-penicillamine, and azathioprine are not associated with any psychiatric findings (Rogers 1985). Methotrexate, given orally or parenterally in low weekly doses (i.e. 20 mg or less) for rheumatic diseases can be associated with headache, light-headedness, and mood swings.

Non-steroidal anti-inflammatory drugs

These are a class of agents distinguished by their ability to inhibit cyclo-oxygenase and thereby impair prostaglandin biosynthesis. Prostaglandins are synthesized in the brain and can modulate the effect of various neurotransmitters. In addition, these non-steroidal agents can cross the blood–brain barrier, which may explain their analgesic and antipyretic effects. These might be mediated by a direct effect on hypothalamic structures.

Indomethacin is the most potent inhibitor of cyclo-oxygenase among the non-steroidal anti-inflammatory drugs: it caused central nervous disturbances, including headache, light-headedness, dizziness, and vertigo, in almost one-half of the patients studied. These manifestations of toxicity can also occur, to a lesser extent, with virtually any of these drugs. More profound changes, including depression, confusional states, visual hallucinations and suicides, are reported with indomethacin (Tollefson and Garvey 1982). Interestingly, indomethacin contains an indole moiety similar to that in serotonin, which may explain its frequent psychiatric manifestations. Sulindac, structurally related to indomethacin, has been linked, in case reports, to delirium (Thornton 1980) and paranoid psychosis (Kruis and Barger 1980). Tolmetin reportedly induced mania in a patient with a prior history of psychiatric disturbances (Sotsky and Tossell 1984). We speculate that its pyrrole moiety, also found in porphobilinogen, which is excreted in acute porphyria, a condition associated with intermittent psychosis, may be responsible.

Propionic acids such as ibuprofen, naproxen, and fenoprofen have been associated with depression and cognitive dysfunction, especially in elderly people (Goodwin and Regan 1982). Insomnia and nightmares have been described with several of the non-steroidal, anti-inflammatory drugs.

Salicylates and salsalate in the therapeutic range have been associated with visual hallucinations (Greer *et al*. 1965), and confusion, especially in elderly individuals (Anderson *et al*. 1976), in whom these can be misinterpreted as signs of Alzheimer's disease (Vivian and Goldberg 1982). Tinnitus and impaired hearing have also been encountered, even with subtherapeutic blood concentrations of salicylates.

Treatment

When should a patient with rheumatological disease be referred to a psychiatrist or other mental health professional? When functional impairment seems out of proportion to the underlying medical disorders, when there are cognitive changes, hallucinations, delusions, bizarre behaviour, confusional states, or when there are questions about the use of psychotropic medications as in patients with depression or significant anxiety states, referral to a psychiatrist should be considered. It becomes more urgent when the possibility of suicide or psychosis arises.

Support groups should be considered for patients who are more isolated as a result of their disorders and seeking more information than can reasonably be provided by the rheumatologist, or more social support than that currently available.

Managing depression

The word depression, to the confusion of some, has been used to describe conditions ranging from grief and adjustment reactions to the more serious and potentially life-threatening major depressive disorder (unipolar depression). In general, antidepressants are indicated for treatment of the more severe depressions.

By current consensus as defined in the Diagnostic and Statistical Manual of Mental Disorders IV (1995) , a major depressive syndrome consists of at least a 2-week period during which at least five of the following nine symptoms will have consistently occurred, representing a change from a previous level of functioning.

1. depressed mood;
2. markedly diminished interest or pleasure;
3. ignificant weight loss or gain;
4. insomnia or hypersomnia;
5. agitation or psychomotor retardation;
6. fatigue;
7. feelings of worthlessness or excessive guilt;
8. diminished ability to think or concentrate;
9. suicidal ideas.

A further stipulation is that one of the five symptoms is either (1) depressed mood, or (2) loss of interest or pleasure.

A contemporary epidemiological study has found that the lifetime prevalence of major depressive disorder in the normal population is 4.4 per cent. If treated with antidepressant medications, approx. 60 to 70 per cent of patients will respond with resolution of the episodes within a few weeks. If untreated, it is estimated that a typical episode of major depression will last from 6 to 12 months and then usually resolve spontaneously. However, depression may persist and present as a chronic depressive disorder in 20 per cent of individuals.

Accurate diagnosis of depression in medically ill patients is confounded by the overlap between symptoms of the medical disorder and symptoms of depression, particularly the neurovegetative symptoms. For example, fatigue, insomnia, and anorexia may all be directly related to medical conditions such as rheumatoid arthritis. For this reason, some investigators have suggested substituting the more cognitive/affective symptoms, such as guilt and low self-esteem, for the neurovegetative ones when diagnosing depression in patients with medical illness.

Psychological interventions and the course of rheumatic diseases

The belief that one can manage and control a specific stressful situation (so-called self-efficacy) is probably an effective antidote to the helplessness so often found at the core of depression in chronically ill patients. In a longitudinal study, Parker *et al*. (1988) demonstrated that increases in self-efficacy could be taught to patients with rheumatoid arthritis, in turn leading to improvements in pain, and in psychological and health status. Using a self-management course for patients with arthritis that emphasizes cognitive, behavioural, and stress management skills, together with mutual support groups, Strauss *et al*. (1986) demonstrated improvements in health status and specifically in measures of arthritis activity.

Muller *et al*. (1987) made a meta-analysis of 15 studies on the effects of psychoeducational interventions on disability, pain, and depression in individuals with rheumatoid arthritis or osteoarthritis. The results indicate that patient education and group psychotherapeutic support can improve the health status of patients with chronic arthritis.

The participation of patients with lupus in a specifically designed self-help course correlated with subsequent lower levels of depression, and greater enabling skills and use of relaxation and exercise (Braden *et al*. 1993).

Psychotropic medications

Mood stabilizers

Lithium is the mainstay of treatment for manic depressive or bipolar affective disorder. It is helpful as an adjunct in treatment-resistant depression and as a prophylaxis against corticosteroid-induced mania. Alternative drugs recently introduced for treatment of acute mania, such as carbamazepine and valproic acid, may

also prove useful in this disorder (Small 1990).

Electroconvulsant therapy

This remains an effective treatment for major depression but is indicated only if the antidepressants are unsuccessful or not tolerated, or if there is an acute risk of suicide that necessitates immediate treatment.

Antidepressants

The four major categories of antidepressants include the serotonin-specific reuptake inhibitors (**SSRI**s), tricyclics, monamine oxidase inhibitors, and others. In general, the therapeutic response rate for all of the antidepressants is similar—approx. 70 per cent in major depression.

SSRIs

Over the past decade or so a number of new antidepressant agents have been introduced. The most important and widely used of these groups are the SSRIs. They are generally now the first-line agents used in the treatment of depression. Fluoxetine (Prozac) was introduced first, in 1987. It has received the most attention and controversy. Two others in the same class, sertraline (Zoloft) and paroxetine (Paxil), quickly followed. More recently, fluvoxamine (Luvox) has been added to the list (Wilde *et al*. 1993). The SSRIs have been so widely used and accepted because of their relatively limited side-effects. They have a highly specific action on serotonin reuptake and, as such, lack the anticholinergic side-effects, sedation, orthostatic hypotension, and cardiotoxicity that can be problematical with other antidepressants. Because of their potentially stimulating properties, they are the only antidepressants given in a single dose in the morning (Table 7). Nausea, jitteriness, insomnia, and headache are among the more common side-effects. Approximately 30 per cent of men will experience sexual dysfunction in the form of delayed ejaculation. All of the SSRIs have long half-lives, with fluoxetine the longest because of the activity of its metabolite. For this reason, medically ill patients should probably be treated with a non-fluoxetine SSRI.

Agent	Dosage range (mg)	Distinction
Fluoxetine (Prozac)	10–40	Longest half-life
Sertraline (Zoloft)	50–100	
Paroxetine (Paxil)	20–40	
Fluvoxamine (Luvox)	20–40	

Table 7 Serotonin-specific reuptake inhibitors

As with virtually all of the other antidepressants, there is an approximately 2- to 6-week lag time at therapeutic dosage before the antidepressants begin to exert their maximum effect on depression.

Tricyclics

Tricyclic antidepressants have been used in clinical practice for over 30 years. There is no evidence that one is more effective than another. Table 8 lists the tricyclic antidepressants and their usual ranges of dosage. The tertiary amines imipramine, amitriptyline, and doxepin are converted to their demethylated secondary amine metabolites desipramine, nortriptyline, and protriptyline, respectively, which retain pharmacological effectiveness and are used clinically as separate agents.

Tertiary amines
Imipramine (100–200 mg)
Amitriptyline (100–200 mg/day)
Doxepin (100–200 mg/day)
Trimipramine (100–200 mg/day)
Clomipramine (100–225 mg/day)
Demethylated secondary amines
Desipramine
Nortriptyline (75–150 mg/day)
Protriptyline (15–40 mg/day)

Table 8 Tricyclic antidepressants and usual dosage ranges

The tricyclics, and for that matter all of the antidepressants, are thought to exert their antidepressant effects through modulation of the availability of biogenic amines at receptor sites in the brain, particularly serotonin and noradrenaline, and through their effect on the responsiveness of adrenergic and serotonergic receptors. The side-effects of these medications also arise from receptor blockade. For example, sedation stems from antihistaminic effects, postural hypotension from a-adrenergic receptor blockade, and blurred vision, dry mouth and constipation from the anticholinergic effects. The demethylated secondary amines, desipramine and nortriptyline, have been increasingly used, especially in elderly patients, because of their relatively low anticholinergic side-effects as well as less sedation and orthostatic hypotension.

A clinical advantage of the tricyclics is the ready availability of methods for measuring their plasma concentration, which can be used to determine compliance and to titrate the dosage for maximum therapeutic effect.

It is useful to start at a low dosage, generally between 10 and 25 mg, about an hour before bedtime, and then to increase gradually (approximately every 3 days) by the same amount, working up to a total dosage range of 50 mg/day. The 10-mg dosage increment is usually prudent in geriatric practice, while an increment of 25 to 50 mg is usually tolerated in the patient under 60 years of age. Once having gradually worked up to 50 mg it is usually easier to increase the dosage the rest of the way up to 150 mg/day or higher.

Patients respond to tricyclics at different dosages. This is probably due to individual biological variability and the rate at which the individual metabolizes the antidepressant. If a lower dosage does not result in remission or results in subtherapeutic blood concentrations, the dosage may need to be increased to 200 to 300 mg. Many patients with chronic rheumatic diseases may benefit from lower doses from the sedative and analgesic effects of the tricyclics.

Other antidepressants

Bupropion (Wellbutrin), a structurally unique compound, has minimal anticholinergic and antihistaminic properties. Its major contraindication has been in patients with seizure disorder. On the plus side, unlike the SSRIs, it does not disrupt ejaculation and other sexual function. Maprotiline (Ludiomil) and trazodone (Desyrel) have

side-effects similar to the tricyclic antidepressants. Venlafaxine (Effexor) is a relatively new addition that has been promoted on the basis of its combined serotonin and norepinephrine reuptake blockade (Feighner 1994). This can be associated with sustained increases in blood pressure in approx. 5 per cent of cases when dosages were maintained above 200 mg/day.

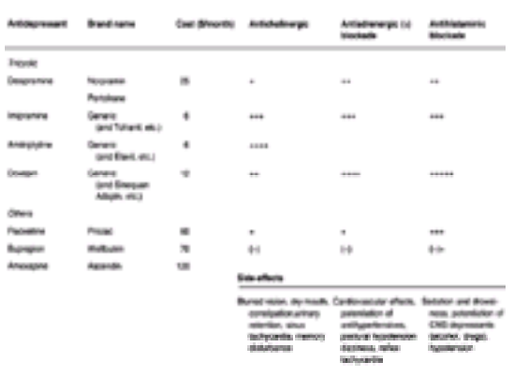
Another antidepressant recently approved by the United States Food and Drug Administration is nefazodone (Serzone) (Anton *et al*. 1994). Like the SSRIs and venlafaxine, the more common side-effects are nausea, lack of appetite, insomnia, nervousness, and sexual dysfunction. These do not have the prominent anticholinergic side-effects and hypotension so characteristic of the tricyclics. They also have the advantage of being non-lethal in overdoses, in contrast to the tricyclics.

Monoamine oxidase inhibitors

The monoamine oxidase inhibitors (imipramine, desipramine tranylcypromine, phenelzine) were the first effective antidepressant drugs but because of tyramine-induced hypertensive episodes their use was limited. These drugs appear to have a special role in the treatment of atypical depressions characterized by hypersomnia, hyperphagia, and a reversed diurnal variation with symptoms typically worse at night than in the morning. They can often be effective in tricyclic-resistant patients and for the treatment of panic disorders. Concomitant ingestion of tyramine-containing foods, sympathomimetic agents, and meperidine (Demerol) must be avoided.

Treatment strategies

The choice of an antidepressant is often determined by its side-effects. Currently SSRIs are likely to be the first drug of choice because of their relatively benign side-effects. Table 9 compares the most commonly used antidepressants in relation to a variety of factors: cost, receptor blockade, and side-effects.



Antidepressant	Brand name	Cost ($/month)	Anticholinergic	Antiadrenergic (α) blockade	Antihistaminic blockade
Tricyclic					
Desipramine	Norpramin	25	+	++	++
	Pertofrane				
Imipramine	Generic (and Tofranil, etc.)	6	+++	+++	+++
Amitriptyline	Generic (and Elavil, etc.)	6	++++		
Doxepin	Generic (and Sinequan Adapin, etc.)	12	++	++++	+++++
Others					
Paroxetine	Paxil	60	+	+	+++
Bupropion	Wellbutrin	70	(+)	(+)	(+)
Amoxapine	Asendin	120			

Side-effects		
Blurred vision, dry mouth, constipation, urinary retention, sinus tachycardia, memory disturbances	Cardiovascular effects, potentiation of antihypertensives, postural hypotension, dizziness, reflex tachycardia	Sedation and drowsiness, potentiation of CNS depressants (alcohol, drugs), hypotension

Table 9 Cost, blockade, and side-effects of commonly used antidepressants

The most frequent reason for failure to respond to antidepressant treatment is inadequate dosage, so monitoring of the plasma concentration is an important step in guiding the clinician (Table 10).

Drug	Therapeutic plasma concentration (mg)
Amitriptyline	150–250*
Doxepin	100–250*
Imipramine	200–300*
Amoxapine	150–500
Desipramine	150–300
Maprotiline	100–300
Nortriptyline	50–150
Protriptyline	100–250
Bupropion	50–100
Fluoxetine	100–300
Sertraline	100–300
Paroxetine	100–300
Trazodone	800–1600

*Drug and metabolite.

Table 10 Antidepressant plasma concentrations

After a satisfactory response has been obtained, the patient is generally continued on the medication for another 9 months before the dose is gradually tapered off. It is now clear that patients who have had a history of two or more depressions in the past should be kept on a maintenance dosage equal to the initial therapeutic dosage (Kupfer 1991).

How does one proceed with the 10 to 40 per cent of patients who fail an initial trial of the medication? The main causes of a poor response are inadequate dosage, non-compliance, often associated with inability to tolerate side-effects, and inadequate duration of treatment. There are no precise guidelines in determining how to proceed after an initial full trial of an SSRI or tricyclic fails. It is important to allow an adequate trial of 4 to 6 weeks. Most clinicians choose another agent with a different pattern of side-effects or biochemical profile, for example in a non-responder to an inhibitor of serotonin reuptake. If the second trial of an antidepressant fails, then the usual choice is between a newer agent, a combination of agents such as an SSRI boosted by a lower dosage of a tricyclic, or with a booster effect from lithium, or a monoamine oxidase inhibitor. For unclear reasons, individual patients may respond to one but not another of these agents. The addition of lithium is an adjunct may be effective in a subgroup of patients, especially those with a family history of bipolar illness.

Analgesic effects of antidepressants

Since their introduction over 30 years ago, tricyclics and, to a lesser extent, monoamine oxide inhibitors have found a role in the treatment of chronic pain syndromes, particularly for neuropathic pain, but also in the treatment of headaches, arthritis, low back, and facial pain. Approximately 50 to 60 per cent of patients with a chronic pain syndrome will benefit from the adjunctive analgesic effects of the antidepressants (Blackwell 1987). The analgesic action of the antidepressants cannot be entirely explained as a placebo effect, nor can it be indirectly related to the drug's effect on depression. In a review of 28 placebo-controlled studies, Magni (1991) concludes that antidepressants are helpful in a wide range of chronic pain syndromes, including fibromyalgia and rheumatoid arthritis, in addition to neuropathic pain, headache, migraine, and facial pain. Several studies of the use of antidepressants in chronic pain have shown pain relief in the absence of measurable changes in depression. Furthermore, structural congeners of imipramine such as carbamazepine and phenothiazine, which are not antidepressants, also have analgesic properties. Animal studies using antidepressants have also demonstrated some reduction in pain behaviour (Butler *et al*. 1985). Although the mechanisms of analgesia of the antidepressants remain unknown, their regulatory effects on serotonin and other biogenic amines, and their potential effect on opiate receptors, provide a theoretical basis.

Most of the antidepressants appear to have these analgesic effects. One must weigh the side-effects of the particular drug against the clinical setting. A patient with chronic pain and insomnia, for example, could benefit from a sedating antidepressant such as amitriptyline or trimipramine. These are generally described once daily, an hour or two before bedtime, and generally at a dose that is half of that used for the treatment of depression. It is wise to start with a low dose, either 10 or 25 mg, and gradually increase, similar to the approach in depression.

Antidepressants in rheumatoid arthritis

In the early 1960s both anecdotal reports and uncontrolled studies suggested that antidepressants might have a role in providing pain relief to patients with rheumatoid arthritis. In the first controlled study of antidepressants in chronic pain, McDonald Scott (1969) demonstrated a significant improvement in rheumatoid arthritis with the use of imipramine. At a dosage of 75 mg daily, imipramine was compared to placebo in a double-blind, cross-over study of 22 patients with arthritis. Patients with a history of depression or other psychiatric disturbance were excluded. A significantly higher proportion of patients receiving imipramine improved both in subjective ratings of pain and in grip strength, and decreasing morning stiffness in 55 per cent of patients with various chronic arthritides.

Subsequent studies in patients with arthritis have suggested that not all antidepressants have similar analgesic properties. For example, Ganvir *et al*. (1980) failed to demonstrate a significant analgesic effect with 25 mg of clomipramine daily as compared to placebo, although they subsequently questioned whether the dosage was adequate. More recently, Frank *et al*. (1988b) compared amitriptyline, desipramine, trazodone, and placebo in a 32-week, double-blind, cross-over trial in 47 patient with rheumatoid arthritis. Although all of these drug regimens, including placebo, produced significant changes in pain measures, only amitriptyline was associated with a significant reduction in the number of painful and tender joints. Intriguingly, a study of the effects of amitriptyline or imipramine in rats with adjuvant-induced arthritis has suggested that the tricyclic antidepressants may have direct anti-inflammatory effects because of a reduction in the physical signs of inflammation (Butler *et al*. 1985).

Managing anxiety disorders

Anxiety disorders are common in the general population and may be more frequent in patients with rheumatic disorders. Acute anxiety reactions may benefit from short-term use of benzodiazepines if they do not interfere with balance or cognitive function. It is important to be alert to more specific kinds of anxiety disorder, such as panic disorder with agoraphobia, obsessive-compulsive disorder, and social phobia. Panic disorder with agoraphobia may benefit from a benzodiazepine, such as aiprazolam or clonazepam, or a tricyclic with antipanic effects, such as imipramine or desipramine. Pharmacotherapy should be given in conjunction with cognitive and behavioural approaches to the phobias. Obsessive-compulsive disorder specifically benefits from clomipramine (Anafranil) and the SSRIs. Patients with social phobia seem to do best with either imipramine or monoamine oxidase inhibitors. All of these disorders are significant problems in their own right and should be differentiated from the more non-specific and transient adjustment reactions seen with anxiety.

Patients needing 'anxiolytic' medications such as benzodiazepines who have a history of alcohol or substance abuse generally are better treated with buspirone (Buspar), which is non-addicting.

Managing sleep disturbances

These are common, particularly in patients with systemic lupus, rheumatoid arthritis, and fibromyalgia (see Chapter 5.7.1). Some patients with lupus may develop insomnia secondary to the use of prednisone. When prednisone is producing manic or hypomanic symptoms associated with loss of sleep, a mood stabilizer such as lithium should be considered. The pain and inflammation of rheumatoid arthritis may be associated with more frequent night-time awakenings. Anti-inflammatory medication is probably more helpful than sedative hypnotic medication at night.

Summary

Finally, some essential clinical guidelines that may be helpful when evaluating rheumatological patients for psychiatric illness are shown in Table 11.

1. Neuropsychiatric illness is a common manifestation of many rheumatic diseases, particularly systemic lupus, where it may be associated with multiorgan involvement or as an isolated condition.
2. The neuropsychiatric manifestations of patients with systemic lupus on corticosteroids are much more likely to be due to the disease than to the steroids.
3. End-organ failure, for example renal or hepatic, and occult infection in the immunocompromised host should be excluded as causes for organic mental syndromes.
4. Consider potential drug toxicities, especially corticosteroids, as a cause of organic brain syndromes.
5. Depression complicating chronic rheumatic disease should be identified and when appropriate treated with antidepressant drugs.

Table 11 Important points to remember

Chapter References

Adams, D. H. *et al*. (1987). Neurological complications following liver transplantation. *Lancet*, **i**, 949–51.

Adams, W. R., Rose, C. D., Eppes, S. C., and Klein, J. D. (1994). Cognitive effects of Lyme disease in children. *Pediatrics*, **94**, 185–9.

Adelman, D. C., Salteil, E., and Klinenberg, J. R. (1986). The neuropsychiatric manifestations of SLE: an overview. *Seminars in Arthritis and Rheumatology*, **15**, 185–99.

Ader, R. (1991) *Psychoneuroimmunology* (2nd edn). Academic Press, New York.

Anderson, R. J. *et al*. (1976). Unrecognized adult salicylate intoxication. *Annals of Internal Medicine*, **85**, 745–8.

Angelopoulos, N., Drosos, A. A., Kostouitsag, E., and Liskos, A. (1988). Personality and psychopathology in patients with primary Sjögren's syndrome. *Terapevticheskii Arkhiv*, **60**, 49–52.

Anton, S. I. *et al*. (1994). Long-term treatment of depression with nefazodone. *Psychopharmacology Bulletin*, **30**, 165–9.

Baildam, E. M., Holt, P. J., Conway, S. C., and Morton, M. J. (1995). The association between physical function and psychological problems in children with juvenile chronic arthritis. *British Journal of Rheumatology*, **34**, 470–7.

Baum, K. A., Hopf, U., and Nehrig, C. (1993). Systemic lupus erythematosus: neuropsychiatric signs and symptoms related to cerebral MRI findings. *Clinical Neurology and Neurosurgery*, **95**, 29–34.

Bell, C. L. *et al*. (1991). Magnetic resonance imaging of central nervous system lesions in patients with lupus erythematosus. Correlation with clinical remission and antineurofilament and anticardiolipin antibody titers. *Arthritis and Rheumatism*, **34**, 432–41.

Bentson, J. *et al*. (1978). Steroids and apparent cerebral atrophy on computed tomography scans. *Journal of Computer Assisted Tomography*, **2**, 16–23.

Bertoli, M., Romagnoli, G. F., and Margreiter, R. (1988). Irreversible dementia following cyclosporin therapy in a renal transplant patient. *Nephron*, **49**, 333–4.

Blackwell, B. (1987). Antidepressants as adjuncts in chronic idiopathic pain management. *International Drug Therapy Newsletter*, **22**, 1–4.

Blalock, S. J. and DeVellis, R. F. (1992). Rheumatoid arthritis and depression: an overview. *Bulletin of Rheumatic Diseases*, **41**, 6–8.

Boston Collaborative Drug Surveillance Program (1972). Acute adverse reactions to prednisone in relation to dosage. *Clinical Pharmacology and Therapeutics*, **13**, 694–8.

Boumpas, D. T. *et al*. (1991). Pulse cyclophosphamide for severe neuropsychiatric lupus. *Quarterly Journal of Medicine*, **81**, 975–84.

Bousser, M. G., Rougement, D., Youl, B. D., and Wechsler, B. (1988). Neurological manifestations of Behçet's disease. *Journal of Vascular Malformations*, **13**, 231–4.

Braden, C. J., McGlone, K., and Pennington, F. (1993). Specific psychosocial and behavioral outcomes from the systemic lupus erythematosus self-help course. *Health Education Quarterly*, **20**,

29–41.

Butler, S. H., Weil-Fugazza, J., Godefray, F., and Besson, J. (1985). Reduction of arthritis and pain behavior following chronic administration of amitriptyline or imipramine in rats with adjuvant-induced arthritis. *Pain*, **23**, 159–75.

Carbotte, R. M., Denburg, S. D., and Denburg, J. A. (1986) Prevalence of cognitive impairment in systemic lupus erythematosus. *Journal of Nervous and Mental Disease*, **174**, 357–64.

Carbotte, S. M. *et al*. (1992). Fluctuating cognitive abnormalities and cerebral glucose metabolism in neuropsychiatric systemic lupus erythematosus. *Journal of Neurology and Neurosurgery*, **55**, 1054–9.

Cassileth, B. R. *et al*. (1984). Psychological status in chronic illness: a comparative analysis of six diagnostic groups. *New England Journal of Medicine*, **311**, 506–11.

Crotty, M., McFarlane, A. C., and Brooks, P. M. (1994). Psychosocial and clinical status of younger women with early rheumatoid arthritis: a longitudinal study with frequent measures. *British Journal of Rheumatology*, **33**, 754–60.

Daly, D. (1942). Central nervous system in acute disseminated lupus erythematosus. *Journal of Nervous and Mental Disease*, **102**, 461–3.

Daud, A. B. and Norudin, R. N. (1988). Solitary paraventricular calcification in cerebral lupus erythematosus: a report of two cases. *Neuroradiology*, **30**, 84–5.

David, J. *et al*. (1994). The functional and psychological outcomes of juvenile chronic arthritis in young adulthood. *British Journal of Rheumatology*, **33**, 876–81.

Diagnostic and Statistical Manual of Mental Disorders IV (1995). American Psychiatric Press, Washington DC.

Denburg, S. D. *et al*. (1993). Cognitive deficits in systemic lupus erythematosus. *Rheumatic Disease Clinics of North America*, **19**, 815–31.

Denburg, S. D., Behmann, S. A., Carbotte, R. M., and Denburg, J. A. (1994). Lymphocyte antigens in neuropsychiatric systemic lupus erythematosus. *Arthritis and Rheumatism*, **37**, 369–75.

Dougherty, P. M. and Dafny, N. (1988). Cyclosporine affects central nervous system opioid activity via direct and indirect means. *Brain and Behavioral Immunology*, **2**(3), 242–53.

Dougherty, P. M., Aronowski, J., Drath, D., and Dafny, N. (1987). Evidence of neuro-immunologic interactions: cyclosporine modifies opiate withdrawal by effects on the brain and immune components. *Neuroimmunology*, **13**(3), 331–42.

Ellison, D., Gatter, K., Heryet, A., and Esiri, M. (1993). Intramural platelet deposition in cerebral vasculopathy of systemic lupus erythematosus. *Journal of Clinical Pathology*, **46**, 37–40.

Falk, W. E., Mahnke, M. W., and Poskanzer, D. C. (1979). Lithium prophylaxis of corticotrophin-induced psychosis. *Journal of the American Medical Association*, **241**, 1011–12.

Fallon, B. A. and Nields, J. A. (1994). Lyme disease: a neuropsychiatric illness. *American Journal of Psychiatry*, **151**, 1571–83.

Feighner, J. P. (1994). The role of venlafaxine in rational antidepressant therapy. *Journal of Clinical Psychiatry*, **55** (Suppl. A), 62–8.

Feinglass, E. F., Arnett, F. C., Dorsch, C. A., Zizic, T. M., and Stevens, M. B. (1976). Neuropsychiatric manifestations of systemic lupus erythematosus: diagnosis, clinical spectrum, and the relationship to other features of the disease. *Medicine*, **55**, 323–39.

Fleming, J. L., Weisner, R. H., and Shorter, R. G. (1988). Whipple's disease: clinical, biochemical, and histopathologic features and assessment of treatment in 29 patients. *Mayo Clinical Proceedings*, **63**, 539–51.

Frank, R. G. *et al*. (1988*a*). Depression in rheumatoid arthritis. *Journal of Rheumatology*, **15**, 920–5.

Frank, R. G. *et al*. (1988*b*). Antidepressant analgesia in rheumatoid arthritis. *Journal of Rheumatology*, **15**, 1632–8.

Ganvir, P., Beammont, G., and Seldrup, J. (1980). A comparative trial of clomipramine and placebo as a adjunctive therapy in arthralgia. *Journal of International Medical Research*, **8** (Suppl. 3), 60–6.

Good, M. I. and Shader, R. I. (1982). Lethality and behavioral side effects of chloroquine. *Journal of Clinical Psychopharmacology*, **2**, 40–7.

Goodwin, J. S. and Regan, M. (1982). Cognitive dysfunction associated with naproxen and ibuprofen in the elderly. *Arthritis and Rheumatism*, **25**, 1013–15.

Greer, H. D., Ward, H. P., and Corbin, K. B. (1965). Chronic salicylate intoxication in adults. *Journal of the American Medical Association*, **193**, 555–8.

Gringas, M. (1976). A clinical trial of tofranil in rheumatic pain in general practice. *Journal of International Medical Research*, **4** (Suppl. 2), 41–9.

Hanly, J. G. *et al*. (1992a). Cognitive impairment in patients with systemic lupus erythematosus. *Journal of Rheumatology*, **19**, 562–7.

Hanly, J. G., Walsh, N. M., and Sangalang, V. (1992b). Brain pathology in systemic lupus erythematosus. *Journal of Rheumatology*, **19**, 732–41.

Hanly, J. G. *et al*. (1993). Cognitive impairment and autoantibodies in systemic erythematosus. *British Journal of Rheumatology*, **32**, 291–6.

Hay, E. M. *et al*. (1992) Psychiatric disorder and cognitive impairment in systemic lupus erythematosus. *Arthritis and Rheumatism*, **35**, 411–16.

Heisel, J. S. (1972). Life changes as etiologic factors in juvenile rheumatoid arthritis. *Journal of Psychosomatic Research*, **16**, 411–20.

Hopkinson, N. and Powell, R. (1990). 'Occult' SLE amongst patients admitted to acute psychiatric wards. (Abstract.) *Arthritis and Rheumatism*, **33**, S102.

Jaworski, T. M., Bradley, L. A., Heck, L. W., Roca, A., and Alarion, G. S. (1995). Development of an observation method for assessing pain behaviors in children with juvenile rheumatoid arthritis. *Arthritis and Rheumatism*, **38**, 1142–51.

Katz, P. P. and Velin, E. H. (1995). The development of depressive symptoms among women with rheumatoid arthritis. The role of function. *Arthritis and Rheumatism*, **38**, 49–56.

Kershner, P. and Wang-Cheng, R. (1989). Psychiatric side effects of steroid therapy. *Psychosomatics*, **30**, 135–9.

Khamashta, M. A., Cuadrado, M. J., Taub, N. A., Hunt, B. J., and Hughes, G. R. (1995). The management of thrombosis in the antiphospholipid antibody syndrome. *New England Journal of Medicine*, **332**, 993–7.

Kiecolt-Glaser, J. K. and Glaser, R. (1989). Psychoneuroimmunology: past, present, and future. *Health Psychology*, **8**, 677–82.

Kim, R. C., Collins, G. H., and Parisi, J. E. (1982). Rheumatoid nodule formation within the choroid plexus. Report of a second case. *Archives of Pathology and Laboratory Medicine*, **106**, 83–4.

Konttinen, Y. T., Santavirta, N., Honkanen, V., Sandelin, S., Schauman, L., and Gronblad, M. (1991). Systemic lupus erythematosus patient guide: influence on knowledge of the disease. *Annals of Rheumatic Disease*, **50**, 900–2.

Kruis, R. and Barger, R. (1980). Paranoid psychosis with sulindac. *Journal of the American Medical Association*, **243**, 1420.

Krupp, L. B. *et al*. (1991). Cognitive functioning in Lyme borreliosis. *Archives of Neurology*, **48**, 1125–9.

Kupfer, D. J. (1991). Long-term treatment of depression. *Journal of Clinical Psychiatry*, **52** (Suppl.), 28–34.

Liang, M. H. *et al*. (1984). The psychosocial impact of systemic lupus erythematosus and rheumatoid arthritis. *Arthritis and Rheumatism*, **27**, 13–19.

Logigian, E. L., Kaplan, R. F., and Steere, A. C. (1990). Chronic neurologic manifestations of Lyme disease. *New England Journal of Medicine*, **323**, 1438–44.

McCune, W. J., Macguire, A., Aisen, A., and Gebarski, S. (1988). Identification of brain lesions in neuropsychiatric systemic lupus erythematosus by magnetic resonance scanning. *Arthritis and Rheumatism*, **31**, 159–66.

McDonald Scott, N. A. (1969). The relief of pain with an antidepressant in arthritis. *Practitioner*, **202**, 802.

McNamara, M. E. (1991). Advances in EEG-based diagnostic technologies. In *Medical-psychiatric practice*, Vol. 1, (ed. A. Stoudemire and B. S. Fogel), pp. 163–92. American Psychiatric Press,

Washington DC.

Magni, G. (1991). The use of antidepressants in the treatment of chronic pain. A review of the current evidence. *Drugs*, **42**, 730–48.

Malinow, K. L., Molina, R., Gordon, B., Selnes, O. A., Provost, T. T., and Alexander, E. L. (1985). Neuropsychiatric dysfunction in primary Sjögren's syndrome. *Annals of Internal Medicine*, **103**, 344–50.

Meyerowitz, S., Jacox, R. R., and Hess, D. W. (1968). Monozygotic twins discordant for rheumatoid arthritis: a genetic, clinical, and psychological study of 8 sets. *Arthritis and Rheumatism*, **11**, 1–21.

Mohan, D., Mohandas, E., and Rajat, R. (1981). Chloroquine psychosis: A chemical psychosis? *Journal of the American Medical Association*, **73**, 1073–6.

Muller, P. D., Laville, E. A., Biddle, A. K., and Lorig, K. (1987). Efficacy of psychoeducational interventions on pain, depression and disability in people with arthritis: a meta-analysis. *Journal of Rheumatology*, **15**, 33–9.

Mustakillio, K. K., Pitkonn, T., and Pihkanen, T. A. (1962). Chloroquine psychosis? *Lancet*, **ii**, 1387–8.

Neuwelt, C. M., Lack, S., Kaye, B. R., Ellman, J. B., and Borenstein, D. G. (1995). Role of intravenous cyclophosphamide in the treatment of severe neuropsychiatric systemic lupus erythematosus. *American Journal of Medicine*, **98**, 32–41.

Omdal, R., Husby, G., and Mellgren, S. I. (1995). Mental health status in systemic lupus erythematosus. *Scandinavian Journal of Rheumatology*, **24**, 142–5.

Osler, W. (1903). On the visceral manifestations of the erythema group of skin diseases. *Transactions of American Physicians*, **18**, 599–624.

Parker, J. C. *et al*. (1988). Pain management in rheumatoid arthritis patients: a cognitive behavioral approach. *Arthritis and Rheumatism*, **31**, 593–601.

Peterson, M. G., Horton, R., Engelhard, E., Lockshin, M. D., and Abramson, T. (1993). Effect of counsellor training on skills development and psychosocial status of volunteers with systemic lupus erythematosus. *Arthritis Care Research*, **6**, 38–44.

Rimon, R. H. (1989). Connective tissue diseases. In *Psychosomatic medicine: theory, physiology, and practice*, Vol. 2, (ed. S. Cheren), pp. 565–609. International Universities Press, Madison CT.

Ritchlin, C. T. *et al*. (1992). Quantitative electroencephalography. A new approach to the diagnosis of cerebral dysfunction in systemic lupus erythematosus. *Arthritis and Rheumatism*, **35**, 1330–42.

Rogers, M. P. (1985). Rheumatoid arthritis: psychiatric aspects and use of psychotropics. *Psychosomatics*, **26**, 915–25.

Rogers, M. P. *et al*. (1980). Effect of psychological stress on the induction of arthritis in rats. *Arthritis and Rheumatism*, **23**, 1337–42.

Rogers, M. P. *et al*. (1992). I-123 iodoamphetamine SPECT scan in systemic lupus erythematosus patients with cognitive and other minor neuropsychiatric symptoms. A pilot study. *Lupus*, **1**, 215–19.

Rubbert, A. *et al*. (1993). Single photon emission computed tomography analysis of cerebral blood flow in the evaluation of central nervous system involvement in patients with systemic erythematosus. *Arthritis and Rheumatism*, **36**, 1253–62.

Shadick, N. A. (1992). Neurologic syndromes of Lyme disease. *Rheumatology Reviews*, **1**, 155–60.

Shadick, N. A. *et al*. (1994). The long-term clinical outcomes of Lyme disease. A population-based retrospective cohort study. *Annals of Internal Medicine*, **121**, 560–7.

Shiowaza, S., Kuroki, Y., Kim, M., Hipohata, S., and Ogino, T. (1992). Interferon-alpha in lupus psychosis. *Arthritis and Rheumatism*, **35**, 417–22.

Sigal, L. H. (1987). The neurologic presentation of vasculitic and rheumatologic syndromes. *Medicine*, **66**, 157–180.

Sigal, L. H. (1995). Pseudo-Lyme disease. *Bulletin of Rheumatic Diseases*, **44**, 1–3.

Small, J. G. (1990). Anticonvulsants in affective disorders. *Psychopharmacology Bulletin*, **26**, 25–36.

Solomon, G. F. (1969). Stress and antibody response in rats. *International Archives of Allergy*, **35**, 97–104.

Sotsky, S. M., and Tossell, J. W. (1984). Tolmetin induction of mania. *Psychosomatics*, **25**, 626–8.

Spiegel, D., Bloom, J. R., Kraemer, H. C., and Gottheil, E. (1989). Effect of psychosocial treatment on survival of patients with metastatic breast cancer. *Lancet*, **ii**, 888.

Stewart, A. I. *et al*. (1989). Functional status and well-being of patients with chronic conditions: results from the medical outcomes study. *Journal of the American Medical Association*, **262**, 907–13.

Stoudemire, A., Linfors, E., and Houpt, J. L. (1983). Central nervous system sarcoidosis. *General Hospital Psychiatry*, **5**, 129–32.

Strauss, G. D. *et al*. (1986) Group therapies for rheumatoid arthritis: a controlled study of two approaches. *Arthritis and Rheumatism*, **29**, 1203–9.

Tarter, R. E., Edwards, N., Hays, A., and Van Thiel, D. M. (1990). Neuropsychiatric dysfunction in a patient with Whipple's disease: effects of antibiotics treatment. *Psychosomatics*, **31**, 225–30.

Thornton, T. L. (1980). Delirium associated with sulindac. (Letter.) *Journal of the American Medical Association*, **243**, 1630–3.

Tollefson, G. D. and Garvey, M. J. (1982). Indomethacin and prostaglandins: their behavior relationships in an acute toxic psychosis. *Journal of Clinical Psycho-pharmacology*, **2**, 62–4.

Tsai, C. Y. *et al*. (1994). Cerebrospinal fluid interleukin-6, prostaglandin E_2 and autoantibodies in patients with neuropsychiatric systemic lupus erythematosus and central nervous system infections. *Scandinavian Journal of Rheumatology*, **23**(2), 57–63.

Vivian, A. S. and Goldberg, I. B. (1982). Recognizing chronic salicylate intoxication in the elderly. *Geriatrics*, **37**, 91–7.

Wamboldt, F. W., Weiler, S. J., and Kalin, N. H. (1984). Cyclosporin-associated mania. (Letter.) *Biological Psychiatry*, **19**, 1161–2.

Wilde, M. I., Plasker, G. L., and Benfield, P. (1993). Fluvoxamine. An updated review of its pharmacology and therapeutic use in depressive illness. *Drugs*, **46**, 895–924.

Wolfe, F., and Howley, D. J. (1993). Depression is not more common in rheumatoid arthritis: a 10 year long study of 6,153 patients with this disease. *Journal of Rheumatology*, **20**, 2025–31.

Wolkowitz, O. M. *et al*. (1990). Cognitive effects of corticosteroids. *American Journal of Psychiatry*, **147**, 1297–303.

2.1 Molecular genetics and its relevance in rheumatology

Patricia Woo

Introduction

The term 'new genetics' was first coined by editor of the *American Journal of Human Genetics*, commenting on an article that described a new DNA analytical approach to mapping the human genome. In general the term is used to describe the study of inheritance at the molecular level, including that of single-gene disorders, which can be traced clearly through families (as in the haemoglobinopathies), and of common (probably polygenic) disorders, such as coronary heart disease, mental illness, and autoimmune diseases. An understanding of underlying genetic factors is important because epidemiological studies of putative environmental factors have not revealed the cause of common diseases such as diabetes and rheumatoid arthritis. Although there is no clear evidence of their Mendelian inheritance, the inflammatory rheumatic diseases do cluster in families. Presumably this reflects the action of environmental factors against a background of genetic susceptibility. If we were able to define the main genes involved, and determine how their products differ from those of unaffected individuals, we would be in a better position to understand how these conditions arise, and eventually how to prevent or manage them.

Because of the successful control of environmental diseases in the developed world, diseases that are wholly or partially controlled by genes are assuming increasing importance in paediatric and adult medicine. Considerable resources have been invested not only in defining genes that may cause or confer susceptibility to human diseases but also in mapping the human genome. 'Reverse genetics' has proved a powerful tool (see below) for pinpointing the gene(s) involved with a particular disease. The genome mapping project will facilitate the identification of the disease-associated gene(s), once the relevant region of the chromosome has been identified. Understanding the biological activity of specific disease-associated genes requires knowledge of their control, functions, and products. Understanding the disease requires additional knowledge of the mutations and how they disrupt the function of the gene in question.

The human genome, gene structure, and expression

Structure of genes

Biologically important proteins are synthesized from a specific code carried within the DNA inside the nucleus of a cell. The code consists of different combinations of four nucleotides. This information is transcribed by copying the DNA into a messenger RNA (**mRNA**), which passes into the cytoplasm (Fig. 1). The mRNA is then associated with ribosomes and amino acids are then synthesized according to the mRNA sequence, that is, the mRNA forms a template for the translation of the genetic code into protein (Fig. 2).

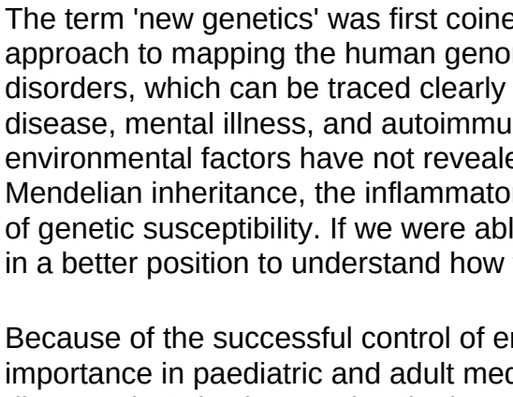


Fig. 1 Transcription and mRNA processing (adapted from Weatherall 1991).

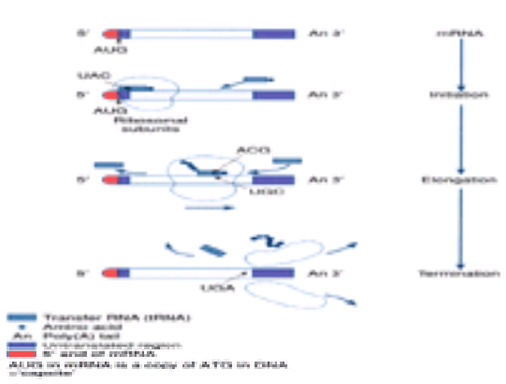


Fig. 2 Translation of mRNA (adapted from Weatherall 1991). AUG, initiation codon for methionine; UGA, termination codon.

A gene contains DNA nucleotide sequences that have their coding regions (exons) interrupted by sequences of unknown function (introns or intervening sequences). Though the role of introns is unclear, it is thought that they may have evolutionary importance. The direction of transcription of the coding sequence is usually controlled by a stretch of sequences upstream—the 5'-promoter region—containing the triplets ATG, which initiate transcription. Termination of transcription is denoted by another triplet, TAA, TAG, or TGA, at the 3' end of the gene. Within the promoter region, there are also sequences that regulate the gene's transcription rate (see below).

Control of gene expression

Every cell has a full complement of all the genes within the genome. Not all of them are actively expressed; the stage of development, the sex, and the cell types are

some of the important determining factors in expression. The mechanisms by which gene expression are regulated in eukaryotes (multicellular organisms) are not completely understood. In general, genes are switched on by the attachment of nuclear proteins to DNA with the appropriate recognition sequence, followed by the unfolding of the helical structure of the double-stranded DNA and also of chromatin. It is thought that this process allows the enzyme RNA polymerase II to gain access to the start site of the gene, the 'cap site', and initiate transcription by synthesizing a mirror image of the DNA coding strand, that is, a nuclear RNA. This RNA strand is edited or spliced to a final template (mRNA) for the synthesis of amino acids (see Fig. 1).

Transcription factors are DNA-binding proteins that mediate transcription of the DNA sequence (Mitchell and Tjan 1989). Certain common structures of these proteins have emerged; there are four types of transcription factors described so far. The first type contains a helix–loop–helix structure, as in homeoboxes of *Drosophila* DNA. The second type is a group of proteins with 'zinc fingers'—loops of the protein with a characteristic amino-acid sequence bound centrally by a zinc ion (Fig. 3). The first such protein described was the ribosomal transcription factor TFIIIA. These 'fingers' bind to the major groove of the DNA double helix. The third group of factors are the so-called leucine zippers (Fig. 4). These are dimeric proteins bound together at the N-terminal, which is leucine-rich. The C-terminals of these polypeptides bind DNA along the helical grooves. These proteins can be homo- or heterodimers, as in the oncogene complexes c- *fos*/c-*Jun*. The fourth type of proteins constitutes the NFkB family, which has in common 'rel' domain(s) that bind DNA as well as other nuclear proteins. The mechanism by which the DNA strand is activated is unclear. It may be a change in the tertiary structure of the DNA that will allow contact with other transcription factors crucial to the recruitment of RNA polymerase, as well as further protein interactions.

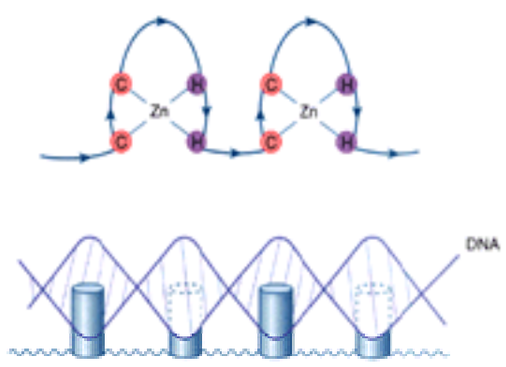


Fig. 3 Zinc-finger DNA-binding motif. C, cysteine; H, histidine.

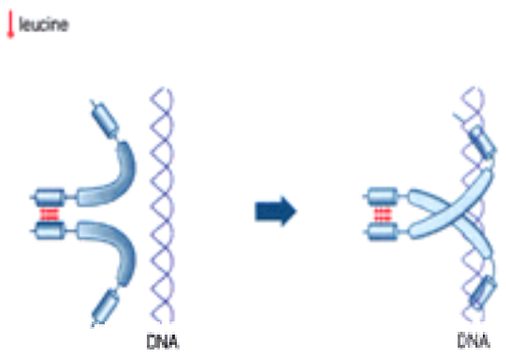


Fig. 4 Leucine zipper.

Transcription factors may be produced constitutively or may require an external stimulus such as cytokines and growth factors to be synthesized and/or activated. The inducible proteins tend to bind to enhancer elements in the 5' region of the gene. The enhancer elements are so called because they increase or facilitate transcription. The effect of the binding of enhancer proteins to DNA may be to cause further unfolding of the chromatin to allow easier access to DNA by RNA polymerase, or they may interact directly with the transcriptional apparatus, which consists of a series of essential DNA-binding proteins in addition to RNA polymerase II.

During development, sections of the chromatin are methylated to prevent activation of the genes within this region (Cedar 1988). Methylation is thought to both change the chromatin structure and prevent the binding of transcription factors. Recent work by Bird and colleagues shows that methylated CpG islands bind a protein, MeCP1, that can prevent binding of activated transcription factors (Meehan *et al*. 1992). Mutations in these regions can lead to demethylation, and activation of 'silent' genes. Another general mechanism of suppressing whole chromosomes is sex determination. Recent work has located a region of the X chromosome that can inactivate the second X chromosome in mice to prevent expression of double copies of genes on the X chromosome (Ashworth *et al*. 1991). The X-inactivation gene has now been identified (Penny *et al*. 1996). The third method of suppression of gene expression is gene-specific and relates to sequences, usually 5' to the gene, but possibly within the introns or the 3'-untranslated region, that interact with cell-specific nuclear factors. In this way, tissue-specific expression is precisely regulated.

The control of gene expression is not confined to transcription. The stability of mRNA and the storage and secretion of the newly synthesized protein can also be regulated. For example, the half-lives of mRNAs for interleukin (**IL**) 1 and tumour necrosis factor-a are very short because the AT repeats in their 3'-untranslated regions confer instability, presumably by allowing processing enzymes to act. The length of the poly(A) tail of the mRNA confers stability to the half-life of mRNA. But this is only one of the stabilizing mechanisms, as there are mRNA species without poly(A) tails. This area of investigation is expanding but there are few answers as yet.

Control of the protein secretion rate is another method of regulation by external agents. For example, apolipoprotein (**apo**) B is regulated by the fatty acid oleates, albumin and insulin in tissue-cultured cells via alteration of its secretion rate (Pulinger *et al*. 1989). Secretion of C-reactive protein is increased by the presence of IL-1 in tissue culture (Macintyre *et al*. 1985).

So far it has been assumed that one DNA template produces one peptide, but there are exceptions to this rule. As described above, the DNA exons are separated by intervening sequences or introns, and these allow different assemblies of exons, as with genes in the immunoglobulin superfamily (Hood *et al*. 1985). Also, mRNAs can have different lengths if there are several start sites for transcription; the mRNAs for calcitonin and calcitonin gene-related peptide are good examples (Edbrooke *et al*. 1985). Alternatively the mRNA can be edited, as demonstrated for apoB by Powell *et al*. (1987). A stop codon is introduced through a C®U conversion in the pre-mRNA of the intestine, resulting in a shorter peptide (B48 from B100), which is also functional. Another example of post-transcriptional modification is a chimerical protein in the absence of a fusion mRNA, for example, one of the subunits of glucose 6-phosphate dehydrogenase (Borst *et al*. 1989).

An apparent exception to Mendelian inheritance has been described. The expression of certain genes is determined by whether they are inherited from the male or female parent, a phenomenon known as parental imprinting (Swain *et al*. 1987). A good example is the inheritance of Angelman syndrome ('the happy puppet') from the mother only. Little is known at present about the molecular mechanisms involved in this fascinating phenomenon.

Coordinated gene expression

Activation and suppression of genes by external signals are the central mechanisms whereby growth, differentiation, metabolism, and inflammation occur. The pathways of extracellular control can be one of three types: endocrine signalling (control molecules acting from a distance), paracrine signalling (where target cells are close to the signalling cells), and autocrine signalling (where cells respond to their own secreted polypeptides). All these extracellular regulatory polypeptides can act in one of two ways: diffusion across cell membranes (e.g. steroids, which are lipid-soluble) or interaction with specific receptors (e.g. all cytokines and growth

factors). Receptor-binding proteins rely on a series of intracellular messenger molecules to be activated once they have bound to their specific receptors, a process called signal transduction. These second-messenger systems are all dependent on protein kinases, transferring the phosphate group of ATP on to threonine, serine, or tyrosine of cytosolic proteins. Many of these are likely to be precursors of DNA-binding proteins. The study of signal transduction is important, providing insight into how genes are coordinated by external signals. Although there are millions of genes, there are clearly common transduction pathways utilized by different external signals. It follows therefore that if there is mutation in the regulatory molecules, be they growth factors, receptors, intracellular messengers, or DNA-binding proteins, the consequences for cell growth and differentiation can be very serious.

The concept of gene families

Many human genes have been shown to belong to gene families. The criteria for family membership are DNA-sequence and/or protein-structural homologies, and similarity in function. There are many good examples of these gene families; in the study of inflammation and rheumatic diseases, the supergene family of immunoglobulins and related genes is one. Diversity of immunoglobulin genes is achieved by shuffling or rearrangement of the many exons by the B lymphocytes, and immunoglobulin- and T-cell-receptor genes are unique in this method of generating diversity. The membrane molecule has a characteristic structure within this gene family (Fig. 5). Related genes include those coding for several leucocyte surface antigens, growth-factor receptors, as well as the T-cell-receptor genes. T-cell receptors recognize antigens coded for by the major histocompatibility complex (**MHC**) from antigen-presenting cells (see Chapter 3.3). It is interesting that the MHC genes also belong to this immunoglobulin superfamily. Thus the proteins involved in antigen presentation and recognition are all part of this family.

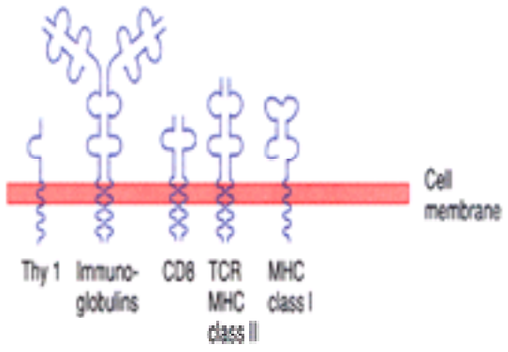


Fig. 5 The immunoglobulin superfamily. Thy 1/CD8, leucocyte antigens; TCR, T-cell receptor.

Another example is the group of growth/cytokine receptor families: the tumour necrosis factor/Fas (CD95) receptor family (see Chapter 3.1) and the IL-6/growth hormone-receptor families. The former use similar ligand-binding domains and the latter have common signalling domains.

There are increasing numbers of examples of gene families, for example the transforming growth factors and bone morphogenetic proteins (see Chapter 2.4), and the proteinases within the serpin family (Carrell *et al*. 1989) and the transcription factor families. All these examples strongly suggest that specialized genes are evolved from a primordial or ancestral gene.

The human genome

It is important for us to define the genes that are involved in human disease, and an estimate of the number of functional genes or gene clusters has been calculated (see Weatherall 1991). There are approx. 3×10^9 nucleotide pairs in the human haploid genome, but half would be repetitive sequences. Of the rest the coding ratio may be approx. 1:30 in each gene cluster, and it was estimated by Weatherall that the total number of clusters (of about 15 genes or so) will be 3000 to 15 000. Thus it is a feasible proposition to attempt to sequence the functional genome, as is the policy of the human genome projects in Europe and the USA. These projects should expedite the location of disease-associated genes once close linkage has been established by reversed genetics (see below).

Principle of reverse genetics

This has been a very useful approach to the identification of genes that cause clinical disease in single-gene disorders such as muscular dystrophy, cystic fibrosis, and retinoblastoma, to name but a few. The first step in the chromosomal localization of the gene is to study gene markers in affected families, multicase families or sibling pairs, and establish linkage between the gene markers and the disease gene—that is, cosegregation of the gene and disease in family studies over three or more generations. This method is often called gene mapping: the technique first used was restriction fragment length polymorphism but now DNA amplification by polymerase chain reaction (see Chapter 4.7). There are clusters of nucleotides of CG repeats, known as microsatellites, scattered on all chromosomes. Some 300 primer pairs are currently available to provide markers at 10 cM apart. This technology is semi-automated and a physical map of the human genome has been produced (Adams *et al*. 1995).

Once the region of the chromosome has been identified, the disease gene can be identified by assessing the segregation of known genes in this region with disease—the candidate-gene approach. Techniques used include extensive DNA cloning to define the gene, its sequence, its derived-protein sequence, and finally its function. At present the identification of a disease gene is very time-consuming, labour-intensive, and expensive; the genome mapping project will speed up this process.

Reverse genetics has been also very useful in unravelling genetic variability in clinical disorders, as well as providing a measure of genetic contributions to common diseases that are usually polygenic (see below).

Examples of single-gene disorders in rheumatology

It is clear that gene action is regulated by a series of steps that finally lead to a protein product. Mutations in single-gene disorders can act at all these steps (Table 1). Most single-gene disorders characterized so far are the result of the synthesis of a structurally abnormal protein, reflecting single-base mutations in the parent gene. Collagen accounts for approx. 25 per cent of total body protein. It is essential for the integrity of cartilage, bone, tendon, blood vessels, heart valves, and many other tissues. Genetic disorders of collagen leading to clinical diseases such as Ehlers–Danlos and Marfan syndromes are clearly set out in Chapter 2.2. The mutations include deletions of parts of the gene, and nucleotide change(s) leading to critical amino-acid substitutions.

Reduced output of gene product
Transcription
Insertions
Deletion or partial deletion
Promoter-box mutations
Other regulatory mutations
Processing of mRNA
Mutations involving splicing
Poly(A) addition-site mutations
Translation
Mutations involving initiation, elongation, or termination
Synthesis of an abnormal gene product
Mis-sense mutations (point mutations)
Fusion genes
Insertions
Deletions
Elongated products
Chain-termination mutations
Frameshift mutations
Defective post-transcriptional modification

Reproduced from Weatherall (1991) with permission.

Table 1 Types of mutations that underlie single gene disorders

Amino acid substitutions in the structural protein can lead to abnormal metabolites, and some of these single-gene disorders are seen in the paediatric rheumatology clinics. For example, the storage diseases such as Hurler's, Scheie, other mucopolysaccharidoses, and mucolipidoses are all the result of defective enzymes, leading to an accumulation of abnormal metabolites in tendons and subcutaneous tissues over joints (see Chapter 5.17.4).

Apart from the production of an abnormal protein, many genetic disorders are caused by the reduced production of a protein, that is, reduced gene expression. The a- and b-thalassaemias are good examples (Weatherall *et al*. 1988). This type of defect has been shown particularly in subgroups of systemic lupus erythematosus, where the absence of the genes for complement C2 or C4 are clear disease-susceptibility factors because of the defective handling of immune complexes in these patients. In C1 esterase-inhibitor deficiency, the defect can be a combination of the production of an abnormal non-functional protein and reduced production of the normal protein. The clinical picture is hereditary angioedema but the patients may also have features of systemic lupus.

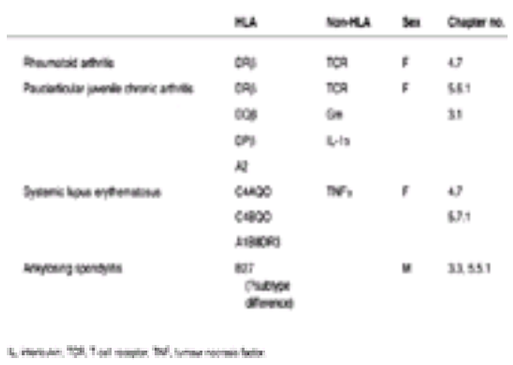
Even if a single-gene defect has been defined for a particular disease, the phenotypic expression, as a result of the different types of mutations and deletions of that gene, can be heterogeneous, and therefore so can the resultant clinical disease manifestations. Good examples include the thalassaemias, haemophilia A, and cystic fibrosis. In these single-gene disorders, other genes and environmental factors may also alter the phenotypes, but the main influence is the mutant gene.

Genetics of common diseases (see review by Lander and Schork 1994)

In contrast to the single-gene disorders, common conditions, such as heart disease, diabetes, rheumatic diseases, psychiatric illness, and cancer, appear to be the result of a complex interaction between environmental factors and several genes at different loci. In order to understand the underlying cause of such diseases, we need to define both environmental and genetic factors. In the case of rheumatoid arthritis, which is prevalent among many racial groups, epidemiological studies have clearly shown that it is not like an infectious disease with a single causative agent. Twin studies of the *HLA-DR* locus have shown that the contribution from this is about 15 per cent of the total genetic susceptibility to rheumatoid arthritis. Thus rheumatoid arthritis is an example of polygenic disease, where it is likely that the combined effect of several genes can allow a number of environmental triggers to produce a similar clinical condition. The drive to unravel the genes that are important in polygenic diseases is because their identification will allow us to understand the biochemical basis of why these diseases occur (e.g. complement deficiencies and immune-complex disease). As a result, more rational methods of disease prevention and management can be developed.

There are two general approaches to defining the genes involved in a particular disease: the candidate-gene approach and the establishment of linkage using large families—the techniques of reverse genetics as mentioned above. These are formidable tasks, especially when gene mapping will be confounded by factors such as genetic heterogeneity, incomplete penetrance, and gene interaction. Once again, as in single-gene disorders, this task will be a great deal easier once we have a reasonable map of the human genome.

As far as the inflammatory rheumatic diseases are concerned, the candidate-gene approach has been favoured historically and considerable effort has been expended in the analysis of the MHC genes and, more recently, the T-cell-receptor genes. The hypothesis is that the trimolecular complex of HLA, peptide, and T-cell receptor in presenting antigen is central to the pathogenesis of disease. More detailed descriptions of current findings can be found in Chapter 3.3 and Chapter 4.7. Table 2 illustrates the polygenic influences in four groups of diseases. Examples of the linkage approach include mapping of the gene for familial susceptibility to psoriasis to 17q (Tomfohrde *et al*. 1994) and familial Mediterranean fever to chromosome 16 (Pras *et al*. 1992).



	HLA	Non-HLA	Sex	Chapter no.
Rheumatoid arthritis	DRβ	TCR	F	4.7
Pauciarticular juvenile chronic arthritis	DRβ	TCR	F	5.6.1
	DQβ	Gm		3.1
	DPβ	IL-1α		
	A2			
Systemic lupus erythematosus	C4AQO	TNFα	F	4.7
	C4BQO			5.7.1
	A1B8DR3			
Ankylosing spondylitis	B27		M	3.3, 5.5.1
Chalkyce difference				

IL, interleukin; TCR, T-cell receptor; TNF, tumour necrosis factor.

Table 2 Polygenic influences in rheumatic diseases

A more recent area of study concerns candidate genes other than those coding for the antigen-presenting trimolecular complex. For example, the inflammatory cytokines are mediators of the inflammatory and immune responses seen in inflammatory rheumatic diseases, particularly rheumatoid arthritis and systemic juvenile chronic arthritis (see Chapter 3.1). Two approaches are used in the study of these candidate genes. One is to define genetic polymorphisms in multicase families, both in the coding and control regions of the genes. An example is the search for polymorphisms of cytokine genes and their associations with autoimmune diseases. This is reviewed in Chapter 3.1. The regulation of these genes appears to be variable according to the genotype, thus providing evidence for the concept that these genes can be modifiers of disease and contribute to its severity rather than being crucial to its development.

A brief discussion of the regulation of gene expression has already been given above. In rheumatic diseases, the interaction between environmental infectious agent(s) and genes may be at the level of the antigen presentation and antigen recognition, that is, the trimolecular complex of MHC, peptide, and T-cell receptor. But intracellular perturbation of the transcriptional apparatus of important cellular genes can also be important. Molecular virology has shown that a number of viruses use cellular transcription factors for their own expression (e.g. human immunodeficiency virus is activated by NFkB). They can activate cellular transcription factors that will also act on other cellular genes; for example, human T-lymphotrophic virus (**HTLV-1**) activates IL-2 through NFkB. In addition, their own transcription/transactivating factors could act on cellular genes; for example, the Tax protein of HTLV-1 can transactivate the genes for IL-2 and IL-2 receptors (Greene *et al*. 1986).

Tumour viruses have long been known to cause neoplastic proliferation of cells. They can achieve this either by insertion into the host genome, causing increased expression of adjacent host-cell genes (insertional mutagenesis), or they may carry genes with transforming properties (viral oncogenes of retroviruses). The viral oncogenes have homologous DNA within the cell and these are part of the its normal transcriptional apparatus. A good example is the NFkB family of DNA-binding proteins, related to the v-*rel* oncogene family (Kieran *et al*. 1990; Ruben *et al*. 1991), which are responsible for transcription of the k light chain of immunoglobulin, MHC, acute-phase genes such as those for serum amyloid A and angiotensinogen, and cytokines. Oncogene products may be involved in many aspects of cellular regulation. They may be growth factors or their receptors (e.g. c-*sis* and the a-chain of platelet-derived growth factor; c-*fms* and the receptor for colony-stimulating factor 1), enzymes associated with these receptors, tyrosine kinase (e.g. c-*fms* and v-*fms*), or DNA-binding proteins. Some of these cellular oncogenes, such as Bcl-2, are important in delaying programmed cell death (apoptosis), and others such as p53 and *myc* hasten apoptosis. The membrane protein Fas also enhances apoptosis when bound by its ligands. Since cell proliferation is related intimately to apoptosis, irregularity in apoptosis of immune-reactive cells can lead to disease. When a single base mutation was found in the Fas–protein in MRL-lpr/lpr mice, which possess features of systemic lupus and rheumatoid arthritis (Watanabe-Fukunaga *et al*. 1992), apoptotic defects were suspected as a mechanism for autoimmune diseases such as systemic lupus and rheumatoid arthritis. Recently, human lymphoproliferative disease has been found in children with mutations in Fas (Rieux-Laucat *et al*. 1995). Thus the role of the oncogenes in apoptosis needs to be characterized in autoimmune disease.

Conclusion

The study of common polygenic diseases is a formidable challenge to the researcher, but the potential benefits are enormous. Recombinant DNA technology is a powerful tool for unravelling this problem. Understanding the function of the products of the genes involved will allow more specific management the diseases. Environmental risk factors can be studied again in the light of genetic subgroups, and preventative measures should then become possible.

Chapter References

Adams, A. *et al*. (1995). The genome directory. *Nature*, **377** (Suppl. to 28 September issue).

Ashworth, A., Rastan, S., Lovell-Badge, R., Kay, G. (1991). X chromosome difference in viability of XO humans and mice. *Nature*, **351**, 406–8.

Borst, P., Benne, R., and Tabak, H. F. (1989). A fused chimeric protein made in human cells. *Cell*, **58**, 421–2.

Carrell, R. W., Avlak, K. S., and Owen, M. C. (1989). The molecular pathology of the serpins. *Molecular Biology Models*, **6**, 35–42.

Cedar, H. (1988). DNA methylation and gene activity. *Cell*, **53**, 3–4.

Edbrooke, M. R. *et al*. (1985). Expression of the human calcitonin-CGRP gene in lung and thyroid carcinoma. *EMBO Journal*, **4**, 715–24.

Greene, W. C. *et al*. (1986). Transactivator gene of HTLV-1 induces IL2 receptor and IL2 cellular gene expression. *Science*, **232**, 877–80.

Hood, L., Kronenberg, M., and Hunkapiller, T. (1985). T cell receptors and the immunoglobulin supergene family. *Cell*, **40**, 225–9.

Kieran, M. *et al*. (1990). The DNA binding subunit of NF_B is identical to KBF1 and homologous to the *rel* oncogene product. *Cell*, **62**, 1007–18.

Lander, E. S. and Schork, N. J. (1994). Genetic dissection of complex traits. *Science*, **265**, 2037–48.

Macintyre, S. S., Kushner, I., and Samols, D. (1985). Secretion of C-reactive protein becomes more efficient during the course of the acute phase response. *Journal of Biological Chemistry*, **260**, 4169–73.

Meehan, R. R., Lewis, J. D., and Bird, A. P. (1992). Characterisation of MeCP2, a vertebrate DNA binding protein with affinity for methylated DNA. *Nucleic Acids Research*, **20**, 5085–92.

Mitchell, P. J. and Tjan, R. (1989). Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science*, **245**, 371–8.

Penny, G. D., Kay, G. F., Sheardown, S. A., Rastan, S., and Brockdorft, N. (1996). Requirement for Xist in X chromosome inactivation. *Nature*, **379**, 131–7.

Powell, L. M., Simon, C. W., Pease, R. J., Edwards, Y. H., Knott, T. J., and Scott, J. (1987). A novel form of tissue specific RNA processing produces Apo B 48 in intestine. *Cell*, **50**, 831–40.

Pras, E., Aksentijevich, I., Gruberg, L., Kastner, D. L., and Pras, M. (1992). Mapping of a gene causing familial Mediterranean fever to the short arm of chromosome 16. *New England Journal of Medicine*, **326**, 1509–13.

Pulinger, C. R., North, J. D., Peng, B. B., Rifici, V. A., de Brito, A. E. R., and Scott, J. (1989). Apo-lipoprotein beta gene is constitutively expressed in HepG2 cells—regulation of secretion by oleic acid, albumin, and insulin and measurement of the mRNA half life. *Journal of Lipid Research*, **30**, 1065–77.

Rieux-Laucat, F. *et al*. (1995). Mutations in fas association with human lymphoproliferative syndrome and autoimmunity. *Science*, **268**, 1347–9.

Ruben, S. M. *et al*. (1991). Isolation of the *rel*-related human cDNA that potentially encodes the 65 kDa subunit of NF_B. *Science*, **251**, 1490–3.

Swain, J. L., Stewart, T. A., and Leder, P. (1987). Parental legacy determines methylation and expression of an autosomal transgene: a molecular mechanism for parental imprinting. *Cell*, **50**, 719–27.

Tomfohrde, J. *et al*. (1994). Gene for familial psoriasis susceptibility mapped to the distal end of human chromosome 17q. *Science*, **264**, 1141–5.

Watanbe-Fukunaga, R., Brannan, C. I., Copeland, N. G., Jenkins, N. A., and Nagata (1992). Lymphoproliferative disorder in mice explained by defects in Fas antigen that mediates apoptosis. *Nature*, **355**, 314–7.

Weatherall, D. J. (1991). *The new genetics and clinical practice*, (3rd edn). Oxford University Press.

Weatherall, D. J., Clegg, J. G., Higgs, D. R., and Wood, W. G. (1988). The haemoglobinopathies. In *The metabolic basis of inherited disease*, (6th edn), (ed. C. R. Scriven, A. L. Beaudet, W. S. Sly, and D. Valle), pp. 2281–366. McGraw-Hill, New York.

2.2 Molecular abnormalities of collagen and connective tissue

F. Michael Pope

Introduction

Collagen is a major connective-tissue protein family with crucial mechanical and scaffolding functions. Other important components of the extracellular matrix include fibrillin, elastin, microfibrillar-associated protein, and certain other skeleton-associated products such as PAX1 to –9, SOX9, FGFR1 to –3, and bone morphogenetic protein, which all have tissue-regulating properties. These various substances occur in a variety of connective tissues such as bone, cartilage, muscle sheaths, ligaments, joint capsules, skin, vascular structures (e.g. veins, arteries, heart valves), lungs, pleuroperitoneal linings, intestinal walls, hernial sacs, and glomeruli (Pope and Nicholls 1986; Pope *et al*. 1988a, Pope *et al*. 1988b). Furthermore, there are numerous subtle molecular interactions with other connective tissue elements such as laminin, elastin, heparin or chondroitin sulphate, other proteoglycans, and fibronectin. Each of these substances can react with other proteins and also bind to a variety of cellular components. These are more specific interactions between types I, II, III collagens and types V, VI, IX, and X collagen, which together form compound fibrils. They also probably influence embryogenesis and fetal development so that errors and mutations interfering with their quantity and quality cause congenital fetal deformities. Thus, mouse embryos deficient in type I collagen may be genetic lethals with death as early as the 12th day, whilst less destructive errors cause disorders such as osteogenesis imperfecta. Various mouse models have been created, including those with errors in collagens I, II, III, V, IX, and X. Mice with rhizomelic dwarfism and cleft palates caused by *COL2A1* mutations have been known since 1990.

Since the review by Vuorio and de Crombrugghe (1990), another six collagen proteins and 10 equivalent genes have been described. Currently this family contains at least 19 proteins and 32 genes (Prockop and Kivirriko 1995). Mutations have been known in some for 15 years or more; thus *COL1A1* and *COL1A2* mutations of type I collagen cause either osteogenesis imperfecta or Ehlers–Danlos syndrome type VII (Byers 1990; Pope 1991). Collagen type II (*COL2A1*) mutations clearly cause certain chondrodysplasias (Kuivaniemi *et al*. 1991); type III collagen (*COL3A1*) mutations usually cause the arterial form of Ehlers–Danlos syndrome (type IV) or occasionally types I or III. An X-linked, a5(IV) collagen, basement-membrane protein coded by the *COL4A5* gene is implicated in X-linked Alport syndrome (Barker *et al*. 1990). In other cases, *COL4A1*, *–4A2*, *–4A3*, and *–4A4* genes have been implicated in autosomal-dominant and -recessive Alport syndrome. Collagen V (*COL5A1*) has recently been implicated in Ehlers–Danlos syndrome types I and II, which are most probably allelic, whilst abnormalities of the *COL11A1* and *–11A2* genes have been implicated in the Stickler syndrome (Nicholls *et al*. 1994; Li *et al*. 1995; Loughlin *et al*. 1995; Snead *et al*. 1994; Vikkula *et al*. 1995). There is also biochemical overlap between the Stickler syndrome and multiple epiphyseal dysplasia (Vikkula *et al*. 1995), with mutational position and the clinical phenotype rather closely related. *COL11A1* and *–11A2* have also been implicated in multiple epiphyseal dysplasia. Defects of type X collagen cause Schmid-type metaphyseal chondrodysplasia (Wallis *et al*. 1994). Nevertheless, mutant mice null for *COL10A1* show normal growth and development of long bones. This suggests that *COL10A1* mutations act as dominant negatives (Rosati *et al*. 1994). Type VII collagen mutations frequently cause dystrophic epidermolysis bullosa (Dunnill *et al*. 1994a, Dunnill *et al*. 1994b; Eady and Dunnill 1994; Eady *et al*. 1994). Recent new collagens include the bullous pemphigoid antigen gene (*BPAG2/COL17A1*). Here mutations have already been detected in generalized atrophic benign epidermolysis bullosa (McGrath *et al*. 1995). Another interesting substance is MARCO, a novel macrophage receptor with a large intracellular collagenous domain, one of the scavenger receptor-protein family. Diseases caused by other similar genes are highly probable. Other important candidate genes for connective tissue disease include those for collagen VI, present in skin, blood vessels, and fibrocartilage, and for type VIII, which is an endothelial collagen. Collagens IX, X, XI, and XII are important components of cartilage and bone, whilst type XIII collagen is keratinocyte-associated (Vuorio and de Crombrugghe 1990). Subsets of common disorders such as osteoporosis, osteoarthritis, and congenital berry aneurysm (Pope *et al*. 1988a; Ala-Kokko *et al*. 1990) are sometimes caused by mutations of collagen types I, II, and III, and resemble those of osteogenesis imperfecta, certain chondrodystrophies, and vascular Ehlers–Danlos syndrome. Other diseases likely to be implicated include Ehlers–Danlos syndrome types I, II, and III (the benign hypermobile syndrome), Ehlers–Danlos syndrome VIII, and dystrophic epidermolysis bullosa, which is usually caused by *COL7A1* mutations. Defects of cartilage or bone morphology caused by connective-tissue gene mutations include achondroplasia, pseudoachondroplasia, and the spondyloepiphyseal dysplasias. Some forms of pseudoxanthoma elasticum may be caused by certain elastin and/or collagen errors. The Marfan syndrome is caused by abnormalities of fibrillin 15 (Godfrey *et al*. 1994). In the same year as Blanton *et al*. (1990) and Kainulainen *et al*. (1990) linked it for the first time to chromosome 15, the fibrillin –5 and –15 genes were cloned. Subsequently, numerous mutations have been detected and the genomic organization of fibrillin 15 established (Corson *et al*. 1993; Nijbroek *et al*. 1995). In contrast, the fibrillin gene on chromosome 5 causes a slightly different disorder (congenital contractual arachnodactyly). There have also been rapid advances in the analysis of defects of elastin and various other microfibrillar elements that cause diseases such as Williams syndrome. The clinical range of the diseases is wide and includes various inherited single-gene diseases, the common forms of which cause significant morbidity and mortality. For example, osteoporosis affects up to 50 per cent of otherwise normal adult women by the age of 75 years (Lewis 1981). Similarly, osteoarthritis is discovered in 10 per cent of routine autopsies, whilst aneurysms in the circle of Willis are found in 2 per cent of autopsies and cause 4000 deaths per annum in the United Kingdom (Mortality Statistics 1977). Similar gene defects may also participate in other common abnormalities, such as varicose veins, inguinal and femoral hernias, and in common forms of familial joint hypermobility syndromes, which affect up to 10 per cent of Caucasians in the United Kingdom.

The molecular analysis of human genetic disease proceeds in one of two ways. Most often a faulty candidate gene is suspected from structural (histological) finding or the protein chemistry, and is subsequently cloned, sequenced, and then analysed for mutations. This approach has been impressively successful in inherited abnormalities of collagens I, II, and III, but in other cases all obvious candidate genes have been excluded, leaving the cause completely unknown. Other diseases worthy of consideration include common forms of myopia, corneal abnormalities, emphysema, mitral-valve prolapse, periodontal disease, and certain disorders of the myotendinous junctions, any or all of which might be caused by connective tissue mutations.

Modern recombinant DNA technology

The recent spectacular advances in recombinant DNA technology include methods for cutting DNA into fragments suitable for insertion into transmissible, extrachromosomal DNA elements (plasmids), which are easily grown in bacteria, thus facilitating the purification and amplification of any desired DNA fragment. Subsequently, progressively larger DNA fragments have been amplified whilst vectors such as l-phages, cosmids, and yeast artificial chromosomes (Watson *et al*. 1987) have increased the length of clonable DNA from tens to hundreds of kilobases (gene sequencing). Similarly, the preferential amplification of any DNA segment by the polymerase chain reaction (**PCR**) (Saiki *et al*. 1986) has revolutionized human molecular genetics.

Two approaches commonly have been used. The first, reverse genetics, identifies the gene locus and then works backwards from gene sequencing to protein structure and function. Early examples were the cloning of genes for Duchenne muscular dystrophy (Monaco *et al*. 1986), cystic fibrosis (Rommens *et al*. 1989), and neurofibromatosis type I (Fountain *et al*. 1989a; Fountain *et al*. 1989b), and also chronic granulomatous disease. Another striking example was in Huntington's disease, which was extraordinary difficult to clone despite being first mapped in 1984 (Gusella *et al*. 1985). Similarly, Alzheimer's disease was first mapped in 1987 but now has several different causative genes (St George-Hyslop *et al*. 1987). Other examples include schizophrenia (McGillivery *et al*. 1990), myotonic dystrophy, fragile X, and the class of neurological defects characterized by unstable triplet repeats. The Marfan syndrome, which was only mapped to chromosome 15 in 1990 (Kainulainen *et al*. 1990), is an excellent example in the genetics of connective tissues. In the alternative approach a candidate protein is identified in various ways and the gene subsequently cloned, followed by mRNA, cDNA, or genomic analysis.

Collagen genes

The gene analysis of collagen or other connective tissue proteins is amenable to both the approaches outlined above. In the classical approach the candidate collagen protein is identified and cloned from cDNA or genomic libraries. Alternatively, a suitable autosomal-dominant disease such as osteogenesis imperfecta can be tested for linkage to *COL1AI* or *–1A2* gene markers or to other restriction fragment length polymorphisms (**RFLP**) nearby (Sykes *et al*. 1990).

The conventional approach has defined at least 30 genes coding for some 20 collagen (compound) proteins. In humans these are rather widely distributed, lying on chromosomes 1, 2, 6, 7, 10, 12, 13, 17, 21, and the X chromosome. There are multiple genes on chromosomes 2 (three), 6, 13, and 21 (two on each) (Vuorio and de Crombrugghe 1990). All code for extended triple helices of Gly–XY polymers (where X and Y are often proline or hydroxyproline). Thus collagen types I, II, III, V, and XI, which form the so-called interstitial collagens, have N- and C-globular extensions with uninterrupted Gly–XY triple helices coded by 52 exons. All other collagens have either globular N- and/or C-terminal extensions, with varyingly large globular interruptions of the triple helix (Kuivaniemi *et al*. 1991). Bending of the rigid helix at such globular interruptions may be important in different tissues such as basement membranes, cartilage, and cornea, and probably allows interaction with other associated macromolecular and cellular components. Cartilage also contains conventional, cross-striated, interstitial type II collagen fibres.

Historically the genes for collagens I, II, III, and V were identified in numerical order of discovery (the numbering system originally arose from protein analysis 20 years earlier). Thus the a1(I) and a2(I) genes *COL1A1* and *COL1A2* of type I collagen were originally cloned in chickens and then in man. These vertebrate proteins have well-defined N- and C-terminal extensions essential for fibril alignment and formation, and the central triple helices are coded by regular GlyXY subunits in cassettes of exon (coding) sequences, which are always multiples of 9 bp, commonly 45, 54, 63, 81, 108, and 162 (Prockop and Kivirriko 1984). Such base-pair combinations suggest that there is duplication of the original 9-bp multiples in various combinations and permutations (Vuorio and de Crombrugghe 1990). Alternatively, an ancestral 54-bp exon with surrounding introns could produce 108 and 162 bp by intron loss and 45 and 99 bp by recombination.

Thus type I collagen is a heteropolymer of two distinct a-chains that spontaneously self-assemble in a 2:1 ratio to form $a1(I)_2a2$ triple helices. Although, rarely, there may be $a1(I)_3$ trimers in skin, a2 trimers cannot assemble and are thermodynamically unstable. Types II and III collagens form homotrimers: $a1(II)_3$ and $a1(III)_3$, respectively. The latter has two cysteine bonds in the a-helix, as well as the normal inter- and intrachain disulphide bonds within the C-propeptides found in other interstitial collagens. Type II collagen is confined to cartilage and the vitreous humour of the eye, whilst type III collagen is widely distributed in skin, ligaments, tendons, pleuroperitoneal linings, the intestinal wall, and blood vessels (arteries, veins, and capillaries) particularly. It plays an especially important part in arterial strength and stability. These four collagens all have a central helix of $(GlyXY)_{333}$ with proline or hydroxyproline (10 per cent) and lysine or hydoxylysine (4 per cent) in the second and third positions (Prockop and Kirivikko 1984). The Gly–XY cassettes are crucial to the normal biophysical properties and correct winding of the collagen triple helix (as will be self-evident from patterns of mutations). The 3' end of the gene codes for a C-terminus of between 243 and 247 amino acids (depending upon collagen type). It has a short telopeptide sequence but is mostly globular with a highly conserved exon structure for all interstitial collagens (Yamada *et al*. 1984; Vuorio and de Crombrugghe 1990). It is coded by exons 49 to 52; exons 49 and 50 vary in size and exons 51 and 52 are constant for the various interstitial genes. In most of the interstitial families, exon 49 codes for between 45 and 63 bp but in a2XI it has reduced to only 15 bp.

The N-propeptides are more divergent, with substantial differences. *COL1A1*, *–1A2*, *–2A1*, and *–3A1* genes have cysteine-rich regions missing in *COL1A2*. The triple helical region of the N-propeptide also varies, being most complex in the *COL2A1* gene, where there are several extra exons.

Other similarities at the 5' and 3' ends of the gene include regulatory sequences upstream of the gene, such as CCAAT between -80 and -100 in the *COL1A1* and *–1A2* genes, respectively. There are also multiple binding sites upstream of this site for proteins such as transforming growth factor-b. There is also a highly conserved enhancer sequence within the first intron of *COL1A1, COL1A2*, and *COL3A1* collagens, sited between +418 and +1524 of *COL1A2* and +820 and +1602 of *COL1A1*. Deletion of certain promoter sequences causes positive or negative effects (Vuorio and de Crombrugghe 1990). There may also be enhancer sequences analogous to those b-globins, downstream of the 3'-coding sequences. The genes for non-interstitial collagens, whilst somewhat different, also have family resemblances to the interstitial collagens. Thus the *COL4* genes (*4A1, 4A2, 4A3, 4A4*, and *4A5*), which all contain several globular interruptions within the triple helix, are highly homologous to each other. They all possess long, C-terminal, globular domains (NC1) with smaller N-terminal globular domains (NC2). Electron-microscopic and biochemical evidence shows that N- and C-terminal interactions between adjacent type IV molecules produce a chicken-wire model (Vuorio and de Crombrugghe 1990). There are more than 20 globular interruptions to the Gly–XY triple-helical repeat, 12 of which are single amino-acid deletions whilst others are globular inserts of up to 24 amino acids, depending upon the particular chain type. Nevertheless, exon sizes conform reasonably well to the 9-bp model. The *COL4A1* and *COL4A2* genes run in opposite orientation and direction from head-to-tail contiguous promoter sequences.

Type VI collagen is highly disulphide-linked and has a short triple helix with globular ends forming microfibrils of antiparallel dimers that are laterally associated. Three distinct genes *COL6A1, 6A2*, and *6A3* code for a1, a2, and a3(VI) collagens. They are situated respectively on chromosomes 2 and 21. Again the helical exons are multiples of 9 bp.

Type VII collagen (*COL7A1*) has globular N- and C-terminal extensions with a central helix of 320-kDa monomers packed as antiparallel fibrils. The C-terminus forms almost half of the molecule. Collagen VIII is an endothelial protein, collagens IX, X, and XI are all components of cartilaginous matrix, whilst the vitreous humour of the eye contains collagens II, V, and XI. In the Stickler syndrome, abnormalities of the vitreous structure and function are combined with myopia and osteoarthritis (degenerative joint disease) caused by mutations of the *COL2A1* or *COL11A1* and *–A2* genes. Fibrillin 15 and collagen VI have also been identified in the vitreous humour. Other important interactions include the decoration of collagen I/III by collagen V. There are strong hints that disturbances of such ordered structures produce distinctly abnormalities. A good example is the cauliflower fibrils that occur in the skin of patients with Ehlers–Danlos syndrome I or II. In other cases, certain molecular components that might fit into the gaps between the quarter-staggered fibrils include calcium (collagen I), collagen V (collagen I and III), or proteoglycans or glycosaminoglycans and other compound fibres. Similar constraints also apply to the NC1 (N-propeptide) of collagen IX. Here, site-directed mutagenesis of this region impairs chain registration, selection, or helical stability. The function of the C- and N-propeptide in those collagens with significant non-helical interruptions is clear, as evidenced by the fibrillogenesis of collagens IV, VI, VII, VIII, and X.

Collagens IX, X, and XI also form components of cartilaginous matrix. In that regard the vitreous humour has proved to be an especially interesting tissue and disorders of the vitreous, of which Stickler's syndrome is a good example, are caused by either homogeneous or heterogeneous anomalies of collagen II, V, and XI. Here the homo- or heterogeneous primers of each or all components are all capable of producing a disorder of fibrillogenesis and collagen triple helices, and resultant vitreous disorder or fragility. The latter usually occurs in combination with myopia or osteoarthritis or degenerative joint disease. Furthermore, other unexpected components have also been identified in the vitreous, such as fibrillin 15 and collagen VI, and defects of both these genes are potentially pertinent to eye disorders. Other important mechanisms include the decoration of collagen I/III fibrils by collagen V (see above). Compound heterozygosity of collagen type I and III and *COL1A1*, *–1A2*, and *–3A1* mutations and collagen I/II and III genetic compounds *COL1A1, 1A2, 3A1*, and *2A1* could produce clinical disease, though no convincing examples have so far been observed. Lastly, enzyme deficiencies such as of lysylhydroxylase (Wenstrup *et al*. 1989) and lysyloxidase (Byers *et al*. 1980) have been described, though the procollagen peptidase deficiency in humans originally postulated by Lichtenstein *et al*. (1974) has only very recently been identified (Smith *et al*. 1992; Nusgens *et al*. 1992). The animal equivalent, in contrast, is very well known (Lapiere *et al*. 1971). The disturbed morphology of the human fibril is virtually identical to that of the cow, sheep, and cat, and is produced by persistence of the N-terminal propeptide sequence in each component of the triple helix of collagen type I. Here the inheritance is autosomal-recessive. In contrast the autosomal-dominant, N-terminal, structural mutants *of COL1A1* and *COL1A2*, although allowing PNa components to persist in one or two a-chains, produce much less severe structural consequences than the triple a-chain enzyme defect.

The molecular analysis of *COL2A1* mutations lagged initially because of the inherent difficulty in culturing chondrocytes and obtaining cartilage samples from which the mutant proteins could be isolated. Those mutations characterized to date closely resemble type I and III mutants, both in the pattern of the protein changes (Murray *et al*. 1989) and from gene-sequencing data (Lee *et al*. 1988; Tiller *et al*. 1990; Kuivaniemi *et al*. 1991). They are usually helical glycine substitutions or deletions. Collagen type II defects cause a variety of inherited disorders including spondyloepiphyseal dysplasia congenita, Stickler syndrome, and osteoarthritis (Palotie *et al*. 1989; Anderson *et al*. 1990; Knowlton *et al*. 1990).

Type IX and XII collagens are highly homologous in both exon structure and protein folding, whilst collagen type XIII is secreted by keratinocytes but also occurs in skin, bone, intestine, and muscle. Collagens IV, VI, IX, X, XII, and XIII have all been cloned and sequenced. Type VIII and XIII still await more detailed analysis. Interesting supramolecular interactions of collagen types (at the fibrillar level) have been observed. Thus collagens I and III and I and V form compound fibrils, whilst collagens II and XI coassociate with collagen IX, which coats the surface of type II. Similarly, collagen XII decorates the surface of type I collagen. Such complex combinations presumably interact with themselves and other connective components. Collagens XIV to XVII have also been recently described. The latter is the so-called bullous pemphigoid-associated antigen, mutations of which produce a relatively mild variant of dystrophic benign epidermolysis bullosa.

Collagen folding and self-assembly

The collagen triple helix forms an unique protein structure containing multiple repeating polymers of (GlyXY) $_n$. The flexible lysine bonds are interspersed with inflexible proline or hydroxyproline to confer a specific helical conformation as defined by hydrogen bonding and water bridges linking individual a-chains to each other (Prockop and Kivirriko 1984; Byers 1990). Furthermore, for proper triple-helix formation it is essential that each of the three a-chains is in correct register. If not, then redundant loops and mismatches could seriously impair the processing of the procollagen extensions. When the triple helix is correctly wound it can be incorporated into mature collagen fibrils. This process may be faulty at several biosynthetic steps: for example, in precursor pro-a-chains with N- and C-terminal propeptides the charged interactions between the C-propeptides are largely stabilized by interchain disulphide bonds. This ensures the correct alignment of individual a-chains to one another, so ensuring that the first few GlyXY triplets adjacent to the C-terminus are correctly aligned (in the three separate a-chains). It is also essential that each component of the triple helix be properly aligned if proper superhelical winding is to take place. Helical winding next propagates towards the N-terminal end, twisting the three chains into a tightly wound, rod-like structure (Byers 1990).

There are also several important post-translational modifications, such as changing of GlyXPro to GlyXhydroxyproline by the enzyme prolyl-4-hydroxylase (Kivirriko and Myllyla 1982). This stabilizes the collagen helix at ambient body temperature, delaying crystallization of the 'in-register', C-propeptide-associated (potential) triple helix. Other important steps include limited lysyl hydroxylations, some of which have galactose or glucosylgalactose residues added. Hydroxylysine stabilizes the cross-links between collagen monomers. The chains cannot be efficiently folded without hydroxylation but the act of folding inhibits this modification, the process therefore being self-regulatory. It having been wound into an optimal configuration, various cleavage enzymes (Byers 1990) dramatically alter the solubility of the precursor molecule. Spontaneous polymerization into fibrils (of individual triple-helical collagen a-helices) then follows in a highly ordered, structured fashion. Quite possibly the self-assembly of collagen fibrils is also affected by other components of the extracellular matrix.

The information outlined above has been accumulated from basic biochemistry, biophysics, molecular biology, and studies of certain specific collagen mutations. It applies especially to the interstitial collagen types I [a1(I) and a2(I)], II (a1(II) $_3$, and III (a1(III)$_3$ proteins. Very probably, similar considerations and constraints apply to other interstitial collagens, such as collagen V (a1- and a2-chains) and collagen XI (a1- and a2-chains). It is also now apparent that defects of *COL5A1*, *–5A2*, *–11A1*, and *–11A2* also produce clinically detectable disease. More recently the *COL5A1* gene has been linked to both Ehlers–Danlos syndromes I and II (Burrows *et al*. 1996; Loughlin *et al*. 1995), and a convincing mutation of *COL5A1* demonstrated in a single patient with Ehlers–Danlos syndrome II (Nicholls *et al*. 1994). Latterly a translocation through the middle of the *COL5A1* gene has also been shown to cause Ehlers–Danlos syndrome II (Torriello *et al*. 1996). Similarly, *COL4* mutations are being categorized in Alport syndrome. Deletions and substitutions of *COL4A5* cause sex-linked recessive Alport syndrome (Barker *et al*. 1990). It is obvious that much still remains to be discovered. In the context of rheumatology there are at least three other cartilage collagen genes (apart from collagen II) with the potential to produce cartilage-related inherited diseases.

Mutation of collagen genes and resulting proteins

Mutations of the genes coding for interstitial collagens type I, II, and III are extremely diverse, and, with certain rare exceptions, virtually private to each family. Amino acid substitutions are particularly disruptive when altering triple-helical glycines, but both the location and type of substitution are important. To date only occasional second- or third-position (X + Y) mutations have been detected and even then their effects are not completely clear. Splice-junction point mutations (Watson *et al*. 1987; Weatherall 1991) are an important mechanism for collagen deletions (analogous to haemoglobin and other splice mutations). Other rarer abnormalities include large (Superti-Furga *et al*. 1988) and small inframe deletions, also nonsense or frameshift mutations (Nicholls *et al*. 1984b; Pihlajaniemi *et al*. 1984). The latter are less common than helical substitutions (Byers 1990; Beighton *et al*. 1991; Kuivaniemi *et al*. 1991). All dramatically affect protein structure or secretion, although overmodified secreted products are, in general, common in type I defects (Bonadio and Byers 1985) whilst overmodified retained protein is common in *COL3A1* mutations (Pope *et al*. 1988a, Pope *et al*. 1988b). Generally speaking the clinical phenotype is much more variable in mutations of collagen type I, ranging from virtual normality (with mild osteoporosis) to congenital fractures with lethal disease in babies (Sillence *et al*. 1984). Defects of collagen type III, on the other hand, whilst more often lethal, exert their main effects in late adolescence, adult or middle life when lethal or life-threatening arterial rupture is commonplace (Pope *et al*. 1988a, Pope *et al*. 1988b; Pope *et al*. 1996). Beighton *et al*. (1992) have attempted to correlate clinical chemical and molecular phenotypes.

Although there are strong clinical (phenotypical) similarities between the different subgroups [except for Ehlers–Danlos syndrome type VII, in which exon-6 splicing defects of a(I) and a2(I) have so far been consistent] (Weil *et al*. 1989a; Weil *et al*. 1989b; Weil *et al*. 1988; Nicholls *et al*. 1991; Pope 1991; Vasan *et al*. 1991), most families with mutations of collagen type I and III (with only one or two rare exceptions) have private mutations (Byers 1990; Kuivaniemi *et al*. 1991). To date nearly all type I mutations have been observed in osteogenesis imperfecta/Ehlers–Danlos syndrome VII; only 5 per cent are recurrent between families. No *COL3A1* mutants have recurred (i.e. so far each family is unique). The molecular analysis of *COL2A1*, which (as outlined above) had previously lagged behind the others, has recently advanced spectacularly with more than 50 recorded mutations (Vikkula *et al*. 1994). Again, private mutations are the norm. As in collagen type I and III mutants, many are glycine substitutions or deletions, with similarly disturbed protein patterns (Murray *et al*. 1989) and gene-sequencing data (Lee *et al*. 1988; Tiller *et al*. 1990; Kuivaniemi *et al*. 1991). There is also impressive evidence that associates *COL2A1* mutations to various inherited disorders such as spondyloepiphyseal dysplasia congenita, the Stickler syndrome, and osteoarthritis (Palotie *et al*. 1989; Anderson *et al*. 1990; Knowlton *et al*. 1990;).

There have been equally spectacular advances in the analysis of Alport syndrome. These are in the linked *COL4A5* gene and also in the autosomal *COL4A3* and *COL4A4* genes. Similarly, gene linkage and mutational analysis of *COL7A1* mutants have rapidly advanced, with linkage of the gene to both the autosomal-dominant and -recessive epidermolysis bullosa dystrophica, and *COL7A1* mutations in both variants. Studies include combinations of cDNA and genomic sequencing of amplified gene fragments (Al-Imara *et al*. 1992; Dunnill *et al*. 1994a; Hovnanian *et al*. 1994). Usually cDNA is amplified as four or five fragments, whilst the 118 *COL7A1* genomic exon sequences are amplified with 70 pairs of separate primers followed by a combination of HDE, SSCP or DGGE analysis. Various null alleles (usually stop-codon mutants), but relatively few glycine helical substitutions or splice junction deletions, have been detected.

There has also been rapid progress in the analysis of collagen IX and X mutants, which cause disorders such as spondylometaphyseal and Schmid-type chondrodysplasias, respectively. Other important recent advances include the identification of a collagen Va1 and -a2 error in either mice or man (*COL5A1* and *COL5A2* mutants), each of which produces specific connective-tissue diseases. For example, the common disease (Ehlers–Danlos syndrome II) is caused by an exon skip of the *COL5A1* (Nicholls *et al*. 1994), whilst the mouse model is homozygous for an N-terminal deletion of the *COL5A2* gene (Andrikopoulos *et al*. 1995). Recently, any doubts about the sporadic nature of the original *COL5A1* mutant have been resolved with the demonstration of linkage to *COL5A1* in a very large British pedigree (Loughlin *et al*. 1995). Also recently, we have linked *COL5A1* to families with mixed Ehlers–Danlos syndrome I/II phenotypes, suggesting that Ehlers–Danlos syndrome I and II are allelic (at least some of the time) (Burrows *et al*. 1996).

Type I collagen mutations

Type I collagen is widespread in bones (particularly the metaphyses and diaphyses of long bones), where it is the major matrix protein. It also occurs in tendons, joint capsules, the dermis, and blood vessels. Not surprisingly, clinical abnormalities include weakened bones with variably diminished bone matrix proportional to the severity of the osteogenesis imperfecta (Fig. 1), which ranges from mild, barely detectable familial osteoporosis to lethal, short-limbed dwarfism (Pope *et al*. 1988b). Sometimes ligaments, joint capsules, scleral thickness, dentine, heart valves, and dermal thickness are affected as judged by joint laxity, blue sclerae, dentinogenesis imperfecta, and prolapsed or incompetent heart valves, respectively. Very severe joint laxity, with or without skin fragility though usually without osteoporosis, is more typical of Ehlers–Danlos syndrome VII, which is accompanied by short stature. In heterozygotes, mutations outside the main collagen helix produce either mild osteogenesis imperfecta or osteoporosis, but may be very severe in homozygotes (Byers 1990). One unusual patient with Marfan syndrome with an abnormal a2(I) collagen protein had a second-position mutation (Phillips *et al*. 1991) affecting her father and herself [a second, somewhat similar patient was briefly mentioned (Godfrey *et al*. 1990)]. It is unclear whether the father in the former family is a true Marfan or a mosaic (Phillips *et al*. 1991). Furthermore, the general relevance is also unclear as no other Marfan family has been linked to the *COL1A2* or any other collagen gene. There are three possible explanations: the a2 genotype produces a phenocopy closely similar to authentic Marfan syndrome, the a2 association is coincident, or the a2 association is genuine and Marfan syndrome is allelically heterogenous. There is very little evidence of the last, however, and with one exception all Marfan syndrome families so far tested have been linked to fibrillin (Blanton *et al*. 1990; Kainulainen *et al*. 1990; Kainulainen *et al*. 1991; Tsipouras *et al*. 1991; Dietz *et al*. 1995; Kainulainen *et al*. 1994; Nijbroek *et al*. 1995). The one exception is highly controversial (Dietz *et al*. 1995; Boileau *et al*. 1995): a very large French family is linked to 3p24.2–p25 (Collod *et al*. 1995) but whether or not the family has Marfan syndrome is hotly debated (Dietz *et al*. 1995; Boileau *et al*. 1995).

Clinical features of osteogenesis imperfecta

The clinical and pathological changes are those of an inherited osteoporosis, and in the worse forms of the disease the bone matrix is severely disorganized and

depleted (Pope *et al*. 1989) (Fig. 1). The periosteum persists normally under these conditions but there is often minor distortion and disorganization of the epiphyseal plates. Radiological changes range from mild osteoporosis with occasional fractures to a widespread and severe skeletal abnormality with absent skull calcification, multiple rib fractures, polyspondyly with spinal collapse, and distorted, multiply fractured bones that are either modelled or unmodelled depending upon clinical subclassification (Pope *et al*. 1989) (Fig. 2).

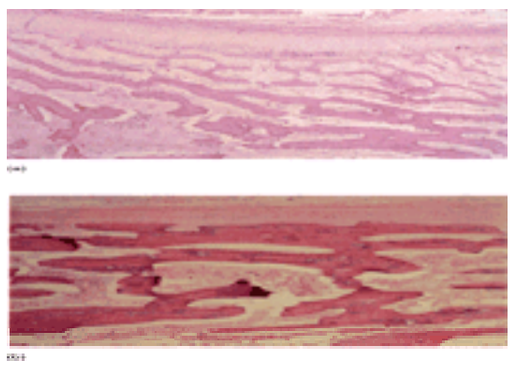


Fig. 1 Histological features of fetal bone in mild osteogenesis imperfecta (a) compared with normal control (b) at 24 weeks' gestation. The trabecular bone is obviously thinner in the osteogenesis imperfecta sample. Original magnification × 4.

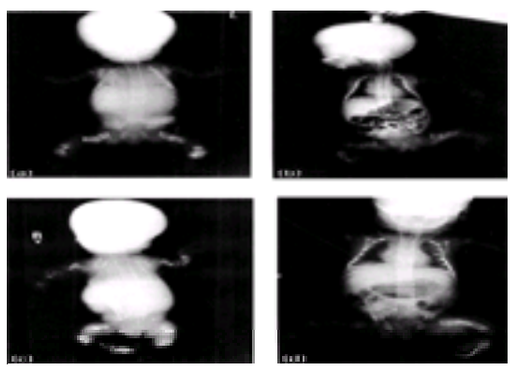


Fig. 2 Radiological phenotype of babies affected by osteogenesis imperfecta ranging from (a) lethally affected, generalized involvement of ribs and unmodelled long bones (Sillence type IIa) to various combinations of unmodelled long bones and nodular or unbroken ribs (b–d). The least severe phenotype (sometimes called Sillence type IIc or III) has gracile, generally unfractured bones but considerable distortion and fractures of the lower limbs (d).

The clinical phenotype reflects these facts, so that affected patients vary from normal-looking children with occasional fractures beginning in infancy to severely crippled, short-limbed dwarves (dying either *in utero* or perinatally). Their variability is recognized by the Sillence classification (Sillence *et al*. 1979), which separates two mild and autosomal-dominant disorders (types I and IV) from two lethal or severely crippling disorders (types II and III) (Fig. 3). Both autosomal-dominant and -recessive inheritance occurs.

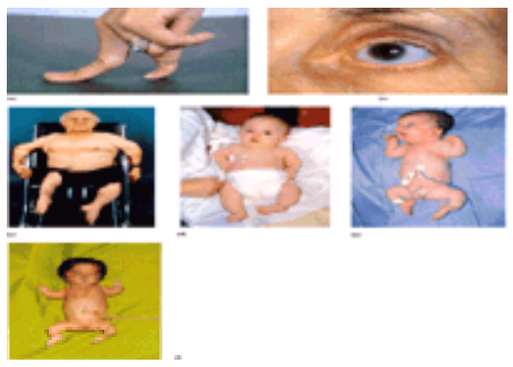


Fig. 3 (a) Many patients with Sillence type I osteogenesis imperfecta have generalized joint laxity (shown here at finger tips); (b) here, this was associated with premature corneal arcus caused by the osteogenesis imperfecta. (c) Severely crippling Sillence type II osteogenesis imperfecta—this patient has a relatively normal face, shortened upper-limb segments, a broadened and distorted chest, and convex, distorted lower limbs: he had a Gly–Cys mutation at position 415 of the collagen a1(I) chain. (d) Severely affected patient with Sillence type IIb osteogenesis imperfecta. (e) Lethally affected patient with Sillence type IIa osteogenesis imperfecta.

Types I and IV osteogenesis imperfecta

Generally both are mild disorders presenting in infancy or childhood with occasional fractures. Although the defect can present at birth, childhood fractures that improve at puberty are more usual. Type I osteogenesis imperfecta differs from type IV only by the deep blue sclerae, which in patients with type IV are pale grey or white. Common features include autosomal-dominant inheritance (multiple affected generations with transmission through either sex), variably short stature, and childhood fractures improving at adolescence but recurring after the menopause. There are occasional, rarer variants in which fractures are more common in females but rare in males. Rather than improve (in number) in adolescence they continue unabated throughout adulthood. In these families, males have fractures only rarely. Many different mutations have now been observed (Fig. 4). Deafness or dentinogenesis imperfecta segregate separately, sparing some families whilst rampant in others. In families with dentinogenesis imperfecta the primary dentition is uniformly opalescent (slate or brownish coloured), whilst the changes in the secondary dentition are usually patchy with the lower incisors and first permanent premolars particularly prone. Dental radiographs characteristically show short, tapering roots and expanded crowns. Histologically, there are sparse or agglomerated dentinal tubules, both by light and electron microscopy (Fig. 5). The dentinogenesis imperfecta/osteogenesis imperfecta disorder is separate from isolated dentinogenesis imperfecta, which affects the primary and secondary dentition equally. More rarely it is associated with minor joint laxity similar to Ehlers–Danlos syndrome III (Komorowska *et al*. 1989) and is also probably a connective tissue disorder. The gene has been located to chromosome 4.

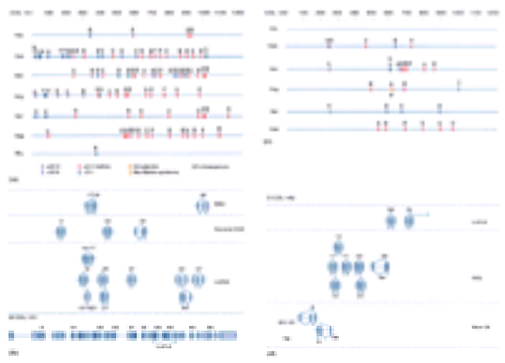


Fig. 4 Distribution of type I collagen mutations arranged by glycine substitutions for COL1A1 and COL1A2 (a and b) and by exon skips (c and d). (Updated from information by Byers (1990), Kuivaniemi *et al.* (1991), and Cole and Dalgleish (1995).) Different colours indicate the severity of osteogenesis imperfecta and exon skips are denoted by superscripts. OI = osteogenesis imperfecta, ED = Ehlers-Danlos syndrome, op = osteoporosis.

Fig. 5 (a) Typically patchy distribution of dentinogenesis imperfecta in adolescent teeth of mild Sillence type I osteogenesis imperfecta; the lower incisors and premolars are particularly affected. (b) Dental radiographs in typical osteogenesis imperfecta showing shortened, tapering roots, obliterated pulp cavities, and tulip-like crowns. (c) This mother and baby both have mild Sillence type I osteogenesis imperfecta. Note mother's triangular head. Both individuals have blue sclerae.

Sillence type II osteogenesis imperfecta is a severely crippling and often lethal disease (Fig. 2(a, b) and Fig. 3(b). Affected babies die either *in utero* or perinatally from a variety of causes such as cerebral haemorrhage, strangulated hernias, or chest infections. Some of them occasionally have severe congenital heart disease. The majority who die early do so from chest infections and pulmonary restriction, which complicate the widespread fractures. Sillence described three main types, IIa, IIb and IId, of which the first two have unmodelled, rectangular limb bones, especially the femurs and humeri. Osteogenesis imperfecta type IIa has characteristically widened and multiply fractured ribs, which have a generalized, feathered appearance. Osteogenesis imperfecta type IIb, in contrast, shows either nodules (from callus formation) or unbroken but thinned and slightly misshapen ribs. After slight modification of Sillence's classification by Thompson *et al*. (1987) (Fig. 2(a,b)), this type was called osteogenesis imperfecta III instead of type IIc as originally proposed by Sillence. Osteogenesis imperfecta IIc (Thompson *et al*. 1987) describes babies with very much more severely twisted (sometimes modelled) limb bones, distinct from osteogenesis imperfecta IIa and b. Similarly, the ribs are thinned, nodular, or sometimes contain multiple fractures but are always thinner and more distorted than those of IIa and b. With the thin-boned infantile variant, osteogenesis imperfecta III, most will survive to adulthood. Here there are slightly distorted ribs that are gracile rather than widened, whilst there are well-modelled limb bones often with angulated, U-shaped tibiae. This subset might be better named Sillence IId osteogenesis imperfecta in babies but osteogenesis imperfecta type III in surviving adults (and some certainly die in infancy or early childhood). Osteogenesis imperfecta type IIa is invariably lethal, as are many forms of IIb and IIc, so there are insufficient data to classify adult survivors. Probably osteogenesis imperfecta type III would be an appropriate designation for most of them. In practice the Sillence IIa–c classification is a useful guide for what is actually a continuous spectrum of clinical changes with broad-boned lethal osteogenesis imperfecta at one extreme and thin-boned osteogenesis imperfecta at the other. Such clinical variability is quite consistent with a wide range of individual and unique collagen mutations. Inheritance is usually autosomal-dominant, and most affected babies have arisen from sporadic new mutations, some of which (10 per cent) have parents with somatic gonadal mosaics. Multiple reoccurrences in this circumstance simulate autosomal-recessive inheritance (Byers 1990). Authentic autosomal-recessive inheritance is also well documented (Nicholls *et al*. 1984; Byers 1990) but distinguishable from mosaicism only if the gene defect is fully sequenced. There is also suggestive evidence that some osteogenesis imperfecta type IId phenotypes are unlinked to both *COL1A1* and *COL1A2*. Proteins other than type I collagen must therefore participate.

Collagen type I mutations in osteogenesis imperfecta

Bone matrix protein is made largely of type I collagen, with small amounts of type V and with smaller traces of type III. The last comes from vascular structures such as arteries, veins and bone marrow, stroma, and capillaries (Pope *et al*. 1988a, Pope *et al*. 1988b). To date nearly 150 mutations of type I collagen have been detected, which include point mutations, deletions, insertions, and a number of splicing mutations, as outlined below (Byers 1990; Kuivaniemi *et al*. 1991; Cole and Dalgleish 1995) (see Fig. 4(a,b)). The Sillence classification only roughly correlates with the molecular pathology. Thus milder, Sillence type I and IV osteogenesis imperfecta is caused by silent alleles, point mutations or exon skipping of *COL1A1* or *–1A2* in relatively harmless parts of the molecule. Lethal and severely crippling osteogenesis imperfecta mostly arises from glycine substitutions of either *COL1A1* or *–1A2* genes. These severely disrupt the collagen helix either by producing homozygous, C-terminal mutations or much more commonly by forming disruptive helical substitutions that consequently impair helical conformation (or even produce kinking distant to the substitution site). A good example is Gly–Cys748 of *COL1A1* in which processing of the distant N-propeptide is subsequently severely compromised (Vogel *et al*. 1988). Thus glycine substitutions are often position-specific, either interfering with helical folding (as measured by proteinase susceptibility) or impairing protein crystallization with normal folding. Mutations also produce the so-called protein suicide effect, in which the presence of one or more abnormal proa chains within triple helices causes both normal and abnormal products to be intracellularly degraded (Prockop and Kivirriko 1984; Prockop 1990). Thus either one or two mutant molecules in a heterotrimer of wild-type and mutant products destroy their companion normal components too. Only wild-type a1(I)2a2 heterotrimers (12 per cent of the whole) survive to be successfully secreted. Occasionally both mutations (shortened and normal heterotrimers) escape. Domains of micro-unfolding present as ladders of varying molecular size after proteinase digestion. This implies heterogeneity of protein function within the triple helix.

Point mutations (see Fig. 4(a))

Gly–Cys substitutions clearly illustrate how mutational location affects clinical phenotype. Thus two similar Gly–Cys substitutions, both within the C-terminal, CNBr fragment a1(I)CB6, cause widely disparate clinical phenotypes. The one causes very mild, autosomal-dominant Sillence type I whilst the other causes lethal, Sillence type II osteogenesis imperfecta. The effect of the two errors correlates precisely with their location. Both produce a novel, higher molecular-weight, disulphide-linked dimer running above normal collagen a1(I) chains. Both individual peptide maps showed a pattern derived from collagen a1(I) chains, with a1(I)CB6 running as a dimer; after reduction the dimer disappeared whilst a1(I)CB6 reappeared (Steinmann *et al*. 1984; Pope *et al*. 1988b) (Fig. 6). The lethal a1(I) dimer melted at 38°C whereas the dimer from the mild osteogenesis imperfecta melted normally at 41°C (Fig. 7). This disturbed helicity in the lethal but not the mild mutation. At the time the location of both mutations to *CB6* belied their very different clinical phenotypes. Possible explanations included a Gly–Cys substitution in the severely affected patient and a second- or third-position Arg or Ser–Cys substitution in the milder case. Alternatively, a second silent mutation might have produced a combined lethal effect. Cloning and sequencing the relevant gene fragments showed that both mutations were GCG–TCG; the lethal at position 988 and the mild at position 1017 only 30 bp (10 amino acids) away. The mild mutation substituted the first glycine outside the collagen triple helix (with unimportant biophysical effects), whilst the 988 mutation changed the tenth glycine from the C-terminus of the a1(I) helix and is consequently lethal. Whereas the 988 substitution produced a protease-susceptible collagen that melted at 2°C lower than normal, the 1017 mutation produced an a1(I) collagen protein that melted normally (Fig. 7). Subsequently, at least 13 other cysteine substitutions have been identified at scattered locations ranging from near the N-terminus to as far 3' as from the C-terminus (see Fig. 4(a)). Substitutions at the C-terminus of amino acid 690 to the end of the helix have been lethal, with those between amino acids 391 and 559 severely crippling. In contrast, N-terminal ('left-hand') mutations between 94 and 223 have been uniformly mild (Byers 1990; Kuivaniemi *et al*. 1991; Cole and Dalgleish 1995). This is consistent with two other mutations that we have studied, the one at residue 19 close to the N-terminus producing premature osteoporosis, and the other at amino acid 415 producing a mixed Sillence type III/IV phenotype. Those cysteine substitutions of a2(I) collagen produce less predictable phenotypes. Thus, one at position 259 had an osteogenesis

imperfecta type III phenotype whilst at position 472 there was a genetic lethal and at position 646 a severe type IV phenotype (Byers 1990; Kuivaniemi *et al*. 1991). Such differences are partly due to ascertainment bias. For example, whether severe crippling affects survivors or premature lethality kills younger babies is partly determined by randomly encountered respiratory pathogens or other complications such as respiratory restriction, cerebral haemorrhage, or congenital heart disease.

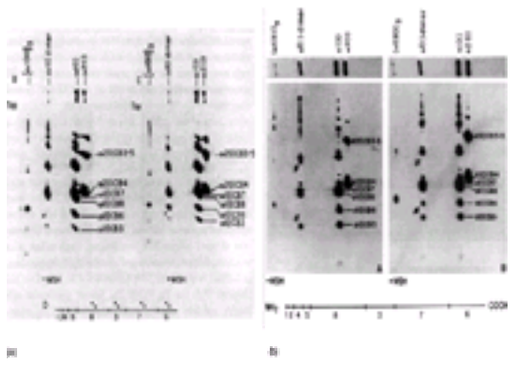


Fig. 6 Two-dimensional gel electrophoresis of radiolabelled collagens from (a) a lethally affected patient described by Steinmann *et al*. (1984) and (b) a mildly affected patient described by Steinmann *et al*. (1986). The lethally affected individual has tilted 'smiling' of all CNBr peptides and the mildly affected has horizontal, 'non-smiling' gels, yet both have mutations localized to a1(I)CB6 and produce an additional, higher molecular-weight dimer running heavier than the a1(I) chains. This is obvious when the gels are compared before (-MSH) and after (+MSH) reduction with mercaptoethanol. The order of peptides from N- to C-termini is shown diagrammatically.

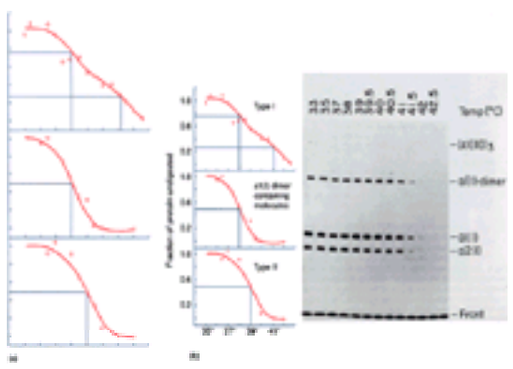


Fig. 7 Comparison of the melting profiles of the lethal and mild osteogenesis imperfecta (from Steinmann *et al*. 1984 and Steinmann *et al*. 1986). The lethally affected patient has differing melting curves for the cysteine-containing and normal a1(I) chains (shown as a graph). (a) The cysteine-containing a1(I) chain melts at 2°C earlier than its wild-type; in contrast, the cysteine-containing chain from the mildly affected individual melts identically to the wild-type a1(I) chains. The mutational position therefore influences the proteinase susceptibility of the two different mutant chains. Whereas the melting curve is plotted as a ratio in (a), similar data are shown at each time point in (b).

Other substitutions (see Fig. 4)

To date, there are more than 150 mutations of type I collagen a1 or a2 genes (*COL1A1* and *COL1A2*), including those already discussed. The same broad generalization applies, that is, alterations closest to the C-terminus are the most disruptive (Byers 1990; Kuivaniemi *et al*. 1991). Glycine to cysteine, alanine, serine, arginine, valine, and aspartate have all been identified in the *COL1A1* gene and glycine, valine, arginine, and aspartate in the *COL1A2* gene. Severity varies with the position and also the type of substitution (Fig. 4(a)). For example, Gly–Asp mutations are universally lethal, even at position 97, which lies close to the N-terminus of a1(I) collagen up to position 883 much nearer the C-terminus. Similarly, the two recorded a2(I) Gly–Asp substitutions at positions 547 and 976 are also lethal. Somewhat paradoxically, since they are such small amino acids Gly–Ser mutations can be lethal if located at the C-terminus. Gly–Val substitutions (also rather rare) are lethal wherever they occur. Up to 1991, no Gly–Glu substitutions had been recorded for either type I gene nor have any appeared subsequently. On the other hand, several have now been identified in the *COL3A1* gene (see below). Clinical severity is therefore determined by helical position, the nature of the substitution, and other less tangible influences such as local interaction with neighbouring molecules or local domain effects, whereby neighbouring changes can have widely disparate clinical consequences (Byers 1990). Many questions still remain and the distant consequences of specific mutations (perhaps caused by three-dimensional protein-folding effects) also require clarification. For example, as mentioned earlier, the Gly–Cys748 mutation introduces a local kink at this position, yet mysteriously interferes with the function of the N-propeptide more than 500 amino acids away (Vogel *et al*. 1988).

Similarly, studies of collagenase A and B fragments have identified areas of micro-unfolding between 637 and 775 of collagen a1(I). Two nearby Gly–Ser substitutions in positions 598 and 631 show an adjacent region between amino acids 500 to 631 (Westerhausen *et al*. 1990). Very probably others lie elsewhere in the collagen helix.

Other important questions include the mechanisms for chain selection. Thus, why should a1(I) $_2$a2 be formed in preference to a1 trimer? Presumably, but for no clear reason, a2 is required for bone calcification and organization but is less essential in tissues such as skin or sclerae.

Exon-skipping deletions (Fig. 4(b))

Most vertebrate structural genes, including those for collagen types I and III, are interrupted by non-coding introns of between 80 and 10000 bp. Collagen introns lie at the lower end of this scale. The primary mRNA transcript is transcribed from genomic DNA. It is then very precisely processed and the 52 constituent exons rejoined correctly. These include the 42 triple-helical exons, the four C-terminal exons, and the six exons of the N-propeptide. The 5' and 3'ends of each intron have invariant splice-junction donor sites (at the 3' end of the 5'-most exon) and acceptor sites at the 5' end of the 3'-most exon. The former is usually GT (in genomic sequence) and the latter AG. Some pre-mRNA transcripts are spliced into several independent transcripts by exon shuffling. Examples include pituitary hormones, viral proteins, and the immunoglobulin variable regions (Watson *et al*. 1987). Mutations of the conserved junctional sequences disrupt correct splicing of those complicated genes, with many introns. Collagen demonstrates many such examples. Single nucleotide changes can destroy a normal splice site or activate hidden (cryptic) alternatives. Thus shortened, extended, extra, or missing exons result. If the 3' polyadenylation sequence is missing, the translation continues to the next downstream 3' equivalent. Such variability is a potent means of evolutionary change.

The gradient of collagen exon-skipping deletions (Fig. 4(b)) lies along the collagen genomic sequence in a similar fashion to point mutations (Byers 1990; Kuivaniemi *et al*. 1990, Kuivaniemi *et al*. 1991). Thus several C-terminal *COL1A1* mutations have been lethal (exons 43 and 44), whilst mid-gene defects are severe or lethal (exons 27 and 30). In contrast, N-terminal skipping can be either mild (17) or lethal (14), and similar considerations apply to the *COL1A2* exon skips. Up to 1991 there had been 18 *COL1A1* and *COL1A2* exon skips and now (in 1995) there are 36 skips (Byers 1990; Kuivaniemi *et al*. 1991; Cole and Dalgleish 1995). We have observed an atypical, large family with autosomal-dominant osteogenesis imperfecta who have mis-splicing of exon 9 of the *COL1A2* gene with a shortened a2-protein (Nicholls *et al*. 1991). Protein mapping was unhelpful but amplification by PCR showed there was an 11-bp insertion spanning the intron–exon junction of both exon and intron 9. This caused mis-splicing of exon 9, with a shortened a2(I) protein missing 18 amino acids.

Null alleles commonly cause Sillence type I osteogenesis imperfecta (Willing *et al*. 1992) and mouse models with *COL1A1* mutations have a phenotype similar to human osteogenesis imperfecta (Khillan *et al*. 1991).

Mutation with milder phenotypes that are neither deletions nor caused by exon skipping

Many Sillence type I osteogenesis imperfecta patients produce less type I collagen than normal. This is easily detectable by polyacrylamide gel electrophoresis as a relatively higher abundance of type III collagen protein (Fig. 8(a)). Type I collagen is normally secreted and not overmodified, and has no structural abnormalities. Theoretically, mutations affecting gene expression, translational efficiency, stability of proa chains, and molecular assembly could all produce relatively low amounts of type I collagen protein, though in practice stop-codon mutations are common. Steady-state levels of cytoplasmic mRNA are often lowered but may also be normal (Byers 1990). Gene linkage to *COL1A1* is commonplace in up to 95 per cent of this phenotype (Sykes *et al*. 1990). Occasionally, point mutations occur (as in our Gly–Cys family). Willing *et al*. (1990) observed complete failure of incorporation of the proa1(I) chain in a patient with a small deletion, causing abnormal chain elongation. Occasionally, these are structural *COL1A1* alterations occurring with a very similar protein analysis to those causing lethal, Sillence type II osteogenesis imperfecta (Fig. 8(b)). Presumably, abnormalities of protein migration indicate more subtle underlying changes that determine other crucial differences in the clinical phenotype. Null mutants are distinguishable from wild-type genes by reverse transcriptase-PCR of allele-specific, intragenic cDNA polymorphisms. This allows confident discrimination between expressed or non-expressed alleles (Willing *et al*. 1990).

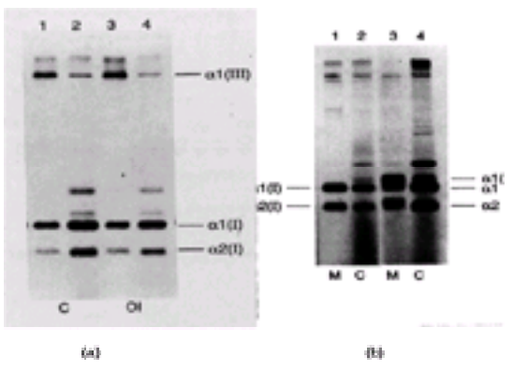


Fig. 8 Radiolabelled collagen profiles. (a) Low type I collagen producing mild osteogenesis imperfecta: the intensity of the unreduced a1(III) band distinguishes affected individuals [relatively more type III collagen secreted into the medium (track 3) from unaffected (track 1)]. There is no difference in the cell-layer collagens (tracks 2 and 4). C, control; OI, osteogenesis imperfecta. (b) Double a1(I) components [one a1(I)H, of higher molecular weight] that occur in some rare, mild OI families. In this instance a double a1(I) band is secreted into the medium (M) whereas in the cell layer(C) mainly a1(I) is retained (tracks 3 and 4). The equivalent band retained in the cell layer of the control (track 2) is type V collagen.

COL1A2 mutations that lower the synthesis of a2(I) will also affect the ratio of collagen III to I. Here the a1:a2 ratios will exceed 2, in contrast to depletion of a1(I), with ratios near to 1. Those a chains not incorporated into heterotrimers could either form a1(I) trimers [a1(I) chains] or self-degrade (a2 chains). Exceptionally chains with apparent a2 dimers can form anomalously. It is unclear whether these are stabilized by inter- or intramolecular cross-links, but the clinical phenotype is not osteogenesis imperfecta.

A unique, 4-bp deletion in *COL1A2* was clinically silent in the heterozygote male whilst the carrier female manifested premature osteoporosis. In contrast, the homozygous affected child had very severe, Sillence type III osteogenesis imperfecta. Both heterozygous parents also showed increased joint laxity and slightly thin bones (Nicholls *et al*. 1984b). The 4-bp deletion produced a frameshift with a nonsense downstream sequence resulting in a defective sequence of the last 31 amino acids. It is also possible that deletions at a2 might cause entirely different (non-osteogenesis imperfecta) clinical phenotypes.

Gonadal mosaicism

There are now many examples of single *COL1A1* or *COL1A2* mutations causing lethal osteogenesis imperfecta. Usually, these are sporadic, new, autosomal-dominant mutations with negligible risks for recurrence. Occasionally, there are extraordinary families with four or more affected children but normal parents, suggesting the possibility of germ-line mosaicism (Byers 1990). Such families were previously considered autosomal-recessive Compound heterozygosity for two separate collagen mutations is also possible. Thus both Sillence osteogenesis imperfecta types IIb and c were considered in this category. Most are gonadal mosaics, as judged by the clinical expression of a milder (non-lethal) phenotype in the carrier parent, the properties of the mutant protein and expression of the mutant gene in fibroblasts, hair bulbs, saliva, lymphocytes, seminal fluid, or ovarian biopsies. A good example is Gly–Cys at position 472 of the *COL1A2* gene. Here a man with clinically mild osteogenesis imperfecta fathered separate children with osteogenesis imperfecta from two different mothers. The mutant gene occurred in 33 per cent of sperm, 67 per cent of lymphocytes, and 100 per cent of dermal fibroblasts. A second example is Gly–Cys904 in *COL1A1*. Here, the overmodified collagen was protease-susceptible at a low denaturation temperature. The proband's mother was a mixed somatic/gonadal mosaic with mild, Sillence type I osteogenesis imperfecta (including a triangular face and blue sclerae). A third example is Gly–Arg550. Here a lethally affected fetus had a mildly mosaic father with the osteogenesis imperfecta type IV phenotype.

Mouse models

Mice with severe disruption or severely shortened forms of collagen type I reproduce a clinical phenotype very similar to or identical with osteogenesis imperfecta (Jaenisch *et al*. 1983; Khillan *et al*. 1990).

Osteogenesis imperfecta with neither *COL1A1* nor *COL1A2* mutations

We have observed several examples, some with Sillence type IId osteogenesis imperfecta. One osteoporotic but otherwise normal mother partially expressed a doublet of normal and overmodified mutant a1(I) chains; she has had three children lethally or severely affected by osteogenesis imperfecta. In a second family of osteogenesis imperfecta, two clinically normal parents have produced six consecutively affected children with type IId (D. A. Gibbs *et al*., 1991; personal communication). Here, both an overmodified and a normal protein are produced. Mapping data imply a defect at the 3' end of either the *COL1A1* or *COL1A2* gene, and the protein overmodification was sufficiently distinctive for prenatal diagnosis (Fig. 9(a)) In the absence of such changes, ultrasonographic monitoring of limb length can be very useful (Fig. 9(b)). Both compound heterozygosity and autosomal-recessive inheritance are unlikely (with odds of 4096/1 against), and cloning and sequencing of the a1(CB6) region and the homologous region of *COL1A2* followed by allele-specific hybridization of parental tissues seems to exclude collagen type I, with the implication of mutations of genes other than *COL1A1* or *COL1A2*.

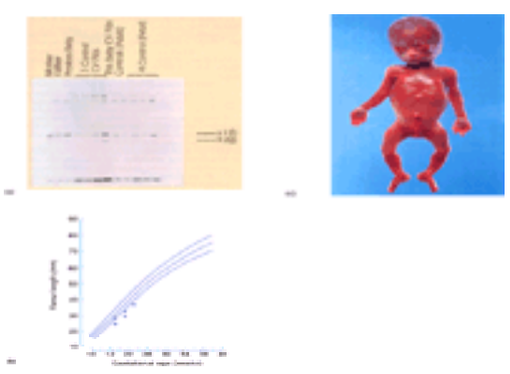


Fig. 9 (a) Patterns of overmodified a(I) and a2(I) collagen profiles in two consecutively affected siblings. The seriously previously affected baby is compared with the current pregnancy in various parental normal and fetal controls. (b) The ultrasound calculation of femoral length was used as a back up. It consistently fell near or below the third percentile. (c) The terminated fetus had obvious deformities of his lower limbs.

Osteogenesis imperfecta with normal collagen type I protein

Some patients with osteogenesis imperfecta have no obvious abnormalities of a1(I) or a2(I) collagen proteins but still have abnormal *COL1A1* and *–1A2* genes. Genetic linkage studies are a valuable alternative in suitable families and will often pinpoint the mutant gene. Several RFLPs are now available, including *Msp*I and two *Rsa*I sites for a1(I) and *Eco*RI, *Msp*I, *Stu*I, and *Rsa*I for the *COL1A2*. All of them are suitable for analysis by Southern blotting or PCR (Sykes *et al*. 1990), with at least a 95 per cent probability that over 90 per cent of families with osteogenesis imperfecta will be linked to either *COL1A1* or *COL1A2* markers. Most Sillence type IV pedigrees have segregated with *COL1A2* markers, whilst families with Sillence type I osteogenesis imperfecta are three times more likely to be linked to a *COL1A1* marker.

Linkage in autosomal-recessive osteogenesis imperfecta

Such linkage data have been consistent for autosomal-dominant families but not for autosomal-recessive inheritance. For example, we have studied an unusual Irish 'traveller' family with three consecutive children lethally affected by type IId osteogenesis imperfecta. Here, two first cousins had married. In this family, linkage data suggested that neither the *COL1A1* nor *COL1A2* gene was to blame yet the protein mapping data showed clear overmodification of collagen type I peptides. The latter usually would indicate C-terminal mutations of either the *COL1A1* or *COL1A2* genes. To date this paradox remains unresolved. We have also observed at least three similar Asian Indian or Middle European families. Possible explanations include unusual somatic mosaicism, compound heterozygosity, or the involvement of unidentified matrix gene other than *COL1A1* or *COL1A2*.

Ehlers–Danlos syndrome type VII

Whether collagen type I (*COL1A1*/*COL1A2*) mutations cause osteogenesis imperfecta or Ehlers–Danlos syndrome VII depends entirely upon the location of the error. Type VII is a distinct subset of Ehlers–Danlos syndrome typified by extreme early joint laxity, increased skin fragility (similar to Ehlers–Danlos syndrome I/II), and misshapen collagen fibrils by transmission electron microscopy (Fig. 10(b)). Ehlers–Danlos syndrome VII overlaps clinically those with osteogenesis imperfecta mutations caused by N-terminal helical mutants, but patients with Ehlers–Danlos syndrome VII rarely break bones whilst the osteogenesis imperfecta group lacks skin fragility. There are distinctive genetic subsets of Ehlers–Danlos syndrome VII: one group with structural abnormalities of type I collagen (*COL1A1* and *COL1A2*), and another in which the enzymes responsible for peptide cleavage (the N- and C-propeptidases) are faulty. The former (structural) group usually has errors in exon 6 of either the *COL1A1* or *COL1A2* genes (Fig. 10a). The latter group has specific errors in the N-terminal propeptidase that cleaves the N-propeptide from both pro-a1(I) and pro-a2(I) chains. Very probably (see below), intraexonic mutations of exon 6 and other more complex deletions of this region can also produce similar disorders. Here the subtleties of the protein chemistry will depend upon whether the N-proteinase or pepsin/trypsin cleavage sites, or lysine hydroxylation sequences, are omitted. (Fig. 10(c))

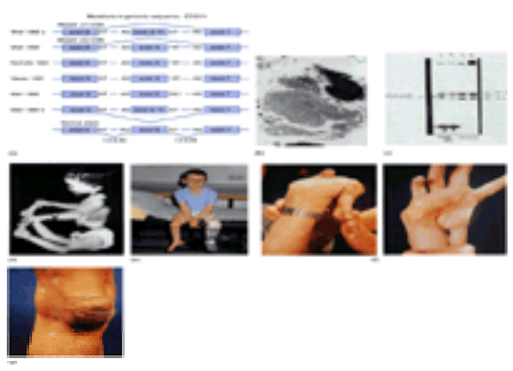


Fig. 10 (a) Diagrammatic representation of certain published mutations in Ehlers–Danlos syndrome VII affecting the splicing of exon 6 of either the *COL1A1* or *COL1A2* genes. (b) Collagen fibril patterns in our Dutch patient (Pope *et al*. 1991). (c) Radiolabelled collagen profiles from Ehlers–Danlos syndrome VII patients (tracks 1 and 2) compared with normal control (tracks 3 and 4) with persistent pNa2 chains after pepsinization. (d) Affected patient with Ehlers–Danlos syndrome VII (after McKusick 1972)—note extreme rotation of hips and knees. (e) British patient with *COL1A1* splice-junction mutation. Note early premature cutis laxa and joint laxity sufficiently severe to require bracing of the unstable knees. (f) Laxity of finger joints and (g) abnormal scarring of knees and shins from the Dutch patient whose fibres are shown in (b) and (c). These changes resemble Ehlers–Danlos syndrome I/II (see also Fig. 20 and Fig. 21).

Ehlers–Danlos syndrome type VIIC also closely resembles the peptidase deficiency of cows, sheep, and Himalayan cats (Pope *et al*. 1988b) in which failure to remove any precursor N-propeptides causes lethal or severe skin fragility. Skin from these animals has very characteristic hieroglyphic fibrils when examined by transmission electron microscopy. Recently, equivalent human mutations have shown identical changes (Nusgens *et al*. 1992; Smith *et al*. 1992). Other notable features of these patients with Ehlers–Danlos syndrome VIIC include identical biochemical changes, with disordered removal of the faulty N-propeptides, and various novel clinical features including premature, generalized cutis laxa with severe blepharochalasis, severe joint dislocation particularly of the hips, and unusually wrinkled skin of the dorsum of the hands and feet as well as the palms and soles.

Types VIIA and B are possibly the most individual of the phenotypes of Ehlers–Danlos syndrome. Biochemically the molecular defect is remarkably consistent and so far, in published families, has involved a splicing failure of exon 6 of either the *COL1A1* or *COL1A2* collagen (Weil *et al*. 1989a, Weil *et al*. 1989b; Nicholls *et al*. 1991; Cole *et al*. 1993) (Fig. 10(a)). This specifically deletes the peptidase cleavage site from one or two of the three a-chain components of collagen type I heterotrimers, with consequent slippage and disruption of collagen helical packing and faulty formation of fibrils, visible as highly irregular, angulated fibrillar patterns (Fig. 10(b)). At the protein level, inefficient processing of procollagen to collagen causes persistence of pNa1 or pNa2 molecules after pepsinization (Fig. 10(c)).

The clinical features include extreme, persistent joint laxity, presenting as congenital dislocation of the hips in infancy and later joint laxity in childhood (Fig. 10(d,e,f)), short stature, scarring of the knees (Fig. 10(g)), forehead, and shins as severe or even worse than in Ehlers–Danlos syndrome I, with a typical criss-cross patterning of the palms. Inheritance is autosomal-dominant. Surprisingly, fractures and osteoporosis are rare, although they commonly accompany nearby N-terminal helical, *COL1A1* or *–1A2*, point mutations in severely loose-jointed variants of osteogenesis imperfecta with osteoporosis and short stature (see Fig. 4(a)). Mutated sequences 5' of exon 6 might also cause phenocopies of Ehlers–Danlos syndrome VII. As mentioned earlier, more subtle mutations within exon 6, such as insertions, deletions or point mutations, might also produce similar phenotypes, especially if highly conserved residues are changed. Perhaps mutation of the adjacent exons (3, 4, and 5) might also produce similar diseases but with variable severity of structural or fibrillar packing. Certainly, mutations of the first few helical exons of the *COL1A1*, *–1A2* or *–3A1* genes regularly produce the expected skeletal or arterial phenotypes of osteogenesis imperfecta and Ehlers–Danlos syndrome IV, which nevertheless are distinct from Ehlers–Danlos syndrome VII. Mutations of the N-terminal propeptidase of *COL3A1* could presumably still produce hieroglyphic fibrils but with vascular rather than predominantly ligamentous pathology (Karl Kadler, personal communication).

Chondrodystrophies and chondrodysplasias

Until very recently the mechanisms of the chondrodystrophies have been extremely difficult to unravel (Maroteaux 1970; Beighton 1995). They are very heterogeneous, with more than 100 recognized non-allelic disorders. Bone and cartilage matrix is also biochemically diverse and contains numerous structural proteins including collagen types I, II, III, V, VI, IX, X, and XI, proteoglycans such as chondroitin sulphate, proteoglycan link protein, hyaluronic acid, and certain other minority components such as osteopontin, osteonectin, and the Gla proteins. All are therefore potential candidate genes. As recently as 1989 knowledge of the molecular pathology of these disorders was primitive. In the first version of this chapter (1993), a few *COL2A1* (cartilage collagen II protein) mutations had been detected in disorders such as achondrogenesis, Kniest dysplasia, certain spondyloepiphyseal dysplasias, and Stickler syndrome (Ahmed *et al*. 1990). Since 1994 the field has exploded to include other unexpected genes such as the segmentation genes, *HOX, PAX* and *SOX*, certain connective tissue components such as morphogenetic proteins (e.g. bone morphogenetic protein 5), and other skeletal regulators such as the fibroblast growth factors and their receptors, the *FGFR* genes. Other important substances have included glycosaminoglycan sulphation proteins and the sulphate transport proteins (see below). Simultaneously there has been

consolidation and expansion in the analysis of more conventional connective-tissue scaffolding genes such as *COL9A1*, *–2*, and *–3* (collagens IX) and *COL10A1* (collagen X), all of which cause skeletal dysplasia. Other unexpected candidates are the thrombospondin gene family, a good example of which is the *COMP5* gene on chromosome 19, which causes multiple epiphyseal dysplasia. In men and mice the *COL11A1* and *–11A2* genes have been separately implicated, causing autosomal-recessive chondrodysplasias (with abnormal cartilage, limbs, ribs, mandible, and trachea) in mice, and, in humans, autosomal-recessive Stickler syndrome with spondyloepiphyseal dysplasia, osteoarthritis, and deafness but without ophthalmic changes, caused by homozygous Gly–Arg substitutions (the so-called otospondylomegaepiphyseal dysplasia syndrome). Here there was severe, degenerative osteoarthritis of hips, knees, and elbows; other features included mid-facial hypoplasia with shortened noses and supraorbital ridges, and sensory neural deafness. Some Stickler families unlinked to *COL2A1* or *–11A2* have been linked to *COL11A1* (Snead *et al*. 1994) whilst in others the gene locus is unknown. The phenotype of the *COL11A1* families includes vitreoretinal detachment, premature myopia, joint laxity (lessening with age), and skeletal disproportion, but without short stature or facial hypoplasia. There has been rapid progress in work on *COL2A1* mutations, with convincing abnormalities now demonstrated in achondrogenesis, Kniest syndrome, the Stickler syndrome, and certain spondyloepiphyseal dysplasias including Strudwick-type spondyloepimetaphyseal dysplasia. The field has also been greatly enriched and illuminated by the creation of transgenic mouse models for several chondrodystrophies. Occasionally these have provided unforeseen and unexpected information. Thus, mice with *COL11A1* deletions have convincing chondrodysplasia, with shortened snouts, stunted limbs, and tracheomalacia (Li *et al*. 1995). Mice homozygous for *COL5A2* mutations have kyphoscoliosis and cutaneous fragility, and resemble humans with Ehlers–Danlos syndrome type II (Andrikopoulos *et al*. 1995). [*COL10* knock-outs are clinically normal (Rosati *et al*. 1994), although human, autosomal-dominant negative, glycine substitutions in *COL10A1* cause Schmid metaphyseal dysplasia. Here a faulty protein is much more disruptive than a missing component.]

Other important clues include the location of chondrocalcinosis with osteoarthritis to chromosome 8q and the implication that proteoglycan sulphation is disturbed in diastrophic dwarfism and certain types of achondrogenesis (Hastbacka *et al*. 1994; Superti-Furga 1995). Studies of mutations in the *FGFR* gene family have revolutionalized the molecular pathology of the craniofacial dysplasias of Apert, Crouzon, and Pfeiffer syndromes as well as of apparently unrelated disorders such as achondroplasia, hypochondroplasia, pseudoachondroplasia, and thanatophoric dwarfism, which are therefore allelic to the craniofacial dysplasias and one another.

Collagen mutations causing chondrodysplasias

The major connective tissue protein of cartilage collagen is coded by the *COL2A1* gene. It is also present in the vitreous humour of the eye. Other important structural proteins include collagen IX (a1, a2, and a3 chains), collagen X, which forms in hypertrophic cartilage in the metaphyses, and more unexpected elements such as cartilage-specific matrix protein, a member of the thrombospondin family. Cartilage is therefore a complex connective tissue that is phylogenetically ancient. Whilst typical in containing cell matrix and fibrous elements, it differs from other tissues such as skin, ligament, and bone in having a higher proportion of cells and matrix relative to the fibrous elements. During fetal life the entire protoskeleton is initially cartilaginous, remaining so until birth, when only fragments have differentiated into more mature bone. Later, except for certain articular surfaces, most of original cartilaginous protoskeleton changes into bone, a continuing process that in humans is not complete until well into adolescence. Even then, cartilaginous elements persist in the ears, nose, ribs, articular surfaces, and respiratory tracts, and at other unexpected sites such as the vitreous humour, where a more gelatinous form of cartilage matrix persists and is crucial for the optical properties of the posterior eye. Since it adheres closely to the retina, abnormalities and fibrous contractions result in retinal detachment and blindness.

Cartiliginous matrix has several variants including hyaline and fibrocartilage. The latter includes white fibrocartilage, which is rich in collagen, and yellow elastic cartilage, which is rich in elastin. Articular hyaline cartilage lines the articular ends of bones within synovial cavities. Here it provides a robust and stable surface lubricated by synovial fluid. Collagen bundles align tangentially near the articular surface but thicken if running vertically or more deeply. Chondrocytes tend to be flattened superficially but form distinctive vertical columns as they approach the bony metaphyses. Chondrocytes are firmly embedded in the cartilaginous matrix containing several collagenous proteins, such as collagens II, VI, IX, X, and XI; whole genes have all been cloned and sequenced. Many other cartilaginous components have also been located, cloned, and sequenced. This has allowed rapid progress in the molecular pathology of various chondrodysplasias and, since 1992, specific mutations in *COL2A1, COL5A2, COL11A1*, *–11A2*, *–9A1*, *–9A2*, and *–10A1* genes have been shown to produce a wide range of skeletal disorders (Table 1).

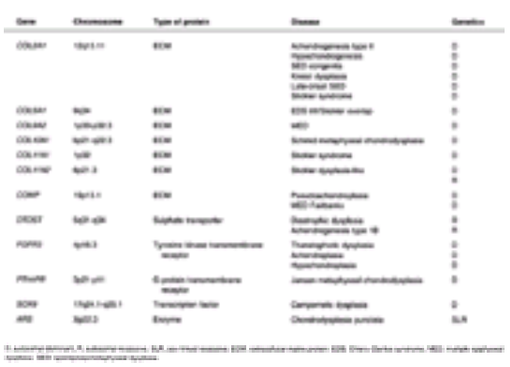


Table 1 Collagen gene mutations and their skeletal manifestations

Abnormalities of other components combine to produce an impressive range of chondrodysplasias. Thus, abnormalities of the *COMP5* gene produce a cartilage-specific glycoprotein that has a five-armed structure similar to thrombospondin. Here the clinical phenotype causes spondyloepiphyseal or metaphyseal dysplasias, similar or identical to those caused by *COL9A1* and *–9A2* mutations, respectively. Probably shared structural interactions of the two molecules produce common organizational changes and pathology. Mutation of *COL10A1* produces Schmid metaphyseal dysplasia in which the hypertrophic cartilage of the metaphyses is abnormal.

Since 1993 there have been other quite unexpected advances, implicating components that are shared by bone and cartilage alike, such as the *FGFR* genes. These bind certain fibroblast growth factors that orchestrate the expression and differentiation of numerous other mesenchymal growth factors. Their receptors are diverse but homologous and include FGFR-1, -2, and -3. All have extracellular, immunoglobulin-like, binding regions with transmembranous tyrosine kinase-like domains and intracellular sequences that transmit certain cytoplasmic signals. Mutations of either the IgG or transmembranous regions of FGFR-1, -2, and -3 have been implicated in achondroplasia (FGFR-3), Crouzon, Apert, Pfeiffer, and Jackson–Weiss syndromes (FGFR-1 and -2), and disorders such as achondroplasia, hypochondroplasia, and thanatophoric dwarfism. Although unexpected, this provides a convincing rationale for disorders in which the orchestration and timing of skeletal function are impaired. These molecules are quite separate from, and upstream (in hierarchical terms) of, disorders caused by faulty structural proteins such as collagens II, X, and XI, of which collagen II only had been implicated as recently as 1993. The three years to 1996 produced a revolution in the molecular pathology of skeletal dysplasia in which a whole hierarchy of faulty products clearly participates (including defects of pattern, structural, organizational and mechanical errors, or changes in structural elements with scaffolding functions).

Hyaline cartilage is translucent though slightly opaque because of scattering by transmitted light. In polarized light it is fibrillar and, by electron microscopy, contains permeating fibrils of 10 to 20 nm. The *COL2A1* gene has the typical organizational structure of a fibrillar collagen comprising a central helix with N- and C-terminal extensions (Vuorio and de Crombrugghe 1990). The C-terminus is conventionally organized and coded by four exons (49–52), with exons 51 and 52 identical (in size) to those of other fibrillar collagens. The N-terminus is more variable, especially in the short triple helix, which in the *COL2A1* gene has several additional exons (and one non-helical interruption). Part of exon 3 and two extra exons (4B and 5B) code for the 79 amino-acid sequence of the N-terminus. The central triple helix is coded by exons 7 to 48 inclusive, and exons 49 to 52 encode the C-propeptide. The gene of 28 Kbp maps to chromosome 12.

COL2A1 polymorphisms

Various intragenic and flanking polymorphisms have been described (Francomano *et al*. 1987, Francomano *et al*. 1988; Wordsworth *et al*. 1988). They include *Hind*III and *Pvu*II variants detectable with a 9.2 kb *Eco*R1 fragment. In a large, autosomal-dominant, Finnish pedigree the *HindIII* RFLP was tightly linked to familial osteoarthritis Palotie *et al*. 1989. Similarly, a *Bam*H1, intragenic, RFLP is more common in Caucasian patients with osteoarthritis than in normal controls (Hull and Pope 1989), although in other examples *COL2A* is not involved (Baldwin *et al*. 1995). The *Eco*R1/*Bam*H1 hypervariable region in the 4.3-kb, *Eco*R1 3' flanking region of *COL2A1* has a useful variable repeat, which is deleted or duplicated in 30 bp multiples. More recently, the 5' *Hin*fI RFLP has also been informative in

spondyloepiphyseal dysplasia congenita (Anderson *et al*. 1990).

Many suitable markers for the *COL5A1*, *–5A2*, *–9A1*, *–9A2*, *–9A3*, *–11A1*, and *–11A2* genes have also been recently identified, all of which can now be tested in skeletal abnormalities. Here the 'educated guess' strategy is useful when the candidate gene is unknown, but other biochemical and structural information provides strong hints. Otherwise a systematic search of the genome becomes necessary and, in that regard, the recent Human Genome Mapping Initiative has produced a very dense map of human gene markers such as $(CA)_n$ sequence-tagged sequences. Any of these markers can firmly locate faulty human loci. A good recent example is the unexpected identification of the *COMP5* locus on chromosome 19 in multiple epiphyseal dysplasia (Briggs *et al*. 1995). Here, two separate lines of research converged: one had located the megaepiphyseal dysplasia locus whilst the other pursued the basic structure, function, and biology of the *COMP5* gene and protein. No doubt many other examples are destined to appear in the future.

Studies of human disease

Type II collagen predominates in cartilage but also forms the vitreous humour (of the eye). Several lines of evidence combine to suggest that cartilage collagen genes can cause specific skeletal dysplasias in man and other animals. Thus McKusick (1972), in categorizing inherited defects of connective tissue as a specific subset of human disorders, noted that chondrodystrophies were unusually diverse and variable examples. They form natural subgroups, which Spranger (1985) has very clearly delineated and which are likely to share common abnormalities in relevant cartilage components. For example, the distribution of collagen type II in articular cartilage, epiphyseal plates, and the vitreous humour strongly implies specific pathological patterns. Similarly, achondrogenesis I and II, hypochondrogenesis, and spondyloepiphyseal dysplasia are excellent candidates for *COL2A1* mutations as they all share the features of severe joint degeneration, variable skeletal deformities, and myopia. Furthermore, all have disordered cartilage organization with degenerative joint disease. The Kniest/Stickler and Wagner syndrome(s) also share similar clinical features. All show severe degeneration of hyaline cartilage and progressive myopia with or without retinal detachment. Other milder disorders with similar mechanisms include familial myopia with premature osteoarthritis and subsets of common osteoarthritis. The combined anatomical, biochemical, and molecular evidence implicates *COL2A1* mutations in most of these diseases.

To detect the mutations, three main approaches were used: (a) the ascertainment of suitable autosomal-dominant or -recessive pedigrees for studies of linked genetic markers; (b) the detailed analysis of cartilage collagen from tissue samples such as iliac crest or chondrocortical junctions or femoral heads; and (c) the amplification by PCR of mutant alleles detected by (a) or (b), followed by indirect or direct sequencing of mutated gene fragments. This strategy has been extremely efficient and informative, and is now largely comparable in outcome with similar studies of type I and III collagen mutations, differing only in that usually protein has been purified from tissue samples rather than cultured cells. Chondrocytes are more difficult to maintain in culture, but can be successfully grown in agarose suspension (Benya and Shaffer 1982; Horton 1988; Byers 1990). Evidently, many lethal or severe *COL2A1* mutants secrete the faulty product and, like patients with severe osteogenesis imperfecta or vascular Ehlers–Danlos syndrome IV, cannot incorporate mutant or even wild-type protein into their hyaline cartilage, instead retaining or degrading the faulty/wild-type heterotrimers intracellularly (Eyre *et al*. 1986). Another useful and important technique is the amplification of *COL2A1* gene transcripts from non-collagenous tissues such as lymphocytes or fibroblasts. This has obviated both the need for difficult cartilage biopsies and the vagaries of chondrocyte culture. Disadvantages include the artefactual amplification of non-cartilage sequences in genes that are alternatively spliced in different tissues and also the lack of any supporting studies of cartilage histology or protein chemistry.

The 'illegitimate transcription' of the very small traces of *COL2A1* mRNA produced by immortalized lymphocytes or cultured skin fibroblasts can be used for the amplification by PCR of *COL2A1* cDNA. Thus cartilage gene sequences can be usefully analysed from non-cartilaginous samples such as skin or blood. This greatly simplifies the opportunities for testing affected individuals from whom cartilage samples are difficult to obtain (Chan and Cole 1991; Chan *et al*. 1993; Cole *et al*. 1993). Similarly, other minority cartilage components, such as the *COL5A1*, *–A2*, *–11A1*, and *–11A2* genes, can be explored in other cartilaginous disorders such as the Stickler syndrome.

Evidently, the hierarchical nature and complexities of skeletal development, growth, and repair also requires other developmental orchestrators such as the *HOX* and *PAX* genes. These are similar to the so-called segmentation genes of *Drosophila*. Others include bone morphogenetic protein (transforming growth factor-b gene family), the FGFR receptor and fibroblast growth factor genes, and other unexpected components such as the thrombospondins. The development and maintenance of bone and artery is very probably just as complex but is much less well understood.

Evidence that *COL2A1* mutations cause chondrodystrophies

Wordsworth *et al*. (1988) observed discordant segregation between RFLP markers for *COL2A1* and diseases such as achondroplasia, pseudoachondroplasia, multiple epiphyseal dysplasia, autosomal-recessive spondyloepiphyseal dysplasia tarda, diaphyseal aclasis, and the trichorhinophalangeal syndrome. He concluded that *COL2A1* was not a common candidate gene for these chondrodysplasias. Unfortunately, families with autosomal-dominant spondyloepiphyseal dysplasia tarda or metaphyseal chondrodysplasia were uninformative (so that linkage could be neither confirmed nor rejected). We now know that chondrodysplasias that share the features of distorted hyaline cartilage, epiphyseal dysplasia, premature osteoarthritis, and vitreoretinal detachment or myopia frequently have *COL2A1* mutations.

The latter form two distinct but related groups. Firstly, the Stickler/Wagner group is autosomal-dominant. Clinical signs include vitreoretinal degeneration, myopia, cataracts and retinal detachment with premature osteoarthritis, epiphyseal dysplasia, and facial abnormalities such as mid-facial hypoplasia with a large philtrum and micrognathia with occasional cleft palate. These disorders are heterogeneous, some patients being marfanoid whilst others are short (McKusick *et al*. 1990), though joint laxity is abnormal in both groups (see Fig. 12(a–d)) The Kniest syndrome is also part of the Wagner/Stickler phenotypical complex. There are at least two types (of Kniest syndrome): one relatively mild and autosomal-dominant (or sporadic), the other genetically lethal and sporadic. The milder overlaps with metatrophic dysplasia, which also shows disproportionate short stature, prominent eyes, and mid-facial hypoplasia (Jones 1988; McKusick *et al*. 1990). Radiographs show epiphyseal irregularity, metaphyseal flaring, late ossification of the femoral heads, and platyspondyly. Histologically the cartilage resembles Swiss cheese (McKusick *et al*. 1990). Other generalized abnormalities include inguinal and umbilical hernias, kyphoscoliosis, odontoid hypoplasia, and retinal detachment and myopia. There was clear linkage of *COL2A1* markers such as *Hind*III RFLP and the variable tandem repeat at the 3_ end of the gene (Knowlton *et al*. 1990). Two families showed tight linkage to *COL2A1* whilst another did not (Nunez *et al*. 1985). Schwartz *et al*. (1989) were the first to show non-allelic genetic heterogeneity in both Stickler and Wagner syndromes. They suggested that the Wagner syndrome (retinal or vitreoretinal degeneration with minimal systemic features) was unlinked to *COL2A1*, whilst the Stickler syndrome was clinically and genetically more heterogeneous (and only sometimes linked).

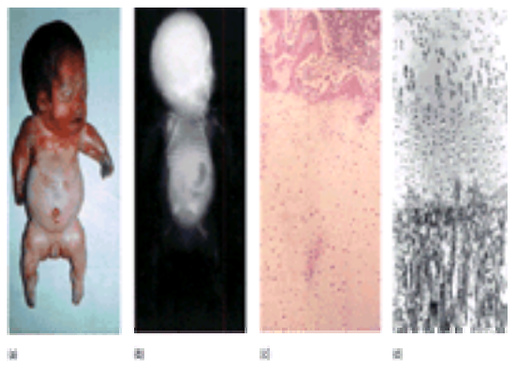


Fig. 11 (a) Typical clinical appearance of patient with lethal achondrogenesis II—note the rhizomelia and abdominal distension with the constricted small chest. (b) The radiographs show shortened ribs, metaphyseal flaring, and shortened long bones. (c) The histological appearance of the femur shows bulbous hyperplasia of the epiphyses and disorganization of cartilage columns, including a fibroblastic inclusion at the epiphyseal plate, compared with the normal control. (d) The bony trabeculae are also disorganized in the achondrogenesis bone. Original magnification × 157.

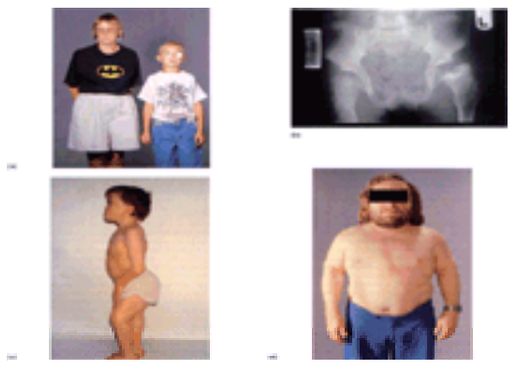


Fig. 12 Collagen II associated mutations. (a) Affected mother and son with *COL2A1*-linked Stickler syndrome. Note the high myopia. (b) Premature degenerative joint disease in similar patient. (c) Another Stickler variant; note the lumbar lordosis and mid-facial hypoplasia. (d) Spondyloepiphyseal dysplasia with rhizomelic shortening of the upper segments of the upper limbs.

Subsequently, we have described genetic heterogeneity as well as distinctive vitreoretinal features that allow us to separate *COL2A1* families from all of the others. Sometimes non-*COL2A1* families are linked to relevant genes such as *COL11A1*, whilst in others the candidate genes are so far unknown. Nevertheless, it is noteworthy that vitreous humour contains heterologous molecules made of collagens II, Va2, XIa1 or XIa2 chains, which can be interchangeable. Consequently the *COL5A1*, *–5A2*, *–1A1*, and *–11A2* genes are also strong candidates for the Stickler/Wagner complex (Snead 1995; Vikkula *et al*. 1995).

Spondyloepiphsyeal dysplasia congenita is also linked to *COL2A1* (see Fig. 12(d)). Here, notable clinical signs include odontoid hypoplasia, ovoid vertebrae with narrowed spaces, kyphoscoliosis, lumbar lordosis, chest deformities, poorly mineralized epiphyses, poorly calcified femoral heads, and myopia. It is a clear autosomal-dominant and has a good prognosis. This contrasts with achondrogenesis, which is a severe, short-limbed dwarfism with highly disorganized cartilage matrix causing death *in utero* or in the neonatal period (McKusick *et al*. 1990). Affected infants are born prematurely with large heads, a short trunk, and severe micromelia (Fig. 11).

Achondrogenesis types I and II (Langer–Saldino; Parente–Fraccaro) are distinguished by rather subtle radiological features that are separated into types IA, IB, and II (similar to the Sillence type II osteogenesis imperfecta classification) by subtle differences in rib and limb morphology and fracture pattern. Generally, types IA and IB have stubby ribs and distorted limbs, whilst babies with type II have straighter limbs and ribs (McKusick *et al*. 1990). Grebe-type achondrogenesis is a separate, extremely severe subtype. McKusick accepts Borochowitz's suggestion that further subdivides the Parente–Fraccaro classification. Achondrogenesis–hypochondrogenesis is at the milder end of this spectrum.

Histologically the affected bone shows disorganized endochondral ossification and severely distorted growth plates. (Fig. 11(c,d)). Cartilage vascularity is increased and chondrocytes are dilated, with granular inclusions within the endoplasmic reticulum. These features strongly suggest a morphological, structural disturbance of collagen II. Other mechanisms are also possible. For example, Superti-Furga (1994) and Superti-Furga (1996) have identified an abnormality in the metabolic activation of sulphation in achondrogenesis type IB, and also demonstrated a number of mutations of this gene

Finally, linkage to *COL2A1* has been demonstrated in three exceptional osteoarthritis families. In particular, two large Finnish families with premature osteoarthritis (without chondrodysplasia) had combined lod scores exceeding 3 (Palotie *et al*. 1989). This suggests that *COL2A1* mutations can cause non-syndromic osteoarthritis without chondrodysplasia. Subsequently, Knowlton *et al*. (1990) also showed tight linkage to the *COL2A1 Hind*III RFLP in a second osteoarthritis family. Many affected family members had both Heberden's nodes and typical radiographic signs of osteoarthritis in hips, shoulders, wrists, and hands. There were also flattened metatarsal heads with irregular flattening of vertebral end-plates, suggesting a mild chondrodysplasia. Another mild chondrodysplasia with premature osteoarthritis (Naquamaland hip dysplasia) was also linked to *COL2A1* (Sher *et al*. 1991), but subsequently these workers excluded *COL2A1* as the candidate for Benkes familial osteoarthritis (Beighton *et al*. 1994). However, in both American and South African families the extrapolation to osteoarthritis 'vulgaris' is unproven, in contrast to the Finnish evidence. Nevertheless, both studies illustrate the potential for analysing the pathogenesis of osteoarthritis. Further support has been forthcoming from associated studies: for example, certain RFLPs for *COL2A1* are twice as frequent in patients with osteoarthritis as in normal controls (Hull and Pope 1989). Contrastingly, in another study of segregation in eight osteoarthritic patients with a slightly different phenotype including Heberden's nodes, the evidence for *COL2A1* skewing was borderline. Most probably, therefore, osteoarthritis is genetically heterogenous with a common histological expression but variable mechanisms, some of which are *COL2A1*-related. Any of the other structural genes expressed in cartilage would also be strong candidates in the causation of osteoarthritis (Dieppe *et al*. 1987).

Collagen protein analysis in the chondrodysplasias

Over the past 10 to 15 years the analyses of type I and III collagen mutations have been very straightforward since such abundant radiolabelled proteins can readily be separated, mapped, and melted from cultured skin fibroblasts. Both secreted and intracellularly retained mutant proteins from the majority of cell lines of severe osteogenesis imperfecta or arterial Ehlers–Danlos syndrome can be characterized (Pope *et al*. 1988a, Pope *et al*. 1988b; Byers *et al*. 1990; Kuivaniemi *et al*. 1991) Until recently, similar analyses of the chondrodysplasias were greatly hindered by the technical difficulties of chondrocyte cell culture. Not only was primary culture difficult but the cells were unstable in agarose suspension (Benya and Shaffer 1982; Horton 1988). Chondrocytes are now readily derived from simple iliac-crest biopsies (Chan *et al*. 1993), and they can be routinely redifferentiated with semisolid alginate beads. Furthermore, immortalized lymphocytes (Chan and Cole 1991) or cultured skin fibroblasts produce traces of *COL2A1* mRNA from which cDNA can be amplified. Tissue from biopsies of the iliac crest or chondrocostal junction can also be used to purify collagen type II proteins, which can then be analysed conventionally by peptide mapping or temperature denaturation in a similar fashion to radiolabelled type I or III collagens in osteogenesis imperfecta and Ehlers–Danlos syndrome (Byers 1990; Kuivaniemi *et al*. 1991; Tiller *et al*. 1995).

Two very important pioneering papers clearly illustrate this approach. Godfrey and Hollister (1988) described a lethally affected, short-limbed fetus with collagen type II achondrogenesis/hypochondrogenesis. Cartilage was histologically abnormal, with disorganized growth plates and overvascularized, fibrotic, hypercellular cartilage similar to that illustrated in Fig. 11(c). Collagen extracted from cartilage of the femoral head was then analysed either by a-chain or gene analysis of *COL1A1* or *COL1A2* mutations in osteogenesis imperfecta (see above). Proteolytic thermal denaturation studies were also undertaken, and the extractability of cartilage collagen and the amounts of hexosamine and proteoglycan examined. Affected patients possessed both normal and overmodified a1(II) collagen detectable as abnormal doublets (Fig. 13(a)). Similarly, CNBr cleavage in one and two dimensions confirmed overmodification of the whole collagen a1(II) helix as measured by doublets in one-dimensional gels and 'smiling' or tilted bands in the two-dimensional gels (Fig. 13(a,b)). Furthermore, only collagen type II was overmodified whilst collagen a1(I), a2(I) or a1(III) chains (as expected) were not. By extrapolation to analysis of lethal osteogenesis imperfecta this located the mutation at the C-terminus of a1(II) collagen. Such accurate mapping also provided a means of cloning and sequencing the mutant gene just as in osteogenesis imperfecta and Ehlers–Danlos syndrome IV mutations. Even though defective at the C-terminus, much of the mutant protein was still secreted and could easily be isolated from cartilage biopsies. In contrast, the patient studied by Eyre (1986) could not secrete the faulty mutant collagen II protein.

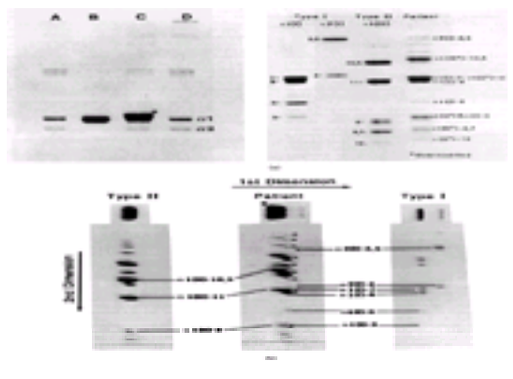


Fig. 13 (a) Collagen a1(II) chains derived from a patient with achondrogenesis: compared with the control a1(II) chains (track B) the mutant produces normal and overmodified (*) forms; similarly (right panel) all a1(II) cyanogen bromide fragments are overmodified (Godfrey and Hollister 1988). (b) The two-dimensional peptide

maps of forms excised from (11a) showed typical 'smiling' of all cyanogen bromide peptides in an affected patient. These are analogous to the a1(I)CB6 mutations causing osteogenesis imperfecta as illustrated in Fig. 7(b), and normal and overmodified species together form a diagonal 'smile'.

Murray *et al*. (1989) then described very similar studies in nine patients with spondyloepimetaphyseal and spondyloepiphyseal dysplasias. As expected, all were dwarfed and had abnormal epiphyses, varying metaphyseal irregularities, flattened vertebral bodies, and myopia. Here, costal cartilage was sampled from the more mildly affected children under general anaesthesia. Using standard procedures with 4M GuCl and proteinase inhibitors, type II collagen was extracted with pepsin or acetic acid, and whole collagen a1(II) chains or CNBr-cleaved a-chains characterized by polyacrylamide gel electrophoresis. All affected patients had abnormal mobility of collagen a1(II) chains with variable overmodification of CNBr-derived peptides. As expected, and by analogy with *COL1A1*, *–1A2*, and *–3A1* mutants (Bonadio and Byers 1985; Byers 1990; Kuivaniemi 1995), the extent of overmodification varied with position, being least severe near the N-terminus (mild) and worse towards the C-terminus (increasingly severe). Thus the least affected patients were the least modified and vice versa. These pioneering studies paved the way for more detailed mutant analysis by RFLP analysis combined with protein mapping, gene amplification, and sequencing (see also Fig. 4 and Fig. 19). Very substantial progress has occurred since the first edition of this chapter.

Fig. 14 Diagram of COL2A1 (type II collagen) mutations, substitutions (grey), deletions (blue), duplications (yellow), and 'stops' in red.

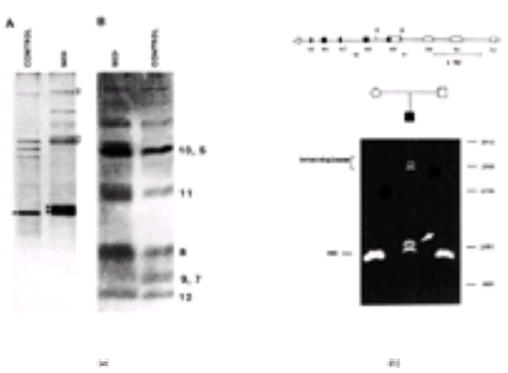


Fig. 15 Collagen a1(III) chains and CNBr peptide pattern of extracted cartilage components from a control and a patient with spondoepiphyseal dysplasia (Tiller *et al*. 1990). (a) Whole a1(III) chains from the patient are overmodified compared with the control and after cleavage with CNBr show overmodification of each CNBr fragment—very similar to Fig. 15(a). This is analogous to the overmodified a1(I) components shown in Fig. 6(a). (b) Amplification of the relevant portions of exons 48–49 showing normal and lengthened mutant fragments with heteroduplex formation between wild-type and mutant copies; at the bottom is a sequence of mutated and wild-type fragments showing 30 amino acids (90-bp duplication) of the mutant allele, situated in exon 48.

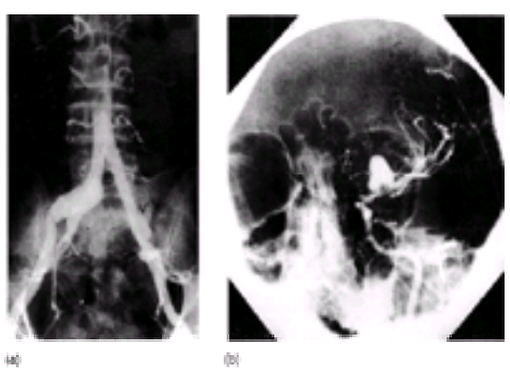


Fig. 16 (a) Aortogram showing fusiform right femoral aneurysm, and (b) large middle-cerebral aneurysm from two different patients with Ehlers–Danlos syndrome IV.

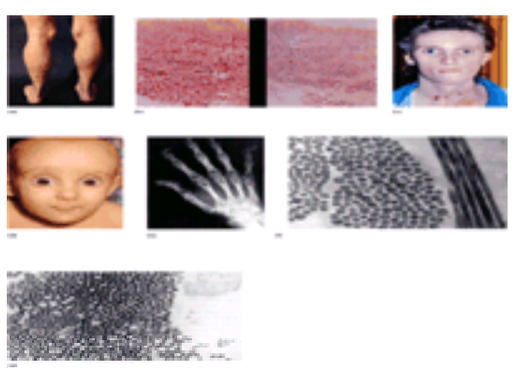


Fig. 17 (a) Typically thin skin of patient with Ehlers–Danlos syndrome IV, showing prominent capillary and venous pattern with some venous varicosity. (b) Histological appearance of the thinned skin (R) shows collagen depletion and elastin proliferation compared with the control (L). (c, d) The facial phenotype is especially characteristic with large eyes, thin lips, and lobeless ears in the acrogeric arterial type. This is equally easily recognized in adults (1c) and babies (1d). (e) Hand radiographs often show a subtle acro-osteolysis (arrowed). (f, g) Transmission electron micrographs showing irregularity of a collagen fibril diameters in (f) skin and (g) arterial walls.

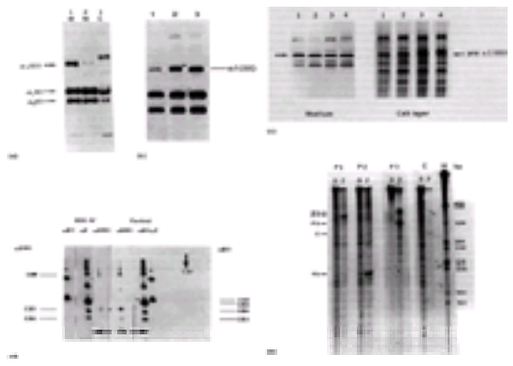


Fig. 18 (a) Typically poor secretion and intracellular retention of pepsinized collagen a1(III) chains in an Ehlers-Danlos syndrome IV mutant. Here the medium collagens (track 1) are diminished and an overmodified form is retained within the cell layer compared with the normal control in track 3. (b) Sometimes, instead of being poorly secreted (as in track 1) the mutant form is secreted as an a1(III) doublet (as in track 2) in which the secreted mutant product is starred; at other times (not shown) a normal and a shortened form of the protein are simultaneously secreted. (c) Overmodification of retained (cell layer) pro-a1(III) chains is variable (cell layer; tracks 2,3,4) according to the helical position of the mutation (the nearer the C terminus the greater the overmodification and the more retarded the band); the effect upon collagen secretion is remarkably similar (medium; tracks 2,3,4). (d) Two-dimensional electrophoresis of collagen a1(III) chains: in this instance there is an a1(III) CB5 doublet in the mutant (left hand) a1(III) track compared with the normal (right hand) sample; as expected, a1(I) and a2(I) tracks are identical. (e) Chemically cleaved cDNA/cDNA hybrids from positive mutant/wild-type standard (C) and three mutant/wild-type hybrids P1 –3. Under normal conditions the wild-type hybrids migrate at the top of the gel and are uncleaved. Those hybrids with mismatches will allow chemical cleavage at the point of incompatibility. In this instance, three different point mutations have been cleaved as has a small deletion (P3) of 27 bp, which produce two different-length cleavage fragments. The known positive control, glycine to aspartate 883, produces a faint but distinct cleaved fragment. The test mutant cleavages in P1 and P2 are at residue 910 in the case of P1, migrating later on the gel and smaller in the case of P2. Molecular-weight markers allow accurate estimations of fragment sizes.

Fig. 19 Diagrams of the distribution of currently known *COL3A1* mutations (Kuivaniemi *et al*. 1995; Pope *et al*. 1996): (a) point mutations; (b) splicing and deletion mutations.

Molecular analysis of the chondrodysplasias (see Table 1)

When this section was first drafted in 1991 very few *COL2A1* mutations had been identified and no other connective-tissue components had been implicated. Subsequent progress in the past 3 years has been exceptional: thus defects have been identified in other cartilage genes such as *COL5A2*, *–11A1*, and *–11A2* as well as potential defects in the *–9A1*, *–9A2*, and *–10A1* genes (see Table 1).

Several *COL2A1* mutations have been detected in disorders such as spondyloepiphyseal dysplasia, achondrogenesis, and familial osteoarthritis (Fig. 14). The patterns and mechanisms of mutation are again closely analogous to those that cause osteogenesis imperfecta and Ehlers–Danlos syndrome IV. Subsequently, mutations of other genes such as *COL9A1*, *–9A2*, *–11A1*, *–11A2*, *–5A2*, and other cartilage components have been observed in a variety of disorders.

Until 1991 only three mutations had been published (Fig. 14). The first was in a large family with an obviously abnormal type II collagen protein. Here Southern blotting of the proband's genomic DNA showed a 400-bp deletion in the 3.7-kb *Eco*R1 fragment next to the hypervariable region at the 3' end of *COL2A1* (Lee *et al*. 1988). This region coded for the last four exons (45–48) of the collagen II triple helix and all four exons (49–52) of the C-terminal propeptide. The investigators specifically used PCR amplification of this region. By consecutive amplification of exons 47–51, 47–52, 49–51, and 49–52 the deletion breakpoint was precisely located to exon 48. Sequencing confirmed a 390-bp deletion spanning the middle of intron 47 to the 5' splice-site of intron 48. This deleted all 108 bp of exon 48, consequently omitting 36 amino acids from the collagen II triple helix (residues 964–999). Such shortened molecules induce protein suicide just as in osteogenesis imperfecta and Ehlers–Danlos syndrome mutants (Prockop and Kivirriko 1984), and are abnormally retained intracellularly. As in the stoichiometry of collagen III homotrimers, only one-eighth of the collagen a1(II) chains can form normal, wild-type homotrimers that are correctly secreted. The remaining seven-eighths are intracellularly retained and/or degraded. Protein data from this patient are shown in Fig. 15.

The second early mutation lay very close to the first. Here the affected child was sporadic (Tiller *et al*. 1990). CNBr peptide mapping accurately mapped the mutation to the C-terminus and all peptides were consequently overmodified (Fig. 15(a)). The mutation lay between exons 45–52. (Fig. 15(b)). Amplification of exon 48 showed double fragments from the affected patient with normal/mutant heteroduplexes. Sequencing showed an intraexonic 45-bp (15 amino acid) duplication within exon 48. Anomalous homologous meiotic recombination had produced one allele with an extra 15 amino acids. In general the repeating GlyXY motif of collagen provides numerous potential sites for such recombinational asymmetry.

Subsequently, several other glycine substitutions have been identified and are irregularly distributed along the collagen II helix (see Fig. 14). For example, substitutions at positions 247, 274, 293, and 997 (all Gly–Ser) have all produced the clinical phenotype of spondyloepiphyseal dysplasia. In other examples, Arg75–Cys at position G^{+5}IVS20 causes skipping of exon 20, whilst more complex deletions/duplications, 964–999 and 970–984, respectively, also produce spondyloepiphyseal dysplasia. A Gly–Ser154 mutation has recently been observed in Strudwick-type spondyloepimetaphyseal dysplasia, which is more severe than spondyloepiphyseal dysplasia and also shows disproportionate, short stature. The three published examples included Gly–Val292, Gly–Cys709, and Gly–Cys302, respectively, all of which have abnormal collagen a1(II) a-chains and abnormal CNBr peptides similar to those of Fig. 13 and Fig. 15 (Tiller *et al*. 1995). Spondyloepiphyseal dysplasia has also been caused by *COL11A2* mutations. Tiller's patients were unusual in having *COL2A1* Gly–Cys substitutions. Transgenic mice heterozygous for *COL2A1* Gly85–Cys have a lethal chondrodysplasia (Garofalo *et al*. 1991).

Achondrogenesis

Achondrogenesis, hypochondrogenesis, and Kniest dysplasia form one disease spectrum. Clinical changes range from lethal, asphyxiating chondrodystrophy with very short limbs and multiple fractures (type IA achondrogenesis) to a milder disorder with disproportionate short stature, cleft palate, mid-facial hypoplasia, and a flattened nasal bridge (Kneist dysplasia). Infants with achondrogenesis usually die *in utero* or perinatally whilst Kniest dysplasia, although deforming, is more benign.

Experiments by Godfrey and Hollister (1988) accurately located (Fig. 13) a mutation to the C-terminus which was subsequently cloned and sequenced. The original strategy was to make cosmid libraries from the size-fractionated genome, that is DNA. The library was screened with a normal *Eco*R1, 3.7-kb genomic subclone within the mutated region. The proband was heterozygous for a *Hind*III *COL2A1* RFLP. Exons 51 and 52 were amplified with intron-specific primers (just as in the cases of spondyloepiphyseal dysplasia). Two unusual variations were detected. The first (in exon 51) was a neutral change of serine to glycine, caused by a *Pvu*II RFLP segregating in the patient's healthy relatives. The frequency in the normal population was 30 per cent. The second changed Gly–Ser (GGC–AGC) at position 943 of the collagen type II triple helix. It is therefore analogous to *COL1A1* osteogenesis imperfecta (lethal) mutations of Gly–Ser between positions 913 and 1003 (Byers

1989), all of which seriously impair the normal formation, secretion, and melting/denaturation profiles of the collagen type I helix. Very probably the collagen II helix was similarly disturbed. Subsequently a Gly–Glu at position 853 of collagen type II was described by Bogaert *et al*. (1992) in a patient with hypochondrogenesis. The field has now been advanced very rapidly by the combination of conventional protein analysis, peptide mapping, reverse-phase high-performance liquid chromatography, and genomic and cDNA PCR amplification (Tiller *et al*. 1993; Mortier *et al*. 1995; Tiller *et al*. 1995).

Since the first Gly–Ser *COL2A1* substitution (Vissing *et al*. 1989), numerous others have followed (see Fig. 14) including G310–D, G574–S, G691–R, G817–V, G891–A, and G943–S (two separate examples). Although without a consistent pattern, this implies much domain-related variability in the collagen II triple helix. The clinical phenotype of matrix cartilage disorganization correlates poorly with mutational position or type. Thus Gly–Ser574 is as damaging as two similar changes at positions 943. Contrastingly, serine substitutions of position 493 cause non-lethal and relatively mild spondyloepiphyseal dysplasia. Thus the phenotype/genotype correlations in achondrogenesis/spondyloepiphyseal dysplasia are more blurred than equivalent *COL1A1* mutations of osteogenesis imperfecta. In general, smaller substitutions such as Gly–Ser or Gly–Cys have clearer gradients of severity generally (worse at the C-terminus and mildest at the N-terminus).

In Kniest dysplasia there is a cluster of deletions close to the N-terminus (deletions of 94–108, 102–108, 124–141, 142–156, 274–279, and 361–378). The same phenotype is also caused by changes at positions 1007–1112 (at the other end of the molecule). On the other hand, Gly103–D is much more severe than its N-terminal location would suggest. Probably, shortening of the collagen triple helix is severely disruptive wherever it occurs and whilst deletions are rare in *COL1A1*, *–1A2*, and *–3A1* they are common in *COL2A1*.

The Stickler syndrome

Stickler syndrome is both clinically heterogeneous and distinctive (Fig. 12(a–c)). As originally described by Stickler it combines premature myopia with retinal detachment and early joint degeneration. Joint hypermobility is more variable. The severity of myopia varies and Stickler syndrome overlaps clinically with the Wagner syndrome, which has similar retinal, facial, and skeletal abnormalities. The Pierre-Robin anomaly with mandibular hypoplasia and cleft palate also overlaps with Stickler syndrome. Snead *et al*. (1994) have described two classes of vitreoretinal degeneration and recognize two specific phenotypes. One segregates with *COL2A1* whilst the other does not. Sometimes the latter is linked to *COL11A1* (see below).

Francomano *et al*. (1987, 1988) first showed linkage of the *COL2A1* gene to Stickler syndrome, which was also later confirmed by Knowlton *et al*. (1990). Allelic heterogeneity of Stickler syndrome was first described by Schwartz *et al*. (1989) and Fryer *et al*. (1990). Subsequently, progress has been rapid. British studies by Snead (1994) showed two sorts of vitreous morphology. One is always linked to the *COL2A1* gene whilst the other is unlinked and also heterogenous. The non-*COL2A1* phenotype includes mid-facial hypoplasia, tall or short stature, very high myopia, and marfanoid-like skeletal and facial features. Recent advances in the biology and biochemistry of vitreous composition clearly implicate protein and extracellular matrix components other than collagen II. The latter coassembles and coassociates with other collagens such as a1(II), a2(II) and a2(V) to form compound fibrils, whilst collagen VI and fibrillin 15 have also been deleted in the vitreous. Genes for all of these elements are therefore potential candidates for Stickler syndrome and related disorders. Gene linkage to either *COL11A1* or *–11A2* has been identified in certain families with Stickler syndrome, whilst *COL11A2* mutations have been described in both autosomal-dominant and -recessive osteochondrodystrophies with Stickler syndrome (Vikkula *et al*. 1995). The corresponding homozygous mutation caused severe spondyloepiphyseal dysplasia and sensorineural deafness without vitreous changes (otospondylomegaepiphyseal dysplasia syndrome). There was a homozygous glycine substitution of *COL11A2*. Although none of the *11A2* families had myopia or vitreoretinal degeneration, there was severe degenerative joint disease of the hips, knees, elbows and shoulders, and mid-facial hypoplasia with a depressed nasal bridge. *COL11A1* mutations in mice cause an autosomal-recessive chondrodystrophy with shortened snouts, cleft palate, protruding tongue, and shortened limbs. Airway obstruction, tracheomalacia, and hypoplastic lungs cause increased perinatal mortality from respiratory failure (Li *et al*. 1995). Homozygous stop-codon mutations in residue 570 of the mouse *COL11A1* gene have also been produced. The equivalent human heterozygotes show Stickler syndrome with vitreoretinal degeneration (Snead *et al*. 1994). In other Stickler families, stop-codon mutations are common. Thus, the first three-generation family with retinal and joint degeneration linked to *COL2A1* and every other published Stickler (*COL2A1*) mutant have stop-codon defects. These include several Arg–STOP abnormalities and others created by unexpected 1-bp deletions (Vikkula *et al*. 1994).

Mutations of other collagens that cause Stickler syndrome

In addition to collagen II a1(II), cartilage contains collagens a1(V), a2(V), a1(XI), and a2(XI) chains. All are interstitial collagens with similarly organized triple helices and variably large N- and C-terminuses. In both the vitreous and articular cartilage, certain components are interchangeable. For example, collagen Va2 chains can substitute for collagen XI a2 in collagen XI trimers. Thus collagens II, V, and XI are very important structural components of cartilage, which also contains collagen IX (coded by the *–9A1*, *–9A2*, and *–9A3* genes, respectively). Furthermore, hypertrophic cartilage is rich in collagen X. Latterly, mutations of such minority components have also been identified in diseases such as multiple epiphyseal dysplasia and Schmid metaphyseal dysplasia (see below).

Collagen XI mutations

There are now several convincing illustrations of the role of collagen XI in chondrodysplasia. Thus *COL11A2* has been linked to autosomal-dominant Stickler syndrome without ophthalmic changes (i.e., no myopia, retinal or vitreous pathology). The facial changes closely resemble those in Stickler's original descriptions but segregated irregularly with cleft palate and early arthropathy. The mutation changes GT to AG at the donor splice site causing an exon skip. A second homozygous Gly–Arg mutation in *COL11A2* was demonstrated by the same investigators in a more severe, autosomal-recessive chondrodysplasia characterized by severe degenerative joint disease (hips, knees, elbows, and shoulders), facial hypoplasia, and sensorineural hearing loss (Vikkula *et al*. 1995). Similarly, mouse homozygous knock-outs for *COL11A1* showed limb shortening, metaphyseal widening, and abnormal cartilage composition in the limbs, ribs, and trachea (Li *et al*. 1995). Perhaps collagen XI influences fibril formation, cartilage cohesion, and growth-plate organization.

Snead distinguishes Stickler syndrome type I from type II by the vitreoretinal phenotype. Type I Stickler syndrome always has an associated congenital vitreous abnormally with premature medium to high myopia (8–18 dioptres), whilst type II has a lesser non-congenital myopia that is of later onset. Stickler syndrome type I is always linked to *COL2A1* whilst the type II never is. Furthermore, Snead (1994) has recently identified linkage to the *COL11A1* gene in a family with Stickler syndrome type II. Here the clinical phenotype includes medium to low myopia, retinal detachment, variable joint laxity (regressing with age), minor skeletal disproportion, and cleft palate.

Collagen V, which substitutes for collagen XI in heterotypic fibrils of the vitreous, is also a strong candidate gene. Recently, a *COL11A2* mouse knock-out model has been described (Andrikopoulos *et al*. 1995). Affected homozygotes had severe kyphoscoliosis but without limb shortening. Both skin and cornea showed fibril disorganization. In this context, *COL11A2* has already been linked to variants of Ehlers–Danlos syndrome II (Burrows *et al*. 1995; Loughlin *et al*. 1995) and an exonic deletion of *COL11A2* has already been identified in a patient with Ehlers–Danlos syndrome II with corneal flattening (Nicholls *et al*. 1994). Other possible candidates include other cartilage-specific collagens such as the *– 5A1*, *–5A2*, *–9A1*, *–9A2*, and *–9A3* genes (although so far this has caused only Ehlers–Danlos syndrome-like disorders). Snead (1994) has recently suggest that Wagner and Stickler syndromes cannot be distinguished by the type or location of the *COL2A1* mutation. Instead he suggests that the vestigial vitreous of the Stickler syndrome type I-linked group is caused by dominant negative disruption of the collagen triple helix by either exonic deletions or nonsense Gly substitutions of *COL2A1*. The latter might be analogous to the collagen I-depleted bones in osteogenesis imperfecta and the collagen III-depleted arteries and skin of Ehlers–Danlos syndrome IV.

Other cartilage collagens and other extracellular matrix molecules

Collagen IX

Collagen IX is a fibril-associated (FACIT) collagen that decorates the surface of collagen II fibrils. Two mouse models have recently been created: one is a dominant-negative construct (Nakata *et al*. 1993) and the second a gene knock-out (Fassler *et al*. 1994). The latter produced vary mild disease whilst the former was severely crippling. In man the *COL–9A2* locus on chromosome 1p32 is linked to multiple epiphyseal dysplasia (Briggs *et al*. 1994). Probably the clinical phenotype is more easily recognized in man than mice. Multiple epiphyseal dysplasia is an autosomal-dominant disease in which mild short stature combines with early-onset osteoarthritis. It includes the Ribbing and Fairbank subtypes, the latter more severe than the former. Symptoms include pain and stiffness of large joints, with the onset varying between early childhood and mid-adolescence depending upon clinical severity. Pseudoachondroplasia, part of the same clinical spectrum and overlapping with multiple epiphyseal dysplasia, has much shorter hands and very stubby fingers. Cartilage biopsies show chondrocytes with unusual inclusions. Both mild and severe pseudoachondroplasia have now been linked to chromosome 19 and, subsequently, mutations of the *COMP5* gene have been identified at this locus (Briggs *et al*. 1995). Multiple epiphyseal dysplasia has also been linked to a second locus on chromosome 19 (Oehlmann *et al*. 1994).

Collagen X

Collagen X is a short-chain collagen with a helical centre and large, globular, non-collagenous regions at the N- and C-termini. It is expressed only in hypertrophic chondrocytes. The protein assembles into novel hexagonal structures and probably coassociates with collagen type II. Its actual function is unclear, though possibilities include a role in vascular invasion or matrix mineralization. Mice with dominant-negative (loss of function) mutations show defective compression of growth-plate cartilage. Contrastingly, mice homozygous for *COL10A1* knock-outs (Rosati *et al*. 1994) have virtually normal phenotypes. Presumably other collagens compensate for loss of function whilst disruption of function is very much more disturbing. Human *COL10A1* with Schmid-type metaphyseal chondrodysplastic mutants have been described both by Warman *et al*. (1993) and Wallis *et al*. (1994). The phenotype includes short stature, normal spines and hands, together with coxa vara and distorted distal femoral metaphyses. A 13-bp deletion disrupted the C-terminal propeptide. This closely resembles the *COL1A2* homozygote in severely deforming osteogenesis imperfecta. The *COL10A1* mutation is a frameshift causing rearrangement of the 60 terminal amino acids of the C-propeptide, which is also truncated by nine amino acids. Other *COL10A1* mutations have subsequently been located to the same region (Dharmavaram *et al*. 1994; McIntosh *et al*. 1994)

Osteoarthrosis (osteoarthritis)

In an American family, generalized osteoarthritis and a minor chondrodysplasia has been linked to *COL2A1*. Here, Knowlton *et al*. (1990) identified the mutant allele by segregation analysis with a 7-kb *Hind*III marker. The whole 28 kb of *COL2A1* was screened from exons 2B to 52, and an Arg(TGT)–Cys(TGC) mutation detected at position 519 of collagen type II (within the helix).

Although considered causative in this instance, such second-position (X or Y) helical changes are generally of doubtful structural significance in *COL1A1* and *–1A2*. In other collagens such as *COL4A5* there are exceptions to this rule. Allele-specific oligonucleotide hybridization with either the mutant or wild-type sequences showed assortment of the mutant sequence in family members with osteoarthritis whilst the wild-type allele was consistently transmitted to their normal relatives. Nevertheless, this may be a neutral polymorphism with an unidentified mutant sequence existing elsewhere such as exon 1, or the *COL2A1* promoter is not completely excluded. Alternatively, the change in the helical position of cysteine is pathogenetic and has produced unexpected conformational changes. More detailed mutational mapping in patients with osteoarthritis will clarify this matter. This study confirmed the pioneering studies of Palotie *et al*. (1989), who had shown linkage to *COL2A1* in a very large Finnish family with osteoarthritis. It also was consistent with the preliminary study of Hull and Pope (1989). Wordsworth *et al*. (1988) found only borderline evidence of association with *COL2A1* in a somewhat different osteoarthritis phenotype with prominent Heberden's nodes, with a non-statistically significant excess of certain haplotypes. Very probably, osteoarthritis is clinically heterogeneous so that collagen genes such as *COL2A1*, *–11A1*, *–11A2*, *–5A1*, *–5A2*, or even other constituents of extracellular matrix or bone/cartilage, are also candidates. For example, early onset of osteoarthritis with chondrocalcinosis has recently been linked to chromosome 8 (Baldwin *et al*. 1995).

*Type III collagen (*COL3A1*) mutations*

The clinical effects of type III collagen mutations closely mirror the distribution and functions of this protein. Thus, skin, blood vessels (especially arteries), pleuroperitoneum, ligaments, tendons, and the walls of gastrointestinal tracts are especially fragile and prone to damage. Often the clinical consequences can be life-threatening (Pope *et al*. 1988a, Pope *et al*. 1988b; Pope *et al*. 1996). All affect tissues in which collagen III is abundant: for example, it comprises 70 per cent of arterial collagen but rather less (35 per cent) of the cutaneous protein. Even though all such tissues are delicate, it is arterial fragility that is especially dangerous, and, unfortunately, small, medium or even larger arteries can suddenly burst at any time. In practice such catastrophes are most common in middle age, though not unknown even in adolescence. The collagen-depleted vessels are delicate and intolerant of sudden changes in arterial pressure; thus, isometric activities such as weightlifting and sprinting are potentially hazardous and should be forbidden. Pregnancy is also potentially hazardous. Rupture of the ascending or descending aorta with mediastinal bleeding can occur at any age, whilst medium-sized arteries, such as the renal, splenic, iliac, axillary, femoral, popliteal, and anterior and posterior tibial, are also fragile, especially after the age of 30 years (Fig. 16(a)). Unexpected or atypical abdominal pain should always be regarded with deep suspicion but cautiously investigated, preferably non-invasively. The internal carotid artery and intracranial arteries are commonly prone to dissection and ectasia, whilst multiple premature intracranial aneurysms are also recorded. Rupture of the internal carotid artery within the cavernous sinus is also a well-known feature (Fig. 16(b)) (Graf 1965; Pope *et al*. 1991). Cerebral aneurysms occasionally even occur in infancy and may be sporadic or familial (Pope *et al*. 1991). Easy bruising is especially common in childhood (Roberts *et al*. 1984). The external phenotype varies from classical, acrogeric Ehlers–Danlos syndrome IV with typical facial features and short stature to non-specific changes indistinguishable from benign hypermobile syndrome/Ehlers–Danlos syndrome III (Fig. 17). The latter can be very heterogeneous with various subtle subsets, one of which (with lumbosacral striae, mitral-valve prolapse, and aortic rupture) resembles the Marfan syndrome. Ehlers–Danlos syndrome III is distinguished by the normal lenses and bodily proportions in which the span equals the height, in contrast to Marfan syndrome where span exceeds height.

Typically, the cutaneous features of Ehlers–Danlos syndrome IV include thin, translucent skin with prominent capillaries, especially over the anterior upper chest, shoulders, and upper back (Fig. 17(a)). The unusually thinned skin of the hands and feet looks prematurely aged and resembles steroid atrophy. The face is often Madonna-like in younger women, showing large, attractive eyes, little subcutaneous fat, lobeless or small-lobed ears, and a thin, straight nose and lips (Pope *et al*. 1988; Pope *et al*. 1996) (Fig. 17(c,d)) Similar facial patterns can also occur rarely in Ehlers–Danlos syndrome I/II and merge with those of the normal population. Nevertheless, the special facial features when combined with short stature and slimness typify vascular Ehlers–Danlos syndrome IV. The facial features are often recognizable in family photographs, and are useful for family clinical analysis. Clubfoot deformities are common and, although improving in childhood, frequently persist into adult life. Subtle erosions (acro-osteolysis) of the finger tips are also common in acrogeric Ehlers–Danlos syndrome (Fig. 17(e)). Affected children occasionally present with congenital dislocations of the hip that can be confused with Ehlers–Danlos syndrome VII or the Larsen syndrome. Generalized laxity of joints is generally less impressive in Ehlers–Danlos syndrome IV. Low birth weight with small stature is common but non-specific. Histological examination of the skin typically shows dermal thinning, collagen depletion, and elastic proliferation (Fig. 17(b)), which also accompanies types I and II. Transmission electron microscopy shows an irregular distribution of collagen fibril diameters in skin and arterial samples (Fig. 17(e, f)), implying a role for type III collagen in the regulation of collagen fibril thickness. The latter also depends upon the relative proportions of collagen I, III and V and other unspecified components. In collagen III deficiency, collagen fibrils are irregular in size and shape, whilst in Ehlers–Danlos syndrome I/II, which is often caused by abnormalities of collagen V, cauliflower fibrils are common.

Biochemical and genetic changes

Collagen type III comprises only 15 per cent of total collagen protein synthesized by skin fibroblasts. It is measured chemically either as mercaptoethanol-treated procollagens or as unreduced or reduced a-chain monomers, which migrate separately from a1(I) collagen chains. Two-dimensional electrophoresis [similar to that of a1(I) and a2(I) collagens] is also an option, whilst the two most C-terminal peptides [a1(III)CB5 and CB9] comigrate. Mercaptoethanol-reduced monomers of CB5 + 9 cannot be separated, whilst if the disulphide bond of CB9 (which possesses a Cys residue) remains intact, CB9 runs as a dimer whilst CB5 is separated as a monomer. Unfortunately, the smaller peptides CB3, 4, 6, and 8 are less easily resolved. Vertebrate collagenase, which cuts collagen type III into N-terminal three-quarter and C-terminal one-quarter fragments, is also useful for molecular mapping. Because type III collagen is much less abundant than type I, peptide mapping of type III has lagged behind that of type I, in which protein mapping and subsequent gene amplification is relatively simple. To date, less than between one-third and one-half as many *COL3A1* mutants as *COL1A1*/*1A2* defects have been published.

In classical, acrogeric Ehlers–Danlos syndrome IV (Fig. 17(c,d)), little or no type III collagen is secreted. Instead, overmodified mutant protein is retained or degraded intracellularly. Occasionally, diminished or modified protein is secreted without intracellular retention (Fig. 18(a, c)). These mutations have mechanisms different from those that cause retained or secreted, overmodified type III collagen, which frequently have impaired helical winding, similar to certain osteogenesis imperfecta mutants. Exceptionally, in lethal Ehlers–Danlos syndrome IV, the secretion of type III collagen varies between 50 and 80 per cent of normal (Nicholls *et al*. 1988).

Peptide mapping of normal collagen III by one- or two-dimensional electrophoresis (Fig. 18(c,d)) is usually straightforward. Protein extracts can also be stained with silver or Coomassie blue. Very occasionally, type III collagen is normal in classical Ehlers–Danlos syndrome IV, although CB5 and 9 are normal (Pope *et al*. 1988a, Pope *et al*. 1988b; Pope *et al*. 1991). Perhaps the error is located in an N-terminal peptide such as CB3, or similar fragments such as CB1, 2 or 7, which are too small to be detected. Other possibilities are electrophoretically silent mutations or non-allelic phenocopies caused by other mutant genes. The degree of protein overmodification is both position- and residue-dependent. Just as for *COL1A1* and *COL1A2* mutants, abnormalities nearest to the C-terminus produce the most overmodified protein. 'Smiling' gels similar to those that can accurately locate abnormalities of a1(I), a2(I) or a1(II) should also discriminate between a1(III)CB5 and a1(III)CB9 mutations. A good example is the Ehlers–Danlos syndrome IV family with variable type III secretion mentioned earlier (Nicholls *et al*. 1988) which, when mapped by either one- and two-dimensional CNBr peptide gels produced a shortened form of a1(III)CB5. PCR showed an in-frame intraexonic deletion of 27 bp (caused by spliced mispairing) (Richards *et al*. 1991). Single-base substitutions, gene deletions, and splicing abnormalities are common, whilst other mechanisms such as base insertions are less frequent. Null alleles (many with premature chain termination) may also be expected. So far we have identified more than 20 glycine substitutions or exon skips (Fig. 19). Mutational position and type correlate reasonably well with clinical severity (Pope *et al*. 1996). Already several other

splice-junction mutations have been described.

With the methods outlined above the analysis of abnormalities in collagen III protein is simple. If the intracellularly retained protein is so very sparse it cannot be easily mapped (by immunoprecipitation, CNBr cleavage or proteinase digestion), then other approaches become necessary. Thus, specific cDNA or genomic fragments can be amplified by PCR with strategically placed primers, combined with mapping of S1 nuclease (Kuivaniemi *et al*. 1991) or RNAase A (Grange *et al*. 1988), or by the chemical cleavage of mismatched cytosines or guanosines by OsO_4 or hydroxylamine (Cotton *et al*. 1989). This strategy previously identified a 100-bp deletion (Cole *et al*. 1990) and we have also used it successfully. Since glycine (GGX) mutations are relatively common, cleavage with hydroxylamine is potentially more useful than with OsO_4, which is also technically more demanding. Cleavage accurately locates mismatches and requires direct or indirect sequencing to identify the actual mutant sequence. The method is particularly relevant for those patients with Ehlers–Danlos syndrome IV who have normal secretion and peptide maps of collagen III. We have previously identified several different Ehlers–Danlos syndrome mutants by this method (Fig. 18(e)). Alternatively, the whole cDNA can be amplified in several convenient portions and subsequently automatically sequenced (by machine), which theoretically shortens the process to as little as 7 days whilst conventional analysis takes at least 6 months. This includes 4 to 6 weeks each for primary cell culture and for protein mapping. The latter step can be omitted and cDNA or genomic mutational analysis and sequencing undertaken instead. Since each mutation is unique and common disorders such as cerebral or aortic aneurysms are also very varied, the analysis of individual mutations requires regional or nation-wide resources and may not be cost-effective in preventing cerebral aneurysms.

The current situation

In 1991, 15 *COL3A1* mutations had been identified (Kuivaniemi *et al*. 1991) but by 1995 nearly three times this number have now been recorded. Many of the 20 that we have sequenced lie within CNBr peptides 5 and 9. For technical reasons, mutations are more easily identified here but ascertainment (of the acrogeric clinical phenotype) also produces clustering. In contrast to osteogenesis imperfecta, in which the clinical gradient is well defined (least severe at the N-terminus and greatest at the C-terminus), *COL3A1* mutants are uniformly disastrous wherever they lie. A chemical gradation is still obvious, as C-terminal mutants are chemically more disruptive and may also induce more severe clinical phenotypes such as fragile arteries. Yet the nature of arterial rupture is such that even a single episode is potentially lethal. This contrasts totally with osteogenesis imperfecta in which the most severe clinical phenotype arises following the most numerous and continual fractures. The clustering of mutants with acrogeric Ehlers–Danlos syndrome IV who have a very high risk of arterial rupture has several unexpected exceptions, such as Gly–Glu625, which includes mild generalized joint laxity and late-onset osteogenesis imperfecta with Heberden's nodes but with apparently normal arteries. Originally we had included Gly–Asp883 with mild external features without arterial ruptures in this category; unfortunately, however, the affected father and daughter (aged 57 and 25 years respectively) recently both suddenly died of arterial rupture.

Mutations identified to date (Fig. 19)

Amongst the published defects (Kuivaniemi *et al*. 1991, 1995; Pope *et al*. 1996) are point mutations Gly619, Gly790, Gly883, Gly910, Gly1018, and several splice-junction mutations of exons 16, 20, 25, 27, 41, and 42 (Kontusaari *et al*. 1990; Kuivaniemi *et al*. 1990). All have changes at or near the crucial splice donor sites with complete exons always misspliced in-frame. The IVS20 mutation described by Kontusaari *et al*. (1990) is a typical example. Here the proband, a 34-year-old male, died of acute abdominal, retroperitoneal, and mediastinal haemorrhage. Other family members such as his brother and father had also died prematurely of arterial rupture.

We have recently observed several other unusual examples including two exon-24 skips caused by very similar splice-junction errors. In one affected father and daughter, the secretion of collagen III was reduced but a slightly shortened and an overmodified protein was secreted. The clinical phenotype included extreme fragility of medium-sized arteries. The proband eventually died from sudden aortic rupture but had previously undergone a mid-thigh amputation to control torrential bleeding from a routine operation to strip varicose veins. He also had suffered several previous pneumothoraxes. His affected daughter had survived avulsion of her common iliac artery and vein following a minor fall. The external clinical phenotype was non-specific but included large eyes and lobeless ears. Previously the proband had a provisional diagnosis of the Marfan syndrome but had a normal span/height ratio and no arachnodactyly or lens dislocations. In contrast, the second exon-24 skip showed a typical acrogeric facial phenotype but virtually normal patterns of type III collagen. Presumably the mutant splice-junction site was mostly ignored and normal collagen III mRNA spliced instead.

There was similar clinical variability in three different splice-junction skips at exon 37, of which two were acrogeric. Two had suffered fatal, late arterial rupture, one with several previous arterial tears that had been successfully repaired and the other with numerous simultaneous tears. So far, the youngest patient, although acrogeric, is free of arterial damage. Various other exon skips include missplicing of exon 43 in a two-generation family, two of whom had carotid–cavernous sinus aneurysms and various combinations of left renal and splenic aneurysms. Here the phenotype included large prominent eyes, slightly thin skin, and increased elastic proliferation on histological examination. Other mutants are detailed in Pope *et al*. (1996) and Fig. 19.

Mutation analysis

Our current strategy is to amplify *COL3A1* cDNA in four or five overlapping fragments. As we do not use genomic amplification (of the 52 exons), certain subtle in-house mutants may be missed. Stop-codon mutants are common in *COL2A1* and *COL7A1* mutants (see above and below). cDNA amplification of *COL3A1* mRNA is followed by density-gradient electrophoresis of the 3' coding sequences covering most of CB5 and 9. Originally, PCR products were then indirectly cloned and sequenced. Latterly, we use direct sequencing of double-stranded, PCR-amplified DNA (which is faster, easier, and more efficient). Early errors detected by others include deletions of exon 20 with a complicated mix of misspliced products such as exons 19, intron 20, and exon 19 in variable proportions (ranging from 24 to 132 nucleotides) (Tromp *et al*. 1989). Genomic clones showed a G–A splice-junction substitution in the first base of intron 20. Some variably spliced products deleted exon 20 (54 bp) but others included variably-sized intron sequences transcribed from the first, illegitimate, consensus splice-junction sequence.

Pepsin-resistant type III collagen secreted by cultured fibroblasts was decreased, whilst skin and arteries were deficient in type III collagen. The G–A mutation destroyed a *Bst*U1 site allowing the detection of affected family members. One of the proband's three children and his clinically normal sister were also mutants. Allele-specific oligonucleotides with either normal G or the mutant A sequence were also useful. Here, mutant DNA hybridizes to the mutant (A) oligonucleotide whilst wild-type sequences hybridize with the normal (G) in DNA dot-blots. Deletions of exons 16, 25, 27, 41, and 42 were mentioned in the first edition of this book, and more recently skipping of exons 7, 14, 24, 37, 43, and 45 has also been reported (Pope *et al*. 1996).

Gene deletions or insertions

Just as in osteogenesis imperfecta, there may be occasional large, and also smaller, deletions ranging from 27 to 7500 bp (7.5 kb) that are distinct from exon skips. Superti-Furga *et al*. (1988) have described two remarkable examples. One had rather mild Ehlers–Danlos syndrome IV and had successfully completed a normal pregnancy. Her fibroblasts produced low amounts of normal and shortened a1(III) homotrimers, whilst mixed heterotrimers were retained intracellularly. Southern blotting showed a 7.5-kb deletion extending from intron 33 to exon 48. This was an in-frame deletion between residues 586 and 999. A second, slightly smaller deletion induced more severe disease in an affected father and son. The father suffered a fatal intra-abdominal haemorrhage and the surviving son secreted two forms of type III procollagen/collagen, the mutant being shorter than usual. Typical clinical features of Ehlers–Danlos syndrome IV included thin skin, lobeless ears, and acrogeria. The protein deletion was estimated at 220 amino acids, based upon molecular weight. More detailed DNA analysis showed a 3.3-kb genomic deletion from exon 9 to 24 with a complex variable tandem repeat causing potential asymmetrical recombination at meiosis (Lee *et al*. 1991a). This is also a useful population and linkage marker with a three-allele deletion pattern (Superti-Furga *et al*. 1990). In both, the deletion breakpoints are now defined (Lee *et al*. 1991b; Vissing *et al*. 1991).

A third (2 kb) deletion was observed by McGookey *et al*. (1989) (Milewicz and Durie 1994). Here a female with sporadic Ehlers–Danlos syndrome (type unspecified) produced traces of procollagen type III, which, after pepsin isolation, ran as two closely apposed type III homotrimers. The cell-layer collagens were not described, although mixed heterotrimers were probably retained or degraded intracellularly. The proband's clinically normal father also secreted traces of the mutant homotrimer but nearly normal amounts of the wild-type protein. Genomic DNA probed with a *COL3A1* cDNA deleted a 5' *Hind*III fragment. The father's lymphocyte cDNA did not contain the deletion (as judged by Southern blotting). He was probably a mixed gonadal–somatic mosaic and the mutation had occurred after embryonic divergence of fibroblast and lymphocyte precursors. This was the first published example of gonadal mosaicism in Ehlers–Danlos syndrome IV. Just as in osteogenesis imperfecta, this mechanism is relatively common, especially in families with multiple affected siblings and clinically normal parents.

The mutation has subsequently been more fully characterized. Other smaller deletions have also been described. For example, as mentioned earlier (Richards *et al*. 1991), we have studied a three-generation Belgian family with variable production of collagen type III. Protein mapping showed a deletion in a1(III)CB5 (Fig. 18(d))

(Nicholls *et al*. 1988). This was caused by a novel, in-frame, 27-bp deletion within the 108-bp exon 37 close to the C-terminus, the so-called spliced, mispaired, homologous intronic sequences (Richards *et al*. 1991). Similar mechanisms also occur in globin and other genes (Richards *et al*. 1991). Although undetectable in whole a-chains, peptide mapping with CNBr was unambiguous.

Mutations such as point mutations, deletions or insertions are easily demonstrated by allele-specific amplification or hybridization. The strategy works well for carrier genomic sequences prepared from chorionic villus biopsies, amniocentesis or ordinary blood samples of appropriate family members.

Genetics of Ehlers–Danlos syndrome

Ehlers–Danlos syndrome is clinically heterogeneous and characterized especially by cutaneous fragility with variable permutations of skin thinning, ligamentous laxity, short stature, spinal deformity, vascular fragility or retinal detachment. By certain combinations of clinical and biochemical characteristics there are at least nine distinct genetic subsets, of which types IV, V, VI, and VII have specific biochemical abnormalities. Types I/II (see below) can also now be added, as mutations of collagen type V cause at least some of them. Collagen proteins are faulty in types I, II, III, IV, and VIIA and B, whilst types VI, VIIC, and IX are caused by enzyme mutations. Ehlers–Danlos syndrome type VII is caused by either structural or enzyme abnormalities. With rare exceptions the causes of Ehlers–Danlos syndrome types III and VIII are so far unknown, whilst the existence of Ehlers–Danlos syndrome V is controversial. Thus lysyloxidase deficiency is a strong candidate, usually causing either cutis laxa with bladder diverticula or Menkes syndrome. Collagen III mutations cause vascular Ehlers–Danlos syndrome and N-terminal collagen I mutations cause autosomal-dominant forms of Ehlers–Danlos syndrome VIIA and B. The latter patients differ from those with Ehlers–Danlos syndrome types I and II only in stature; they are usually petit(e) and in more severe cases have greater looseness of joints and ligaments, especially of fingers, shoulders, knees and hips, together with late scoliosis. Collagen V abnormalities (see below) cause at least some types of Ehlers–Danlos syndrome II, which is probably allelic with Ehlers–Danlos syndrome I. Other candidates for this phenotype include the *COL5A1*, *–11A1*, and *–11A2*. Cutaneous fragility is the norm in Ehlers–Danlos syndrome I but can be as marked in VIIA and B. (Fig. 20(a–f)) Compound heterozygosity for collagens I, III, V, and XI is also possible, as are double heterozygotes for collagen, elastin, fibrillin, lysyloxidase, lysylhydroxylase, and other components of the extracellular matrix.

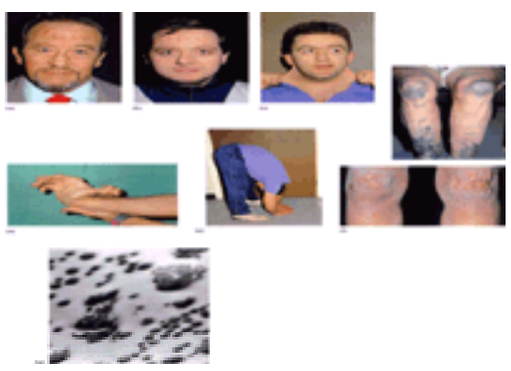


Fig. 20 (a, b) Facial features of affected father and son with Ehlers–Danlos syndrome I, both showing typical facial scarring. (c) There is also often abnormal cutaneous extensibility or epicanthic folds. (d, e) Joint laxity is both widespread and commonplace. (f) Scarring of the shins and knees is also usual, closely resembling that of certain patients with Ehlers–Danlos syndrome VI (see Fig. 10(e)). (g) Cauliflower fibrils are typical (as viewed by transmission electron microscopy × 62 500).

Ehlers–Danlos syndrome VI is caused by specific enzymatic abnormalities, especially lysylhydroxylase deficiency (from homozygous or compound heterozygous gene mutation) causing Ehlers–Danlos syndrome VI. The latter differs from other subtypes of Ehlers–Danlos syndrome by its inheritance, which is either autosomal-recessive or compound. In contrast to other subtypes there is severe infantile or adolescent scoliosis and spectacular generalized joint laxity. The lysylhydroxylase gene lies on chromosome 1p36 (Hautala *et al*. 1992). Known mutations include homozygous termination-codon duplication rearrangements and both complex and single exon skips (Pousi *et al*. 1994; R. Myllyla and F. M. Pope, unpublished observations). Ehlers–Danlos syndrome type VIIC closely resembles types VIIA and B but produces much more disorganization of collagen fibril, as illustrated by hieroglyphic fibrils in transversely sectioned electron micrographs. Here the pNa1 and pNa2 propeptide extensions remain uncleaved so that the entire triple helix is faulty. All mutant molecules then misassemble to form flanged fibres that appear in transverse sections as hieroglyphs. In humans the clinical phenotype includes spectacular cutis laxa, joint laxity, and short stature in adults, differing only slightly from the types VIIA and B phenotypes in which either the pNa1 or the pNa2 chains are faulty. Here, family trimers incorporate either one or two abnormal chains. Lysyloxidase deficiency can potentially cause several osteogenesis imperfecta disorders, but it is unclear whether the families with X-linked Ehlers–Danlos syndrome V originally described by Beighton were lysyloxidase-deficient. Lysyloxidase deficiency also causes a syndrome of cutis laxa and extreme joint laxity, with bladder diverticula, a hydronephrosis, and chronic renal failure, which overlaps with both Ehlers–Danlos syndrome III and VI but is a much more severe and a separate disease. Deficiency of procollagen peptidase produces autosomal-recessive Ehlers–Danlos syndrome VIIC, as described above. Transmission electron microscopy shows angulated irregularities of collagen fibril shape in Ehlers–Danlos syndrome VII (most extremely in Ehlers–Danlos syndrome VIIC), with size irregularities in *COL3A1* deficiency (Ehlers–Danlos syndrome IV), some Ehlers–Danlos syndrome III, and hereditary aneurysms and spectacular disorganization (so-called cauliflower fibres in both Ehlers–Danlos syndrome I and II and occasionally pseudoxanthoma elasticum). There are animal models for *COL5A1* and *COL11A2* deficiency in which mutant mice show fibril irregularities in both homozygotes and heterozygotes. Cauliflower fibrils have not been observed, although there is an Ehlers–Danlos syndrome I/II rabbit model that has spectacular cauliflowers, fused fibres, and considerable bending and fraying of the affected structures (in skin at least) (Fig. 20(g)).

Collagen type V mutants

Three distinctive and separate pieces of evidence implicate collagen V (a1 and a2 chains) in the aetiology of Ehlers–Danlos syndrome I/II. Firstly, screening of numerous families with syndrome types I and II has shown occasional abnormalities of structural proteins (Nicholls *et al*. 1994). One patient had mixed clinical features of Ehlers–Danlos syndrome I/VII as judged by short stature, extensive skin fragility (of the forehead, knees and shins), and generalized joint laxity (Fig. 21(a,b)). Electron microscopy of the skin showed slight irregularity of fibril shape without cauliflower fibrils (Nicholls *et al*. 1994). Subsequently, an exonic deletion and genomic splice-junction mutation of *COL5A1* was detected. In a second patient with similar extreme clinical features but unexplained recurrent intraperitoneal bleeding, collagen V chains were both overmodified. By analogy with osteogenesis imperfecta either *COL5A1* or *–5A2* are implicated, but so far the mutation is unidentified. Two very large British families of Ehlers–Danlos syndrome II have also been linked to *COL5A1*. Thus Wordsworth and his colleagues (Loughlin *et al*. 1995) recently showed linkage in a large, four-generation, British family with syndrome type II and a lod score 11.9. We have studied a family with mixed syndrome I/II with a lod score of 4.6 (Burrows *et al*. 1996). Recently a translocation deletion causing both Ehlers-Danlos syndrome and hypomelanosis of Ito has been described (Torriello *et al*. 1996). Mouse mutants, heterozygous and homozygous for a *COL11A2* knock-out, demonstrate the equivalent phenotype consistent with human Ehlers–Danlos syndrome II. Thus heterozygotes had abnormal histological appearances of skin and mild ligamentous laxity, whilst homozygotes had much more extreme skin changes with severe spinal deformities. Electron microscopy showed misassembled fibres of skin and cornea (the fibres were both enlarged and irregular).

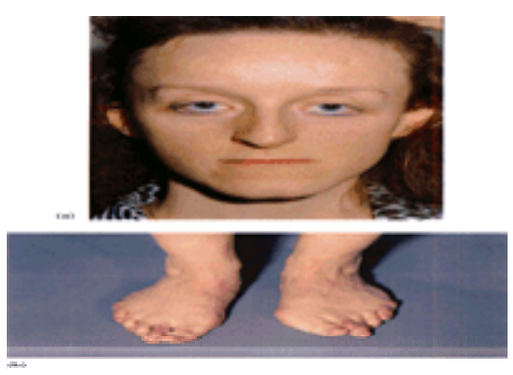


Fig. 21 (a, b) Face and feet of a patient with Ehlers–Danlos syndrome with collagen V mutation. The eyes have flattened corneas but the generalized physical

features strongly resembled those shown in Fig. 10 and Fig. 20 (i.e. were phenotypically similar to Ehlers–Danlos syndrome I, II, and VII).

Collagen type IV mutations

Understanding of the genetics and molecular mechanisms of Alport syndrome has also progressed very rapidly since 1991. Then, X-linked Alport syndrome with haematuria, thinning of the glomerular basement membrane, and deafness had been linked to Xq21.3–Xq22.2 and to the *COL4A5* gene (Hostikka *et al*. 1990), with a lod score of 9.1. There are six other types of Alport syndrome (McKusick *et al*. 1990): type I (classic juvenile with deafness), type II (X-linked juvenile with deafness), type III (X-linked adult with deafness), type IV (X-linked adult without deafness), type V (autosomal-dominant with deafness and thrombocytaemia), and type VI (autosomal juvenile with deafness). All share a symptom complex of deafness, haematuria, and thickening or splitting of the glomerular basement membrane through the lamina densa, but except for the type III disorder are caused by autosomal genes. Overall this is a common disease with an overall frequency of 1/5000.

By 1991 the first *COL4A5* mutation had been discovered by Southern blotting, which was used to study one family with a 15-kb genomic deletion of 5 including exons 5 to 10. The collagen a5(IV) protein was shortened by 240 amino acids. Soon after a second family was studied in which the third exon from the C-terminus contained a Cys–Ser mutation (Zhou *et al*. 1990). This probably impaired folding of the C-terminal globular domain, which is heavily reliant upon specific disulphide bridges. Consequently, normal helix formation and intermolecular assembly must be severely compromised. All families with X-linked Alport syndrome presumably have mutations of *COL4A5*, but, up to 1991, the mechanism of autosomal-dominant Alport syndrome was unclear. The other members of the type IV collagen gene family (*COL4A1–4A4*) are obviously strong candidates. Since then, more than 150, all mostly private, mutations have been described, including large deletions, insertions, glycine substitutions, and stop codons. Autosomal variants have been caused by both *COL4A3* and *COL4A4* mutations, and there are very probably *COL4A1* and *–4A2* mutants as well. The basement membrane components are at least tissue-specific.

Collagen type VII mutations

Collagen VII is a basement membrane-associated protein that forms distinctive, cross-striated fibrils, easily visible by transmission electron microscopy. The protein is found in stratified squamous epithelium and amnion, from which it is usually purified. The fibrils attach epidermal basement membrane to underlying dermal connective tissue, inserting both into the lamina densa and into the anchoring plaques of the underlying stroma, thereby anchoring the basement membrane to the underlying reticular dermis. Faults in the organization or composition of these structures severely disrupt the stable adhesion of epidermis to dermis. Such faulty tissue spontaneously splits at the level of the basement membrane/reticular dermis, with tense and painful, haemorrhagic, subepidermal blisters that eventually heal with scarring. There are both autosomal-dominant and -recessive forms. In the former (see below), painful blisters occur over extra-articular sites, such as the knees, elbows, knuckles, and lateral malleoli. In severe, autosomal-recessive dystrophic epidermolysis bullosa, blistering extends to the whole integument, including nails, trunk, and face, and also to buccal, oesophageal, and genitourinary mucous membranes (see below) (Fig. 22(a–d)).

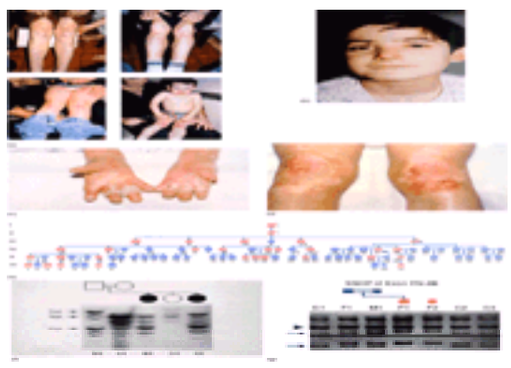


Fig. 22 (a) Montage of family with autosomal-dominant dystrophic epidermolysis bullosa (DEB) showing typical scarring and blistering of knees, shins, and the hands. (b, c, d) Severe recessive (R) DEB showing facial scars (b), mittening of the hands (c), and severe scarring of the knees (d). (e) Typical multigenerational pedigree of autosomal-dominant DEB. (f) Segregation of the variable random sequence D351100 with RDEB. The combination of parental alleles 4 and 6 carries both mutant genes. (g) Single-stranded confirmational analysis of exon FN4B of the *COL7A1* gene showing segregation of severe DEB with the abnormal component.

Collagen VII protein has C-terminal and N-terminal globular ends, of which the former is the larger. There is also a central interrupted triple helix. Individual molecules stack in antiparallel arrays. The C-terminus accounts for half the molecular weight of the entire protein. Western blots with various polyclonal and monoclonal antibodies to the globular and helical epitopes detect a 290-kDa protein with 140-kDa C-terminus.

Clinical phenotype

Epidermolysis bullosa dystrophica has autosomal-dominant (mild) (Fig. 22(a,e)) and -recessive (lethal) forms (Fig. 22(b–d)), both characterized by subepidermal blisters that split through the lamina of the basement membrane (Eady 1990). Skin sections stained with antitype VII antibodies stain the basement membrane either poorly or not at all. In some cases (e.g. the epidermolysis bullosa dystrophica inversa type), mutant type VII collagen protein is retained intracellularly. Such changes are analogous to those found in *COL1A1*, *–1A2*, *–2A1*, and *–3A1* mutants in diseases such as osteogenesis imperfecta, chondrodysplasia or vascular Ehlers–Danlos syndrome (see above), all of which also frequently impair helical stability or secretion of the mutant molecules. In the case of epidermolysis bullosa dystrophica and in contrast to such as Ehlers–Danlos syndrome and osteogenesis imperfecta, many are caused by stop-codon rather than helical glycine substitutions, or exon skips. Electron microscopy of the skin from such *COL7A1* mutations shows depleted or even no anchoring fibrils at all (largely dependent upon the type and clinical severity of the defect).

Since 1991 there has been impressive progress in exploring the molecular structure and analysis of *COL7A1*. Not only has the cDNA and genomic organization of the gene been completed (Parente *et al*. 1991; Christiano *et al*. 1992) but it has also been convincingly implicated in the causation of both subgroups of epidermolysis bullosa dystrophica. Thus the linkage of autosomal-dominant and -recessive dystrophic epidermolysis bullosa to *COL7A1* at 3p21 was rapidly followed by various mutations causing both dominant and recessive epidermolysis bullosa dystrophica (Fig. 22(f)). NC1, non-collagenous, helical, and NC2 mutations have all been described. They include exon skips, chain-termination codons, and helical glycine substitutions. Combinations include heterozygous dominant, homozygous recessive, and double heterozygous mutations (Hallopeau–Siemens inversa).

Gene structure and organization

Anchoring fibrils contain predominantly collagen VII protein (Lunstrum *et al*. 1986; Saiki *et al*. 1986;), although interactions with other matrix components such as collagen V and fibrillin are also possible. The collagen VII gene (*COL7A1*) is compact (32 kb) but complex, with 118 exons separated by rather small introns. The 9.2-kb mRNA encodes a 340-kDa protein (Parente *et al*. 1991; Christiano *et al*. 1992; Greenspan 1993). The NC1 domain is large, containing several distinct elements such as cartilage matrix protein 9, fibronectin III, and von Willebrand factor domains, and Cys/Pro sequences. Next is a triple helical collagenous domain and lastly the *C2*-coding region, the whole gene containing 118 separate exons (Christian *et al*. 1994a). Individual molecules stack antiparallel producing the typical cross-banded anchoring fibril joined at the epidermal end to the lamina densa and at the dermal end to so-called anchoring plaques.

Mutations of COL7A1

Since 1991 there has been very rapid progress both in the identification of the gene locus and in the analysis of mutants. This may be summarized as follows. Firstly, five independent studies have confirmed the linkage of *COL7A1* markers to all autosomal-dominant and -recessive variants of dystrophic epidermolysis bullosa. With regard to the dominant form, linkage was independently observed by American, British, and Dutch groups (Ryynanen *et al*. 1991, 1992; Al-Imara *et al*. 1992; Ryynanen *et al*. 1992; Gruis *et al*. 1992). The Americans tested both Finnish and American pedigrees whilst the British and Dutch looked at native families only.

Various markers close to the *COL7A1* locus at 3p21 were used initially, sometimes supplemented with intragenic *COL7A1* RFLP. Two different groups have studied linkage of *COL7A1* to severe autosomal-recessive disease (dysmorphic epidermolysis bullosa including so-called Hallopeau–Siemens (mitis) and inversa subtypes (Hovnanian *et al*. 1992). These were grouped together as generalized recessive dystrophic epidermolysis bullosa by the British investigators (Dunnill *et al*. 1994a). Fifty-two French and Turkish families with at least one living offspring were studied by Hovnanian *et al*. (1992), whilst Dunnill *et al*. (1994a) studied a mixed population of British, Italian, Irish, and South African families. Lod scores were 4.94 (British) and 3.97 (French). Both groups studied mutations on the assumption that N-terminal stop codons were likely (as proposed originally by Hovnanian). Four patients with heterozygous stop codons changing CpG sequences such as CGA(Arg)–UGA(STOP) were detected in the first eight exons of *COL7A1* to be tested by Hovnanian *et al*. (1992). Dunnill *et al*. (1994) also opted to screen the NC1 coding region in their British/Irish/Italian/South African patients. They found four stop-codon mutants, three of which were identically heterozygous but possessed different second mutations, whilst the fourth was homozygous for a different stop codon in the fibronectin III repeat at 4BN to the other three (Fig. 22(b–g)). The two homozygously affected children had mutilating epidermolysis bullosa and with C–U transition changing Arg to STOP. This abolished an Xho site for which both parents were obligate heterozygotes and which the children were homozygous. The carrier parents had clinically detectable, Hallopeau–Siemens epidermolysis bullosa inversa.

Two other mutations were also published at about this time. The first was a homozygous Met–Lys substitution in the non-collagenous NC2 domain producing relatively mild dystrophic epidermolysis bullosa mitis (Christiano *et al*. 1993). The second was a homozygous stop codon in the NC1 domain caused by a single base-pair insertion deletion (Hilal *et al*. 1993). This produced autosomal-recessive disease and severe mittening deformities of the hands and feet—the so-called Hallopeau-Siemens pattern. Subsequently, Dunnill *et al*. 1994*a*; *b*; 1995*a*; *b*; have described a nonsense mutation in the NC1 domain at position 578 in two unrelated individuals, one of whom had a second mutation deleting residue 789 and also creating a premature stop codon. In contrast to osteogenesis imperfecta/Ehlers–Danlos syndrome IV mutants, which show an increasing gradient of severity from N- to C-terminus, the pattern in severe recessive/compound epidermolysis bullosa dystrophica is the reverse. Thus, homozygous stop-codon mutants in the NC1 domain are very severe whilst the homozygous inversion of NC2 is much less so. Presumably, stop-codon homozygosity that produces no mutant products is less disruptive to the interactions of anchoring fibrils with other components and the clinical disorder much milder as a consequence. This has strong parallels with homozygous *COLX* mouse knock-outs in which the clinical phenotype is barely noticeable whilst single, dominant, *COLX* Gly mutations produce notable abnormalities in mice and men. In other words, loss of function can be much less damaging than altered function. Since dystrophic epidermolysis bullosa is relatively rare (with fewer than 100 families with dominant or recessive forms in the United Kingdom) (McGrath *et al*. 1993, 1995c), but with equal numbers of sporadic cases, the immediate future strategy will be to identify as many mutations as possible to optimize genetic counselling and prenatal diagnosis. Combined evidence from worldwide research should then provide a phenotype/genotype map similar to those already available for osteogenesis imperfecta, Ehlers–Danlos syndrome and the chondrodystrophies, before the development and exploration of animal models, then the development of gene therapy (a specific lesion of stratified squamous epithelium of this type should be one of the more easily soluble models for successful gene therapy) (Christiano *et al*. 1994b, Christiano *et al*. 1995, Christiano *et al*. 1996a, Christiano *et al*. 35 b).

In contrast to *COL7A1*, homozygous *COL1A1, COL1A2*, and *COL3A1* mutations are either extremely rare or missing. Possibly, in these instances, double stops would be genetically lethal as the affected proteins are more physically widespread and indispensable. Single stops of *COL2A1* frequently cause the Stickler syndrome. Overall, progress since 1994 has been very rapid, coinciding with technical improvements in PCR, heteroduplex analysis, and the demonstration of the full automated genomic sequence of *COL7A1*. The gene is compact but complex, though examination of the whole, 32-kb, 118-exon gene is now feasible. To date, many premature stop-codon mutants have been identified in autosomal-recessive, Hallopeau-Siemans epidermolysis bullosa and dystrophic epidermolysis bullosa mitis. Premature stops or glycine substitutions have been equally frequent in heterozygous dominant dystrophic epidermolysis bullosa. Other mechanisms of similar basement membrane non-connected genes include defects of laminin 3 and the bullous pemphigoid antigen BP 180 (McGrath *et al*. 1995a, McGrath *et al*. 1995b).

Marfan syndrome (Fig. 23)

The Marfan syndrome is characterized by long extremities, fingers and feet (arachnodactyly), tall stature, pectus deformities of the chest, facial and optical abnormalities such as dislocated lenses, high myopia, high-arched palate, and mandibular hypoplasia. There is also a distressingly common predisposition to sudden aortic rupture, particularly in middle adulthood (Fig. 23) (Marfan 1896). Mitral-valve prolapse, joint laxity, and cutaneous striations are also common. None of these individual physical features in themselves is specific. It is their subtle combination that makes the disorder so recognizable. Beighton *et al*. (1988) have clearly defined major and minor diagnostic criteria, various permutations of which allow a reasonably rational diagnosis to be made or suspected. Major criteria include disproportionately tall stature with arachnodactyly and skeleton deformity, lens dislocations, and aortic-valve disease and dilation.

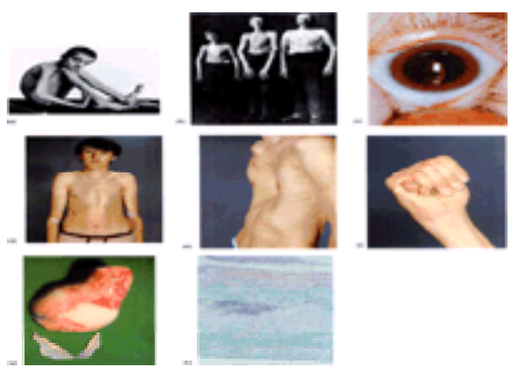


Fig. 23 Marfan syndrome phenotypes: (a) Marfan's original patient; (b) original family described by Weve (1931)—note chest deformities and elongated fingers and arms; (c) typical dislocated lens from individual with classical Marfan syndrome; (d, e) pectus excavatum and carinatum from two different patients with Marfan syndrome; (f) positive thumb sign (the grasped thumb protrudes beyond the width of the palm; (g) dilated aortic root from typical Marfan patients; (h) medial elastic degeneration of Marfan syndrome.

Although Marfan syndrome is one of the original 'inherited defects of connective tissue' (McKusick 1956; McKusick 1972), until quite recently its molecular genetics was obscure. Nevertheless, it is a clear autosomal-dominant (Weve 1931) (Fig. 23(b)) with complete penetrance and a prevalence of approx. 1/25 000.

Since mid-1991, complete molecular understanding of the disease has largely been achieved. Firstly, the abnormal gene was located by an international consortium whose combined evidence (Blanton *et al*. 1990) suggested chromosome 15 as the most likely gene locus. Next, specific gene linkage to *D15S45* was identified by Peltonen and her colleagues. Extended studies of 17 authentic families with classical Marfan syndrome by five different American and Northern European research groups quickly showed allelic homogeneity for the fibrillin 15 locus (Kainulainen *et al*. 1991). The only exception is a very large French family linked to 3p21 (Collod *et al*. 1995). Here the Marfan phenotype is disputed (Boileau *et al*. 1995; Dietz *et al*. 1995). Fibrillin 15 is a completely convincing candidate whose properties are consistent with known physical properties of Marfan syndrome. For example, the tissue distribution of fibrillin matches the known physical signs of the syndrome, such as the clinical and structural abnormalities of the aortic media, and faulty suspensory ligaments of the lens, skeleton and skin. Other preliminary evidence included poor immunofluorescent staining of the dermis and cultured skin fibroblasts from most Marfan patients (Godfrey *et al*. 1990).

Subcloning and sequencing of the human fibrillin 15 gene

Maslen *et al*. (1991) elucidated both the cDNA sequence structure and chromosomal location of a fibrillin on chromosome 15, whilst Lee *et al*. (1991*b*) also located a second fibrillin gene on chromosome 5, hinting at another on chromosome 2. Very quickly, Dietz *et al*. (1991) detected the first specific mutations in the Marfan syndrome. Soon after the Marfan-like syndrome, congenital contractural arachnodactyly, was linked to the fibrillin gene on chromosome 5 (Tsipouras *et al*. 1992). Congenital contractual arachnodactyly has many Marfan-like features but without vascular fragility. Fibrosis and contractures of certain ligamentous and cartilaginous structures such as the small flexor tendons of the hands or feet, with crumpling of the external ear, are prominent. Such clinical variability probably reflects the tissue distribution of particular fibrillin proteins.

Molecular structure of the fibrillin 11 gene

The two American groups who originally characterized the fibrillin 11 gene adopted slightly different approaches. Maslen *et al*. (1991) had screened a placental cDNA library with universal degenerate oligonucleotides of a 20-bp fibrillin peptide sequence. They located overlapping cDNA clones containing 6.9 kb of a predicted 10-kb

mRNA. Sequencing showed several novel features, including both 6- and 8-Cys motifs and an 184 amino-acid carboxyterminus. The overall 394 amino-acid sequence was homologous to various unexpected but well-known protein motifs, such as the 6-Cys repeat that strongly resembles vertebrate epidermal growth factor and also the *Drosophila* notch protein. The 8-Cys repeats resembled the binding protein for transforming growth factor-b $_1$, which is also homologous to epidermal growth factor. The epidermal growth factor-like repeats implied that the skeletal overgrowth of Marfan syndrome might somehow be mediated by abnormal stimulation of skeletal growth.

The second group (Lee *et al*. 1991b) used fibrillin-specific monoclonal antibodies to purify fibrillin peptide fragments from which sense and antisense oligonucleotide primers were designed. These were then used to select fibrillin cDNA sequences from amplified fibroblast or lymphocyte mRNA/cDNA libraries. When fibroblast cDNA was subtracted with lymphocyte cDNA a 1.6-kb fibroblast cDNA was identified. This was also fibrillin 15 and hybridized as expected to a 10-kb mRNA. Subsequent screening of genomic libraries produced two fibrillin genomic sequences located to chromosome 5q23–31 and 15q15–21. Both contained cysteine-rich, epidermal growth factor-like repeats and were used for linkage to Marfan syndrome and (later) congenital contractural arachnodactyly. Lee *et al*. (1991) also isolated a four-allele $(TA_4)_n$ repeat and a Taq RFLP on chromosome 15 and a seven-allele $(GT)_n$ on chromosome 5. These markers were extremely useful for testing linkage of fibrillin 15 to Marfan syndrome and fibrillin 5 to congenital contractual arachnodactyly.

Other important elements of fibrillin gene structure and organization

The complete genomic cDNA organization of the fibrillin 11 gene has a 11-kb coding region of 65 exons in a 110-kb genomic sequence (Pereira *et al*. 1993). There are numerous independent cassettes with distinctive C- and N-termini, 43 epidermal growth factor-like, calcium-binding motifs, four epidermal growth factor-like, non-calcium-binding regions, seven transforming growth factor-b $_1$-like binding regions and two hybrid motifs. The epidermal growth-factor regions form two subsets, mostly 8-Cys and others with 6-Cys repeats. Calcium binding depends heavily upon highly conserved cysteines at residues 6, 7, 8, 17, 19, and 20, whilst residues 8 to 19 form the b-hydroxylase consensus sequence. Similarly, residue 10 is a b-hydroxylated Asp whilst 15 is an aromatic Tyr (Dietz 1992; Dietz 1993a, Dietz 1993b). Calcium-binding cysteines of epidermal growth factor bridge between residues 6/8, 7/17, and 19/20, and as is evident below, mutations at any of these highly conserved sites are pathological. A quadruple-loop secondary structure results, forming a double 'S' shape upon which calcium binding depends. Various other residues are also strongly conserved, particularly the b-hydroxylase consensus sequence between residues 8 and 19. Yet other cassettes form transforming growth factor-b-like binding-protein domains (Lee *et al*. 1991; Corson *et al*. 1993; Pereira *et al*. 1993). These are interspersed with calcium-binding regions. Combinations of such domains somehow regulate interchain disulphide bonds, whilst the calcium-binding cysteines form intrachain disulphide bonds (Dietz *et al*. 1995). At intervals between the N- to the C-termini, transforming growth factor-b-like cassettes are occasionally coded by two exons but more usually, like the epidermal growth factor regions, by a single exon. The N- and C-termini are each coded by two exons. In all there are 58 cassettes coded by 61 exons whilst transforming growth factor-b domains are coded by cassettes 7, 14, 10, 33, 36, 44, and 50, with calcium regions of epidermal growth factor in between (Corson *et al*. 1993; Pereira *et al*. 1993, Nijbroek *et al*. 1995).

Fibrillin 5 and 15 (Fib1 and Fib2) gene markers in human diseases

Early mutational linkage studies (Dietz *et al*. 1991; Lee *et al*. 1991b; Maslen *et al*. 1991) used two separate but complementary approaches. Firstly, Ramirez's group (Lee *et al*. 1991) tested linkage of the fibrillin 15 gene to several plausible candidate diseases such as Marfan syndrome, familial aortic dissection, mitral-valve prolapse, annuloaortic ectasia, the marfanoid hypermobility syndrome, and congenital contractual arachnodactyly. Only Marfan syndrome linked to fibrillin 15. Next, the Saiki group (Dietz *et al*. 1991, Dietz *et al*. 1992; Maslen *et al*. 1991) used a fibrillin-specific PCR to search for mutations in Marfan syndrome. They also tested families with disorders such as mitral-valve prolapse, medial degeneration of the aorta, and idiopathic scoliosis, all of which were normal. Later fibrillin 15 was renamed fibrillin 1 and fibrillin 5 was renamed fibrillin 2.

Fibrillin-specific mutational analysis

Initial work concentrated on the mutational analysis in Marfan syndrome by PCR amplification of fibrillin 15 cDNA. This project eventually showed that Marfan syndrome, infantile Marfan syndrome, isolated ectopia lentis, and Marfan syndrome families with aortic dilatation were all allelic (and caused by fibrillin 15 mutations). Methods included combinations of cDNA, genomic mutational analysis, protein mapping of ^{35}S-labelled fibrillin proteins, and *in vitro* studies of fibrillin protein assembly using Western-blotted fibrillin and rotary-shadowing electron microscopy to study protein size, migration, and assembly.

Mutational analysis (Table 2)

Early studies

Dietz's originally amplified cDNA sequences corresponded to amino acids 388–408 and 831–849 synthesized from fibroblast mRNA. Later this protocol was extended to include all fragments of the partial 6.9-kb cDNA sequences. Differences in conformation caused by point mutations, exonic deletions or other abnormalities were then sought by single-stranded confirmational analysis followed by *M13* cloning and sequencing. Later (see below) the protocol was streamlined to include the last 3 to 4 kb of the fibrillin-15 cDNA sequence and double-stranded heteroduplex analysis in silver-stained MDE gels followed by direct rather than indirect sequencing.

Table 2 List of Marfan mutations as of early 1996 (after Nijbroek *et al*. 1995)

Mutations were originally found in two separate Marfan syndrome patients. The first had the severe infantile syndrome, severe joint laxity, scoliosis, high-grade myopia, mitral incompetence, and premature dissection of the ascending and descending aorta. The second was only slightly less severely affected. Both had a G–C transversions of nucleotide 716 in codon 239; Pro was substituted for Arg, therefore replacing a basic amino acid with a non-polar one, and impairing protein secondary structure. No similar mutation was detectable in 19 patients with sporadic Marfan syndrome or in 24 randomly affected family members chosen from various Marfan families, nor were any gross gene rearrangements found in 38 unrelated Marfan patients. By the combination of PCR single-stranded or heteroduplex DNA analysis the early mutational analysis was both cost-effective and relatively fast.

In 1991 much still remained to be done. Thus the missing 3.1 kb of the 10-kb mRNA had not been cloned and faster, better methods of mutational screening still needed to be developed. Obvious questions at the time included how often mutations were clustered and private, and whether mutational patterns such as point mutations, stop codons, and exonic and genomic deletions were usual. Rapid and reliable mutational analysis was urgently needed for accurate genetic counselling and prenatal diagnosis. Four years later (1995), by virtue of a large collaborative consortium, mainly of workers in the United States, Finland, and Western Europe, most of these questions now have satisfactory answers.

More recent analysis

The present situation is well summarized in papers by Sykes (1993), Hewett *et al*. (1994), Kainulainen *et al*. (1994), and Nijbroek *et al*. (1995). The latter lists 53 Marfan syndrome mutations, 15 of which are new. They were detected by genomic amplification combined with heteroduplex analysis. PCR primers were chosen for

immediate direct sequencing: 78 per cent were detected by heteroduplex analysis, with 13 novel polymorphisms, four of which were in coding regions. Eight were missense mutations, seven of which were located in the calcium-binding, epidermal growth-factor domain. Another was a missense change in a transforming growth factor-b binding protein-like region. Other mutations included one each in residues 2 and 5, another in the conserved Arg at residue 15, whilst a nearby conformational change in residue 16 substituted Arg for Pro. Contrastingly, in severe Marfan syndrome, there were substitutions in each of the disulphide bridges linking residues 7 to 17 and 19 to 20 with no less than five changes at Cys8 (in residues 129, 926, 1663, 2321, and 2307, each at homologous sites in different cassettes). Other mechanisms included small nucleotide deletions (1, 2, and 3 bp) and amino acid insertions. There were five premature stop-codon mutants, four caused by frame shifts and one by a nonsense mutation. Upstream premature terminations dependent upon 3' exons caused very shortened proteins unable to degrade. Other defects included null alleles with little or no expression of the mutant transcript and double heterozygosity of separate errors in severe infantile Marfan syndrome (Karttman *et al*. 1994). Here W217G and G2627 were derived from the paternal and maternal sides, respectively. The affected child produced no fibrillin microfibrils in culture, whilst fibrillin antibodies failed to stain skin sections.

Mutations that cause infantile Marfan syndrome mainly cluster in exons 24, 26, 27, and 31 between the third transforming growth factor-b $_1$binding-protein domain and the third next epidermal growth factor calcium-binding region (Corson *et al*. 1993; Kainulainen *et al*. 1994; Nijbroek *et al*. 1995). These adjacent domains structurally interact to form a single b-sheet. If intact, the N- and C-termini can somehow influence the behaviour of fibrillin 1 monomers. Nijbroek *et al*. (1995) suggest that the most structurally disruptive mutations are those that least perturb microfibrillar assembly; that is, no microfibrils are better than damaged ones. For example, severe disruption of disulphide bonds in transforming growth factor-b binding-protein regions 3 and 4 are associated with mild Marfan syndrome. Contrastingly, severe disease accompanies non-cysteine substitutions, in-frame deletions or insertions. There have also been two mutations causing truncated polypeptides (Milewicz and Durie 1994; Karttman *et al*. 1994). One is a 366-bp deletion with a shortened but not well-secreted protein (Kainulainen *et al*. 1992). Nijbroek *et al*. (1995) provided a good mutational map, which is shown in Table 2.

Other analytical strategies

Just as in the analysis of collagen I, II, and III mutants, molecular sequencing is often combined with protein studies. In the case of fibrillin, ^{35}S-labelled fibroblast cultures will effectively label profibrillin and fibrillin proteins, which are then visualized by polyacrylamide gel electrophoresis. Generally, three classes of Marfan syndrome mutants are detectable, including a general depression of fibrillin labelling (caused by degraded or null alleles) and shortened proteins such as the 366-bp deletion described by Kainulainen *et al*. (1992). Sometimes there are no visible changes, in some of which abnormal fibrillin 1 protein is masked by fibrillin 2, with which it comigrates. The protocols for pulse-chase analysis profiles have been clearly described by McGookey-Milewicz *et al*. (1992), Raghunath *et al*. (1993), Aoyama *et al*. (1994), and Nijbroek *et al*. (1995). Aoyama *et al*. (1994*a, b*) distinguish five distinct protein phenotypes according to the ratios of secreted (medium) and retained (cell layer) products. The least severe, group V, shows no difference from normal controls. Groups I to IV have variably reduced medium:cell-layer fibrillin ratios, which were lowered or almost equal in groups I and II, respectively. The secretion in the medium in groups III and IV was normal but here cell-layer fibrillins differed, being slightly reduced in group III but subsequently altered in group IV.

Lastly, fibrillin assembly can be observed by rotary-shadowing electron microscopy. This directly visualizes microfibril assembly in cultured dermal or other fibroblasts (Kielty *et al*. 1994, 1995). Interesting perturbations of microfibrillar periodicity and fragmentation have been demonstrated in several circumstances. The normal, 10 to 12 nm, beaded structures formed by trimeric filamentous association is clearly visible. Irregular periodicity, microfibrillar fraying, and missing transverse filaments or clustered beads have all been observed. Electron microscopy of fixed material can also be used for testing structure and function. Interpretation is rather subjective, and disordered patterns can sometimes be produced from normal cell lines and might vary with culture conditions, temperature, and length of incubation (Eladha *et al*. 1995).

Fibrillin 2 mutations

Similar approaches are also valid in clinical molecular analysis of fibrillin-2 mutants. This gene is located on chromosome 5 (Tsipouras *et al*. 1992). Congenital contractual arachnodactyly differs from Marfan syndrome in the presence of joint contractures and crumpled ears (Beale syndrome). Dolichostenomelia and scoliosis also are part of the syndrome whereas aortic rupture/dilatation and lens dislocations are absent.

For fibrillin 2 a combination of protein analysis from cultured fibroblasts and gene amplification works as well as for fibrillin 1. The tissue distribution of fibrillin 2 differs from that of fibrillin 1, being predominantly in the perichondrium, joint capsules, and vertebral columns whilst relatively sparse in the cardiovascular system and missing from the lens suspensory ligaments. Very probably the interactions of the two molecules differ subtly in producing varied clinical phenotypes of the two disorders.

Other genes and Marfan syndrome

Marfan syndrome is homogenous, caused only by mutations of the fibrillin-1 gene on chromosome 15, with two exceptions. The first is an intriguing family with an affected female who synthesized a double collagen a2(I) chain (Phillips *et al*. 1991). Arg(GGC) is changed to Gln(GAC) at position 618 (Phillips *et al*. 1991). This second-position mutation changes a charged Arg residue that is highly conserved in chick, bovine, and human material and also other similar human genes such as *COL1A1*, *–1A2*, *–3A1*, and *–5A2*. Charged amino acids are very highly conserved in collagen sequences and uncharged residues such as Gln may theoretically alter the intermolecular alignment of the a2(I) chain-containing fibrils. The substitution also accounts for the altered electrophoretic mobility of the mutant collagen a2 chain for several reasons. It is still unclear what relevance the mutation has to the Marfan syndrome. Although both the proband and her father carried the mutant *COL1A2* gene, only the proband had convincing clinical changes (without lens dislocation), but the father may have been a mixed somatic–gonadal mosaic. More significantly, no other Marfan syndrome has ever been linked to the *COL1A2* locus.

An unusual French family described by Collod *et al*. (1995) is also interesting. Here, linkage to chromosome 15 was excluded and the locus mapped to chromosome 3p24.2–p25. This is quite near the *COL7A1* locus at 3p21, which was later excluded as a candidate gene. The family contained over 170 affected individuals in at least four generations. Clinical features included small stature, widened arm span, arachnodactyly, scoliosis, pectus excavatum, valve prolapse, and aortic dilatation, dissection and rupture, together with myopia or ectopia lentis. The investigators were convinced that these fitted Marfan syndrome whilst other prominent researchers disagreed. Doubtless neither fibrillin 1 nor fibrillin 2 are the only causes of Marfan syndrome-like symptoms and signs. 'Splitters' have postulated intrafamilial genetic heterogeneity as the likely reason for the vascular pathology of the Marfan syndrome (Dietz *et al*. 1995). The original investigators, in contrast, are convinced that this is Marfan syndrome and that the Berlin nosology needs widening and modification (Collod *et al*. 1995).

Future directions and development

The information outlined above quite clearly illustrates the importance of 'the new genetics'.

Chondrodysplasias

Introduction

There are over 150 distinctive chondrodysplasias (Spranger 1992). They are inherited in autosomal-dominant, autosomal-recessive or X-linked recessive patterns from numerous different genes. The variability of clinical phenotype is exceptional, with almost bewildering numbers of variants. Understanding the mechanisms of mutations should illuminate fundamental questions regarding mechanisms of embryonic development, pattern formation (segmentation), determinants of bone, cartilage and limb shape or stability, and other important molecular interactions of skeletal components one with another. Until 1993, this field had lagged behind rather rapid advances in the unravelling of the molecular mechanisms of other matrix structures such as skin, ligaments, arteries, and bones. In the late 1980s this had begun to change with the initial identification of *COL2A1* mutations in certain chondrodysplasias, such as epiphyseal dysplasia or premature osteorarthrosis with associated vitreous or skeletal deformities. Examples of this group are achondrogenesis, Kniest syndrome, spondyloepiphyseal dysplasias, and the Stickler syndrome. There has been a recent information explosion regarding the mutational mechanisms of widely disparate disorders such as the Crouzon/Apert syndrome, Pfeiffer craniofacial abnormalities, thanatophoric dwarfism, achondroplasia and hypochondroplasia, as well as diastrophic dysplasia—all of which share similar mechanisms. Such abnormalities are caused not only by unexpected structural elements such as *COMP5*, but also by other substances that regulate, orchestrate, and determine skeletal patterning, shape, integrity, and development such as *HOX*, 'spiny hedgehog', fibroblast growth factor and its receptors (types 1–4), bone morphogenetic protein, GMP regulator proteins, growth hormone, parathyroid hormone-like protein, and modifying substances such as sulphate transporters and sulphatases of various kinds. In contrast, mutations of structural proteins such as those in collagens II, IXa1, IXa2, IXa3, X, and XIa1 or IXa2 are relatively rare. Hopefully, cataloguing of mutations in various types of collagen and matrix components will eventually clarify the biology of embryonic, limb, skeletal, and structural

development in very unexpected and enlightening ways.

Since the first version of this chapter was written in 1993, there have been unusually rapid and unforeseen advances in the molecular pathology and biology of these disorders. Skeletal maldevelopment is the result of many of them, although only relatively few are caused by errors in scaffolding components. Errors in the orchestration of connective tissue expression are very important indeed. Horton (1995) has recently tabulated in summary form most of the recent examples, which we have subsequently outlined below.

Hierarchical development of bone and skeleton

Skeletal development follows a orchestrated sequence and pattern that determines overall growth and morphology, and subtle variations in shape or location of particular components. These complicated symmetrical structures need regulating both in time and space. Final bone anatomy results from complex interactions of many different regulatory patterning or structural genes. Until very recently only some of the latter had been amenable to analysis.

Morphogenesis (the integrated determination of shape and pattern) results from a dynamic balance between cartilage or bone deposition and breakdown, so producing a constantly shifting balance between each of the individual constituents. There is also the continuous ebbing and flowing of bone and cartilage deposition and degradation, more particularly in those bones that grow by the balance of continual accretion and removal. Whilst individual bones enlarge they still maintain the shape of their original embryonic cartilaginous templates. There is a hierarchical organization of those factors that control this process. Such factors include the interaction genes and proteins that determine bone shape with factors that regulate bone growth, remodelling, modelling as well as certain crucial structural components including collagens II, IX, X, XI, BMP, and *COMP5* (Kingsley 1994; Erlebacher *et al*. 1995).

Skeletal patterning

Vertebrate skeletons have three major components, the head and neck (craniofacial region), the spine/pelvis (axial), and the limbs (appendicular skeleton). The craniofacial component derives from the embryonic branchial arches, which themselves arise from the neural crest. The axial skeleton (spine, ribs, and pelvis) originates from the sclerotomes whilst the lateral mesoderm dictates which elements eventually become the limbs. Except for the skull and pelvis, which are largely of membranous origin (and which develop directly from mesenchymal condensations), most of the remainder originates as cartilaginous templates that gradually transmute into bone. This sequence includes ordered mesenchymal condensation, chondrocytic differentiation, hypertrophic calcification, and apoptosis (Kraus *et al*. 1993). Vascularization is the trigger for the replacement of the cartilaginous protoskeleton with more mature, osteoblastic bony elements. Eventually, after epiphyseal closure, only traces of cartilage remain distally upon the articular surfaces, although hypertrophic metaphyseal cartilage precedes trabecular bone formation. Skeletal patterning is determined by certain segmentation genes, which turn out to be transcriptional regulators (the so-called homeobox genes). The homeobox genes are highly conserved, pattern-determining controllers acting very early in the embryos of segmented animals (including species as diverse as vertebrates and insects). The functions of particular subsets specifically vary with segmental positions (and are organized and expressed temporally in a head-to-tail distribution with subsets that specifically regulate the face, mid and peripheral elements). As might be expected, mutations of such neuroectodermal patterning genes seriously disrupt the correct sequences and patterns of, for example, branchial arch development and axial morphology. Typical vertebrate examples are mutations in the *HOXA2, HOXA13d*, and the *PAX* superfamily, all of which (see below) cause variants of craniofacial or skeletal patterning and development (Krumlauf 1994).

HOXA2 and *HOXA13d* mutants

HOXA2 acts before the appearance of the second branchial arch. Mutations in mice delete crucial elements of the second arch and instead replaces them with first-arch derivatives that are consequently duplicated (Gendron-McGuire *et al*. 1993). Examples include duplication of the middle-ear ossicles, duplication of the optic capsule, and bilateral clefting of the secondary palate. Such mutations disturb the execution of second-arch development even though the necessary precursors for their later appearance (that is, the neural-crest rhombomeres) are already safely in place (Rijli *et al*. 1993).

HOXA13d mutants

Similar inactivation of the *HOXA13d* gene transmutes the fourth into a third sacral vertebra with consequent limb abnormalities (Dolle *et al*. 1993); these include shrivelling of digits II and V in both the front and hind paws, and also the phalanges of digits III and IV, which are compressed whilst all the corresponding terminal phalanges are virtually normal. Sometimes extra rudimentary digits appear posteriorly (on the ulnar margins of the forelimbs). The affected animals are otherwise normal, in keeping with the late (temporal) activation of this rostrally located (sacral) member of the *HOX* family. Mutations impair the condensation of the cartilaginous precursors of digital bones so that final ossification is abnormal. The *HOX* family can be regarded as a series of morphogenetic clocks by means of which vertebrate development is controlled by the consecutive switching of the four members of the *HOX* gene cluster (*HOXA–D*). These are spatially disturbed in the same nose-to-tail order in which they both appear and function anatomically. Thus, the switching on of any 3' gene immediately precedes its next 5'-most adjacent associate. Mutations of these genes are highly homologous throughout many disparate animal species. Thus, errors of similar determining elements in flies produce spectacular disorganizations such as bithorax and antennopaedia in which segments are duplicated or legs appear where wings should normally be.

PAX mutations

Mutations at the *PAX* loci are homologous to one another (Tassabenji *et al*. 1993). This gene family participates not only in skeletal (embryonic) development but also in carcinogenesis. Mutations produce either gain or loss of function patterns. Thus *PAX3* (loss of function) causes disorders such as Waardenburg syndrome, craniofacial deafness with hand abnormalities or rhabodyomyosarcoma depending on whether this gene is amplified or duplicated. Similarly, *PAX6* gene loss-of-function mutations cause faulty eye development with aniridia and several *PAX* mutations perturb development of the central nervous system. *PAX3* mutations frequently disturb skeletal development, with characteristic facial malformations (broadened nasal bridge and misshapen orbits) combined with skeletal abnormalities such as contractures and hypoplasia of the upper limb bone. The craniofacial/deafness/hand syndrome MIM122880 is also a *PAX* mutant.

Other human homeobox mutants

Autosomal-dominant craniosynostosis (Boston type) is caused by mutations of the *MSX2* gene (Jabs *et al*. 1993) located on chromosome 5 (Warman *et al*. 1993). This is distinct from other craniosynostotic loci such as for Crouzon, Apert, and Jackson–Weiss syndromes, which are caused by mutations of the *FGFR1* or *–2* genes. As might be expected, *MSX2* is expressed precisely in the face, mandible and maxilla, especially in the cranial sutures. The responsible mutation changed histidine to proline at a site that is highly conserved in species as disparate as human, arthropods, sea urchins, and vertebrates. Only Boston-type craniosynostosis is caused by errors in these genes in contrast to others, such as Apert syndrome, in which the *FGFR* genes participate (see above).

Mechanism of action of the homeobox and segmentation genes

Skeletal symmetry largely depends upon certain crucial inductive interactions of one matrix gene with another. Thus limb-patterning signals emanate from the so-called zone of polarizing activity (*ZPA*) genes, which are specifically localized to the mesoderm of the posterior limbs. These entities are controlled by the apical ectodermal ridge of the developing limb. Transplantation of the ridge anteriorly induces a second, duplicated limb bud that is a mirror image of normal. Retinoic acid (vitamin A) has a similar effect. Thus a soluble inducing element spatially mediates *ZPA* activity by specifically inducing *FGFR4*. In turn this unleashes a cascade of regulatory substances that further influence patterns and development. Such inductive patterns execute the spatially and temporally expressed commands of the homeobox genes mentioned above. Thus the orchestration factor released by *ZPA* activity is mediated by the 'sonic' gene. In *HOX13D* mutants, 'sonic hedgehog' then switches on bone morphogenetic protein 2, which is a transforming growth factor-b protein that in turn itself tells cartilage when to appear and develop. 'Sonic hedgehog' factor (Riddle *et al*. 1993) is presumed to act either as a diffusable morphogen in a concentration gradient or by the initiation of cell–cell interactions by downstream elements in the hierarchy preceding the eventual appearance of more permanent scaffolding compounds. *ZPA* liberates a morphogen that also duplicates the action of retinoic acid, which in turn produces a specific signalling molecule. The signalling molecule resembles the segmental polarity regulators that are essential and common in developing embryos. These genes include those such as 'armadillo', 'engrailed', 'gooseberry' and 'wingless', all of which are crucial for the correct sequential development of *Drosophila* embryos. 'Hedgehog' is a classical concentration-dependent morphogen very similar to its equivalent *Drosophila* homologue. It is highly conserved in all animal species and is 78 per cent identical in chicks and flies, implying a very fundamental function in embryogenesis. Its name derives from a character in a computer game, whilst its expression is located to the posterior limb bud precisely congruent with the distribution of *ZPA*.

Local regulation of bone shape

The morphometry of prototype cartilage precursors themselves originates from more blurred mesenchymal condensations. Each step is controlled separately by one or more members of the gene hierarchy families mentioned above. Thus, mouse mutants, such as brachypodia and short-ear, can produce variously misshapen components or malformed skeletal elements. For example, short ear (sc) phenotype radically alters both sternal shape and the morphology of the external ear whilst 'brachypodia' (bp) specifically shortens appendicular long bones. Equally importantly, the equivalent embryonic precartilaginous structures are also similarly distorted. Furthermore, subcutaneous and bp are caused by specific inactivation of either bone morphogenetic protein 5 or GGF5. These are both members of the transforming growth factor-b superfamily that are expressed later than *HOX* patterning genes but before (earlier) than connective-tissue structural genes, which are later players in the hierarchy of development of gene expression. In skeletal development, such substances somehow orchestrate and timetable the development and three-dimensional patterning of precartilaginous mesenchyme. Similarly, the complexities of endochondral ossification necessitate the coordinated and orderly expression of epiphyseal, metaphyseal, and diaphyseal components essential for the synchronization of the cartilaginous epiphyseal plate/metaphysis/diaphysis, which enlarges progressively from infancy to adulthood in a similar manner to the circumferential seasonal accretion of tree rings.

Dysplasias, such as achondroplasia, hypochondroplasia, pseudoachondroplasia, and diastrophic dwarfism, that distort the appendicular skeleton are produced by disturbances of different elements in this complex process.

The functions of fibroblast growth factors (FGF) and their receptors

The *FGF* superfamily specifies proteins that bind to specific transmembrane FGF receptors (FGF1–4) (Grovl and Yaifin 1992). These regulate the proliferation or differentiation of certain mesenchymal and neuroectodermal cells, and are partially tissue-specific. Mutations of the *FGFR* genes produce very specific diseases such as achondroplasia (Lajeune *et al*. 1995), or Pfeiffer or Crouzon syndromes. They are transmembrane glycoproteins that bind fibroblast growth factors. The various homologues have common features such as intracellular tyrosine kinase and extracellular immunoglobulin domains and transmembrane bridges. The immunoglobulin domain that is variably spliced can generate various mutations or the FGFR3 receptors with differing affinity for FGF. The *FGFR3* gene that is expressed in cartilage rather than bone cause achondroplasia. Similarly, Crouzon and Pfeiffer syndromes are caused by mutations of *FGFR1* and *FGFR2* (Jabs *et al*. 1995). The various genes share both overlapping and common patterns of embryonic expression, and strongly influence both embryonic and fetal development. Thus, expression of *FGFR3* is detectable in the neurotubular epithelium, adult neuroglia, and cochlear ducts. On the other hand, *FGFR1* and *–2* occur in precartilaginous bone but also in resting cartilage during endochondral ossification (Peters *et al*. 1993).

Abnormalities of the *FGFR* genes

Recent advances in molecular pathology of certain apparently unrelated chondrodysplasias or craniofacial anomalies have all implicated the *FGFR1*, *–2* or *–3* genes in a manner that is both domain and position related. Thus, achondroplasia is caused by an *FGFR1* mutation, as are the Crouzon, Apert, Pfeiffer, and Jackson–Weiss syndromes. Very recently, both thanatophoric dwarfism and hypochondroplasia have also been shown to have *FGFR* mutants.

Achondroplasia (Fig. 24a–c)

The first mutation in achondroplasia was described by Shiang *et al*. (1994) in 15 out of 16 random achondroplastics, and by Rousseau *et al*. (1994) in 23 individuals. Remarkably the error lay in the transmembrane coding region, nearly always the same base of the same codon. The same single faulty codon was later detected in 153 out of 154 achondroplastics (Bellus *et al*. 1995). The mutation in codon 380 in the immunoglobulin repeat region is a G–A transversion, which in 150 individuals was at position 138. The remaining three had a G–C substitution at the same locus. The 154th case is so far unidentified. No G–T transversions have occurred and may either be genetic lethals or have a different clinical phenotype. So far, more than 99 per cent of achondroplastics have identical, single base-pair mutations. This is quite remarkable and has no precise parallels (so far) in any other genetic disease, but implies extreme chemical instability of this locus.

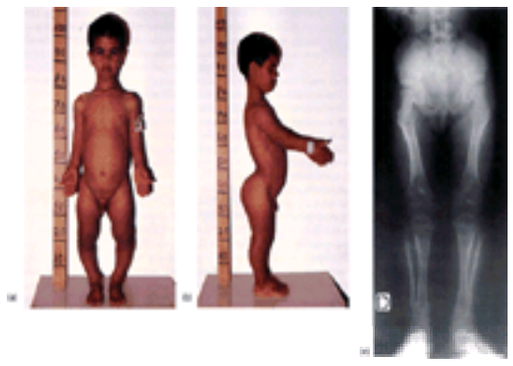


Fig. 24 Phenotypical (a) and (b), and radiographic features (c) of a typical patient with achondroplasia. The resemblance to Fig. 25 is obvious and both defects are caused by abnormalities of the same gene

Hypochondroplasia (MIM146000)

Hypochondroplasia, which is allelic to achondroplasia, is caused by mutations in the tyrosine kinase region of the *FGFR3* gene. Thanatophoric dysplasia is also caused by mutations at a different *FGFR3* locus. Although thanatophoric dysplasia (MIM187600) superficially resembles achondroplasia, there are variously distinctive clinical and morphological differences. Overall it is very much more disabling and severe, and also has associated severe defects in the central nervous system.

The mutation rate in achondroplasia is between 15 to 760 times higher than that of any other known CpG mutation. Other known examples such as globin, cystic fibrosis, haemophilia, and collagen mutations are much lower.

Other *FGFR3* mutations including thanatophoric dwarfism

Thanatophoric dysplasia has several subtly different phenotypes, all of which are severely deforming and genetically lethal. Recently the mutations that cause this disorder have been located to sequences between the Ig2 and Ig3, or between Ig3 and the transmembrane coding region of *FGFR* where the mutations also lie. There is also a second thanatophoric (type II) locus in the distal half of the tyrosine kinase extracellular coding domain. In contrast, many hypochondroplasia mutations lie within in the tyrosine kinase A domain, such as Arg540–Lys, which is a common error at this site. Similarly, Pfeiffer, Apert, Crouzon and Jackson–Weiss *FGFR3* mutants are positional homologues of equivalent *FGFR1* and *FGFR2* mutations.

Thanatophoric dysplasia (Fig. 25(a–e))

This is the most common, lethal, neonatal skeletal dysplasia with a frequency of 1/20 000 live births. Typical features include micromelia, macrocephaly with a frontal bossing, flattened vertebral bodies with disorganized cartilage growth plates, and markedly shortened ribs with abdominal distension. Affected babies usually die from respiratory failure caused by restrictive lung disease from the severe thoracic dysplasia. There are two radiographic subsets, thanatophoric dysplasia I with curved shortened femurs (telephone-receiver pattern) and clover-leaf skull, and thanatophoric dysplasia II with straight legs and severe clover-leafing of the skin. Achondroplastic homozygotes strongly resemble thanatophoric dysplasia I and II heterozygotes. These two disorders have long been considered as allelic. This conclusion is correct as both achondroplasia and thanatophoric dysplasia are caused by *FGFR* mutations (see above and below).

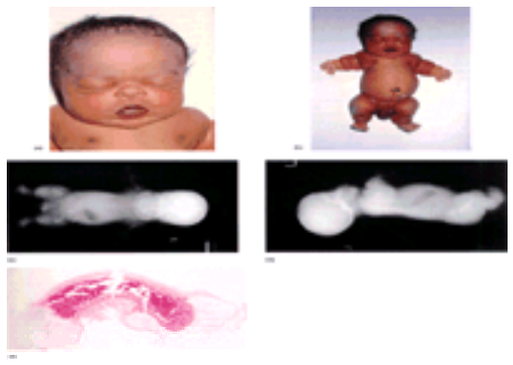


Fig. 25 (a, b) Facial and full-body phenotype of affected baby with thanatophoric dysplasia. (c, d) Babygram radiographs of the same child showing severe thoracic dysplasia, telephone-receiver femurs, and a dysplastic skeleton. (e) Histological appearance of section of femur showing a disturbed diaphysis and disorganized cartilage × 1.5.

Recently the two types of thanatophoric dysplasia have been distinguished by their molecular pathology. Firstly, Tavormina *et al*. (1994) suggested that the *FGFR3* gene functions either in growth-plate development or in calvarial bone fusion of the developing fetus. Those *FGFR* mutations such as Arg248–Cys and Lys650–Glu replace crucial residues and severely disrupt the secondary or tertiary structure and function of *FGRF3* protein as a result. Perhaps such errors severely impair intracellular signalling either by distorting protein structure or by preventing the dimerization of adjacent protein chains by a positive/negative effect. Such mechanisms are very much more disruptive than loss of function; thus simple deletion of one allele (lower dosage) has no clinical effects. The FGF signal-transduction pathway is complex, comprising at least nine ligands and four receptors, of which three are alternatively spliced.

FGFR1 and *FGFR2* mutations

The discovery of *FGFR3* mutations in achondroplasia is paralleled by the discovery of similar changes in the *FGFR2* gene in Crouzon syndrome (Reardon *et al*. 1994). The identification of *FGFR1* mutations in Crouzon, Apert, Pfeiffer, and Jackson–Weiss disorders rapidly followed. A little later, similar abnormalities were observed in thanatophoric dysplasia and hypochondroplasia. Such varied clinical phenotypes are dictated entirely by mutational position and type.

Craniosynostosis

The craniosynostoses include the Apert, Crouzon, Pfeiffer, and Jackson–Weiss syndromes. Common features include premature fusion of skull sutures producing skull distortion, widely spaced, bulbous eyes with hyperteleorism, and mid-facial folds (Muenke *et al*. 1994). The Pfeiffer, Crouzon, and Jackson–Weiss syndromes overlap, but the Pfeiffer syndrome is distinguished by severe hand and foot anomalies such as broadened thumbs, syndactyly, and shortened fingers and toes. Broadened big toes with tarsal and metatarsal syndactyly are also common in the Jackson–Weiss syndrome. Some Pfeiffer syndrome families are linked to chromosome 8 with *FGFR1* but others are not. In contrast both Crouzon and Jackson–Weiss syndromes were initially mapped to the distal long arm of chromosome 10 where *FGFR2* also lies. Subsequently, various *FGFR* abnormalities have been associated with these clinical phenotypes and mutations in three of the four known *FGFR* genes have been detected in the various Crouzon, Jackson–Weiss, and Pfeiffer families.

Examples

A highly specific Pro252–Arg substitution caused by an AC conversion in exon 5 was detected in the extracellular domain of five unrelated families with Pfeiffer syndrome. The Pfeiffer syndrome phenotype included premature skull fusion, widely spaced eyes, and mid-facial hypoplasia. Similarly, the Jackson–Weiss syndrome (MIM123150) combines skull and limb abnormalities with broadened toes and tarsal and metatarsal coalescence as mentioned earlier. Just as in Crouzon syndrome there are mutations in the IgIIIc region of *FGFR2*, whereas the original and subsequent Crouzon mutations included IgIII Tyr328–Cys and Ser347–Cys. The tissue expression and distribution of *FGF1, FGFR2*, and *FGFR3* proteins partially explains the clinical phenotypes of Crouzon and Jackson–Weiss syndromes. Thus, whilst *FGFR1* is expressed throughout active limb bud, *FGFR2* expression at the equivalent stage of embryogenesis is confined to limb-bud ectoderm. *FGFR1* mutants might therefore be expected to have more severe limb deformities. Those *FGFR2* mutations that cause Crouzon/Jackson–Weiss syndrome lie in exon 7, which codes for IgIII. These particular sequences are preferentially expressed in skull *FGFR* proteins and will therefore interfere with the timing of skull suture fusion. Contrastingly, the Pro252–Arg *FGFR1* mutation in Pfeiffer syndrome lies in an exon shared by several isoforms of *FGFR1* protein. The clinical phenotype is therefore less specific. Mutations are located to sequences normally very highly conserved in both man and chickens; so far the function of this domain, whilst important, is completely unknown.

Apert syndrome can also be allelic to Crouzon syndrome (Wilkie *et al*. 1995). Thus identical mutations of the *FGFR1* gene can cause either Pfeiffer or Crouzon phenotypes (Rutland *et al*. 1995). An example is Cys342–Arg previously described in Crouzon syndrome. A Tyr–Cys343 had also been described in Crouzon syndrome. Several Apert mutations mutate either Ser253–Trp or Pro253–Arg (Wilkie *et al*. 1995).

The mutational correlations can therefore be summarized as follows (Mulverhill 1995). *FGFR3* mutations were first demonstrated in achondroplasia (Shiang *et al*. 1994). Rousseau *et al*. (1994) then reported an achondroplasia mutation at nucleotide 1138 changing Gly380 to Arg. Very soon afterwards *FGFR2* mutants were detected in Crouzon syndrome followed by Jackson–Weiss mutants Arg344–Gly in exon 9 of *FGFR2*, which were two amino acids outside the loop-defining cysteine. Next, two *FGFR2* mutations in adjacent residues disrupting the peptide linkage between Ig loops II and III were detected in Pfeiffer syndrome. One form that exhibited mild syndactyly changed Ser252 to Trp whilst the other with more severe deformities and accompanied the Pro253–Arg mutation. Pfeiffer syndrome can be caused by mutations of either *FGFR1* and *FGFR2* genes (e.g., the *FGFR1* Pro252–Arg lies between Ig loops II and III with the *FGFR2* equivalent located to the second half of the IgIII loop). Similarly, Crouzon and Pfeiffer syndromes are allelic and may be caused by identical errors in *FGFR2* such as Cys342–Thy or Cys342–Arg. As Mulverhill (1995) comments, there are only limited numbers of clinical repertoires in response to widely variable molecular changes. Lastly, as mentioned earlier, thanatophoric dysplasia (both types I and II), achondroplasia, and hypochondroplasia are allelic but separately located to protein domains such as the tyrosine kinase and extracellular portions of the *FGFR1* protein. Presumably, epiphyseal growth is both dictated and regulated by particular fibroblast growth factors that in their turn define complex local paracrine/autocrine interactions between the varied mesenchymal components of bone and cartilage, which in turn themselves bind to particular fibroblast growth-factor receptors. These molecules are also influenced by other substances that consequently orchestrate overall skeletal growth. An example of the latter is parathyroid hormone-related protein. This is expressed throughout the epiphyseal plate even though its receptors are confined to the proliferative zone of hypertrophic cartilage where collagen X is also specifically expressed. Knock-out parathyroid hormone-related protein mice have prematurely ossified skeletons. Errors in other important factors such as the sulphatase genes (Hastbacka *et al*. 1995) result in diastrophic dwarfism, whilst faulty sulphate transport causes achondrogenesis (Superti Furga 1994). No doubt many other elements in the jigsaw await discovery.

Chapter References

Ahmad, N. N. *et al*. (1990). A stop codon in the gene for type II procollagen *COL2A1* causes one variant of arthro-ophthalmopathy (the Stickler syndrome). *American Journal of Human Genetics*, **47**, A206.

Ala-Kokko, L., Baldwin, C. T., Moskowitz, R. W., and Prockop, D. J. (1990). Single base mutation in the type II procollagen gene *COL2A1* as a cause of primary osteoarthritis associated with a mild chondrodysplasia. *Proceedings of the National Academy of Sciences (USA)*, **87**, 6565–8.

Al-Imara, L., *et al*. (1992). Linkage of autosomal dominant dystrophic epidermolysis bullosa in three British families to the marker D352 close to the COL7A1 locus. *Journal of Medical Genetics*, **29**, 381–2.

Anderson, I. J., Goldberg, R. B., Marion, R. W., Upholt, W. B., and Tsipouras, P. (1990). Spondyloepiphyseal dysplasia congenita: genetic linkage to type II collagen (*COL2A1*). *American Journal of Human Genetics*, **46**, 896–901.

Andrikopoulos, K., Liu, X., Keene, D. R., Jaenisch, R., and Ramirez, F. (1995). Targeted mutation in the *COL5A2* gene reveals a regulatory role for type V collagen during matrix assembly. *Nature Genetics*, **9**, 31–6.

Aoyama, T., Francke, U., Dietz, H. C., and Furthmayr, H. (1994 *a*). Quantitative differences, intrasequences and extracellular deposition of fibrillin in cultured fibroblasts distinguish five groups of

Marfan syndrome patients and suggest distinct pathogenetic mechanisms. *Journal of Clinical Investigation*, **94**, 130–7.

Aoyama, T., Tynem, K., Dietz, H. C., Francke, U., and Furthmayr, H. (1994 *b*). Missense mutations inframe intercellular processing of fibrillin and microfibril assembly in Marfan syndrome. *Human Molecular Genetics*, **2**, 2135–40.

Baldwin, C. T., Farrer, L. A., Adair, R., Dharmavaram, R., Jimenez, S., and Anderson, L. (1995). Linkage of early onset osteoarthritis and chondrocalcinosis to human chromosome 8q. *American Journal of Human Genetics*, **56**, 692–7.

Barker, D. F. *et al*. (1990). Identification of mutations in the *COL4A5* collagen gene in Alport Syndrome. *Science*, **48**, 224–7.

Beighton, P., ed. (1995) *Heritable disorders of connective tissue*. 5th edn.

Beighton, P. *et al*. (1988). International nosology of heritable disorders of connective tissue (Berlin 1986). *American Journal of Medical Genetics*, **29**,581–4.

Beighton, P. *et al*. (1992). Molecular nosology of heritable disorders of connective tissue. *American Journal of Human Genetics*, **42**, 431–48.

Beighton, P., Cilliers, H. J., and Ramesar, R. (1994). An autosomal dominant (Benkes) premature degenerative osteoarthropathy of the hip joint unlinked to *COL2A1*. *American Journal of Medical Genetics*, **53**, 348–51.

Bellus, G. A. *et al*. (1995). A recurrent mutation in the tyrosine kinase domain of fibroblast growth factor receptor 3 causes hypochondroplasia. *Nature Genetics*, **10**, 357–9.

Bellus, G. A. *et al*. (1995). Achondroplasia is defined by recurrent G-380R mutations of *FGFR3*. *American Journal of Human Genetics*, **56**, 368–73.

Benya, P. D. and Shaffer, J. D. (1982). Dedifferentiated chondrocytes re-express the differentiated collagen phenotype when cultured in agarose gels. *Cell*, **30**, 215–24.

Blanton, S. H. *et al*. (1990). An exclusion map of Marfan syndrome. *Journal of Medical Genetics*, **27**, 73–7.

Bogaert, R. *et al*. (1992). An amino acid substitution (Gly853–Glu) in the collagen a1(II) chain produces hypochondrogenesis. *Journal of Biological Chemistry*, **267**, 22 522–6.

Boileau, C. *et al*. (1995). The question of heterogeneity in Marfan syndrome. *Nature Genetics*, **9**, 230–31.

Bonadio, J. F. and Byers, P. H. (1985). Subtle structural alterations in the chains of type I procollagen produce osteogenesis imperfecta type II. *Nature*, **316**, 636–6.

Briggs, M. D. *et al*. (1994). Genetic mapping of a locus for multiple epiphyseal dysplasia EDM2 to a region of chromosome 1 containing a type IX collagen gene. *American Journal of Human Genetics*, **55**, 678–84.

Briggs, M. D. *et al*. (1995). Pseudoachondrodysplasia and multiple epiphyseal dysplasia due to mutations in the cartilage oligomeric matrix protein gene. *Nature Genetics*, **10**, 330–6.

Burrows, N.P., *et al*. (1996). The gene encoding a1(V) collagen (COL5A1) is linked to Ehlers Danlos syndrome type I/III. *Journal of Investigative Dermatology*, **106**, 1273–6.

Byers, P.H. *et al*. (1980). X-linked anti 'axa' defective collagen crosslink formation due to decreased lysyl oxidase activity. *New England Journal of Medicine*, **303**, 61–5.

Byers, P. H. (1989). Invited editorial: molecular heterogeneity in the chondrodysplasias. *American Journal of Human Genetics*, **45**, 1–4.

Byers, P. H. (1990). Brittle bones, fragile molecules: disorders of collagen gene structure and expression. *Trends in Genetics*, **6**, 293–300.

Chan, D. and Cole, W. G. (1991). Low basal transcription of genes for tissue-specific collagens by fibroblasts and lymphoblastoid cells. *Journal of Biological Chemistry*, **266**, 487–94.

Chan, D., Taylor, T. K. F., and Cole, W. G. (1993). Characterisation of an arginine 789 to cysteine substitution in a1(II) collagen chains of a patient with spondyloepiphyseal dysplasia. *Journal of Biological Chemistry*, **268**, 15 238–45.

Christiano, A. M. *et al*. (1992). The large non-collagenous domain NC1 of type VII collagen is amino terminal DNA chimeric, homologous to cartilage matrix protein the type II domains of fibronectin and the A domains of Von Willebrand factor. *Human Molecular Genetics*, **1**, 475–81.

Christiano, A. M. *et al*. (1993). A missense mutation in type VII collagen in two affected siblings with recessive dystrophic epidermolysis bullosa. *Nature Genetics*, **4**, 62–6.

Christiano, A. M. *et al*. (1994*a*). Structural organisation of the human type VII collagen gene (COL7A1) comprised of more exons than any previously characterised gene. *Genomics*, **21**, 169–79.

Christiano, A. M. *et al*. (1994*b*). Premature termination codons in the type VII collagen gene (COL7A1) underlie severe mutilating recessive epidermolyis bullosa. *Genomics*, **21**, 160–8.

Christiano, A. M. *et al*. (1995). Dominant dystrophic epidermolysis bullosa: a glycine-to-arginine substitution in the triple helical domain of type VII collagen. *Journal of Investigative Dermatology*, **104**, 438–40.

Christiano, A. M. *et al*. (1996*a*). Compound heterozygosity for COL7A1 mutations in twins with dystrophic epidermolysis bullosa: a recessive paternal deletion/insertion mutation and a dominant negative maternal glycine substitution result in a severe phenotype. *American Journal of Human Genetics*, **58**, 682–93.

Christiano, A. M. *et al*. (1996*b*). Glycine substitutions in the triple helical region of type VII collagen result in a spectrum of dystrophic epidermolysis bullosa phenotypes and patterns of inheritance. *American Journal of Human Genetics*, **58**, 671–81.

Cole, W.G. and Dalgleish, R. (1995). Perinatal lethal osteogenesis imperfecta. *Journal of Medical Genetics*, **32**, 284–9.

Cole, W. G., Evans, R., and Sillence, D. O. (1987). The clinical features of Ehlers–Danlos syndrome type VII due to a deletion of 24 amino acids from the proa1(I) chain of type I procollagen. *Journal of Medical Genetics*, **24**, 698–701.

Cole, W.G., *et al*. (1990). Substitution at a splice site in the COL3A1 gene causes exon skipping and generates abnormal type III procollagen in a patient with Ehlers Danlos syndrome type IV. *Journal of Biological Chemistry*, **265**, 17 070–7.

Cole, W. G., Hall, R. K., and Rogers, J. G. (1993). The clinical features of spondyloepiphyseal dysplasia congenita resulting from the substitution of glycine 997 by serine in the a1(II) chain of type II collagen. *Journal of Medical Genetics*, **30**, 27–35.

Collod, G. *et al*. (1995). A second locus for Marfan syndrome maps to chromosome 3p24.2-p25. *Nature Genetics*, **8**, 264–8.

Cooke, R. M. *et al*. (1987). The solution structure of human epidermal growth factor. *Nature*, **327**, 339–41.

Corson, G. M., Chalberg, S. C., Dietz, H. C., Charbonneau, N. L., and Saiki, L. Y. (1993). Fibrillin binds calcium and is coded by cDNAs that reveal a multidomain, structural, alternatively spliced exons at the 5' end. *Genomics*, **17**, 478–84.

Cotton, R. G. H., Rodrigues, N. R., and Campbell, R. D. (1989). Reactivity of cytosine and thymine in single-base pair mismatches with hydroxylamine and osmium tetroxide and its application to the study of mutations. *Proceedings of the National Academy of Sciences (USA)*, **85**, 4397–401.

Dharmavaram, R. M. *et al*. (1994). Identification of a mutation in type X collagen occur in two unrelated families with metaphyseal chondrodysplasia. *Human Molecular Genetics*, **4**, 507–9.

Dieppe, P. (1987). Osteoarthritis and related disorders. In *Oxford textbook of medicine*, Vol. 2 (ed. D. J. Weatherall, J. G. G. Ledingham, and D. A. Warrell), pp. 76–53. Oxford University Press.

Dietz, H. C. *et al*. (1991). Marfan syndrome caused by a recurrent *de novo* missense mutation in the fibrillin gene. *Nature*, **352**, 337–9.

Dietz, H. C., Saraiva, J. M., Pyeritz, R. E., Cutting, G. R., and Francomano, C. A. (1992). Clustering of fibrillin (*FBN1*) missense mutations in Marfan syndrome patients at cysteine residues in EGF like domains. *Human Mutation*, **1**, 366–74.

Dietz, H. C., *et al*. (1993*a*). The skipping of constitutive exons *in vivo*, induced by nonsense mutations. *Science*, **259**, 680–3.

Dietz, H. C. *et al*. (1993*b*). Four novel *FBN1* mutations, significance for mutant transcript level and EGF like domain calcium binding in the pathogenesis of Marfan syndrome. *Genomics*, **17**, 468–75.

Dietz, H. C. *et al*. (1995). The question of heterogeneity in Marfan syndrome. *Nature Genetics*, **9**, 228–9.

Dolle, P. *et al*. (1993). Disruption of the *Hoxd13* gene induces localized heterochrony leading to mice with neotenic limbs. *Cell*, **75**, 431–41.

Dunnill, M. G. S. *et al*. (1994*a*). Genetic linkage to the type VII collagen gene (*COL7A1*) in 26 families with generalized recessive dystrophic epidermolysis bullosa and anchoring fibril abnormalities. *Journal of Medical Genetics*, **31**, 745–8.

Dunnill, M. G. S., *et al*. (1994*b*). A novel homozygous point mutation in the collagen VII gene (*COL7A1*) in two cousins with recessive dystrophic epidermolysis bullosa. *Human Molecular Genetics*, **3**, 1693–4.

Dunnill, M. G. S. *et al*. (1995*a*). Mutation analysis, ultrastructural and immunohistochemical correlations in recessive dystrophic epidermolysis bullosa. *Journal of Investigative Dermatology*, **105**, 449 (Abstr.)

Dunnill, M. G. S. *et al*. (1995*b*). Use of type VII collagen (*COL7A1*) markers in prenatal diagnosis of dystrophic epidermolysis bullosa. *Journal of Medical Genetics*, **32**, 749–50.

Eady, R. A. J. (1990). *Epidermolysis bullosa: a comprehensive review of classification, management and laboratory studies* (ed. G. C. Priestly, M. J. Tidman, J. B. Weiss, and R. A. J. Eady), pp. 1–9. Dystrophic Epidermolysis Bullosa Research Association, Crowthorne, Berks, UK.

Eady, R. A. J. and Dunnill, M. G. S. (1994). Epidermolysis bullosa: hereditary skin fragility diseases as paradigms in cell biology. *Archives of Dermatological Research*, **287**, 2–9.

Eady, R. A. J., McGrath, J. A., and McMillan, J. R. (1994). Ultrastructural clues to genetic disorders of the skin: the dermoepidermal junction. *Journal of Investigative Dermatology*, **103**, 135–81.

Eladha, Z. A., Brunn, T., Furthmayr, H., and Dietz, H. C. (1995). Expression of a mutant fibrillin (*FBN1*) allele upon a normal human or murine genetic background recapitulates a Marfan cellular phenotype. *Journal of Clinical Investigation*, **95**, 874–80.

Erlebacher, A., *et al*. (1995). Toward a molecular understanding of skeletal development. *Cell*, **80**, 371–6.

Eyre, D. R., Upton, M. P., Shapiro, F. D., Wilkinson, R. H., and Vawter, G. F. (1986). Non expression of cartilage type II collagen in a case of Langer–Saldino achondrogenesis. *American Journal of Human Genetics*, **39**, 52–67.

Fassler, R. *et al*. (1994). Mice lacking a a1(IX) collagen develop non inflammatory degenerative joint disease. *Proceedings of the National Academy of Sciences (USA)*, **91**, 5070–4.

Finkelstein, J. E. *et al*. (1991). Analysis of the chondroitin sulphate proteoglycan core protein (*CSPGCP*) gene in achondroplasia and pseudoachondroplasia. *American Journal of Human Genetics*, **48**, 97–102.

Fountain, J. W. *et al*. (1989*a*). Physical mapping of a translocation breakpoint in neurofibromatosis. *Science*, **244**, 1085–7.

Fountain, J. W. *et al*. (1989*b*). Physical mapping of the von Recklinghausen neurofibromatosis region on chromosome 17. *American Journal of Human Genetics*, **44**, 58–67.

Francomano, C. A. *et al*. (1987). The Stickler syndrome: evidence for close linkage to the structural gene for type II collagen. *Genomics*, **1**, 293–6.

Francomano, C. A. *et al*. (1988). The Stickler and Wagner syndromes: evidence for genetic heterogeneity. *American Journal of Human Genetics*, **43** (Suppl.), A83.

Fryer, A. E. *et al*. (1990). Exclusion of *COL2A1* as the candidate gene in a family with Wagner Stickler's syndrome. *Journal of Medical Genetics*, **27**, 91–3.

Gammon, W. R. *et al*. (1992). Non-collagenous NC1 domain of collagen VII resembles multidomain adhesion proteins involved in tissue specific organisation of extracellular matrix. *Journal of Investigative Dermatology*, **99**, 691–6.

Garofalo, S. *et al*. (1991). Reduced amounts of cartilage collagen fibrils and growth plate anomalies in transgenic mice harbouring a glycine to cysteine in the mouse type II procollagen a1 chain. *Proceedings of the National Academy of Sciences (USA)*, **88**, 9648–52.

Gendron-McGuire, M., Malo, M., Zhang, M., and Gridley, T. (1993). *Hoxa2* mutant mice exhibit homeotic transformation of skeletal elements derived from the cranial neural crest. *Cell*, **75**, 1317–31.

Godfrey, M. and Hollister, D. W. (1988). Type II achondrogenesis-hypochondrogenesis; identification of abnormal type II collagen. *American Journal of Human Genetics*, **43**, 904–13.

Godfrey, M. *et al*. (1990). Cosegregation of elastin-associated microfibrillar abnormalities with the Marfan phenotype in families. *American Journal of Human Genetics*, **46**, 652–60.

Graf, C. J. (1965). Spontaneous carotid–cavernous fistulae: Ehlers–Danlos and related conditions. *Archives of Neurology*, **14**, 862–72.

Grange, D. K., Gottesman, G. S., and Marini, J. C. (1988). Detection of point mutations in type I collagen mRNA: an application of an RNA/RNA hybrid system in osteogenesis imperfecta. *American Journal of Human Genetics*, **43** (Suppl.), A85.

Greenspan, D. S. (1993). The carboxyl terminal half of type VII collagen including the non-collagenous NC2 domain and intron/exon organisation of the corresponding region of the *COl7A1* gene. *Human Molecular Genetics*, **2**, 273–8.

Grovl, D. and Yaifin, A. (1992). Complexity of FGF receptors: genetic basis for structural diversity and functional specificity. *FASEB Journal*, **6**, 3362–9.

Gruis, N.A., *et al*. (1992). Genetic linkage between the collagen VII gene (COL7A1) and the autosomal dominant form of dystrophic epidermolysis bullosa in two Dutch kindreds. *Journal of Investigative Dermatology*, **99**, 528–30.

Gusella, J. *et al*. (1985). Deletion of Huntingdon's disease linked G8 (D4S10) locus in Wolf–Hirschorn syndrome. *Nature*, **318**, 75–8.

Hanaford, P. A., Mayhew, M., Baron, M., Winship P. R., Campbell, I. D., and Brownlee, G. G. (1991). Key residues involved in calcium binding motifs in EGF-like domains. *Nature*, **351**, 164–7.

Hastbacka, J. *et al*. (1995). The diastrophic dysplasia gene encodes a novel sulphate transporter: positional cloning by fine-structure linkage disequilibrium mapping. *Cell*, **78**, 1973–8.

Hautala, P. A., Byers, M. G., Eddy, R. L., Shows, T. B., Kivirriko, K. I., and Myllylya, R. (1992). Cloning of human lysyl hydroxylase: complete cDNA derived amino acid sequence and assignment of the gene (*PLOD*) to chromosome 1p36.–36.2. *Genomics*, **13**, 62–9.

Hecht, J. T. *et al*. (1995). Mutations in exon 17B of cartilage oligomeric matrix protein COMP cause pseudoachondrodysplasia. *Nature Genetics*, **10**, 325–9.

Hewett, D., Lynch, J., Child, A., Firth, H., and Sykes, B. (1994). Differential allelic expression of a fibrillin gene (*FBN1*) in patients with the Marfan syndrome. *American Journal of Human Genetics*, **52**, 447–52.

Hilal, L. *et al*. (1993). A homozygous insertion-deletion in the type VII collagen gene (*COL7A1*) in Hallopeau Siemens dystrophic epidermolysis bullosa. *Nature Genetics*, **5**, 287–93.

Horton, W. A. (1988). Approaches to investigating growth plate cartilage in the human chondrodystrophies. *Pathology and Immunopathological Research*, **7**, 85–9.

Horton, W. A. (1995). Molecular genetics of the human chondrodysplasias - 1995 (review). *European Journal of Human Genetics*, **3**, 357–73.

Hostikka, S. L., Eddy, R. L., Byers, M. G., Hoyhlya, M., Shows, T. B., and Tryggvarson, K. (1990). Identification of a distinct type IV collagen a chain with restricted kidney distribution and assignment of its gene to the locus of X-chromosome-linked Alport syndrome. *Proceedings of the National Academy of Sciences (USA)*, **87**, 1606–10.

Hovnanian, A. *et al*. (1992). Genetic linkage of recessive dystrophic epidermolysis bullosa to the type VII collagen gene. *Journal of Clinical Investigation*, **90**, 1032–6.

Hovnanian, A, *et al*. (1994), Recurrent nonsense mutations with the type VII collagen gene in patients with severe recessive dystrophic epidermolysis bullosa. *American Journal of Human Genetics*, **55**, 289–96.

Hull, R. G. and Pope, F. M. (1989). Osteoarthritis and cartilage collagen genes. (Letter.) *Lancet*, **i**, 1337–8.

Jabs, E. W. *et al*. (1993). A mutation in the homeodomain of the human *MSX2* gene in a family affected with autosomal dominant craniosynostosis. *Cell*, **75**, 443–50.

Jabs, E. W. *et al*. (1995). Jackson–Weiss and Crouzon syndromes are allelic with mutations in fibroblast growth factor receptor 2. *Nature Genetics*, **8**, 275–9.

Jaenisch, R. *et al*. (1983). Germline integration of Moloney murine leukaemia virus at the *Mov13* locus leads to lethal mutation and early embryonic death. *Cell*, **32**, 209–16.

Jones, K. L. (1988). *Smith's recognisable patterns of human malformation* (4th edn), pp 278–81. Saunders, Philadelphia.

Kainulainen, K., Pulkkinen, L., Savolainen, A., Kailila, I., and Peltonen, L. (1990). Location on chromosome 15 of the gene defect causing Marfan syndrome. *New England Journal of Medicine*, **323**, 935–9.

Kainulainen, K. *et al*. (1991). Marfan syndrome: no evidence for heterogeneity in different populations and more precise mapping of the gene. *American Journal of Human Genetics*, **49**, 662–7.

Kainulainen, K. *et al*. (1992). Two mutations in Marfan syndrome resulting in truncated fibrillin polypeptides. *Proceedings of the National Academy of Sciences (USA)*, **89**, 5917–21.

Kainulainen, K., Karttunen, L., Puhakka, L., Saiki, L., and Peltonen, L. (1994). Mutations in the fibrillin gene responsible for dominant ectopia lentis and neonatal Marfan syndrome. *Nature Genetics*, **6**, 64–9.

Karttman, L., Raghunath, M., Lonnquist, L., and Peltonen, L. (1994). A compound heterozygous Marfan patient: two defective fibrillin alleles result in lethal phenotype. *American Journal of Human Genetics*, **55**, 1083–91.

Khillan, J. S., Olsen, A. S., Sokolov, B., and Prockop, D. J. (1991). Transgenic mice that express a minigene version of the human gene for type I procollagen (COL1A1) develop a lethal phenotype resembling a lethal form of osteogeneis imperfecta. *Journal of Biological Chemistry*, **266**, 23 373–9.

Kielty, C. M., Phillips, J. E., Child, A,, Pope, F. M., and Shuttleworth, C. A. (1994). Fibrillin secretion and microfibril assembly by Marfan dermal fibroblasts. *Matrix Biology*, **14**, 191–9.

Kielty, C. M., Davies, S. J., Phillips, J. E., Jones, C. J. P., Shuttleworth, C. A., and Charles, S. J. (1995). Marfan syndrome: fibrillin expression and microfibrillar abnormalities in a family with predominant ocular defect. *Journal of Medical Genetics*, **32**, 1–6.

Kingsley, D. M. (1994). What do BMPs do in mammals? *Trends in Genetics*, **10**, 16–21.

Kivirriko, K. I. and Myllyla, R. (1982). Post-translational enzymes in the biosynthesis of collagen: intracellular enzymes. In *Methods in Enzymology*, Vol. 82, Part A, *extracellular matrix*: *structural and contractile proteins* (ed. L. W. Cunningham and M. Frederiksen), pp. 245–304. Academic Press, New York.

Knowlton, R. G. *et al*. (1990). Genetic linkage of a polymorphism in the type II procollagen gene (*COL2A1*) to primary osteoarthritis, associated with mild chondrodysplasia. *New England Journal of Medicine*, **322**, 526–30.

Komorowska, A., Rozynkowa, D, Lee, K. W., Renouf, D. V., MacKenzie, J., and Pope, F. M. (1989). A Polish variant of isolated dentinogenesis imperfecta with a generalised connective tissue defect. *British Dental Journal*, **167**, 239–41.

Kontusaari, S., Tromp, G., Kuivaniemi, H., Ladda, R. L., and Prockop, D. J. (1990). Inheritance of an RNA splicing mutation (G+ $^{1\ 1vs20}$) in the type III procollagen gene (*COL3A1*) in a family having aortic aneurysms and easy bruisability: phenotype overlap between familial arterial aneurysms and type IV Ehlers–Danlos syndrome. *American Journal of Human Genetics*, **47**, **11**2–20.

Kraus, S., Concordet, J-P., and Ingham, P. W. (1993). A functionally conserved homologue of the *Drosophila* segment polarity gene *hh* is expressed in tissues with polarizing activity in Zebra fish embryos. *Cell*, **75**, 1435–44.

Krumlauf, R. (1994). *HOX* genes in vertebrate development. *Cell*, **78**, 191–2.

Kuivaniemi, H., Kontusaari, S., Tromp, G., Shao, M., Sobol, C., and Prockop, D. J. (1990). Identical G $^{+1}$to A mutations in three different introns of the type III procollagen gene (*COL3A1*) produce different patterns of RNA splicing in three variants of Ehlers–Danlos syndrome IV. An explanation of exon skipping with some mutations and not others. *Journal of Biological Chemistry*, **265**, 12 067–70.

Kuivaniemi, H., Tromp, G., and Prockop, D. J. (1991). Mutations in collagen genes. Causes of rare and some common diseases in man. *FASEB Journal*, **5**, 2052–60.

Lajeune, E., Wei, M. H., Bonavanture, J., Munnich, A., Le Merrer, M., and Reiner, D. (1995). *FGFR2* mutations in Pfeiffer syndrome. *Nature Genetics*, **9**, **10**8.

Lapiere, C. M., Lenaers, A., and Kohn, L. D. (1971). Procollagen peptidase: an enzyme excising the coordination peptides of collagen. *Proceedings of the National Academy of Sciences (USA)*, **68**, 3054–8.

Lee, B., Vissing, H., Ramirez, F., Rogers, D., and Rimoin, D. (1988). Identification of the molecular defect in a family with spondyloepiphyseal dysplasia. *Science*, **244**, 978–82.

Lee, B., D'Alessio, M., Vissing, H., Ramirez, F., Steinmann, B., and Superti-Furga, A. (1991 *a*). Characterisation of a large genomic deletion associated with a polymorphic block of repeated dinucleotides in the type III procollagen gene *COL3A1* of a patient with Ehlers–Danlos syndrome type IV. *American Journal of Human Genetics*, **48**, 511–17.

Lee, B. *et al*. (1991*b*). Linkage of Marfan syndrome and a phenotypically related disorder to two different fibrillin genes. *Nature*, **352**, 330–4.

Lewis, A. F. (1981). Fractured neck of femur: changes in incidence. *British Medical Journal*, **283**, 1217–20.

Li, Y. *et al*. (1995). A fibrillar collagen gene *COL11A1* is essential for skeletal morphogenesis. *Cell*, **80**, 423–30.

Lichtenstein, J. R., Martin, G. R, Kohn, I. D., Byers, P. H., and McKusick, V. A. (1974). Defect in conversion of procollagen to collagen in a form of Ehlers–Danlos syndrome. *Science*, **182**, 298–300.

Loughlin, J. *et al*. (1995). Linkage of the gene that encodes the a1 chain of type V collagen (*COL5A1*) to type II Ehlers–Danlos syndrome (EDS II). *Human Molecular Genetics*, **4**, 1649–51.

Lunstrum, G., Saiki, L. Y., Keene, D. B., Morris, N., and Burgeson, R. E. (1986). Large complex globular domains to type VII procollagen contribute to the structure of anchoring fibrils. *Journal of Biological Chemistry*, **261**, 9024–40.

McGillivray, B. C., Bassett, A. S., Langlois, S., Pantzar, T. P., and Wood, S. (1990). Familial 5q11–5q 13.13 segmental duplication cosegregating with multiple anomalies including schizophrenia. *American Journal of Medical Genetics*, **35**, 10–13.

McGookey, D. J., Smith, A., Waldstein, G., and Byers, P. H. (1989). Mosaicism for a deletion in one of the type III collagen alleles indicates that the deletion occurred after identification of cells for recruitment in different cell lineages early in human development. *American Journal of Human Genetics*, **45**, A206.

McGookey-Milewicz, D., Pyeritz, R. E., Crawford, L. S., and Byers, P. H. (1992). Marfan syndrome: defective synthesis secretion and extracellular matrix formation of fibrillin by cultured dermal fibroblasts. *Journal of Clinical Investigation*, **89**, 79–86.

McGrath, J. A., Ishida-Yamamoto, A., O'Grady, A., Leigh, I. M., and Eady, R. A. J. (1993). Structural variations in anchoring fibrils in dystrophic epidermolysis bullosa, correlation with type VII collagen expression. *Journal of Investigative Dermatology*, **100**, 366–72.

McGrath, J. A. *et al*. (1995*a*). Mutations in the 180 Kd bullous pemphigoid antigen (BPAG2), a hemidesmosomal transmembrane collagen in generalised atrophic benign epidermolysis bullosa. *Nature Genetics*, **11**, 83–6.

McGrath, J. A. *et al*. (1995*b*). A homozygous nonsense mutation in the a3 chain of laminin 5 (*LAMA3*) in Herlitz junctional epidermolysis bullosa: prenatal diagnosis exclusion in a fetus at risk. *Genomics*, **29**, 282–4.

McGrath, J. A. *et al*. (1995*c*). First trimester DNA-based exclusion of recessive dystrophic epidermolysis bullosa from chorionic villus sampling. *British Journal of Dermatology*, **134**, 734–7.

McGrath, J. A., Pulkkinen, L., Cristiano, A. M., Leigh, I. M., Eady, R. A. J., and Uitto, J. (1995). Altered laminin 5 expression due to mutation on the b $_3$ chain gene (*LAMB3*). *Journal of Investigative Dermatology*, **104**, 467–74.

McIntosh, I. (1994). Additional mutations of type X collagen confirm *COL10A1* as the Schmid metaphyseal chondrodysplasia locus. *Human Molecular Genetics*, **3**, 303–7.

McKusick, V.A. (1956). *Heritable disorders of connective tissue*. Mosby, St Louis.

McKusick, V. A. (1972). *Heritable disorders of connective tissue* (4th edn). Mosby, St Louis.

McKusick, V. A., Francomano, C. F., and Antonarakis, S. E. (1990). *Mendelian inheritance in man: catalogs of autosomal dominant, autosomal recessive and X-linked phenotypes* (9th edn). Johns Hopkins University Press, Baltimore.

Marfan, A. B. (1896). Un cas de déformation congenitale des quartre membres plus prononcée aux extrémités charactrerisée par l'allongement des os avec un certain degré d'amincissement.

Bulletins et Memoires de la Societé Medical des Hopitaux de Paris, **13**, 220–1.

Maroteaux, P. (1970). Nomenclature internationale des maladies osseuses constitutionnelles. *Annals of Radiology*, **13**, 455.

Maslen, C. L., Corson, G. M., Maddox, B. K., Glanville, R. W., and Saiki, L. Y. (1991). Partial sequence of a candidate gene for the Marfan syndrome. *Nature*, **352**, 334–7.

Milewicz, D. M. and Durie, M. (1994). Severe neonatal Marfan syndrome resulting from a *de novo* 3 bp insertion into the fibrillin gene on chromosome 15. *American Journal of Human Genetics*, **54**, 447–53.

Milewicz, D., Witz, A. M., Smith, A. C., Manchester, D. K., Waldstein, G., and Byers, P. H. (1993). Parental somatic and germ line mosaicism for a multiexon deletion with unusual endpoints on a type III collagen (*COL3A1*) allele produces Ehlers–Danlos syndrome type IV in the heterozygous offspring. *American Journal of Human Genetics*, **53**, 62–70.

Monaco, A. P., Neve, R. L., Colletti-Feener, C., Bertelson, C. J., Kurnit, D. M., and Kunkel, L. M. (1986). Isolation of candidate cDNA's for portions of the Duchenne muscular dystrophy gene. *Nature*, **323**, 546–50.

Mortality Statistics (1977). *Review of Registrar General on deaths in England and Wales*, Table 6, p. 29. HMSO, London.

Mortier, G. R. *et al*. (1995). A radiographic, morphologic, biochemical and molecular analysis of a case of achondrogenesis type II resulting from a substitution for a glycine residue (Gly ^{891} –Arg). in the type II collagen trimer. *Human Molecular Genetics*, **4**, 285–8.

Muenke, M. *et al*. (1994). A common mutation in the fibroblast growth factor receptor 1 gene in Pfeiffer syndrome. *Nature Genetics*, **8**, 269–74.

Mulverhill, J. J. (1995). Craniofacial syndromes: no such thing as a single disease. *Nature Genetics*, **9**, 101–3.

Murray, L. W., Bautista, J., James, P. L., and Rimoin, D. L. (1989). Type II collagen defects in the chondrodysplasias. I. Spondyloepiphyseal dysplasia. *American Journal of Human Genetics*, **45**, 5–15.

Nakata, K. *et al*. (1993). Osteoarthritis associated with mild chondrodysplasia in transgenic mice expressing a1(IX) collagen chains with a central deletion. *Proceedings of the National Academy of Sciences (USA)*, **90**, 2870–4.

Nicholls, A. C., Craig, D., and Pope, F. M. (1984 *a*). An abnormal collagen a chain containing cysteine in autosomal dominant osteogenesis imperfecta. *British Medical Journal*, **288**, 112–13.

Nicholls, A. C. *et al*. (1984*b*). The clinical features of homozygous a2(I) collagen deficient osteogenesis imperfecta. *Journal of Medical Genetics*, **21**, 257–62.

Nicholls, A.C., *et al*. (1988). Linkage of a polymorphic marker for the type III collagen gene (COL3A1) to atypical autosomal dominant Ehlers Danlos syndrome type IV in a Belgian kindred. *Human Genetics*, **79**, 276–8.

Nicholls, A. C., Oliver, J., Renouf, D. V., McPheat, J., Palan, A., and Pope, F. M. (1991). Ehlers–Danlos syndrome type VII: a single base change that causes exon skipping in the type I collagen a2 chain. *Human Genetics*, **87**, 193–8.

Nicholls, A. C., McCarron, S., Narcisi, P., and Pope, F. M. (1994). Molecular abnormalities of type V collagen in Ehlers–Danlos syndrome. *American Journal of Human Genetics*, **55** (Abstr. 1358), 223.

Nicholls, A.C., Oliver, J.C., McCarron, S., Harrison, J.B., Greenspan, D.S., and Pope, F.M. (1996). An exon skipping mutation of a type I collagen gene (COL5A1) in Ehlers Danlos syndrome. *Journal of Medical Genetics*, **53**, 943–6.

Nijbroek, G. *et al*. (1995). Fifteen novel *FBN1* mutations causing Marfan syndrome detected by hetero duplex analysis of genomic amplicons. *American Journal of Human Genetics*, **57**, 8–21.

Nunez, A. M., Francomano, D., Young, M. F., Martin, G. R., and Yamada, Y. (1985). Isolation and partial characterisation of genomic clones coding for a human pro ak(II) collagen chain and demonstration of restriction fragment length polymorphism at the 3' end of the gene. *Biochemistry*, **24**, 6343–8.

Nusgens, B. V. *et al*. (1992). Evidence for a relationship between Ehlers–Danlos type VIIC in human and bovine dermatosporaxis. *Nature Genetics*, **1**, 214–17.

Oehlmann, R., Summerville, G. P., Yeh, G., Weaver, E. J., Jimenez, S. A., and Knowlton, R. A. (1994). Genetic linkage of multiple epiphyseal dysplasia to the pericentromeric region of chromosome 19. *American Journal of Human Genetics*, **54**, 3–10.

Palotie, A. *et al*. (1989). Predisposition to familial osteoarthritis linked to the type II collagen gene. *Lancet*, **i**, 924–7.

Parente, M. G. *et al*. (1991). Human type VII collagen: cDNA cloning and chromosomal mapping of the gene. *Proceedings of the National Academy of Sciences (USA)*, **88**, 6931–5.

Pereira, L. *et al*. (1993). Genomic organisation of the sequence coding for fibrillin, the defective gene product in Marfan syndrome. *Human Molecular Genetics*, **2**, 961–8.

Peters, K., Ornitz, D., Werner, S., and Williams, L. (1993). Unique expression patterns of the FGF receptor 3 gene during mouse organogenesis. *Developmental Biology*, **155**, 423–30.

Phillips, C. L., Shrago-Howe, A. W., Pinnell, S. R., and Wenstrup, R. J. (1991). A substitution at a non-glycine position in the triple-helical domain of proa2(I) collagen chains present in an individual with a variant of the Marfan syndrome. *Journal of Clinical Investigation*, **86**, 1723–8.

Pihlajaniemi, T. *et al*. (1984). Osteogenesis imperfecta: cloning on a pro a2(I) collagen gene with a frameshift mutation. *Journal of Biological Chemistry*, **259**, 12 941–4.

Pope, F. M. (1991). Ehlers–Danlos syndrome. Rheumatic manifestations of haematological disease. *Baillière's clinical rheumatology*, **5**, 321–49.

Pope, F. M. and Nicholls, A. C. (1986). Collagen genes and proteins in human diseases. In *Recent advances in dermatology*, Vol. 7, (ed. R. H. Champion), pp. 25–32. Churchill Livingstone, Edinburgh.

Pope, F. M. *et al*. (1988*a*). Type III collagen mutations in Ehlers–Danlos syndrome type IV and other related disorders. *Clinical and Experimental Dermatology*, **13**, 285–302.

Pope, F. M., Daw, S. C. M., Narcisi, P., Richards, A. R., and Nicholls, A. C. (1988 *b*). Prenatal diagnosis and prevention of inherited abnormalities of collagen. *Journal of Inherited Metabolic Diseases*, **12** (Suppl.), 135–73.

Pope, F. M. *et al*. (1991). Fragile cerebral arteries are caused by type III collagen mutations. *British Journal of Neurosurgery*, **5**, 551–74.

Pope, F. M., Narcisi, P., Nicholls, A. C., Germaine, D., Pals, G., and Richards, A. J. (1996). COL3A1 mutations cause various clinical phenotypes including aneurysm and vascular rupture. *British Journal of Dermatology*, **135**, 163–8.

Pousi, B., Hautala, T., Heikkinen, J., Pajunen, L., Kivirriko, K. I., and Myllala, R. (1994). Alu–Alu recombination results in a duplication of seven exons in the lysyl hydroxylase gene in a patient with the type VI variant of Ehlers–Danlos syndrome. *American Journal of Human Genetics*, **55**, 899–906.

Prockop, D. J. (1990). Mutations that alter the primary structure of type I collagen. The perils of a system for generating large structures by the principle of nucleated growth. *Journal of Biological Chemistry*, **265**, 15 349–52.

Prockop, D. J. and Kivirriko, K. I. (1984). Inheritable disorders of collagen. *New England Journal of Medicine*, **311**, 376–86.

Prockop, D. J. and Kivirriko, K.W. (1995). Collagens, molecular biology of diseases, and potentials for therapy. *Annual Reviews of Biochemistry*, **64**, 403–34.

Raghunath, M., Superti-Furga, A., Godfrey, M., and Steineimann, B. (1993). Decreased extracellular deposition of fibrillin and decorin in neonatal Marfan syndrome fibroblasts. *Human Genetics*, **90**, 511–15.

Reardon, W., Winter, R. M., Rutland, P., Pulleyn, L., Jones, B. M., and Malcolm, S. (1994). Mutations in the fibroblast growth factor receptor 2 gene cause Crouzon syndrome. *Nature Genetics*, **8**, 98–103.

Richards, A. J., Ward, P. N., De Paepe, A., Narcisi, P. M., Nicholls, A. C., and Pope, F. M. (1991). A 27 base pair deletion from one allele of the type III collagen gene *COL3A1* in a large family with non acrogeric Ehlers–Danlos syndrome IV. *Human Genetics*, **88**, 325–30.

Riddle, R. D., Johnston, R. L., Laufer, E., and Tabin, C. (1993). Sonic Hedgehog mediates the polarizing activity of the *ZPA*. *Cell*, **75**, 1401–16.

Rijli, F. M., Mark, M., Lakkaraju, S., Dierich, A., Dolle, P., and Chambon, P. (1993). A homeotic transformation is generated in the rostral branchial region of the head by disruption of *Hoxa-2* which acts as a selector gene. *Cell*, **75**, 1333–4.

Roberts, D. L. L., Pope, F. M., Nicholls, A. C., and Narcisi, P. (1984). Ehlers–Danlos syndrome IV mimicking non accidental injury in a child. *British Journal of Dermatology*, **11**, 341–5.

Rommens, J. M. *et al*. (1989). Identification of the cystic fibrosis gene: chromosome walking and jumping. *Science*, **245**, 1059–65.

Rosati, R. *et al*. (1994). Normal long bone growth and development in type X collagen null mice. *Nature Genetics*, **8**, 129–35.

Rousseau, F. *et al*. (1994). Mutations in the gene encoding fibroblast growth factor receptor 3 in achondroplasia. *Nature*, **371**, 252–4.

Rutland, P. *et al*. (1995). Identical mutations in the *FGFR2* gene cause both Pfeiffer and Crouzon syndrome phenotypes. *Nature Genetics*, **9**, 173–6.

Ryynanen, M., Knowlton, R. G., Parente, M. G., Chung, L. C., Chu, M. L., and Uitto, J. (1991). Human type VII collagen: genetic linkage of the gene (*COL7A1*) on chromosome 3 to dominant dystrophic epidermolysis bullosa. *American Journal of Human Genetics*, **49**, 797–803.

Ryynanen, M., *et al*. (1992). Genetic linkage of type VII collagen (COL7A1) to dominant dystrophic epidermolysis bullosa in familes with abnormal anchoring fibrils. *Journal of Clinical Investigation*, **89**, 974–80.

Saiki, R. K. *et al*. (1986). Enzymatic amplification of b globin gamma sequences and restriction site analysis for the diagnosis of sickle cell anaemia. *Science*, **230**, 1350–4.

St George-Hyslop, P. H. *et al*. (1987). The genetic defect causing familial Alzheimer's disease maps on chromosome 21. *Science*, **235**, 885–90.

Schwartz, R. C. *et al*. (1989). Non allelic heterogeneity in the vitreo-retinal degenerations of Stickler and Wagner types and evidence for intragenic recombination at the *COL2A1* locus. *American Journal of Human Genetics*, **45**, A128.

Sher, C., Ramesar, T., Martell, R., Leammonti, J., Tsipouras, P., and Beighton, P. (1991). Mild spondyloepiphyseal dysplasia (Naquamaland type): genetic linkage to the type II collagen gene. *American Journal of Human Genetics*, **48**, 518–24.

Shiang, R. *et al*. (1994). Mutations in the transmembrane domain of *FGFR3* cause the most common genetic form of dwarfism: achondroplasia. *Cell*, **78**, 335–42.

Sillence, D. O., Senn, A., and Danks, D. M. (1979). Genetic heterogeneity in osteogenesis imperfecta. *Journal of Medical Genetics*, **16**, 101–16.

Smith, L. T. *et al*. (1992). Human dermatosporaxis: a form of Ehlers–Danlos syndrome that results from failure to remove the aminoterminal propeptide of type I procollagen. *American Journal of Human Genetics*, **51**, 235–44.

Snead, M. P. (1994). Stickler syndrome: correlation between vitreoretinal phenotypes and linkage to *COL2A1*. *Eyes*, **8**, 609–14.

Spranger, J. W. (1985). Pattern recognition in bone dysplasias. In *Endocrine genetics and genetics of growth* (ed. C. J. Papadatos and C. S. Bartosocas), pp. 315–42. Liss, New York.

Spranger, J. (1992). International classification of the osteochondrodysplasias. *European Journal of Paediatrics*, **151**, 407–15.

Stacey, A., Bateman, J., Choi, T., Mascara, T., Cole, W. G., and Jaenisch, R. (1988). Perinatal lethal osteogenesis imperfecta in transgenic mice bearing an engineered mutant proa1(I) collagen gene. *Nature (London)*. **332**, 131–6.

Steinmann, R., Rao, V. H., Vogel, A., Bruckner, D., Gitzelmann, R., and Byers, P. H. (1984). Cysteine in the triple helical domain of one allele product of the a1(I) gene of type I collagen produces a lethal form of osteogenesis imperfecta. *Journal of Biological Chemistry*, **259**, 11 129–38.

Steinmann, B., Nicholls, A. C., and Pope, F. M. (1986). Clinical variability of osteogenesis imperfecta reflecting molecular heterogeneity: cysteine substitutions of the a1(I) collagen chain producing lethal and mild forms. *Journal of Biological Chemistry*, **89**, 58–64.

Superti-Furga, A. (1994). A defect in the metabolic activation of sulfate in a patient with achondrogenesis type 1B. *American Journal of Human Genetics*, **55**, 1137–45.

Superti-Furga, A. (1996). Achondrogenesis type 1B is caused by mutations in the diastrophic dysplasia sulfate transporter gene. *Nature Genetics*, **12**, 100–2.

Superti-Furga, A., Gugler, R., Gitzelmann, R., and Steinmann, B. (1988). Ehlers–Danlos syndrome type IV: a multi exon deletion in one of the two *COL3A1* alleles affecting structure, stability and processing of type III procollagen. *Journal of Biological Chemistry*, **263**, 6226–32.

Superti-Furga, A. *et al*. (1990). Identification of a highly polymorphic (CA)n(TA)n(GT)n element in intron 24 of the *COL3A1* gene. *Matrix*, **10**, 249 (Abstr.).

Sykes, B. *et al*. (1990). Consistent linkage of dominantly inherited osteogenesis imperfecta to the type I collagen loci: *COL1A1* and *COL1A2*. *American Journal of Human Genetics*, **46**, 293–307.

Sykes, B. (1993). Marfan gene dissected. (Comment.) *Nature Genetics*, **3**, 99–100.

Tassabenji, M., *et al*. (1993). Mutations on the PAX 3 gene causing Waardener's syndrome type I and II. *Nature Genetics*, **3**, 26–30.

Tavormina, P. L. *et al*. (1994). Thanatophoric dysplasia (types I and II). caused by distinct mutations in fibroblast growth factor receptor 3. *Nature Genetics*, **9**, 321–8.

Thompson, E. M., Young, I. D., Hall, C. M., and Pembrey, M. E. (1987). Recurrence risks and prognosis in severe sporadic osteogenesis imperfecta. *Journal of Medical Genetics*, **24**, 390–405.

Tiller, G. E., Rimoin, D. L., Murray, L. W., and Cohn, D. H. (1990). Tandem duplication within a type II collagen gene (*COL2A1*) exon in an individual with spondyloepiphyseal dysplasia. *Proceedings of the National Academy of Sciences (USA)*, **87**, 3889–93.

Tiller, G. E. *et al*. (1993). An RNA splicing mutation $E^{+5IVS20}$ in the type II collagen gene (*COL2A1*) in a family with spondyloepiphyseal dysplasia congenita. *American Journal of Human Genetics*, **56**, 388–95.

Tiller, G. E. *et al*. (1995). Dominant mutations in the type II collagen gene *COL2A1* produce spondyloepiphyseal dysplasia (Strudwick type). *Nature Genetics*, **11**, 87–9.

Torriello, H. V. *et al*. (1996). A translocation interrupts the *COL5A1* gene in a patient with Ehlers–Danlos syndrome and hypomelanosis of Ito. *Nature Genetics*, **13**, 361–5.

Tromp, G., *et al*. (1989). Single base mutation in the type III procollagen gene that converts the codon for glycine 883 to aspartate in a mild variant of Ehlers Danlos syndrome IV. *Nature Genetics*, **264**, 19 312–17.

Tsipouras, P., Sarfrazi, M. K., Devi, A., Weiffenbach, B., and Boxer, M. (1991). Marfan syndrome is closely linked to a marker on chromosome 15q15 to q21. *Proceedings of the National Academy of Sciences (USA)*, **88**, 4486–8.

Tsipouras, P. *et al*. (1992). Genetic linkage of the Marfan syndrome, ectopia lentis and congenital contractural arachnodactyly to the fibrillin genes on chromosome 15 and 5. *New England Journal of Medicine*, **326**, 905–9.

Tynan, K. *et al*. (1993). Mutation screening of complete fibrillin 1 coding sequence: report of 5 new mutations including two in 8 cysteine domains. *Human Molecular Genetics*, **2**, 1873–21.

Vasan, N. S. *et al*. (1991). A mutation in the proa2(I) gene (*COL1A2*). for type I procollagen in Ehlers–Danlos syndrome type VII: evidence suggesting the skipping of exon 6 in RNA splicing may be a common cause of the phenotype. *American Journal of Human Genetics*, **48**, 305–17.

Vikkula, M., Metsaranta, M., and Ala Kokko, L. (1994). Type II collagen mutations in rare and common cartilage diseases. *Annals of Medicine*, **26**, 107–14.

Vikkula, M. *et al*. (1995). Autosomal dominant and recessive osteochondrodysplasias associated with the *COL11A2* locus. *Cell*, **80**, 431–7.

Vissing, H. *et al*. (1989). Glycine to serine substitution in the triple helix domain of the proa1(II) collagen results in a lethal perinatal form of short-limbed dwarfism. *Journal of Biological Chemistry*, **264**, 18 265–7.

Vissing, H., D'Alessio, L. B., Ramirez, F., Byers, P. H., Steinmann, B., and Superti-Furga, A. (1991). A multi exon deletion in the procollagen type II gene is associated with mild Ehlers–Danlos syndrome type IV. *Journal of Biological Chemistry*, **266**, 5244–8.

Vogel, B. E., Doetz, R., Kadler, K. E., Hojima, Y., Engel, J., and Prockop, D. J. (1988). A substitution of cysteine for glycine 748 of the a1 chain produces a kink at this site in the procollagen I molecule and an altered N proteinase cleavage site over 225 nm away. *Journal of Biological Chemistry*, **263**, 19 249–55.

Vuorio, E. and de Crombrugghe, B. (1990). The family of collagen genes. *Annual Review of Biochemistry*, **59**, 837–72.

Wallis, G. A. *et al*. (1994). Amino acid substitutions of conserved residues in the carboxyl-terminal domain of the a1(X) chain of the type X collagen occur in two unrelated families with metaphyseal chondrodysplasia type Schmid. *American Journal of Human Genetics*, **54**, 169–78.

Warman, M., Mulliken, J. B., Hayward, P. G., and Muller, U. (1993). Newly recognised autosomal dominant craniosynostotic syndrome. *American Journal of Medical Genetics*, **46**, 444–9.

Watson, J. D., Hopkins, N. H., Roberts, J. W., Steitz, J. A., and Weiner, A. M. (1987). In *Molecular biology of the gene* (4th edn), pp. 208–11; 222–34; 632–47. Benjamin/Cummings, Menlo Park, California.

Weatherall, D. J. (1991). In *The new genetics and clinical practice* (3rd edn). Oxford University Press.

Weil, D., *et al*. (1988). Identification of a mutation that causes exon skipping during collagen pre-mRNA splicing in an Ehlers-Danlos variant. *Journal of Biological Chemistry*, **263**, 8561–4.

Weil, D. *et al*. (1989*a*). A base substitution in the exon of a collagen gene causes alternative splicing and generates a structurally abnormal polypeptide in a patient with Ehlers–Danlos syndrome type VII. *EMBO Journal*, **8**, 1705–10.

Weil, D. *et al*. (1989*b*). Temperature dependent expression of a collage splicing defect in fibroblasts of a patient with Ehlers–Danlos syndrome type VII. *Journal of Biological Chemistry*, **264**, 16 804–9.

Wenstrup, R. J., Murad, S., and Pinnell, S. R. (1989). Ehlers–Danlos syndrome type VI: clinical manifestations of collagen type VI hydroxylase deficiency. *Journal of Pediatrics*, **115**, 405–9.

Westerhausen, A., Constantinou, C. D., and Prockop, D. J. (1990). A mutation that substitutes valine for glycine a1 637 in type I procollagen gene *COL1A1* and causes lethal OI. *American Journal of Human Genetics*, **47** (Suppl.), A242.

Weve, H. (1931). Uber Arachnodaktylie (Dystrophia). Mesodermalis Typus Marfanus. *Archives Augenheilkunde*, **104**, 1–46.

Wilkie, A. O. M. *et al*. (1995). Apert syndrome results from localised mutations of *FGFR2* and is allelic with Crouzon syndrome. *Nature Genetics*, **9**, 165–72.

Willing, M. C., Cohn, D. H., and Byers, P. H. (1990). Frameshift mutation near the 3' end of the *COL1A1* gene of type I collagen predicts an elongated proa1(I) chain and results in osteogenesis imperfecta type I. *Journal of Clinical Investigation*, **85**, 282–90.

Willing, M. C., Pruchno, A. J., Atkinson, M., and Byers, P. H. (1992). Osteogenesis type I is commonly due to a *COL1A1* null allele of type I collagen. *American Journal of Human Genetics*, **51**, 508–15.

Wordsworth, P., Ogilivie, D., Priestly, L., Smith, R., Wynne-Davis, R., and Sykes, B. (1988). Structural and segregation analysis of the type II collagen gene *COL2A1* in some heritable chondrodysplasias. *Journal of Medical Genetics*, **25**, 521–7.

Yamada, Y., Liau, G., Mudryj, M., Obici, S., and de Crombrugghe, B. (1984). Conservation of the sizes for one but not another class of exons in two chick collagen genes. *Nature*, **310**, 333–7.

Zhou, J. *et al*. (1990). Cloning of the human a5(IV) collagen gene and characterization of mutations in Alport syndrome. *American Journal of Human Genetics*, 47 (Suppl.), A85.

2.3 Articular cartilage

Tim Hardingham

Introduction

Articular cartilage is a unique and long-time enigmatic tissue that is just beginning to reveal some of its secrets. Just a few millimetres thick, cartilage is ingeniously fashioned so as to withstand the considerable biomechanical forces created by walking, running, and jumping, extending over many decades (Muir and Hardingham 1986). Even during gentle walking, the force acting at the hip joint is equivalent to four times the body weight. Articular cartilage (Fig. 1) is unique in that it is a poorly cellular tissue with no basement membrane, no innervation, and no direct blood supply, relying on diffusion for its nutritive requirements. It is increasingly being realized that cartilage failure as seen in diseases such as osteoarthritis is far from being the passive degenerative condition long assumed, but represents an imbalance between dynamic reparative and catabolic processes, which in health proceed in a harmonized and strictly regulated manner (Hardingham 1990; Hardingham and Bayliss 1990).

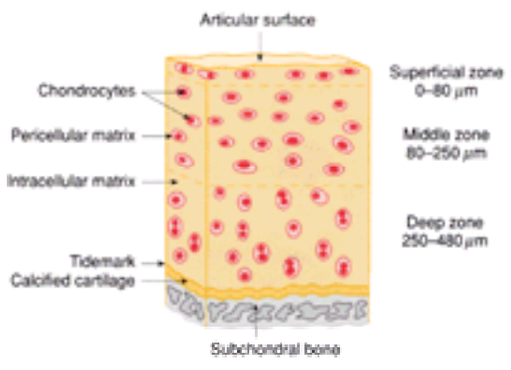


Fig. 1 Structure and organization of articular cartilage. Schematic drawing showing changes in organization from the articular surface to the subchondral bone.

The load-bearing properties of the tissue depend on the structure and matrix. The integrity of the matrix is maintained by the activity of the chondrocytes (Stockwell and Meachim 1979). The matrix contains more than 70 per cent water but, because of the high proteoglycan content, there is only free diffusion of small molecules. Solutes of large molecular weight are excluded from proteoglycan domains and, therefore, the mobility of molecules in the matrix depends on their size. The articular cartilage is bound to the underlying bone through a narrow calcified zone of cartilage. The junction between the calcified and non-calcified zone of cartilage forms the tidemark, which is named for its strong histological staining.

The articular surface is formed by a dense collagen layer of low proteoglycan content, but which is otherwise a part of the cartilage rather than a separate boundary layer. The articular joint and its component structures surrounding the synovial compartment are not bounded by any basement membrane. There are no epithelial or endothelial boundaries. All the structures are of mesodermal origin. The synovial membrane has a thickness of one or a few cells and forms the surface layer of the synovial tissue, but it is otherwise not distinct from it. The deeper synovial tissue is well supplied with blood vessels and contains many cells rich in fat deposits. The synovial tissue becomes a site of major inflammatory reaction in joint diseases such as rheumatoid arthritis.

In chronic rheumatoid arthritis it undergoes massive hyperplasia and forms a complex villous structure extending into the synovial space with a thickened lining cell layer and heavy lymphocytic infiltration with lymphoid follicles. The articular cartilage, synovium, and ligaments thus share a common compartment, with the synovial fluid linking them together. The tissues are therefore exposed to all the components that come into synovial fluid. Inflammatory events in the synovium that lead to the release of inflammatory mediators and cytokines will have consequences on ligaments and on articular cartilage.

The function of the articular joint is to give movement and mobility to the otherwise rigid bony skeleton (Mow and Lai 1980). The joint transmits force from one bone element to another. Bones are linked by ligaments and muscle/tendon elements. The articular cartilage provides two opposing smooth surfaces of extremely low friction for easy articulation, and a deformable and elastic tissue that helps distribute the load more evenly on to the underlying bone.

The different components in the joint thus have complimentary mechanical functions. All these connective tissues respond to the mechanical forces placed upon them and, in general, extra use leads to hypertrophy and mechanical strengthening. This is evident from experimental studies which showed that moderate running exercise in dogs resulted in an increase in articular cartilage thickness and proteoglycan content, and a rise in compressive stiffness (Kiviranta *et al*. 1992). This contrasted with excessively strenuous running exercise, which led to a fall in proteoglycan content, a loss of compressive stiffness, and some remodelling of the subchondral bone; but even severe exercise did not cause degeneration of the articular surface or lead to degenerative joint disease (Arokoski *et al*. 1993). High loading of a healthy joint is thus unlikely alone to initiate joint disease.

Unlike exercise, disuse of the joint leads to atrophy of the articular cartilage (Buckwalter 1995). Even temporary immobilization of a joint can result in a fall in the synthesis and content of proteoglycan; more prolonged disuse can result in the degeneration of articular cartilage and its eventual complete loss and replacement of the joint space with fibro-fatty tissue. The healthy maintenance of the different tissues thus depends on their regular mechanical loading in normal use and this reflects a property of the cells within the tissue that control the production and maintenance of the matrix. It is also clear that changes in the mechanical performance of any one joint tissue will have consequences on the loads experienced by other components causing further adaptive responses. This may be particularly prejudicial to articular cartilage integrity if it results in joint instability and greater impact loading. Therefore, the articular joint and its component tissues are far from being static structures and even in the mature animal they are in a constant dynamic state of adaptation and response. Many of the changes present in joints after many years of chronic joint disease may thus reflect adaptive responses to altered patterns of load distribution in the joint as well as direct effects of the disease processes

themselves.

Cartilage

Cartilage occurs in various forms in the body and with different functions. In embryological development it is the important forerunner of long bone development and it persists in the growth plate as the major site of long bone growth. Cartilage is also formed in the body whenever there is a recapitulation of bone development, such as in fracture healing. Cartilage occurs in other sites of the body: it supports the air passageways as bronchial rings, in the larynx it supports the vocal chords, and it occurs as nasal cartilage. It is also found in considerable amounts as intercostal cartilages and at other sites it is in various ways reinforced by other connective tissue components. The cartilage supporting the external ear contains elastin fibres, and fibrocartilage, such as that found in the menisci of the knee, contains extra type I collagen reinforcement. The intervertebral discs also contain a radially reinforced collagen structure surrounding inner regions more similar to hyaline cartilage that resist compressive loading. A thin hyaline cartilage layer also forms an articular surface at the end plate of the vertebrae where they contact the intervertebral discs. In all these forms, the basic features of cartilage are that the cells within it produce a large expanded proteoglycan-rich matrix, which together with the fibrillar network make a stiff but elastically deformable tissue that is able to withstand repetitive compressive loading.

Physical properties of articular cartilage

The physical properties of articular cartilage depend on the structure and organization of the macromolecules in the extracellular matrix (Hardingham 1990; Hardingham and Bayliss 1990). They largely can be understood in terms of the contribution made by fibrillar and non-fibrillar components (Fig. 2). The collagen fibrillar network is of fine fibres that have no preferred orientation in the mid-zone of the cartilage, although at the articular surface they are parallel to the surface and with one preferred orientation, and they are rather more perpendicular to the surface in deeper zones. The collagen provides an essential framework to the tissue that gives it shape and form. The triple helical collagen molecules are organized into fibrils with overlapping and cross-linking of adjacent molecules, and fibrils laterally associate into longer fibres. The structure of collagen gives it impressive tensile properties and this is utilized in cartilage in a special way to produce a tissue that is not only strong in tension but also resistant to compression. This is achieved by filling the interfibrillar matrix with a very high content of proteoglycan, primarily aggrecan. The aggrecan at high concentration draws water into the tissue as it creates a large osmotic swelling pressure (Fig. 2).

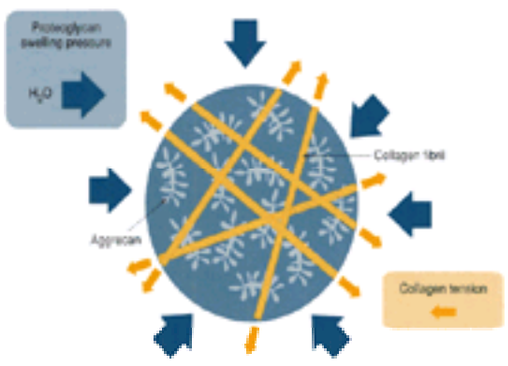


Fig. 2 The combined functions of collagen and proteoglycan in providing the compressive resilience of cartilage. The collagen fibres resist tension and provide the shape and form of the tissue. The proteoglycans (mainly aggrecan) are immobilized and create a large osmotic swelling pressure that extends the collagen network. The resulting tissue is tough, hard-wearing, and deforms elastically under load.

The osmotic pressure is caused by all the negatively charged anionic groups on aggrecan which carry with them mobile counter ions such as Na^+. This creates a large difference in the concentration of ions inside the cartilage compared with outside and also an imbalance amongst the freely diffusible anions and cations. Water is drawn into the tissue as a result of this osmotic imbalance and because aggrecan is too large and immobile to redistribute itself. The water thus swells and expands the aggrecan-rich environment. This places the collagen network under tension and an equilibrium is achieved when tension in the collagen network balances the swelling pressure, i.e. when no more water enters the tissue because the force is insufficient to stretch the collagen network any further. At this equilibrium, with the tissue swollen with water, it has good compressive resilience as any new load on the tissue now places the collagen network under further tension. Another feature of the composite collagen/aggrecan organization is now important, as not only is aggrecan greatly restricted in its ability to move within the matrix and the collagen/aggrecan network is stiff and resists deformation, but aggrecan offers great resistance to any fluid flow and redistribution of water. The tissue thus behaves largely as a stiff elastic polymer to sudden impact loading, but shows some slow inelastic deformation with sustained loads. However, removal of all loads leads to a redistribution of water and a return to the preloading equilibrium position. The articular cartilage thus forms a tough but compliant load-bearing surface and these characteristics depend on the integrity of the collagen network and the retention within it of a high concentration of aggrecan.

Collagens

The collagens are a family of secreted matrix proteins that contain elements of a unique triple-helical peptide structure with a repeating motif, with glycine as every third amino acid and a high frequency of proline, 40 to 50 per cent of which is hydroxylated. The presence of hydroxyproline is essential for the stability of the collagen triple helix and hydroxylation of proline is an early post-translational step in synthesis. The collagen gene family is now known to be quite large, but its members fall into two main groups, those that contain long, uninterrupted, triple-helical 'collagenous' domains that form the main fibrillar collagens, types I, II, III, V, and XI, and those that contain shorter, non-fibrillar-forming collagenous domains that form a variety of beaded filament and network structures (Vuorio and de Crombrugge 1990). A further subgroup has been identified as the fibril-associated collagens with interrupted triple-helical domains, such as type IX, XII, and XIV (Olsen 1995). All the fibrillar collagens pack into fibres with a characteristic overlap between adjacent triple helices that gives a staggered array, which facilitates the formation of a cohesive cross-linked fibre, and accounts for the 67 nm banding pattern visible in electron micrographs of collagen fibres.

The most common collagen in the body is the type I fibrillar collagen that occurs as the main structural element in bone, skin, ligaments, and tendons, often occurring together with type III collagen. In contrast, articular cartilage contains collagens of a more restricted distribution that are specifically associated with cartilage (see Table 1). The major collagen of cartilage is type II, which forms 80 to 90 per cent of the total content. It is a long-chain fibrillar collagen and forms the major fibre network of the tissue. Type II collagen differs from type I in containing three identical chains, whereas type I contains two a_1(I)-chains and one a_2(I)-chain. Type II also contains more hydroxylysine than type I and is generally more glycosylated, with a higher content of galactose–glucose disaccharides attached to the hydroxylysine residues. Type II collagen has a very limited distribution in other tissues; apart from in cartilage and the intervertebral discs, it only occurs in the vitreous chamber of the eye. It forms a long-chain, fibrillar, triple helix that is 285 nm long and has large N- and C-terminal propeptides that are removed enzymatically by specific N- and C-propeptidases prior to fibril formation. The fibres it forms are generally of a smaller diameter than those formed by type I collagen and they are cross-linked at specific segments of the N- and C-terminal regions. These bonds are of a stabilized Schiff base that forms a pyridinolone. It is therefore a permanent covalent linkage and mature collagen fibres cross-linked in this way can only be solubilized and resorbed by extensive proteolytic action to excise the cross-linked segments (Eyre and Wu 1995).

	Type	Molecular composition	Tissue distribution
Class I fibril forming	Type II	[α_1(II)$_3$]	Throughout tissue, major constituent
300 nm, triple helix	Type XI	[α_1(XI) α_2(XI) α_3(XI)]	Throughout tissue (possible core for type II fibrils)
Class II, short helix molecules	Type VI	[α_1(VI) α_2(VI) α_3(VI)]	Concentrated pericellularly
	Type IX	[α_1(IX) α_2(IX) α_3(IX)]	Linked to type II on the outside of fibrils, some pericellular enrichment
	Type X	[α_1(X)$_3$]	Deep calcified zone only (abundant in growth plate cartilage)

Table 1 Collagens found in articular cartilage (modified from Eyre *et al*. 1991)

Two more distinct collagens also occur in cartilage, these are types IX and XI (see Fig. 3). The latter is another long-chain fibrillar collagen comparable with type II. It can account for up to 10 per cent of the total collagen content of cartilage, but is typically about 3 per cent in adult articular cartilage. It is even more glycosylated than type II and is resistant to collagenase. There are two distinct type XI a-chains, but these show structural similarity to type V collagen a-chains, which are also expressed at a low level in cartilage and can be incorporated in hybrid type XI/V molecules. Type XI collagen in cartilage therefore contains a varying proportion of type V a-chains, which is reported to increase with development (Eyre and Wu 1995). Studies on the cross-links that join different peptides have shown that frequently type XI molecules are joined to each other, but some results show that linkages between type XI and type II chains are also present. Some of the collagen fibres in cartilage may therefore be cofibres of type II and type XI, and in these cases type XI may form the core elements, although analysis shows that some type XI collagen retains its N-propeptides in the tissue and this may be more compatible with a fibril surface role for these molecules.

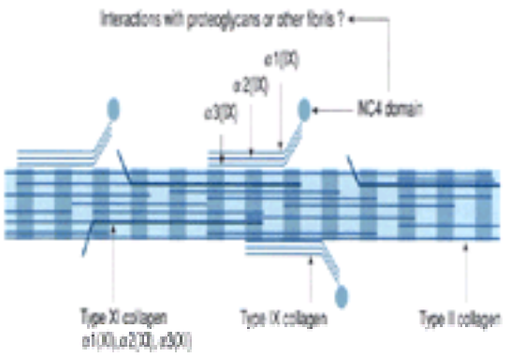


Fig. 3 Schematic showing collagen fibrils in cartilage formed from type II, type IX, and type XI collagen (from Jacenko and Olsen 1995).

Type IX collagen has a special structure with an interrupted collagenous sequence. It has three collagenous segments and four non-collagenous domains, and electron micrographs of the isolated molecule show it to be an elongated molecule with a kink towards one end. Type IX collagen contains three different a-chains and it has been shown that one of the a-chains (a_2 IX) has an attachment site for a chondroitin sulphate chain in the third non-collagenous domain. In fact the a_2-chain is slightly longer at this site than the a_1- or a_3-chains and this extension coincides with the kink in the electron micrograph image.

Localization of type IX collagen shows that it is found on the outside of type II fibrils and imaging has shown globular domains extending away from the surface of type II collagen fibrils at the end of short stalks, which would correspond with the fourth non-collagenous domain and the third collagenous segment lying at an angle with the axis of the first and second collagenous segments. Cross-links have been identified between the third non-collagenous domain and the N-terminal telopeptide region of type II collagen, and also between the second collagenous domain of type IX and the C-terminal region of type II collagen (Wu *et al*. 1992). The cross-linked fragments isolated show that type IX lies antiparallel to type II collagen in the fibres. Cross-links are also formed between type IX molecules. Type IX collagen therefore has the capacity to mediate covalent links between separate type II collagen fibrils. As stromelysin is able to cleave type IX collagen at the second non-collagenous domain and to cleave the telopeptides from type II collagen, it has the capacity to degrade these interfibrillar links (Eyre and Wu 1995). This would have major consequences on the collagen network and could provide a mechanism to facilitate matrix remodelling without great damage to major elements of the type II collagen fibrillar network. Linkage analysis has identified defects in the genes of the fibrillar cartilage collagens II, IX, and XI as being responsible for some inherited diseases affecting the joints, including early-onset osteoarthritis, Stickler's syndrome, and various chondrodysplasias (Jacenko and Olsen 1995).

As with type II collagen, type IX is also expressed in the eye. In studies on avian eyes, it has been shown that its expression used a different gene promoter from that used in cartilage and the protein expressed has an additional peptide within its sequence that enlarges the fourth non-collagenous domain (Nishimura *et al*. 1989). It has also been shown in chickens that the chondroitin sulphate chain synthesized on the third non-collagenous domain in the vitreous chamber is almost 10 times longer than that synthesized in cartilage (Yada *et al*. 1990). The structure of type IX is thus strongly influenced by the tissue in which it is being expressed.

Type VI collagen is found in a number of different tissues in the body. It contains three different a-chains and forms a short, triple-helical segment only 100 nm long. There are large, disulphide-bonded, non-helical extensions that are not removed as propeptides, and fibril formation appears to involve the packing of antiparallel molecules in tetrameric units that form microfibrillar structures lacking the characteristic 67 nm banding of the main fibriller collagens. Type VI collagen lacks the tight structure and extensive cross-linking of the fibrillar collagen and is easily extracted from the tissue in denaturing solvents. The distribution of type VI collagen in most tissues suggests a pericellular role, particularly associated with cell–matrix interactions. There is some evidence in cartilage that it is distributed principally in the pericellular region surrounding the chondrocyte, forming a lacuna bounded by a dense meshwork of fine collagen fibrils that forms a basket or chondron. Type VI collagen contains several peptide sequences (RGD) that form possible sites for binding to cell receptors. It also shows properties of interaction with hyaluronan. Type VI collagen may thus participate in cell–matrix organization in a number of different ways.

Type X collagen is a short-chain collagen consisting of three identical a-chains that form a triple-helical segment of 150 nm in length. It occurs primarily in calcifying cartilage produced by hypertrophic chondrocytes in the growth plate. It has been proposed that type X collagen may have some function associated with calcification, but transgenic mice deficient in type X collagen were found to develop normally and endochondral ossification was not greatly affected. It may therefore have more of a temporary structural role in the growth plate during the transition from cartilage to bone. It is also present in a limited amount in the calcified zone at the junction of articular cartilage with the bone and its expression in articular cartilage appears to be increased in osteoarthritis.

Proteoglycans

The major proteoglycan of cartilage is aggrecan, which is of high molecular mass. Aggrecan contains a large core protein (molecular mass: approximately 250 kDa) but most of its mass is provided by the polysaccharide chains attached to it (see Fig. 2). There are about 100 chondroitin sulphate chains (molecular mass: 10 to 25 kDa each) and up to 50 keratan sulphate chains (molecular mass: 15 kDa each) on each aggrecan molecule and also some *O*-linked oligosaccharides and a few *N*-linked oligosaccharides. Each aggrecan molecule thus has a total molecular mass of 1.5×10^6 to 2.5×10^6 kDa and contains about 10 per cent protein and 90 per cent carbohydrate. The carbohydrate chains are all synthesized on the aggrecan core protein during post-translational glycosylation within the chondrocyte prior to its secretion into the matrix (Hardingham 1986). Proteoglycans follow pathways of synthesis in common with other secreted glycoproteins and the site of major glycosaminoglycan synthesis lies in the vesicular compartments of the medial/trans-Golgi. The mechanisms of biosynthesis are highly efficient and the time taken for protein synthesis in the rough endoplasmic reticulum to glycosaminoglycan chain synthesis and secretion from the chondrocyte is about 20 to 30 min.

The protein core of aggrecan is a single polypeptide chain with several subdomains of different function (Fig. 4) (Hardingham *et al*. 1992; Hardingham and Fosang 1992). At the N-terminal of the protein core there are two globular domains, G1 and G2, separated by a short extended segment (21 nm long). The G1 domain provides the site for aggregation, and has specific properties of binding to hyaluronan and a separate globular link protein. Aggregates are composed of aggrecan molecules bound to hyaluronan via their G1 domains and each hyaluronan chain can have up to a hundred or more aggrecans bound to it. The interaction of link protein with the G1 domain stabilizes its binding to hyaluronan and essentially 'locks' the aggrecan into the aggregate form. The protein structures of the G1 domain and link protein are interesting as they are closely related to each other and contain two structural motifs, one is an immunoglobulin-fold and the other is a disulphide-bonded tandem-repeat structure.

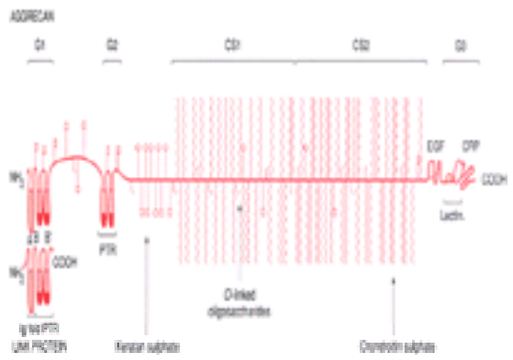


Fig. 4 The structure of the major cartilage proteoglycan, aggrecan. Aggrecan contains three disulphide-bonded globular domains, G1, G2, and G3, and has attached to it chondroitin sulphate, keratan sulphate, and *O*-, and *N*-linked oligosaccharides. Link protein is of similar structure to the G1 domain; both contain an immunoglobulin fold (Ig fold) and proteoglycan tandem repeat (PTR). Aggregation involves G1 binding to hyaluronan and this is stabilized by link protein, which binds to G1 and also to hyaluronan. G2 contains a PTR motif. G3 contains up to three protein motifs: an epidermal growth factor-like sequence (EGF), a mammalian type C lectin sequence (lectin), and a complement regulatory protein sequence (CRP) (see Hardingham and Fosang 1992).

The tandem-repeat structure forms the site for binding to hyaluronan and has specificity for a decasaccharide unit (hyaluronan HA10), whereas association between the G1 domain and link protein is by interaction of the immunoglobulin folds. The protein sequences of link protein are highly conserved amongst different species such that link protein from the cartilage of one species will bind and stabilize aggregate formation by aggrecan from another species. The formation of aggregates is an extracellular process. Aggrecan and link protein are synthesized and secreted together by chondrocytes, but hyaluronan is made separately at the chondrocyte cell surface. Aggregation can thus only occur after aggrecan has entered the matrix and it provides a mechanism for immobilizing it within the cartilage matrix.

The second globular domain also contains a sequence with homology to the tandem repeat found in the G1 domain and link protein, but it does not show any hyaluronan-binding activity (Fosang and Hardingham 1989). The main glycosaminoglycan attachment region of the proteoglycan is a long, extended protein sequence between the G2 domain and the C-terminal G3 domain. It is subdivided into three regions. Close to the G2 domain is a region rich in keratan sulphate which contains a repeating sequence that varies in number in different animal species. The remaining major chondroitin sulphate attachment sequence involves two further patterns of sequence repeats. The attachment of chondroitin sulphate at serine–glycine sequences is restricted in aggrecan to these particular regions. Keratan sulphate has a wider distribution than chondroitin sulphate as it also occurs in the chondroitin sulphate-rich region and close to the G1 and G2 domains, and at some sites chains are *N*-linked rather than *O*-linked.

The globular C-terminal G3 domain contains sequences related to other protein families. One part contains a sequence similar to a family of mammalian type C lectins, such as the asialoglycoprotein receptor. These bind oligosaccharide ligands on glycoproteins and the expressed recombinant G3 domain has some carbohydrate binding properties (Halberg *et al*. 1988). Two other sequences may be variably expressed in the human aggrecan gene: an epidermal growth factor-like repeat and a complement regulatory binding protein sequence (Baldwin *et al*. 1989). The incidence of these variations in aggrecan structure and their functional significance have yet to be determined. There are other proteoglycans structurally related to aggrecan in other tissues, which together form a family of hyaluronan-binding proteoglycans. These include versican in dermis, loose connective tissue, and smooth muscle, and neurocan and brevican in brain tissue. They all contain structurally related G1 and G3 domains although their glycosaminoglycan attachment regions are unrelated.

Aggrecan present in cartilage contains a heterogenous population of molecules differing in size and composition. This has two main causes: first, variations in glycosylation and chain synthesis during biosynthesis result in a polydisperse product with a considerable range of sizes, and second and more importantly, the aggrecan molecules in the cartilage matrix are subject to some proteolytic attack (Hardingham 1986). As the lifetime of each aggrecan in the matrix is quite long (months, or years in mature cartilage), the composition of molecules extracted at any one time represents a cross-section of the population from newly synthesized to the oldest molecules present. The proteolytic action results in a selective loss of C-terminal structures such as G3 and preferential retention of the N-terminal structures. This arises because the N-terminal G1 domain provides the binding to hyaluronan and the mechanism of anchoring aggrecan in the matrix, and it follows that cleavage of the molecule results in the shortened forms containing intact G1 domains being retained in the tissue, whereas the C-terminal fragments are able to diffuse out of the matrix and become lost from the tissue.

The average turnover time of aggrecan appears to become longer with age and its size distribution also becomes steadily biased towards a population containing more partially cleaved molecules. Even in young cartilage there appears to be only a proportion (perhaps 50 per cent) of aggrecan molecules with an intact G3 domain and this appears to be lower in mature tissue. There is also a marked accumulation in older tissue of fragments of low molecular weight comprising only the G1 domain.

In addition to aggrecan, articular cartilage contains two chondroitin sulphate/dermatan sulphate proteoglycans of low molecular weight, decorin and biglycan (Heinegard and Oldberg 1989; Hardingham and Venn 1993). These small proteoglycans account for only about 5 per cent of the total glycosaminoglycans in the tissue, but as they are of small size they are present in molar amounts comparable with aggrecan. In human cartilage decorin is reported to increase in abundance with age whereas biglycan decreases. Both decorin and biglycan are also broadly distributed in other connective tissues besides cartilage. They have similar-sized core proteins and are structurally related, but appear to have quite different properties and presumably different functions. Decorin and biglycan form part of a family of leucine-rich proteoglycans, which contain a common sequence repeat rich in leucine (Kobe and Deisenhofer 1994). Other members of this family include fibromodulin, which also occurs in cartilage, and lumican, a keratan sulphate proteoglycan, which is in the cornea. Whereas decorin and biglycan have chondroitin/dermatan sulphate chains attached towards the N-terminus (one chain for decorin, two for biglycan), fibromodulin has no chondroitin sulphate but has some sulphated tyrosines in the corresponding location, and also some molecules have keratan sulphate attached to *N*-linked oligosaccharide sites within the central leucine-rich repeats. Both decorin and fibromodulin have type II collagen-binding properties (Hedbom and Heinegard 1993) and *in vitro* they delay fibrillogenesis and cause the formation of thinner fibres. *In vivo* decorin has been localized at the d and e bands of type I collagen fibres, whereas fibromodulin binds to the a and c bands (Scott 1995) and these properties are functions of the protein cores. Biglycan in contrast does not have any collagen-binding properties and appears to be distributed close to a cell surface environment (Bianco *et al*. 1990). It may therefore be more important for cell–cell and cell–matrix interactions.

A further interesting property of decorin is its ability to bind transforming growth factor-b (Ruoslahti and Yamaguchi 1991). It may thus function as a matrix repository for this growth factor that might be released as a result of proteinase action fragmenting the decorin protein core. Transforming growth factor-b has also been shown to stimulate selectively the synthesis of decorin and biglycan and modulate the synthesis of the attached chondroitin sulphate/dermatan sulphate chains. The ability of decorin to bind transforming growth factor-b and thereby block this action may thus form part of a feedback control of decorin synthesis and control of other actions of this growth factor. The binding of transforming growth factor-b in the matrix could also provide a store of growth factor that might be released through the action of proteolytic enzymes that attack the matrix following inflammation or tissue injury.

Aggrecan turnover

Proteoglycans in the cartilage matrix are constantly turned over even in mature non-growing tissue. There is a constant slow rate of aggrecan degradation and loss, and its replacement by new synthesis. The mechanisms of turnover, involving biosynthesis by the chondrocytes and degradation in the extracellular matrix, must therefore be co-ordinated so that the tissue content of aggrecan is maintained at a constant level. The chondrocytes are responsible for controlling these events and appear to be sensitive to the aggrecan content of the matrix surrounding them, and some feedback mechanisms enables synthesis and degradation to be coregulated (Lohmander 1988; Morales and Hascall 1989; Caterson *et al*. 1990).

The normal turnover of aggrecan in healthy cartilage appears to involve their proteolytic cleavage in the region close to the G1 domain. This is a most important site of attack as it releases a large glycosaminoglycan-bearing fragment and separates it from its site for aggregation. This is thus an efficient mechanism for mobilizing aggrecan as it involves the minimum of enzyme action. It is important that turnover is an essentially conservative process as cleavage and release of only a small fraction of the tissue content is required at any one time so that the overall tissue content is conserved and the biomechanical properties are sustained.

Several studies of cartilage from different sources have thus established that a major product of normal turnover is a large aggrecan fragment that has lost its G1

domain and is thereby able to diffuse slowly out of the matrix. Investigation of the released fragments from *in vitro* experiments and those released from tissue *in vivo* into synovial fluid has identified cleavage within the interglobular domain at a single predominant site (Sandy *et al*. 1991; Lohmander *et al*. 1993), but attempts to identify the proteinase responsible from tissue extracts have been unsuccessful and in experiments *in vitro* most candidate enzymes cleave at other sites. Neutrophil collagenase, a metalloproteinase, is the only enzyme with any proven activity at this site, although it cleaves more readily at a site common to other metalloproteinases and is predominantly made by neutrophils (Fosang *et al*. 1994). However, some evidence suggests that it may be expressed together with other metalloproteinases by chondrocytes, but it remains to be shown if it is responsible for the action in the tissue. Matrix metalloproteinases are certainly involved at some stage of the degradative process, as specific metalloproteinase inhibitors are effective in reducing aggrecan loss from the cartilage stimulated by cytokines and retinoic acid (Buttle *et al*. 1993). Investigation of the actions of cytokines such as interleukin 1 or tumour necrosis factor-a have shown that they stimulate chondrocytes to degrade their matrix more rapidly and they also cause an inhibition of protein synthesis. They thus have a double action in depleting articular cartilage of its aggrecan content. In the presence of interleukin 1 or tumour necrosis factor-a, the aggrecan fragments released from cartilage are more extensively degraded than those released in normal turnover. The rate of aggrecan release is greatly increased and this results from a large increase in matrix proteolysis with cleavage at several additional sites within the chondroitin sulphate attachment region of aggrecan, which appears to be by the action of the same enzyme that cleaves within the interglobular domain (Ilic *et al*. 1992).

The neutral metalloproteinases are a family of enzymes that are likely to be involved in many aspects of extracellular matrix breakdown and turnover (see Table 2) (Woessner 1991; Murphy 1995). There are now more than 13 cloned enzymes, including collagenases, stromelysins, and gelatinases, which together have a range of complementary actions that enable them to attack and degrade all the major matrix macromolecules. The main matrix metalloproteinases are all secreted by cells in an inactive latent form that requires activation in the extracellular matrix. The process of activation involves a structural rearrangement and proteolytic cleavage with a reduction in molecular weight to give an active enzyme. Activation can be achieved *in vitro* with trypsin, which converts the proenzyme to active enzyme directly, or with APMA (aminophenyl mercuric acetate), which catalyses a protein rearrangement that leads to an autocatalytic cleavage and self-activation. The physiological mechanism of activation has not been fully established, but could involve one metalloproteinase activating others, as in the proposal that gelatinase is the major activator of procollagenase. Other types of proteinase such as plasminogen activator may also catalyse metalloproteinase activation. In addition to activation, control of enzyme activity is also provided by the natural enzyme inhibitors such as TIMP (tissue inhibitor of metalloproteinases), which binds to active enzyme and irreversibly inactivates it. There is normally an excess of TIMP in the extracellular matrix so that any activated enzyme is quickly inhibited. TIMP is produced by all cells that secrete the proenzymes. The extent of proteinase activity in the matrix is thus under tight control and can be regulated in several ways. First, the production of latent proenzymes is varied, second, the production and availability of various activators of the proenzymes can be varied, and third, the production of specific inhibitors can be varied. The metalloproteinases are also inhibited by the general proteinase inhibitors in serum, such as a $_2$-macroglobulin. There are thus mechanisms to prevent proteinases causing more widespread tissue damage should they escape from local control.

Enzyme	MMP family	Matrix substrates
Collagenase 1, 2, 3	MMP-1, -8, -13	Native collagen types I, II, III, VII, and X; also aggrecan and other non-collagenous proteins
Stromelysin 1	MMP-3	Aggrecan, link protein, fibronectin, laminin, gelatins, collagens III, IV, V, and IX, procollagen peptides
Gelatinases, 72 kDa, 92 kDa	MMP-2, -9	Gelatins, collagens IV, V, VII, and X, fibronectin, elastin
Range of activity		pH 5–9
Activation		Secreted latent enzymes require activation by other proteinases
Natural inhibitors		TIMP, TIMP-2, TIMP-3
Major plasma inhibitor		α_2-Macroglobulin
Chemical inhibitors		EDTA, 1,10 phenanthroline

TIMP, tissue inhibitor of metalloproteinase; EDTA, ethylenediamine tetraacetic acid.

Table 2 The neutral metalloproteinases

Growth plate

The cartilage forming the growth plate provides the site for the major longitudinal growth of the long bones of the skeleton (Iannotti 1990; Poole 1991). In macromolecular terms, it has all the attributes of cartilage, containing a dense network of predominantly type II collagen fibres and an aggrecan-rich matrix, but its cells are very different. The growth plate chondrocytes, which are produced by the continuous division of progenitor cells, progress through a rapid programme of differentiation, matrix expansion, hypertrophy, matrix calcification, and further differentiation, or cell death, which is quite unlike other chondrocytes. This process results in endochondral bone formation. As well as being responsible for long bone growth it is a process that also occurs in fracture healing and in osteophyte formation. It therefore has an important role in the growth of the skeleton and many inherited diseases affecting skeletal development are caused by mutations in genes that are active in the growth plate.

Cartilage formation begins during embryonic development as early as 6 weeks, with the condensation and differentiation of mesenchymal cells into chondrocytes to form the main elements of the skeleton. Soon after its formation, the central portion of each cartilage rudiment calcifies and is vascularized, which leads to the formation of a bony diaphysis capped at each end by a cartilaginous epiphysis. Later in development, a secondary centre of ossification develops in each epiphysis, which segregates the region of major cartilage growth from the cartilage that will be retained as the articular surface of the joint (Fig. 5). The transverse plate of cartilage which is sandwiched between the diaphysis and epiphysis remains as the site of growth throughout development, but finally calcifies and closes at skeletal maturity, typically between the ages of 14 and 18 years. The growth plate has a polarized and stratified structure (Fig. 5) that reflects the processes that occur within it. These primarily include a sequence of cell division, matrix expansion, cell hypertrophy, matrix calcification, vascularization, resorption, and replacement by bone and this continues throughout its life. It is thus a dynamic structure, which proliferates chondrocytes at its leading edge and deposits bone at its trailing edge.

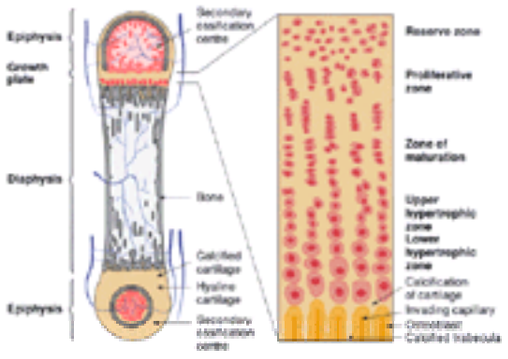


Fig. 5 The growth plate and cartilage development in a long bone at the late embryonic stage. The bone rudiment on the left shows the secondary centres of ossification that separate the growth plate cartilage from the articular cartilage. The expanded growth plate on the right shows the histologically distinct zones of the growth plate cartilage. Blood vessels penetrate the reserve zone of cartilage and the lower calcified zone, but the proliferative and hypertrophic zones are avascular. (Redrawn from Wallis 1993, with permission.)

The cartilage of the growth plate has three histologically and functionally distinct zones (Fig. 5) (Iannotti 1990; Poole 1991; Cancedda *et al*. 1995). At the leading edge of the growth plate is the reserve zone of cartilage. It contains a sparse population of chondrocytes that are only slowly dividing and they reside in a collagen-rich matrix. This is a stable non-dynamic structure that forms the interface between the growth plate and the bone of the epiphysis. Blood vessels penetrate the reserve zone, but do not enter the adjacent proliferative zone. The progenitor cells mark the boundary between the reserve zone and the proliferative zone. These are the stem cells that generate the continuous supply of cells for cartilage growth. Columns of chondrocytes extend away from the stem cells and they deposit and

expand the matrix surrounding them, thus providing the major mechanism to elongate the bone.

This highly active phase of chondrocyte maturation and matrix deposition is followed by cell hypertrophy. This is very evident histologically as the hypertrophic zone (Fig. 5), in which the cells become 5 to 10 times larger and the matrix is diminished prior to its calcification. The rate of cell division is again low in the hypertrophic zone and there is a much lower rate of synthesis of matrix macromolecules than in the proliferative zone. However, there is increased synthesis of some characteristic components of hypertrophic chondrocytes, such as alkaline phosphatase and type X collagen, and there is increased accumulation of intracellular calcium. Matrix vesicles, which are membrane-bound structures of 100 to 150 nm in diameter, are found in the interterritorial matrix of the hypertrophic zone. They are rich in calcium and their abundance increases towards the zone of calcification. It is undoubtedly the function of the chondrocytes in the hypertrophic zone to prepare the matrix for calcification, but the exact sequence of events and the mechanisms involved are a source of considerable debate. It is clear that hypertrophic chondrocytes accumulate calcium and produce matrix vesicles that are rich in calcium. They also produce alkaline phosphatase which can release inorganic phosphate, but what changes occur in the matrix to induce calcification are less clear. The unique expression of type X collagen in this region leads to the expectation that it may have a role in calcification; however, the development of transgenic mice that lack type X collagen but have apparently normal growth plate function suggested that it was not essential for calcification. The lowest calcified region of the hypertrophic zone forms the boundary for the penetration of blood vessels from the diaphyseal side of the growth plate and this forms an advancing front of cartilage matrix resorption and of new bone formation.

The fate of the chondrocytes in the calcified zone of the growth plate as the cartilage matrix is resorbed and bone takes its place is not completely clear (Cancedda *et al*. 1995). There is evidence of cell death amongst the hypertrophic chondrocytes in this region and in one model the resorption of the cartilage matrix and its replacement by bone could be caused by cells brought in by the blood vessels that penetrate into this region. However, there is also evidence that hypertrophic chondrocytes in culture can show some osteogenic potential and express matrix molecules, such as osteonectin, osteopontin, and switch from type II and type X collagen synthesis to type I, and deposit extracellular mineral. It therefore appears entirely feasible that some hypertrophic chondrocytes may survive and become osteoblasts and participate in the deposition of the growing bone. It may be that both these processes occur and that some hypertrophic chondrocytes undergo programmed cell death by apoptosis and are eliminated from the calcified zone as the matrix resorption and blood vessel invasion advances, whereas others may survive and proceed to participate in the generation of the new bone as osteoblasts. It may also be that the proportion of cells following either route varies from one site of endochondral ossification to another.

The chondrocytes in the growth plate are very sensitive to a range of growth factors including those of the insulin-like growth factor family, the transforming growth factor-b family (including bone morphogenic proteins), the fibroblast growth factor family, and vitamin D metabolites (Poole 1991; Cancedda *et al*. 1995). Some of these, such as the transforming growth factor-b family (and bone morphogenic proteins) are most active in promoting chondrocyte differentiation, whereas others, such as the fibroblast growth factors are most active in stimulating cell division, and these have little effect on fully differentiated cells. In contrast, the insulin-like growth factors are very effective in stimulating the synthesis of matrix macromolecules and also suppressing matrix degradation and they thus have a very positive anabolic effect. The effect of growth hormone on cartilage is mainly by stimulating the local production of insulin-like growth factors. The most rapid phase of growth is associated with sexual maturation at puberty and is driven by the hormonal changes and the accompanying increase in systemic growth factor production. Chondrocytes of the growth plate differ from other chondrocytes in that they are sensitive to vitamin D metabolites. These are essential for normal mineralization, and vitamin D deficiency, which causes rickets, is characterized by a slowing down of mineralization and an expansion of the hypertrophic zone of the growth plate. At the cellular level, the chondrocytes fail to accumulate calcium and there is poor production of matrix vesicles. The failure of calcification and bone formation causes the widening of the growth plate, which leads to instability and to deformed growth.

Investigation of many families that show inherited skeletal abnormalities, such as chondrodysplasias, has now identified a range of single gene defects that are responsible (Wallis 1993; Cancedda *et al*. 1995; Olsen 1995). Many of these are in the type II collagen gene in families with Kniest dysplasia, some forms of Stickler's syndrome, with spondyloepiphyseal dysplasia, and with hypochondrogenesis. Others have been identified in type IX collagen, type XI collagen, and in metaphyseal chondrodysplasia type Schmid in type X collagen. The sites of mutations involve substitutions, deletions, and frame shifts and each family group contains a different form. The range of phenotype is also very variable from mild to severe and there is no clear pattern of correlation between the site of the mutation, or its form, and the phenotype that results. Some forms of multiple epiphyseal dysplasia and pseudoachondrodysplasia have been linked to the cartilage oligomeric matrix protein gene; this produces a thrombospondin-like protein in cartilage matrix and its function is not yet understood. Some forms of hypochondrodysplasia are linked to polymorphisms in the insulin-like growth factor 1 gene. There are therefore a range of defects in genes that form part of the matrix or are involved in the processes that lead to the development of the matrix that can result in inherited skeletal abnormalities and with current techniques of molecular biology many more of these will be identified.

The basic process of chondrocyte differentiation, matrix expansion, matrix calcification, and its replacement by bone is not exclusive to the growth plate. A similar process occurs in fracture healing where the initial union in the fractured bone is made by chondrogenic development at the site of periosteal damage, which expands and forms the fracture callous, before it is eventually resorbed as cortical bone and healing becomes complete. The formation of osteophytes in synovial joints close to the articular surface also proceeds through an initial chondrophytic stage, often beginning at the site of the cartilage–bone junction. The differentiation of chondrocytes at this site forms a cartilage growth or chondrophyte, which then calcifies and becomes remodelled into a bony. It is only subsequently that the osteophyte fuses with the main cortical bone. Osteophytes therefore form not as direct outgrowths of the bone, but from separate, small, osteogenic centres that produce through a cartilage-mediated process, similar to that of the growth plate, a bony buttress joined with the main bone shaft.

Age-related change in articular cartilage

Age-related changes in articular cartilage are distinct from those characteristic of osteoarthritis, but give clues to the increasing susceptibility of cartilage to damage in old age. Age determines both the composition of the extracellular matrix as well as the distribution of chondrocytes and their response to external factors such as cytokines (see Table 3) (Hardingham and Bayliss 1990). Zonal changes in the distribution of chondrocytes are seen with ageing although the total number of chondrocytes shows little variation. Chondrocytes in the superficial cells mostly disappear with some increase in the cellular content of the deeper layers. Numerous changes in the extracellular matrix with ageing may compromise function. With old age there is a decrease in hydration of the matrix with a corresponding increase in compressive stiffness. This may have implications for the remarkable property of cartilage to undergo reversible deformation upon loading and may result in increased transmission of forces to subchondral structures. Loss of hydration is probably secondary to changes in the matrix components. Proteoglycan aggregation is adversely affected by proteolytic damage to link protein and a decrease in the available binding sites on hyaluronan; the latter is the result of binding of the free binding region domains that represent the limited digestion product of proteolysis of the glycosaminoglycan attachment region of core protein. Aggregation may also affect pore size distribution and solute permeability. Other changes observed include increasing proteoglycan heterogeneity with a general reduction in proteoglycan size accompanied by an increased ratio of keratan sulphate to chondroitin sulphate. Although the concentration of hyaluronan rises with age, this results from the gradual accumulation of partially degraded hyaluronan, rather than from a higher rate of synthesis. Age-related changes in the response of human cartilage to cytokines, especially to interleukin 1, have been documented with maturation, but little consistent change in response in old age has been noted. However, the reality is liable to be a complex interaction of endogenous anabolic and catabolic factors that are acting in articular cartilage in any age and considerable difficulties confront the investigator attempting to study their action *in vitro*.

Chondrocytes
Loss of most superficial chondrocytes, some increase in the number of deeper cells
Extracellular matrix
Decreased hydration with greater compressive stiffness of the matrix; increase in stable collagen cross-links
Proteoglycans
Aggrecan of smaller average size and higher keratan sulphate/chondroitin sulphate ratio, increased content of free hyaluronan binding region domain (G1); content of decorin may increase and biglycan decrease
Hyaluronan and link protein
Increased hyaluronan content, but hyaluronan chains of shorter length, increased proteolytic damage of link protein

Table 3 Age-related changes in articular cartilage following skeletal maturation

Cartilage response to cytokines and growth factors

Interleukin 1 (**IL-1**) and tumour necrosis factor-a (**TNF-a**) have generally been shown to have major effects on articular cartilage, including the inhibition of proteoglycan synthesis and the stimulation of matrix proteoglycan degradation (Fig. 6). Studies with human cartilage have not found a major effect on matrix degradation in most samples, which is in contrast to results obtained using immature bovine and other animal cartilage (Hardingham *et al*. 1992). Human cartilage studied *in vitro* has been found to respond in three ways: (i) no significant increase in the release of proteoglycan over several days in culture, (ii) a delayed response with increased release after about 6 days in culture, and (iii) acute activation of proteoglycan loss within 24 h. In the latter case, dramatic evidence of catabolism leading to 80 to 90 per cent depletion of proteoglycans within 3 to 4 days in culture has occurred in occasional specimens (about 5 per cent). However, the majority of cartilage explants show no significant increase in release. This probably corresponds to a low level of activation of the catabolic process, rather than no activation at all, because in human cartilage the dense collagenous network is probably less permeable to aggrecan fragments than in young animal cartilages and therefore requires more catabolism to give a significant increase in aggrecan release. Aggrecan biosynthesis is much more sensitive to cytokine action than aggrecan catabolism and the effects are much more consistent amongst cartilages from different animal sources. Recombinant human IL-1 demonstrates a dose-dependent reduction in proteoglycan synthesis in all animal and human cartilage samples tested; recombinant human TNF is also effective, but less potent. The sensitivity of human articular cartilage to IL-1 (and TNF-a), declines with age. The effects on aggrecan synthesis show that there is no lack of penetration of the cytokines to the chondrocytes in human cartilage, but that some steps in the response of the chondrocytes that lead to increased matrix degradation are less stimulated in uman cartilage than that from many animal sources. This might involve the production of less proenzyme, or of less activator of the proenzymes, or an increase in the inhibitors of the enzymes. Any of these effects may suppress the extent of proteolytic action and hence the rate of proteoglycan degradation and release from the matrix. It has been speculated that the effects of cytokines on human articular cartilage in joint disease may result primarily from shutting off aggrecan synthesis, rather than from any increase in matrix degradation.

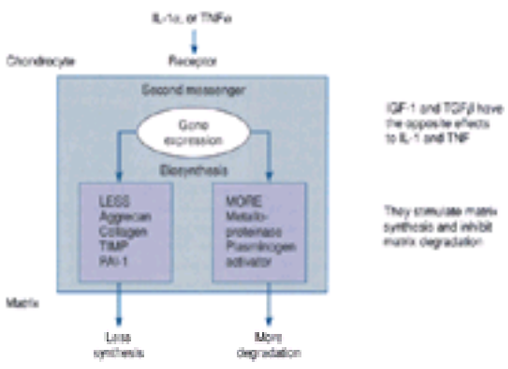


Fig. 6 Effects of cytokines and growth factors on articular chondrocytes.

Of the growth factors, insulin-like growth factor 1 is considered to be the major mediator of cartilage growth. As well as stimulating synthesis of proteoglycans, it may also reduce the rate of proteoglycan catabolism (Fig. 3), although human articular chondrocytes show a decrease in responsiveness to insulin-like growth factor 1 with increasing age. Other growth factors such as epidermal growth factor, fibroblast growth factor, and transforming growth factor beta (**TGF-b**) have been shown to potentiate the effect of insulin-like growth factor 1 on proteoglycan synthesis.

TGF-b has been shown to enhance the synthesis of small proteoglycans (decorin or biglycan) relative to the large aggregating proteoglycan. In explant cultures of cartilage from dogs that have undergone cruciate ligament transection, there was selectively more synthesis and a higher content of small proteoglycans, suggesting that TGF-b may be responsible in part for the increased activation of chondrocytes in experimental osteoarthritis (Venn *et al*. 1995).

In chronic rheumatoid arthritis TNF-a has been successfully identified as a therapeutic target (Elliot *et al*. 1994). Neutralizing, humanized, monoclonal antibody to TNF-a delivered to patients was effective in relieving clinical and laboratory parameters of disease activity, and the benefit was sustained over 1 to 3 months, after which there was a relapse. This suggested that TNF-a was an important cytokine in driving the chronic arthritic process and also that IL-1 production in the joint was largely caused by TNF-a, rather than vice versa. This form of treatment was also effective in animal models of arthritis in suppressing inflammatory cell infiltration into the joint and reducing joint swelling, but it is not yet clear how far it can protect from or reverse the damage to articular cartilage. IL-1 receptor antagonist is a natural inhibitor of IL-1 and forms part of a biological mechanism controlling IL-1 action. Blocking IL-1 action with injections of recombinant IL-1 receptor antagonist has also been investigated in experimental arthritis, but very high doses are required to achieve any effect (Lewthwaite *et al*. 1995). One way in which the action of TNF-a is controlled locally in tissues is by the release from the cell surface of the binding domains of TNF-a receptors. The released domains not only no longer function as receptors, but also act as soluble ligands that bind and inactivate the available TNF-a. The release of such receptors appears to be under the control of a cell surface metalloproteinase. The action of TNF-a and IL-1 on articular chondrocytes is thus likely to be controlled, not merely by their local abundance, but also by a network of interactions, involving changes in the expression of receptors, the release from cells of soluble receptors and IL-1 receptor antagonist, and also their actions are modulated by the competing effects of local and circulating growth factors.

Cartilage in experimental models of joint disease

The early events occurring in cartilage in the context of joint disease remain elusive in the human. In this regard, the availability of animal models of joint disease have proved useful in studying the sequence of response of cartilage to specific insults. Animal models of osteoarthritis have been the most extensively studied with respect to the effects on cartilage (for a comprehensive review see Pritzker 1994). The methods of inducing cartilage damage vary from surgical procedures such as section of the anterior cruciate ligament in the dog (the Pond–Nuki model) and partial meniscectomy in the rabbit, to chemical methods such as employing the metabolic poison iodoacetate. Genetic models of joint instability, such as the STR mouse are also used. Changes observed have to be seen in the context of the species used as well as the age and maturity of the animal, which remain important influences on cartilage behaviour and may account for apparently conflicting experimental results.

The Pond–Nuki model

In this model, the anterior cruciate ligament of the dog is sectioned causing instability and laxity in the operated joint, this alters the biomechanics of the joint and leads to the development of changes reminiscent of human osteoarthritis (Table 4) (McDevitt *et al*. 1977). These early events occur whilst the cartilage still appears macroscopically normal. Initial events after cruciate sectioning include increased hydration of the cartilage matrix, probably as a result of early disruption of the collagen network, seen especially in the superficial zones on electronmicroscopy. This is accompanied by a loss of tensile strength, which is detectable at 3 weeks postoperatively in the surface layers and progresses over 6 months to involve the deeper zones. Although there may be some evidence of cell death in the superficial cartilage, in general the chondrocytes show evidence of increased biosynthetic activity as part of a reparative process accompanied by some signs of cell division. The production of matrix components such as proteoglycan and collagen are increased, although because of increased turnover rates, the matrix content does not rise (Carney *et al*. 1992; Ratcliffe *et al*. 1992). However, under certain circumstances, it appears that the total proteoglycan content may rise and be accompanied by an increase in cartilage thickness (Brandt and Adams 1989). Although no changes in the ability of proteoglycans to aggregate has been detected, some minor changes in structure have been detected including lengthening of the chondroitin sulphate chains and the appearance of neoepitopes reflecting subtle changes in the pattern of sulphation (Caterson *et al*. 1990). Structural damage and fibrillation eventually occur and are mostly confined to the weight-bearing region of the tibial surface. Changes in articular cartilage are accompanied by alterations in subchondral bone including sclerosis and rapid osteophyte development. Initially appearing as ectopic, chondrogenic centres close to the cartilage–bone–synovial junction, these subsequently ossify before merging with the cortical matrix of the main bone shaft. Recent work examining the subchondral bone by computed tomographic microdensitometry has shown osteopenia, with an increase in the intertrabecular distance. It has been suggested that the characteristic osteosclerosis is in fact a relative appearance resulting from an absolute decrease in the underlying bony trabecular (Brandt 1991).

Increased hydration
Increased proteoglycan and collagen synthesis
Increased matrix turnover
Decrease in tensile strength of the superficial zone of cartilage
No depletion of proteoglycans or loss of aggregation
Altered proteoglycan structure with lengthened chondroitin sulphate chains and less keratan sulphate
Increased expression of chondroitin sulphate epitopes (3B3, 7D4)
Some mitotic activity of chondrocytes
Marginal osteophyte formation
Subchondral bone sclerosis
Capsular thickening

Table 4 Experimental osteoarthritis (Pond–Nuki): changes in the joint during the first 3 months following transection of the anterior cruciate ligament in the mature dog

There is some controversy surrounding the long-term fate of this model. Studies following animals for up to 2 years suggested that little progression occurred beyond the changes seen in the first few months following surgery. However, in longer-term experiments, it has been shown that cartilage may be maintained for up to 3 years before progressive loss occurs over the subsequent 9 months (Brandt *et al*. 1991).

Manipulation of the experimental method has made it possible to try and dissect out some of the multiple influences affecting the outcome of the reparative/degradative process following injury in articular cartilage and the results underline the fact that multiple factors determine response (Brandt 1991).

Differences in chondrocyte biosynthesis between control and operated joints seen in explant tissue disappear over the course of a few days (Venn *et al*. 1995). This suggests the action of local factors such as the effect of cytokines or growth factors, and the effect of altered biomechanical forces, which would not be present in culture. When the operated limb is immobilized in flexion following surgery, there is no stimulation of matrix production, appearance of fibrillation, or osteophyte formation. In fact there is a reduction of matrix synthesis similar to that seen with immobilization of normal joints. On the other hand, disruption of the normal proprioceptive control of joint forces, as happens with dogs who have undergone unilateral dorsal root ganglionectomy prior to cruciate transection, leads to rapid progression of cartilage breakdown. This occurred within weeks rather than the years required in neurologically intact animals. Some degree of synovial inflammation may be consistently seen in the Pond–Nuki model, although it appears that this is related to bleeding occurring perioperatively. When the synovitis is controlled by meticulous attention to haemastasis, little difference in the cartilage response is observed between those animals with synovitis and those without. Other external influences such as systemic medication may affect the response of cartilage to injury. Systemic salicylate therapy can accelerate the appearance of cartilage breakdown and may be due in part to inhibition of proteoglycan synthesis, which is independent of the effect on cyclo-oxygenase. These studies serve to underline that the response of cartilage to injury is complex and that outcome can be very heterogeneous depending on which external factors come into play.

Meniscectomy models

It is well known in humans that meniscectomy can precede the appearance of osteoarthritis in the affected limb after a variable period of time, usually many years. Similarly, meniscectomy performed in animal models, most commonly the rabbit, produces changes in articular cartilage that become more severe the more extensive the surgery. Meniscectomy alone induces a progressive but mild lesion, with eventually late compensation. However, section of the fibular collateral and sesamoid ligaments in addition to partial lateral meniscectomy leads to a lesion that progresses rapidly over 3 months, characterized by marked fibrillation, ulceration, increased mitosis with some cell death causing the formation of chondrocyte clusters, and loss of proteoglycan from the cartilage matrix (Colombo *et al*. 1983). Changes seen at the biochemical level with these models are similar to those seen with cruciate ligament transection, with increased hydration of the matrix, increased matrix turnover, and remodelling of bone.

Cartilage lesions resembling osteoarthritis can be induced by surgery to the hind knees of guinea pigs (Meacock *et al*. 1990). The rate of progression varies with the surgical procedure. Partial medial meniscectomy produces rapidly evolving lesions that are clearly evident after 1 week and are often severe by 3 weeks. Initial lesions occur on the medial tibial plateau and comprise focal loss of surface chondrocytes, with loss of proteoglycans. The cell and proteoglycan loss becomes progressively more extensive and is accompanied by surface fibrillation of the cartilage. By seven weeks, osteophyte formation can be observed and, by 20 weeks, involvement of the lateral tibial plateau is evident. Occasional subchondral cysts occur by 25 weeks. Evolution is slower following section of the lateral with or without the medial ligament and resembles the progression of spontaneous lesions (see below). Initial lesions, however, occur earlier than in unoperated animals (6 to 10 weeks in contrast with 12 weeks).

Chemical models

A single intra-articular injection of iodoacetate induces rapid changes in the articular cartilage of the affected joint that progresses over a matter of weeks. These changes include chondrocyte necrosis, proteoglycan depletion, cartilage fibrillation and subsequent loss down to bone, and the formation of osteophytes. As in the dog model, experiments that change the nature of the biomechanical forces acting across the joint affect the rate and extent of damage, with exercise exacerbating the lesions and immobilization having a protective effect, although this is reversed and leads to atrophy if immobilization is prolonged over an extended period (Williams and Brandt 1984).

Genetic models

There are several models of spontaneously occurring osteoarthritis in animals including various strains of mice (C57-black, STR-IN, and STR-ORT), Hartley albino guinea pigs, and hip dysplasia in beagle dogs. The STR mouse is particularly useful for experimental work as there is a 100 per cent incidence in males after 5 months of age, with females somewhat less affected. Changes seen in the joint are those of a mild generalized osteoarthritis, most prominent in the knee joints due to a distal shift of the patella leading to altered joint congruity (Walton and Elves 1979). The disease course follows the stages known from human pathology, with roughening of the cartilage surface, radial and horizontal cartilage fissures, and erosions down to bone accompanied by subchondral bone remodelling (Fig. 2), but the cellular and molecular changes they result from have not been characterized. Spontaneous cartilage degeneration develops in the femoral–tibial joint in male Hartley guinea pigs (Bendele *et al*. 1989). Lesions develop initially on the medial tibial plateau between 60 and 90 days, and by 1 year, bilateral fibrillation over 50 per cent of the medial tibial articular surface is consistently seen. The shoulder and elbow joints are affected less frequently. The early lesions of focal chondrocyte death and matrix degeneration become increasingly severe with age and lesions in some 9 or 12 month animals resemble those present in human osteoarthritis. Animals that are diet restricted, with a 29 per cent decrease in body weight, have a 56 per cent reduction in the severity of lesions, suggesting that body mass in guinea pigs, as in humans, is an important predisposing factor in the development of spontaneous osteoarthritis of the knee.

Markers of cartilage damage in joint disease

Recent advances in the understanding of the biochemistry and physiology of articular cartilage together with the advent of monoclonal antibodies have rendered it possible to measure matrix components in blood, synovial fluid, and urine. In the near future, it should be possible to define the role of measurement of these components as markers of cartilage metabolism in health and disease. To date, various epitopes have been identified in the extracellular matrix using both polyclonal and monoclonal antibodies or separation techniques such as high-pressure liquid chromatography.

Proteoglycan epitopes

Keratan sulphate

Unlike other carbohydrate constituents of proteoglycans, keratan sulphate is restricted to cartilage apart from a small amount found in the cornea. The monoclonal antibody 5D4, which recognizes a highly sulphated epitope present only at the non-reducing end of long keratan sulphate chains, forms the basis of a well-characterized assay (Thonar *et al*. 1992). Serum levels show no diurnal variation and, after skeletal maturity has been achieved, do not vary with age. With the injection of papain into a single joint or chymopapain into an intravertebral disc, the acute catabolism of cartilage is mirrored by a transient two- to eightfold rise in levels of keratan sulphate epitope. A good correlation exists between levels of keratan sulphate epitope and other proteoglycan epitopes found on the core protein.

Serum levels of keratan sulphate epitope have been found to be raised in patients with osteoarthritis and are stable with time, although the large overlap seen in the normal population does not recommend it as a diagnostic test (Sharif *et al*. 1995a). However, levels in synovial fluid are about 10 times serum levels and vary inversely with cartilage mass as measured radiographically (Campion *et al*. 1991).

Chondroitin sulphate epitopes

Recent studies have suggested that early changes in osteoarthritis may be accompanied by subtle changes in the structure of chondroitin sulphate chains. The monoclonal antibodies used recognize specific patterns of sulphation within the chains which are rare in normal cartilage, but which are expressed in osteoarthritic cartilage (Caterson *et al*. 1990). These chondroitin sulphate epitopes are also found in embryonic tissue and may represent the expression of a more primitive type of molecule during the attempted reparative processes. The abundance of these epitopes was investigated in patients following knee injury with cruciate ligament rupture or major meniscal damage (Hazell *et al*. 1995). Analysis of synovial fluids from the trauma joint compared with the contralateral uninjured joint showed that trauma caused an increase in the expression of chondroitin sulphate epitopes, which was greatest in the first few months after injury. In contrast there was a fall in the content of keratan sulphate epitopes. The aggrecan fragments released from cartilage in the injured joint were therefore of altered composition and contained chondroitin sulphate chains of higher epitope content and less keratan sulphate. This change may correlate with increased matrix synthesis and may form part of a reparative/remodelling response by the chondrocytes. The combination of a measure of chondroitin sulphate epitopes in synovial fluid to assess new aggrecan synthesis, and of keratan sulphate epitopes to reflect matrix aggrecan degradation, may provide a useful index of cartilage status during the early stages of joint disease. Studies of early lesions of spontaneous osteoarthritis in the guinea pig suggest that changes in the ratio of 6-sulphation and 4-sulphation of the disaccharides of chondroitin sulphate may be an important early event in osteoarthritis (Osborne *et al*. 1994). The initiation of lesions coincides with a decrease in chondroitin-4-sulphate to below 20 to 25 per cent. Non-load-bearing cartilage and cartilage that does not develop osteoarthritic changes have a relatively high content of the chondroitin-4-sulphate, which does not fall below 35 per cent. It is proposed that chondroitin-4-sulphate may have a more stabilizing influence on cartilage matrix during loading than chondroitin-6-sulphate.

Proteoglycan protein epitopes

Polyclonal antibodies that recognize protein epitopes on proteoglycans released into synovial fluid have been used to study the long-term progress of patients following acute knee injury. It has been shown that acute trauma gives rise to a release of cartilage proteoglycan fragments into the synovial fluid and that the level of the rise is related to the severity of the trauma in the initial phase (Lohmander *et al*. 1989). Levels were still significantly raised more than 1 year after injury compared with controls. Knee pain in the absence of acute trauma or abnormal arthroscopic appearance was not associated with increased levels and suggests a lack of cartilage involvement in these cases.

Studies in patients with rheumatoid arthritis have shown that higher synovial fluid levels of proteoglycan were found in patients with a poorer prognosis at the time when the cartilage height was still normal (Saxne *et al*. 1987; Saxne and Heinegard 1992). In addition, when samples were taken at times of increased disease activity, levels were remarkably constant and suggest a consistent response of cartilage to insult in the individual patient.

Hyaluronan

Various assays are available to measure levels of hyaluronan in tissue fluids. There is evidence relating levels of hyaluronan to disease activity and ongoing joint damage in rheumatoid arthritis (Paimela *et al*. 1991). Increased levels of serum hyaluronan have been found to predict progression of knee osteoarthritis (Sharif *et al*. 1995b). In this study, 94 patients with knee osteoarthritis were followed for 5 years. At entry, hyaluronan levels were significantly related to disease duration, minimum joint space, and previous surgery. Moreover, those with progressive joint disease had significantly higher baseline hyaluronan levels than those with stable disease.

Collagen

The potential exists to monitor the synthesis and degradation of type II collagen found in cartilage by measuring levels of type II procollagen or the urinary excretion of pyridinoline cross-links, respectively. Levels of pyridinoline cross-links have been found to be elevated in patients with rheumatoid arthritis and osteoarthritis, with higher levels in patients with rheumatoid arthritis and a correlation with inflammatory disease activity (Seibel *et al*. 1989). Carboxy-terminal type II collagen, a marker of collagen type II synthesis, is elevated over control levels in osteoarthritic synovial fluid and increases with disease progression. In addition, the levels correlate with body mass index as a measure of obesity in osteoarthritis. In traumatic arthritis synovial fluid, levels of carboxy-terminal type II collagen correlate with the severity of the cartilage lesion as assessed by arthroscopy (Shinmei *et al*. 1995).

Non-collagenous matrix molecules

More has become known about the structure and function of various non-collagenous matrix molecules, which are probably primarily involved in cell attachment and collagen fibre modulation (Heinegard and Oldberg 1989). Although many, such as fibronectin, are widely distributed throughout the connective tissue of the body, others, such as cartilage oligomeric protein and cartilage matrix protein, are more restricted to cartilage and could therefore be potential markers of altered cartilage metabolism. Cartilage matrix protein is not found in articular cartilage and levels could give an indication of the systemic component to raised levels of cartilage-derived molecules that are less specific in their distribution. Recently, a relationship has been found between serum levels of cartilage oligomeric protein and progression of knee osteoarthritis (Sharif *et al*. 1995a). Progression of disease was defined by a 2 mm or greater reduction in joint space or the necessity of knee surgery during a 5-year follow-up. During the first year, serum levels of cartilage oligomeric protein were significantly elevated in 20 of 57 patients who progressed compared with those that did not. Other studies monitoring bone sialoprotein in synovial fluid have suggested that its increased release in rheumatoid arthritis reflects the extent of tissue destruction (Saxne *et al*. 1995).

Patterns of change

The ability of individual markers to detect and monitor disease is likely to limited. More helpful is a number of assays used in combination that may reflect the molecular nature of the lesion and the sum of anabolic and catabolic events (Saxne and Heinegard 1995) (see Table 5).

Tissue component	Epitope assayed degradation	Remodelling/repair
Type II collagen	Pyridinoline	C-propeptide
Aggrecan	Keratan sulphate	3B3, 846 CS epitopes
Matrix protein	Cartilage oligomeric protein	
Proteases	Stromelysin/collagenase	TIMP
Cytokines	IL-1, TNF-α	TGF-β

IL-1, interleukin 1; TNF-α, tumour necrosis factor-α; CS, chondroitin sulphate.

Table 5 Potential markers for monitoring the health of articular cartilage

This approach has been adopted in a study of patients with rheumatoid arthritis with destructive hip disease (Mansson *et al*. 1995). Eighteen patients were studied from a prospective cohort of 180 with early rheumatoid arthritis. Nine of these were patients who had developed early erosive hip disease and nine were age- and sex-matched patients with arthritis that was more slowly progressive. Increased levels of cartilage oligomeric protein were found in all patients with progressive disease, whereas levels of the chondroitin sulfate epitope 846 were increased only in those with milder disease. The levels of type II collagen C-propeptide, a marker of collagen II synthesis, were elevated in both groups. These results suggest that elevated serum levels of cartilage oligomeric protein may indicate an adverse prognosis, whereas elevated 846 epitope may be more favourable. The elevation of type II collagen propeptide in both groups is indicative of a selective synthesis of

collagen compared with aggrecan, which may have implications for the quality of the repair process.

Further work is required before a panel of meaningful cartilage function tests are available. Interpretation of levels depends on the epitope recognized and the characteristics of the assay involved, an understanding of the kinetics and clearance of the epitope and its possible alteration by disease, and the effect of external factors on concentration in body fluids, such as the exclusion volume effect of hyaluronan in synovial fluid. The relevance of levels of the epitope then needs to be evaluated in long-term, prospective, clinical trials stratified for confounding variables such as drug therapy, age, disease duration, and disease severity. With this caveat aside, it should soon prove possible to develop a series of assays that accurately reflect the health of articular cartilage and its response to disease and therapy.

Chapter References

Arakoski, J., Kivaranta, I., Jurvelin, J., Tammi, M., and Helminen, H. (1993). Long distance running causes site dependent decrease of cartilage glycosaminoglycan content in the knee joint of beagle dogs. *Arthritis and Rheumatism*, **36**, 1451–9.

Baldwin, C.T., Reginato, A.M., and Prockop, D.J. (1989). A new epidermal growth factor-like domain in the human core protein for the large cartilage specific proteoglycan. *Journal of Biological Chemistry*, **264**, 15747–50.

Bendele, A.M., White, S.L., and Hulman, J.F. (1989). Osteoarthrosis in guinea pigs: histopathologic and scanning electron microscopic features. *Laboratory Animal Science*, **39**, 115–21.

Bianco, P., Fisher, L.W., Young, M.F., Termine, T.D., and Robey, P.G. (1990). Expression and localization of the two small proteoglycans decorin and biglycan in developing human skeletal and non skeletal tissues. *Journal of Histochemistry and Cytochemistry*, **38**, 1549–63.

Brandt, K.D. (1991). Animal models: insights into osteoarthritis (OA) provided by the cruciate-deficient dog. *British Journal of Rheumatology*, **30** (Suppl. 1), 5–9.

Brandt, K.D. and Adams, M.E. (1989). Exuberant repair of articular cartilage damage. Effect of anterior cruciate ligament transection in the dog. *Transactions of the Orthopedic Research Society*, **14**, 584.

Brandt, K.D., Myers, S.L., Burr, D., and Albrecht, M. (1991). Osteoarthritis changes in canine articular cartilage subchondral bone and synovium fifty-four months after transection of the anterior cruciate ligament. *Arthritis and Rheumatism*, **34**, 1560–70.

Buckwalter, J.A. (1995). Osteoarthritis and articular cartilage use, disuse and abuse: experimental studies. *Journal of Rheumatology*, **22** (Suppl. 43), 13–15.

Buttle, D.J., Handley, C.J., Ilic, M.Z., Saklatvala, J., Murata, M., and Barrett, A.J. (1993). Inhibition of cartilage proteoglycan release by a specific inactivator of cathepsin B and an inhibitor of matrix metalloproteinases. *Arthritis and Rheumatism*, **36**, 1709–17.

Campion, G.V., McCrae, F., Schnitzer, T.J., Lenz, M.E., Dieppe, P.A., and Thonar, E.J.M.A. (1991). Levels of keratan sulfate in the serum and synovial fluid of patients with osteoarthritis of the knee. *Arthritis and Rheumatism*, **34**, 1254–9.

Cancedda, R., Cancedda, F.D., and Castagnola, P. (1995). Chondrocyte differentiation. *International Review of Cytology*, **159**, 265–359.

Carney, S.L. *et al*. (1992). Changes in proteoglycan turnover in experimental canine osteoarthritis. *Matrix*, **12**, 137–47.

Caterson, B. *et al*. (1990). Modulation of native chondroitin sulphate structure in tissue development and in disease. *Journal of Cell Science*, **97**, 411–17.

Colombo, C. *et al*. (1983). A new model of osteoarthritis in rabbits. *Arthritis and Rheumatism*, **26**, 875–86.

Elliot, M.J. *et al*. (1994). Randomised double-blind comparison of chimaeric monoclonal antibody to tumour necrosis factor (cA2) versus placebo in rheumatoid arthritis. *Lancet*, **344**, 1105–10.

Eyre, D.R., and Wu, J.-J. (1995). Collagen structure and cartilage matrix integrity. *Journal of Rheumatology*, **22** (Suppl. 43), 82–5.

Eyre, D.R., Wu, J.J., Woods, P.E., and Weis, M.A. (1991). The cartilage collagens and joint degeneration. *British Journal of Rheumatology*, **30** (Suppl. 1), 10–15.

Fosang, A.J. and Hardingham, T. (1989). Isolation of the N-terminal globular domains from cartilage proteoglycans. *Biochemistry Journal*, **261**, 801–9.

Fosang, A.J. *et al*. (1994). Neutrophil collagenase (MMP-8) cleaves at the aggrecanase site E_{373}-A_{374} in the integlobular domain of cartilage aggrecans. *Biochemistry Journal*, **304**, 347–51.

Halberg, D.F., Proulx, G., Doege, K., Yamada, Y., and Drickamer, K. (1988). A segment of the cartilage proteoglycan core protein has lectin-like activity. *Journal of Biological Chemistry*, **263**, 9486–90.

Hardingham, T. (1986). Structure and biosynthesis of proteoglycans. *Rheumatology*, **10**, 143–83.

Hardingham, T. (1990). Degenerative joint disease. In *The metabolic and molecular basis of acquired disease* (ed. R.D. Cohen, B. Lewis, K.G. Albert, and A.M. Denman), pp. 1851–69. Ballière Tindall, London.

Hardingham, T. and Bayliss, M. (1990). Proteoglycans of articular cartilage: changes in aging and in joint disease. *Seminars in Arthritis and Rheumatism*, **20**, 12–33.

Hardingham, T. and Fosang, A.J. (1992). Proteoglycans: many forms, many functions. *FASEB Journal*, **6**, 861–70.

Hardingham, T.E. and Venn, G. (1993). Chondroitin sulphate/dermatan sulphate proteoglycans from cartilage: aggrecan, decorin and biglycan. In. *Dermatan sulphate proteoglycans*: *chemistry, biology, chemical pathology* (ed. J.E. Scott), pp. 207–17. Portland Press, London.

Hardingham, T.E., Fosang, A.J., and Dudhia, J. (1991). Aggrecan, the chondroitin sulphate/keratan sulphate proteoglycan from cartilage. In *Articular cartilage and osteoarthritis* (ed. K.E. Keuttner, R. Schleyerbach, J.G. Peyron, and V.C. Hascall), pp. 5–19. Raven Press, New York.

Hardingham, T.E., Bayliss, M.T., Rayan, V., and Noble, D.P. (1992). Effects of growth factors and cytokines on proteoglycan turnover in articular cartilage. *British Journal of Rheumatology*, **31** (Suppl. 1), 1–6.

Hazell, P.K., Dent, C., Fairclough, J.A., Bayliss, M.T., and Hardingham, T.E. (1995). Changes in glycosaminoglycan epitope levels in knee joint fluid following injury. *Arthritis and Rheumatism*, **38**, 953–9.

Hedbom, E. and Heinegard, D. (1993). Binding of fibromodulin and decorin to separate sites on fibrillar collagens. *Journal of Biological Chemistry*, **268**, 27307–12.

Heinegard, D. and Oldberg, A. (1989). Structure and biology of cartilage and bone matrix noncollagenous macromolecules. *FASEB Journal*, **3**, 2042–51.

Iannotti, J.P. (1990). Bone disease. *Orthopedic Clinics of North America*, **21**, 1–17.

Ilic, M.Z., Handley, C.J., Robinson, H.C., and Mok, M.T. (1992). Mechanism of catabolism of aggrecan by articular cartilage. *Archives of Biochemistry and Biophysics*, **284**, 115–22.

Jacenko, O. and Olsen, B.R. (1995). Transgenic mouse models in studies of skeletal disorders. *Journal of Rheumatology*, **22** (Suppl. 43), 39–41.

Kiviranta, I., Tammi, M., Jurvelin, J., Arakoski, J., Saamanen, A.-M., and Helminen, H.J. (1992). Articular cartilage thickness and glycosaminoglycan distribution in the canine knee joint after strenuous running exercise. *Clinical Orthopedics*, **283**, 302–5.

Kobe, B. and Deisenhofer, J. (1994). The leucine-rich repeat, a versitile binding motif. *Trends in Biological Sciences*, **19**, 415–21.

Lewthwaite, J. *et al*. (1995). The Role of TNFa in the induction of antigen-induced arthritis in the rabbit and the anti-arthritic effect of species-specific TNFa neutralising monoclonal antibodies. *Annals of the Rheumatic Diseases*, **54**, 366–74.

Lohmander, L.S. (1988). Proteoglycans of joint cartilage. Structure, function, turnover and markers of joint disease. *Baillière's Clinical Rheumatology*, **2**, 37–62.

Lohmander, L.S., Dahlberg, L., Ryd, L., and Heinegard, D. (1989). Increased levels of proteoglycan fragments in knee joint fluid after injury. *Arthritis and Rheumatism*, **32**, 1434–42.

Lohmander, L.S., Neame, P.J., and Sandy, J.D. (1993). The structure of aggrecan fragments in human synovial fluid: evidence that aggrecanase mediates cartilage degradation in inflammatory joint

disease, joint injury and osteoarthritis. *Arthritis and Rheumatism*, **36**, 1214–22.

Mansson, B., Carey, D., Alini, M., Ionescu, M., Rosenberg, L.C., and Poole, R. (1995). Cartilage and bone metabolism in rheumatoid arthritis. *Journal of Clinical Investigation*, **95**, 1071–7.

McDevitt, C.A., Gilbertson, E., and Muir, H. (1977). An experimental model of osteoarthritis; early morphological and biochemical changes. *Journal of Bone and Joint Surgery*, **57B**, 24–35.

Meacock, S.C., Bodmer, J.L., and Billingham, M.E.J. (1990). Experimental osteoarthritis in guinea-pigs. *Journal of Experimental Pathology*, **71**, 279–93.

Morales, T.I. and Hascall, V.C. (1989). Factors involved in the regulation of proteoglycan metabolism in articular cartilage. *Arthritis and Rheumatism*, **32**, 1197–201.

Mow, V.C. and Lai, W.M. (1980). Recent developments in synovial fluid biomechanics. *Society for Industrial and Applied Mathematics Review*, **22**, 275–317.

Muir, H. and Hardingham, T.E. (1986). Biochemistry of connective tissue. In *Copeman's textbook of rheumatic diseases* (ed. J.T. Scott), pp. 177–98. Churchill Livingstone, Edinburgh.

Murphy, G. (1995). Matrix metalloproteinases and their inhibitors. *Acta Orthopaedica Scandinavica*, **66** (Suppl. 26), 55–60.

Nishimura, I., Muragaki, Y., and Olsen, B.R. (1989). Tissue-specific forms of type IX collagen-proteoglycan arise from the use of two widely separated promotors. *Journal of Biological Chemistry*, **264**, 20033–41.

Olsen, B.R. (1995). New insights into the function of the collagens from genetic analysis. *Current Biology*, **7**, 720–7.

Osborne, D., Woodhouse, S., and Meacock, R. (1994). Early changes in the sulfation of chondroitin in guinea-pig articular cartilage, a possible predictor of osteoarthritis. *Osteoarthritis and Cartilage*, **2**, 215–23.

Paimela, L., Heiskanen, A., Kurki, P., Helve, T., and Leirisalo-Repo, M. (1991). Serum hyaluronate level as an predictor of radiologic progression in early rheumatoid arthritis. *Arthritis and Rheumatism*, **34**, 815–21.

Poole, A.R. (1991). Cartilage assembly and mineralization. In *Cartilage: molecular aspects* (ed. B. Hall and S. Newman), pp. 179–211. CRC Press, Boca Raton, FL.

Pritzker, K.P.H. (1994). Animal models for osteoarthritis: processes, problems and prospects. *Annals of the Rheumatic Diseases*, **53**, 406–20.

Ratcliffe, A., Billingham, M.E.J., Saed Nejad, F., Muir, H., and Hardingham, T.E. (1992). Increased release of matrix components from articular cartilage in experimental canine osteoarthritis. *Journal of Orthopedic Research*, **10**, 350–8.

Ruoslahti, E. and Yamaguchi, Y. (1991). Proteoglycans as modulators of growth factor activities. *Cell*, **64**, 867–9.

Sandy, J.D., Neame, P.J., Boynton, R.E., and Flannery, C.R. (1991). Catabolism of aggrecan in cartilage explants: identification of a major cleavage site within the interglobular domain. *Journal of Biological Chemistry*, **266**, 8683–5.

Saxne, T. and Heinegard, D. (1992). Synovial fluid analysis of two groups of proteoglycan epitopes distinguishes early and late cartilage lesions. *Arthritis and Rheumatism*, **35**, 385–90.

Saxne, T. and Heinegard, D. (1995). Matrix proteins: potentials as body fluid markers of changes in the metabolism of cartilage and bone in arthritis. *Journal of Rheumatology*, **22** (Suppl. 43) 71–4.

Saxne, T., Wollheim, F., Pettersson, H., and Heinegard, D. (1987). Proteoglycan concentration in synovial fluid: predictor of future cartilage destruction in rheumatoid arthritis? *British Medical Journal*, **295**, 1447–8.

Saxne, T., Zunino, L., and Heinegard, D. (1995). Increased release of bone sialoprotein into synovial fluid reflects tissue destruction in rheumatoid arthritis. *Arthritis and Rheumatism*, **38** 82–90.

Scott, J.E. (1995). Extracellular matrix, supramolecular organisation and shape. *Journal of Anatomy*, **187**, 259–69.

Seibel, M.J., Duncan, A., and Robins, S.P. (1989). Urinary hydroxy-pyridinium crosslinks provide indices of cartilage and bone involvement in arthritic diseases. *Journal of Rheumatology*, **16**, 964–70.

Sharif, M. *et al*. (1995*a*). Relationship between serum cartilage oligomeric matrix protein levels and disease progression in osteoarthritis of the knee joint. *British Journal of Rheumatology*, **34**, 306–10.

Sharif, M. *et al*. (1995*b*). Serum hyaluronic acid level as a predictor of disease progression in osteoarthritis of the knee. *Arthritis and Rheumatism*, **38**, 760–7.

Shinmei, M., Kobayashi, T., Yoshihara, Y., and Atsuyoshi, S. (1995). Significance of the levels of carboxy terminal type II procollagen peptide, chondroitin sulfate isomers, tissue inhibitor of metalloproteinases and metalloproteinases in osteoarthritis joint fluid. *Journal of Rheumatology*, **22** (Suppl. 43), 78–81.

Stockwell, R.A. and Meachim, G. (1979). The chondrocytes. In *Adult articular cartilage* (ed. M.A.R. Freeman), pp. 115–22. Pitman Medical, London.

Thonar, E.J.M.A. (1992). Serum keratan sulfate: a measure of cartilage proteoglycan metabolism. In *Articular cartilage and osteoarthritis* (ed. E.K. Kuettner, R. Schleyerbach, J.G. Peyron, and V.C. Hascall), pp. 429–44. Raven Press, New York.

Venn, G., Billingham, M.E.J., and Hardingham, T.E. (1995). Increased proteoglycan synthesis in cartilage in experimental canine osteoarthritis does not reflect a permanent change in chondrocyte phenotype. *Arthritis and Rheumatism*, **38**, 525–31.

Vuorio, E. and de Crombrugge, B. (1990). The family of collagenases. *Annual Review of Biochemistry*, **59**, 837–72.

Wallis, G.A. (1993). Here today, bone tomorrow. *Current Biology*, **3**, 687–9.

Walton, M. and Elves, M.W. (1979). Zone thickening in osteoarthrosis, observations of an osteoarthrosis prone strain of mouse. *Acta Orthopaedica Scandinavica*, **50**, 501–6.

Williams, J.M. and Brandt, K.D. (1984). Exercise increases osteophyte formation and diminishes fibrillation following chemically induced articular cartilage injury. *Journal of Anatomy*, **139**, 559–71.

Woessner, J.F. (1991). Matrix metalloproteinases and their inhibitors in connective tissue remodelling. *FASEB Journal*, **5**, 2145–54.

Wu, J.-J., Woods, P.E., and Eyre, D.R. (1992). Identification of the cross-linking sites in bovine cartilage type IX collagen reveals an anti-parallel type II–type IX molecule relationship and type IX–type XI bonding. *Journal of Biological Chemistry*, **267**, 23007–14.

Yada, T., Suzuki, S., Kobayashi, K., Hoshino, T., Horie, K., and Kimata, K. (1990). Occurrence in chick embryo vitreous humour of a type IX collagen proteoglycan with an extraordinarily large chondroitin sulfate chain and a short al polypeptide. *Journal of Biological Chemistry* **265**, 6992–9.

2.4 Bone in health and disease

Roger Smith

Bone is a metabolically active tissue that is formed, removed, and replaced throughout life (Evered and Harnett 1988; Martin *et al*. 1988; Tam *et al*. 1989; Manolagas and Jilka 1995). These processes depend on bone cells, whose activities are determined by many factors. They include genetic, mechanical, nutritional, and hormonal influences and a host of short-acting cellular messages, including growth factors, collectively termed cytokines (Russell 1990). This chapter deals with the normal physiology of bone (Vaughan 1981; Smith 1984; Avioli and Krane 1990; Coe and Favus 1991; Noda 1993; Brighton *et al*. 1994; Smith 1995), and outlines the changes in specific skeletal disorders.

Bone in health

Structure

Bone consists of cells and an extracellular mineralized matrix (35 per cent organic and 65 per cent inorganic). About 90 per cent of the organic component is type I collagen. The remainder includes many non-collagen products of the osteoblast, such as osteocalcin, osteonectin, and proteoglycans (Robey 1994; see below). The mineral is present mainly as a complex mixture of calcium and phosphate in the form of hydroxyapatite.

Two anatomical types of bone may be defined, trabecular (cancellous) and cortical. The proportion of these in bones normally differs: for instance the vertebral bodies are predominantly trabecular and the shafts of the long bones are cortical.

This distribution is relevant both to the functions of the bone and to skeletal disorders such as osteoporosis. In trabecular bone there are more metabolically active surfaces in a given volume than in cortical bone. The basic multicellular units (see below) act on the surfaces of trabecular bone and through resorbing channels (cutting cones) in cortical bones. The fine structure of bone is described by Boyde (1972).

Bone is often thought to be inert because of its structural rigidity and persistence after death, and also to be composed entirely of chalk because it contains 99 per cent of the body's calcium. Although both assumptions are superficially reasonable, neither is correct.

Bone cells

Conventional histological sections of bone demonstrate three types of bone cell, which are clearly different (Fig. 1 and Fig. 2): osteoblasts, which may be plump and apparently active, or flat and apparently inactive (otherwise called bone-lining cells); multinucleated osteoclasts, which most often occupy areas of resorption; and osteocytes within their lacunas in the mineralized bone, apparently in contact with other osteocytes and bone cells through their extensions in the canaliculi (Fig. 2 and Fig. 3). All these cells may be seen together in one microscopic field. They are often in close contact with cells of the bone marrow, which contains their precursors and also brings them into close relationship with the immune system. These cells and the mineralized organic matrix have almost infinite possibilities for exchange of information. It is the complexity of bone that provides both the challenge and the fascination for those interested in its disorders. Histological techniques have been developed to study the sequential events enacted by teams of bone cells in so-called basic multicellular units (Eriksen *et al*. 1989); the techniques of cell biology are used to study the origin and functions of different types of cells and the communications between them (Martin *et al*. 1988). These approaches provide different information: the histology tells us about the temporal events in bone as a tissue whereas the cell biology provides clues about the activities of the different cells, albeit in artificial culture systems.

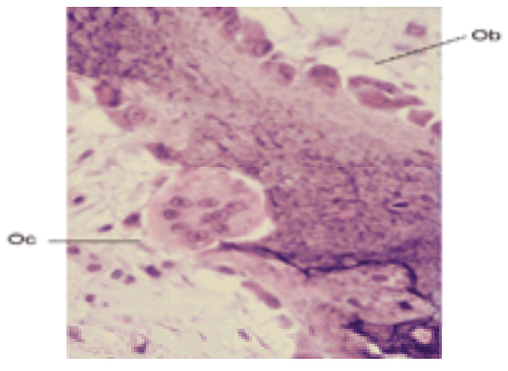


Fig. 1 A multinucleated osteoclast (centre: arrow, Oc) is present in a Howship's resorption lacuna along one edge of a bone trabecula; a row of plump mononuclear osteoblasts lie along the opposite edge (right: arrows, Ob). Haematoxylin and eosin, magnification × 4000.

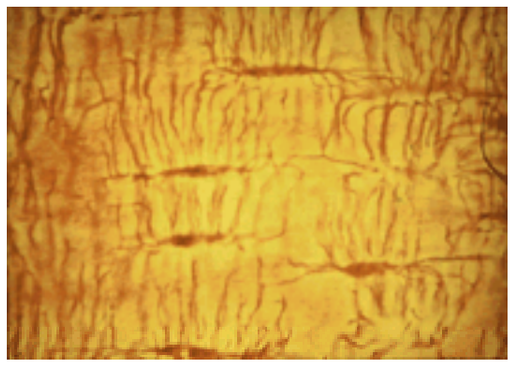


Fig. 2 A ground bone section showing osteocytes with numerous canaliculi, magnification × 400.

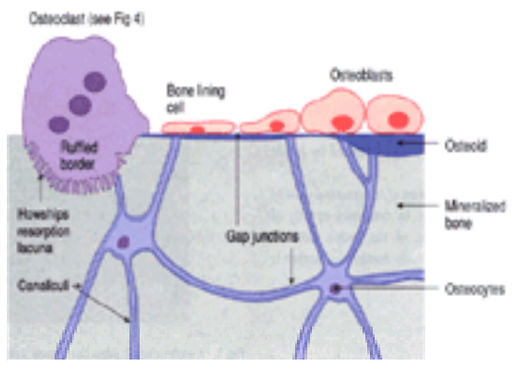


Fig. 3 A diagram of bone structure to show the relationship of different cell types.

The cells of bone occupy a central position in its physiology. Osteoblasts are one particular cell form derived from the mesenchymal stromal cell system; fibroblasts and adipocytes are others (Nijweide *et al*. 1988; Owen and Friedenstein 1988; Manolagas and Jilka 1995). Osteoclasts are haemopoietic in origin. The marrow contains precursor cells for the stromal and haematopoietic system; currently these are referred to as fibroblast colony forming units and granulocyte-macrophage colony forming units, respectively (see below). The osteocytes which are imbedded in the mineralized bone are derived from osteoblasts, as are also the bone-lining cells. All bone cells communicate with each other and control bone modelling during growth and remodelling throughout life. The constant processes of osteoclastic bone resorption and osteoblastic bone formation are closely linked and take place in basic multicellular units. The cellular cycle of such a unit (Fig. 4) begins with activation of multinucleate osteoclasts from their macrophage-like mononuclear precursors; osteoclasts produce resorption (Howship's) lacunas (Fig. 1) on the surface of trabecular bone, or cutting cones in cortical bone (Eriksen *et al*. 1989). These are similar processes. In cancellous (trabecular) bone, the basic multicellular unit may be looked upon as a cortical multicellular unit sectioned through the middle. Resorption is followed by a reversal phase, during which a cement line is deposited, and then the formation by osteoblasts of new bone matrix, which is subsequently mineralized. In the young adult, when the bone mass is constant, the amount of new bone formed equals that resorbed; in childhood more bone is formed than resorbed; and in later years there is an imbalance between these two processes in favour of resorption, leading to osteoporosis. The timescale of the remodelling cycle is approximately established, although estimates differ (Fig. 4).

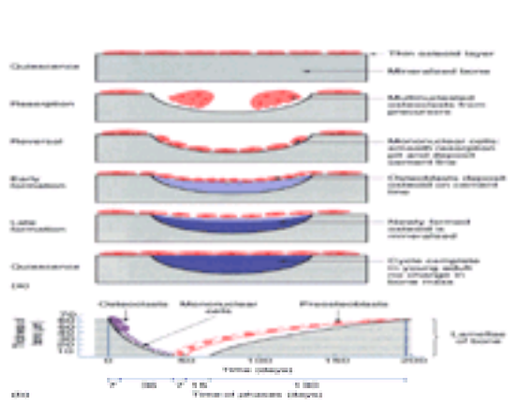


Fig. 4 (a) The remodelling events on the surface of bone during a bone multicellular-unit sequence (modified from Martin *et al*. (1988)). (b) The time sequences of these events from 20 young normal individuals (modified from Eriksen *et al*. (1989)).

Eriksen *et al*. (1989) give details of histomorphometrical findings in normal subjects and in those with metabolic bone disease. They point out that bone is one of the few tissues in which a time marker may be incorporated, and that histomorphometry can separate alteration in cellular activity from changes in cell number. The turnover of bone at the tissue level is determined by the activation frequency of basic multicellular units and the functional rates of individual cells. The mechanism of bone loss is different in different disorders. There are many unsolved mysteries about basic multicellular units. One is what factors lead to activation of the osteoclasts to initiate the resorbing cycle. Another is how cells communicate with each other (see below). However, it is becoming clear that many activities of the osteoclast depend on those of the osteoblast, which is a dominant cell in the skeletal scene.

Osteoblasts, osteocytes, and bone-lining cells

Osteoblasts (Fig. 1) have many important functions (Fig. 5). It is possible that the functions are divided between them. They respond to endocrine factors, both systemic and local (cytokines), and to mechanical stress. They synthesize the organic bone matrix (mainly collagen) and non-collagen proteins, and they control bone mineralization. Importantly, they also appear to direct the activity of other cell types, particularly the osteoclasts. In this respect they may also activate the bone-resorbing cycle.

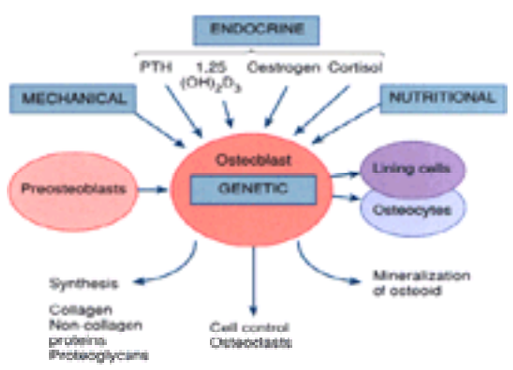


Fig. 5 The central position of the osteoblast in bone physiology. Broad arrows show the origin of osteoblasts from preosteoblasts, themselves derived from stromal

cell (fibroblast colony forming unit) lineage, and of the lining cells and osteocytes.

Osteocytes (Fig. 2) occupy lacunas within the mineralized bone. They communicate with each other by gap junctions via processes within the canaliculi and probably have an important function in the detection of, and response to, mechanical forces within mineralized bone. Bone-lining cells, a form of flattened, inactive osteoblast, cover endosteal surfaces and bone trabeculas, being separated from each other by gap junctions. They may also isolate bone fluid (if this exists) from the general extracellular-fluid compartment.

Osteoclasts

These multinucleated cells are derived from precursors in the haemopoietic system (granulocyte–macrophage colony forming units), but their immediate cell of origin is debated (Chambers 1988). They resorb bone by attaching themselves to its surfaces and forming a seal to isolate their area of activity (Baron *et al*. 1993) (Fig. 6). Within this sealed zone they produce a very acid environment, with the aid of a proton pump linked to the enzyme carbonic anhydrase II, within which digestion of whole bone by lysosomal enzymes occurs. Absence of carbonic anhydrase II is linked to a rare form of osteopetrosis (see below). Osteoclasts have receptors to calcitonin, which directly suppresses their activity, and also possibly to oestrogens (Glass and Broom 1993), but not, so far as is known, to any other hormone. However, they do respond to prostaglandins. The resorptive effects of parathyroid hormone and of 1,25-dihydroxycholecalciferol are probably mediated through the osteoblast.

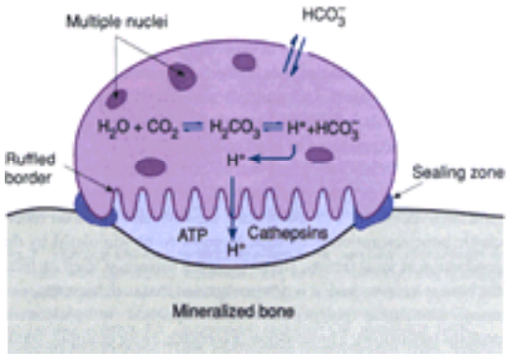


Fig. 6 A diagram of the main features and functions of the osteoclast.

Bone formation

The factors that control bone formation are complex and not fully understood, but must work largely through the osteoblast. Osteoprogenitor cells, the precursors of osteoblasts, are found in the periosteum and the endosteal surfaces close to the bone marrow. The local remodelling stimulus for new bone formation comes from some aspect of bone resorption, which could, for instance, be polypeptide growth factors or morphogenetic proteins liberated by this process (see below).

The activities of bone cells are influenced by cytokines. A cytokine may be defined as a peptide, produced by a cell, that acts as an autocrine, paracrine, or endocrine mediator. In this sense the definition includes a large number of substances of which many can be shown to have effects on bone and cartilage metabolism. Such effects have been demonstrated principally in experimental (and artificial) conditions and their physiological role is largely unknown. Many cytokines have alternative names and multiple actions, with synergism and antagonism; Table 1 provides examples of cytokines identified in bone matrix and produced by bone cells.

Effective on bone	In bone matrix/by bone cells
Interleukins (ILs)	IL-1, IL-6, IL-8
Tumour necrosis factors (TNFs)	TNF-α (cachectin)
	TNF-β (lymphotoxin)
Interferons	
Growth factors	Platelet-derived growth factors
	Insulin-like growth factors (somatomedins)
	Fibroblast growth factors
	Transforming growth factor-α and -β
	Epidermal growth factor
	Bone morphogenetic proteins (see Table 2)
Haemopoietic and colony stimulating factors (CSFs)	
Granulocyte (GCSF)	
Macrophage (MCSF)	MCSF
Granulocyte–macrophage (GMCSF)	GMCSF
Others	Parathyroid hormone-related peptides

Table 1 Cytokines, growth factors, and other mediators with effects on bone and those demonstrated in bone matrix and/or produced by bone cells

The local factors that control osteoblastic activity are probably of considerable importance (Canalis *et al*. 1989). Only some of these have been identified. The osteoblastic activity that follows on osteoclastic bone resorption is thought by some to be stimulated by the production of local factors, either from the osteoclast itself or from the bone it has resorbed. It is now recognized that such bone contains many stimulating polypeptides and important morphogenetic proteins (see below). The situation is complex, as for example, transforming growth factor-b appears to belong to a family of multifunctional regulatory peptides, and bone is probably its most abundant source. Not only do osteoblasts synthesize this factor but they also have high-affinity receptors for it, and are mitogenically stimulated by it. Growth factors may be defined as polypeptides with mitogenic activity, although they also affect the function of differentiated cells. They are produced by cells and, therefore, they may also be included under the general heading of cytokines. Important growth factors in bone matrix that have originated from cells within bone (Table 1) include platelet-derived growth factors, fibroblast growth factors, insulin-like growth factors (or somatomedins), transforming growth factors, and other osteoinductive factors (Canalis *et al*. 1989; Manolagas and Jilka 1995).

Bone resorption

Osteoclasts are controlled by systemic and local hormones but there is no direct evidence that they are influenced by mechanical stress. Calcitonin directly inhibits the osteoclast, temporarily abolishes the active, ruffled border, and suppresses the generation of new osteoclasts. Bone resorption is increased by parathyroid hormone and 1,25-dihydroxycholecalciferol. As the osteoclast does not contain receptors to either of these hormones it is proposed that the resorbing activity is mediated via the osteoblast. Again, the messages that the osteoblasts use to turn on the resorbing activity of the osteoclast are largely unknown. Amongst them are the prostaglandins. The number and activity of the osteoclasts is increased by a variety of cytokines produced by lymphocytes or monocytes (lymphokines and monokines, respectively) and by peptide growth factors (epidermal growth factor, transforming growth factors). They are also stimulated by a group of factors previously called osteoclast-activating factors (see below) (Table 2).

Systemic hormones	Cytokines
Activation	
Parathyroid hormone	Epidermal growth factor
1,25$(OH)_2$ vitamin D_3	Transforming growth factors-α and -β
Thyroxine	Prostaglandins
	Interleukin 3
Parathyroid hormone-related peptide	Granulocyte–macrophage colony-stimulating factor (and other colony-stimulating factors)
	Osteoclast-activating factors
Inhibition	
Calcitonin	Interferon-γ
	Transforming growth factor-β
	Interleukin 6

Table 2 Factors affecting bone resorption by direct or indirect effects on osteoclasts

Communication between bone cells

Many interactions between bone cells can be shown in experimental systems, but their relative importance *in vivo* is unknown.

Figure 7 attempts to bring together those influences on the osteoclast and osteoblast (and between them) that contribute to bone turnover. In addition to illustrating those factors already discussed, it shows the likely importance of colony-stimulating factors identified within the haemopoietic system, and the direct effect of the osteoblast in controlling the local production of collagenase in preparation for osteoclastic bone resorption. Manolagas and Jilka (1995) have recently updated this subject. They emphasize the importance of the close association of the haematopoietic system in the marrow and the bony tissue which surrounds it. The cells of both systems share progenitors, produce and respond to some of the same cytokines and colony-stimulating factors, and are critical for each other's function. The proposed relationship between these cells and the effect upon them of various cytokines is indicated in Fig. 8. It is suggested (Manolagas and Jilka 1995) that the differentiation pathway of osteoclasts is controlled by stromal osteoblastic cells utilizing cytokines such as interleukin 6 and interleukin 1; and further that the production of these and other cytokines by the osteoblast is stimulated also by interleukins, and additionally by systemic hormones such as 1,25-dihydroxycholecalciferol and parathyroid hormone.

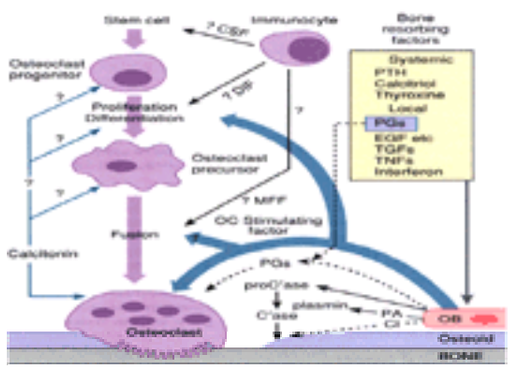


Fig. 7 Humoral and local factors regulating bone resorption (from Martin *et al*. (1988) with permission). Note how many of the arrows still have question marks. Stimulation represented by continuous arrows, inhibition by interrupted arrows. C'ase, collagenase; pro C'ase, procollagenase; CI, collagenase inhibitor; PA, plasminogen activator; OB, osteoblast; EGF, epidermal growth factor; TGFs, transforming growth factors; TNFs, tumour necrosis factors; MFF, macrophage fusion factor; DIF, differentiation-inducing factor; CSF, colony-stimulating factor.

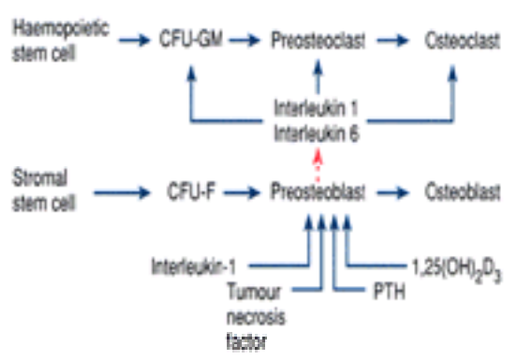


Fig. 8 The likely relationship between bone cells, their precursors, and cytokines. For CFU-GM (granulocyte–macrophage colony forming units) and CFU-F (fibroblast colony forming units) see text (modified from Manolagas and Jilka (1995)).

Bone mass

The eventual size and density of the skeleton and its development during the early years of life is influenced by important genetic factors. This genetic background is modified by mechanical stress, nutrition, the systemic effects of hormones, and by local factors produced by the bone cells themselves (Fig. 9). The relative contributions of these may vary with age (Fig. 10). Recent investigations suggest that the rate of bone loss as well as peak bone mass may be genetically determined, that adequate calcium intake conserves bone mass at all ages, and that mechanical loading of the skeleton is also important at any age.

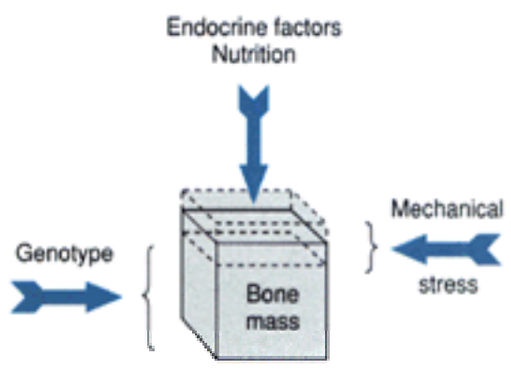


Fig. 9 A diagram of the main factors which influence bone mass (from Heaney (1986) with permission).

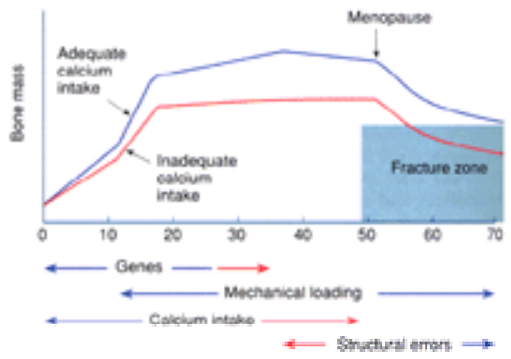


Fig. 10 The relative importance of factors contributing to bone mass with age (from Riggs and Melton (1988) with permission). The postmenopausal rate of bone loss is also influenced by genetic factors and calcium intake.

Genetic factors

Pocock *et al*. (1987) have confirmed the heritability of bone mass in monozygotic twins and Seeman *et al*. (1989) provide evidence that the bone mass of daughters of osteoporotic women is less than that of daughters of non-osteoporotic women. Recent Australian observations on mono- and dizygotic twins and postmenopausal women have shown that bone mass in the populations studied is partly related to particular polymorphisms in the vitamin D receptor gene (Morrison *et al*. 1994), although this has not been confirmed elsewhere (Hustmyer *et al*. 1994). In a disorder such a osteogenesis imperfecta, where there is clear evidence of gene mutations, the heritable component of bone mass is very important (see below). Work on this syndrome has illuminated important genetic mechanisms that account for apparently non-Mendelian forms of inheritance, particularly somatic and germ-line mosaicism (Hall 1988; Sykes 1990; Bernards and Gusella 1994).

Nutrition

It seems common sense that both the size and density of the skeleton should be related to nutrition, particularly of calcium, protein and energy, but this has been difficult to prove (Kanis and Passmore 1989; Nordin and Heaney 1990). Recent twin studies have demonstrated a significantly greater increase in radial and vertebral bone density in those prepubertal twins on calcium supplements (Johnston *et al*. 1992) and current evidence supports the notion that additional calcium is good for the skeleton at all ages (Heaney 1993).

Mechanical factors

The main function of the skeleton is mechanical and it has long been known that within the skeleton bone is laid down along the lines of stress. Thus, at any age, stress-induced changes in bone mineral density are site-specific (Heinonen *et al*. 1993). *In vitro* experiments show that osteoblasts in culture respond to mechanical stress by an increase in cAMP and phosphoinositol, partly mediated by prostaglandins (Sandy *et al*. 1989). Since the osteocyte may be the primary sensor of mechanical loading, attempts are being made in vivo to identify the candidate gene in this cellular response (Mason *et al*. 1995).

Systemic hormones

Sex hormones, testosterone and oestrogen, encourage new bone formation. Growth hormone is an important anabolic skeletal agent during the early years of life, both directly and through the local production of insulin-like growth factors. A number of hormones primarily thought of as resorptive may also have anabolic actions on the osteoblasts. These include parathyroid hormone, which has a long-term anabolic effect increasing the proliferation of osteoblast precursors, and 1,25-dihydroxycholecalciferol.

Bone matrix

The osteoblast synthesizes a large variety of substances that together form the organic bone matrix. These may be divided into collagen and non-collagen proteins, and the proteoglycans (Robey 1989; Robey 1994).

Collagen

Collagen is the major extracellular protein in the body and more than half is within the skeleton (Prockop and Kivirikko 1984). There are many different molecular types, with different functions and different genes (Burgeson 1988; Van der Rest and Garrone 1991; Hulmes 1992; Byers 1993; and Chapter 2.2) (Table 3). The main fibrillar collagens are type I, II, and III; type V and XI are minor fibrillar collagens; type IX, XII, and XIV collagens are associated with fibrils. Collagen in bone is type I. This is a heteropolymer composed of two a_1-chains and one a_2-chain. The general structure of an a chain is $(GlyXY)_{338}$. The a-chains are synthesized as precursors within the osteoblasts and undergo a number of synthetic steps that include posttranslational hydroxylation of proline and lysine residues; certain hydroxylysine residues are further modified into aldehydes and also glycosylated. After removal of their extensions the triple helical molecules form an exact structure with a quarter-stagger overlap, which is subsequently cross-linked (Fig. 11). In type I collagen the so-called hole zones within this structure provide a template for early mineralization. The arrangement of the collagen molecules differ according to their tissue function (Hulmes 1992). Mutations in the collagen genes and defects in posttranslational modification cause many inherited disorders of connective tissue, of which osteogenesis imperfecta is the main example. Recent work has demonstrated the causal importance of mutations in other collagens particularly those which cause chondrodysplasias (type II, IX, and X; see below).

Type	α-Chains	Most common molecular form	Tissue distribution
I	α1(I), α2(I)	$[\alpha 1(I)]_2\,\alpha 2(I)$	Most connective tissues, e.g. bone, tendon, skin, lung, cornea, sclera, vascular system
II	α1(II)	$[\alpha 1(II)]_3$	Cartilage, vitreous humour, embryonic cornea
III	α1(III)	$[\alpha 1(III)]_3$	Extensible connective tissues, e.g. skin, lung, vascular system
IV	α1(IV), α2(IV), α3(IV), α4(IV), α5(IV)	$[\alpha 1(IV)]_2\,\alpha 2(IV)$	Basement membranes
V	α1(V), α2(V), α3(V)	$[\alpha 1(V)]_2\,\alpha 2(V)$	Tissues containing collagen I, quantitatively minor component
VI	α1(VI), α2(VI), α3(VI)	$\alpha 1(VI)\,\alpha 2(VI)\,\alpha 3(VI)$	Most connective tissues, including cartilage
VII	α1(VII), α2(VII)	$[\alpha 1(VII)]_3$ $\alpha 2(VII)$?	Basement-membrane-associated anchoring fibrils
VIII	α1(VIII), α2(VIII)	$[\alpha 1(VIII)]_2\,\alpha 2(VIII)$?	Product of endothelial and various tumour cell lines
IX	α1(IX), α2(IX), α3(IX)	$\alpha 1(IX)\,\alpha 2(IX)\,\alpha 3(IX)$	Tissues containing collagen II, quantitatively minor component
X	α1(X)	$[\alpha 1(X)]_3$	Hypertrophic zone of cartilage
XI	α1(XI), α2(XI), α3(XI)*	$\alpha 1(XI)\,\alpha 2(XI)\,\alpha 3(XI)$	Tissues containing collagen II, quantitatively minor component
XII	α1(XII)	$[\alpha 1(XII)]_3$	Tissues containing collagen I, quantitatively minor component
XIII	α1(XIII)	$[\alpha 1(XIII)]_3$?	Quantitatively minor collagen, found e.g. in skin, intestine
XIV	α1(XIV)	$[\alpha 1(XIV)]_3$?	Tissues containing collagen I, quantitatively minor component

*Closely related to α1(I)

Table 3 The vertebrate collagens

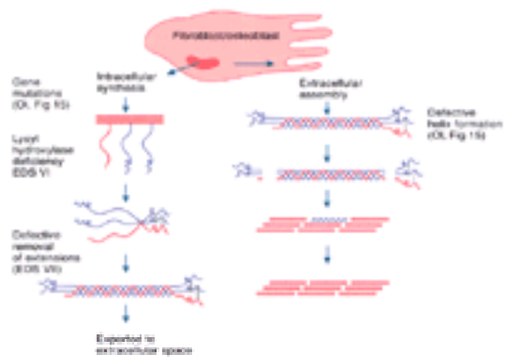


Fig. 11 Diagrammatic representation of the main synthetic pathways of collagen and the effects of some mutations. Type I collagen mutations which cause osteogenesis imperfecta are mirrored by similar mutations in type II collagen causing chondrodysplasias, and defects in minor collagens (type IX and X) (based on Prockop and Kivirikko (1984) with permission).

Non-collagen proteins

Many proteins may be extracted from bone, which differ according to the starting material and the methods used (Triffitt 1987; Termine 1988; Robey 1989; Robey 1994) (see Table 4). These include osteocalcin (Gla protein), sialoproteins, and various phosphoproteins such as osteonectin and osteopontin (Young *et al*. 1993). Within the group of non-collagen proteins it is convenient also to include the bone proteoglycans and, importantly, the bone morphogenetic proteins (Table 4 and Table 5). The complexity of non-collagen substances sequestered in bone matrix is emphasized by Hauschka *et al*. (1988) in their discussion of polypeptide growth factors. Osteoblasts are primarily responsible for the biosynthesis of all of these substances. However, other sources include substances selectively absorbed from the plasma, the products of monocytes, lymphocytes, and other cells within the bone marrow, the capillary network with its associated endothelial cells and basement membranes, and substances derived from the cartilage during endochondral ossification.

Names	Properties
Osteonectin (SPARC protein)	Gene on chromosome 5
	Binds to both collagen and hydroxyapatite
	Synthesized by many cells, including platelets
Sialoprotein I (osteopontin)	In many different bone cells; mRNA also in renal tubules, neuronal cells, etc.
Sialoprotein II (bone sialoprotein)	Gene on chromosome 4 (same as osteopontin)
	Fairly bone specific
Osteocalcin (bone Gla protein, BGP)	Gene on chromosome 1; found in young osteoblasts and odontoblasts
Matrix Gla protein (MGP)	Not unique to bone; mRNA earlier than that for BGP in development
Proteoglycan I (PG-I, biglycan)	Two separate species, but share single heterogeneous core protein
	Not unique to bone
Proteoglycan II (PG-II, decorin)	
Bone morphogenetic proteins (BMPs)	See Table 5
Insulin-like growth factors I and II	

Table 4 Major non-collagen protein products of bone cells

BMP	Chromosome	Comment[a]
1	8	Not a member of the TGF-β gene family
2	20p12	
3	4p14–q21	Dentinogenesis imperfecta type II
4	14	
5	6	
6	6	
7	20	

[a]BMP 2–7 are all members of the TGF-β gene family. There is evidence that BMP 2, 4, and 7 could be involved in the Holt–Oram syndrome. Any BMP genes could be candidate loci for FOP (Wozney 1993). TGF-β, transforming growth factor-β.

Table 5 The chromosomal localization of some bone morphogenetic proteins (BMPs)

The non-collagen proteins extracted from bone also depend on the type of bone and the methods used. Two-dimensional electrophoretic analysis may demonstrate more than 200 of them. Many of these bind calcium and many are phosphoproteins. A large number also originate outside bone. Many alternative names have been used for these substances. Robey (1989, 1994) in reviews of the structure and biochemistry of bone, points out that no unambiguous function has been determined for any bone matrix protein to date. Few, if any, are unique to bone, and bone matrix proteins can be expressed (at least transiently) in many tissues.

Osteonectin

This is the most abundant non-collagen protein produced by human osteoblasts (Young *et al*. 1993). It binds strongly to calcium ions, utilizing high and low affinity binding sites, hydroxyapatite, and native collagen. It is not limited to mineralizing tissue, being also found in human platelets. Although osteonectin mRNA is widely distributed in developing tissues, osteonectin is most abundant in bone.

Bone sialoproteins

Two types of these are now recognized. Their relative abundance varies with species studied. Thus sialoprotein I is a minor component of human bone, but a major contributor to total sialoprotein in rat bone. The protein contains a Arg–Gly–Asp cell-attachment sequence and is therefore called osteopontin. The major human bone sialoprotein is type II, similar to the substance originally described by Herring (for references see Termine (1988) and Robey (1989)).

Bone Gla-containing proteins

There are two of these, osteocalcin (bone Gla protein) and matrix Gla protein. The term Gla refers to the g-carboxylated glutamic acid residues, formed by vitamin K-modulated, posttranslational carboxylation of peptide-bound glutamic acid. These proteins have some sequence homology but are products of different genes. Matrix Gla protein is also a cartilage protein, and is found at an earlier developmental stage than osteocalcin. Despite extensive research on osteocalcin, its function is unknown. Warfarin-treated animals do not show abnormal mineralization. Biosynthesis of osteocalcin is regulated by 1,25-dihydroxycholecalciferol (and no other hormone), which enhances its nuclear transcription and eventually secretion from bone cells. Plasma osteocalcin has been linked to the rate of bone formation or, less specifically, bone turnover.

Bone proteoglycans

Proteoglycans from cartilage have been studied more extensively than those from bone (Hardingham 1981; Hardingham and Fosang 1992; Young *et al*. 1993). Bone matrix proteoglycans differ from them in their small overall size, relatively larger amounts of protein, and longer chondroitin sulphate chains. The two proteoglycans of

fetal bone (I and II) are different gene products. Small proteoglycans such as decorin and biglycan are thought to interact with growing collagen fibrils in a precise manner and to regulate their growth, maturation, and interactions.

Bone morphogenetic proteins

It has been known for many years that demineralized bone matrix contains substances capable of inducing ectopic bone formation. These substances are present in such small amounts that their extraction and isolation has presented great difficulties (Table 5). Wozney *et al*. (1988) reported the cloning and recombinant expression of several molecules that could induce cartilage but not bone; Sampath *et al*. (1987) isolated a 22-kDa protein from bone called osteogenin, capable of inducing new bone formation when implanted subcutaneously with the residue of extracted bone matrix. Bentz *et al*. (1989) isolated an osteo-inductive factor, which was active only together with transforming growth factor-b_1 and -b_2. Martin *et al*. (1989) emphasize the importance of the transforming growth factor-b supergene family in cartilage and bone formation. Subsequent work has shown that a large number of bone morphogenetic proteins exist, many of which belong to the transforming growth factor-b family (Tabas *et al*. 1991; Wozney 1993; Centrella *et al*. 1994) (Table 5). The bone morphogenetic proteins act as differentiation factors concerned with the induction and expression of multiple phenotypes derived from mesenchymal cells. Investigation of recombinant bone morphogenetic proteins has shown that individually and together they are involved in chondrogenesis and osteogenesis. Additionally members of the family are involved in the patterning and development of the human skeleton (Tickle 1994). It is becoming clear that recombinant bone morphogenetic proteins have great clinical potential in the repair of bony defects (Editorial 1992a).

Bone mineral and mineralization

Mineralization occurs against the background of the organic bone matrix. The ways in which mineralization happens continue to be debated (Boskey 1994) but there is good evidence that, in most mineralized tissues, calcifying vesicles derived from chondrocytes or osteoblasts provide a focus for mineralization (Robey 1989). These vesicles are particularly demonstrable in cartilage, but their function in the organized matrix of bone is controversial. The precipitation of calcium within these vesicles may be controlled by the action of a pyrophosphatase, which locally destroys pyrophosphate, itself an inhibitor of mineralization. Alkaline phosphatase is one such pyrophosphatase and is readily demonstrable both in osteoblasts and in mineralizing vesicles.

It is possible, for the purposes of clarity, to consider two types of mineralization—homogeneous nucleation, which occurs in the lumen of the matrix vesicles from amorphous calcium phosphate to form crystalline hydroxyapatite, and heterogeneous nucleation, which is collagen-mediated and may partly rely on absorbed non-collagen proteins as nucleators. After this first phase (mediated either by vesicles or collagen) there is a second phase of rapid spread of mineralization, initially in the hole zones and later the overlap regions of the collagen matrix.

Calcium and phosphorus balance

Calcium is essential for innumerable functions such as reproduction, neurotransmission, hormone action, cellular growth, and enzyme action. The central position of calcium as an ionic messenger continues to be explored. There have been considerable advances in our understanding of the messengers which control cellular processes by generating internal calcium signals. Chief amongst these is inositol triphosphate which is generated via G-protein and tyrosine kinase-linked receptors (Berridge 1993). Much has been written about external calcium balance and the main hormones that control it (MacIntyre 1986; Nordin 1990). Phosphate balance is less well understood. The circulating concentration of plasma calcium is determined by the amount of calcium absorbed by the intestine, the amount excreted by the kidney, and the exchange of mineral with the skeleton. The relative importance of these exchanges differs during growth and in different disorders. Total plasma calcium is closely maintained between 2.25 and 2.60 mmol/l, of which nearly half is in the ionized form (47 per cent ionized, 46 per cent protein bound, and the remainder complexed). The skeleton contains approximately 1 kg (25 000 mmol) of calcium. The main daily fluxes of calcium in the adult are shown in Fig. 12.

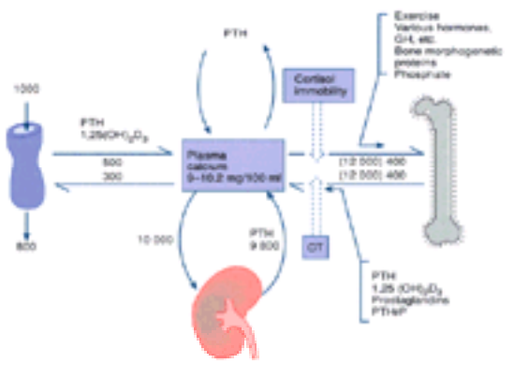


Fig. 12 Factors that control calcium balance; units are in mg per day (to convert to mmol divide by 40), and refer to an adult. The figures in parentheses are an estimate of exchange through the cellular barrier of bone. CT, calcitonin; GH, growth hormone; PTH, parathyroid hormone; PTHrP, parathyroid hormone-related peptide.

Parathyroid hormone

The gene for parathyroid hormone is on chromosome 11. The hormone is synthesized as a large precursor for export (Kronenberg *et al*. 1986) and its secretion is stimulated by a reduction in the plasma concentration of ionized calcium and reduced by an increase in 1,25-dihydroxycholecalciferol. The way in which the parathyroid cells respond to very small changes in the extracellular concentration of calcium has been illuminated by the cloning and characterization of a calcium-sensing receptor from parathyroid-derived cells (Brown *et al*. 1993) (Fig. 13). Mutations within this receptor cause both hypercalcaemia and hypocalcaemia. An increase in parathyroid hormone secretion leads to an increase in calcium absorption through the gut, an increase in calcium reabsorption through the kidney, and an increase in bone resorption. Intestinal calcium absorption is mediated by 1,25-dihydroxycholecalciferol and 1a-hydroxylation is stimulated by parathyroid hormone, so that the effect of parathyroid hormone on intestinal calcium absorption is indirect and mediated by 1,25-dihydroxycholecalciferol. In contrast, the renal effect of parathyroid hormone on calcium reabsorption is direct. The cellular effects of parathyroid hormone are mediated through specific G-protein-coupled receptors in kidney and bone and appear to utilize more than one system (Caulfield and Rosenblatt 1990; Demay *et al*. 1993). Parathyroid hormone encourages osteoclastic bone resorption by its effect on the osteoblast as previously described.

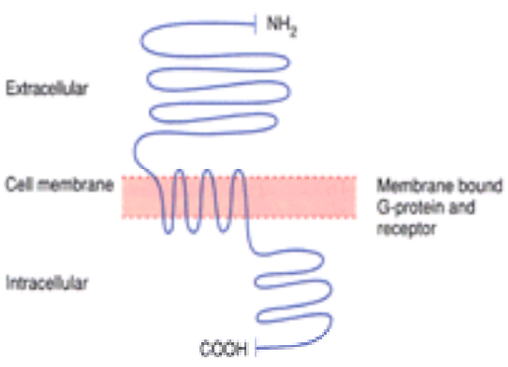


Fig. 13 A simplified diagram of the extracellular Ca^{2+}-sensing receptor from bovine parathyroid (redrawn from Brown *et al*. (1993)).

Vitamin D

Vitamin D (Fraser 1995) is synthesized either as vitamin D_3 (cholecalciferol) within the skin from its precursor 7-dehydrocholesterol under the influence of ultraviolet light (usually as sunlight), or taken in with food, either as vitamin D_3 or D_2 (ergocalciferol). It is subsequently transported to the liver by a binding protein where it undergoes 25-hydroxylation; 25-hydroxyvitamin D ($25(OH)D_3$) is then hydroxylated in the 1 position by the renal 1a-hydroxylase, 1,25-dihydroxycholecalciferol ($1,25(OH)_2D_3$) is the active metabolite of vitamin D, and it has widespread effects (Fig. 14). Its classic effects are on calcium metabolism, where it promotes the synthesis of a calcium-transporting protein within the cells of the small intestine. Its effects are mediated through a widely distributed vitamin D receptor which contains DNA and hormone-binding components. This receptor combines with retinoid X-receptors and vitamin D response elements to produce its effect on target tissues. It is now realized that 1,25-dihydroxycholecalciferol has many effects outside mineral metabolism, concerned with the immune system and the growth and differentiation of a wide variety of cells (Holick 1988; Reichel *et al*. 1989; De Luca *et al*. 1990). Measurement of plasma 25-OHD has proved to be a useful indicator of vitamin D status, and work on 1,25-dihydroxycholecalciferol and its receptors has illuminated the cause of the rarer forms of inherited rickets (see below). Although the kidney is the main sources of 1,25-dihydroxycholecalciferol, it is now clear that this metabolite can be synthesized by a variety of granulomata, providing a partial explanation for the hypercalcaemia of sarcoidosis and (occasionally) lymphomas (Cox and Haddad 1994).

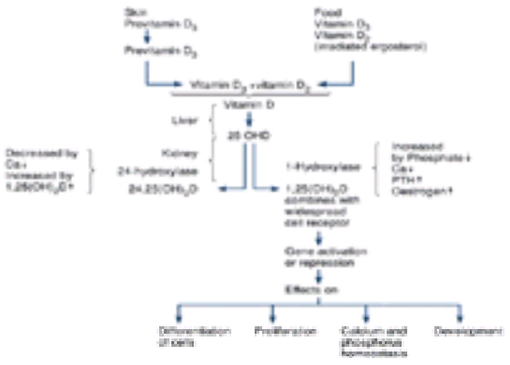


Fig. 14 The synthetic pathways and molecular and cellular effects of $1,25(OH)_2D$.

Calcitonin

Calcitonin (CT), a 32-amino-acid peptide, is one product of the CT gene which also encodes the calcitonin gene-related peptide (CGRP). Experimentally its secretion (by the C cells of the thyroid) is increased by an increase in plasma calcium. The main effect of administered calcitonin is to reduce bone resorption by direct and reversible suppression of the osteoclast (Azria 1989) via a receptor closely related to the parathyroid hormone receptor (Hruska *et al*. 1993). The physiological role of calcitonin is uncertain, although it is thought to protect the skeleton during physiological stresses such as growth and pregnancy.

Parathyroid hormone-related protein (PTHrP)

This hormone (Martin 1990) was discovered by studies on patients with non-metastatic hypercalcaemia of malignancy (see below). It has close sequence homology to parathyroid hormone at the amino-terminal end of the molecule and has very similar effects (Table 6). Its gene is located on the short arm of chromosome 12, which is thought to have arisen by a duplication of chromosome 11, which carries the human parathyroid hormone gene. It has been detected in a number of tumours, particularly of the lung. There is also evidence that it may have a role in fetal physiology, controlling the calcium gradient across the placenta and maintaining the relatively higher concentrations in the fetal circulation. The recent observation that targeted disruption of the PTHrP gene leads to a lethal dysplasia has widened our knowledge of its possible functions (Karaplis *et al*. 1994).

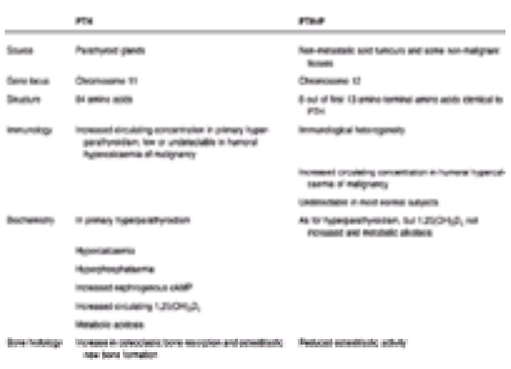


Table 6 Parathyroid hormone (PTH) and parathyroid-hormone related peptide (PTHrP) compared

Other hormones

Apart from the recognized calciotropic hormones, the skeleton is influenced by corticosteroids, the sex hormones, thyroxine, and growth hormone. The main effect of excess corticosteroids (either therapeutic or in Cushing's disease) is to suppress osteoblastic new bone formation. Both androgens and oestrogens promote and maintain skeletal mass. Osteoblasts have receptors for oestogen, although they are not abundant. Thyroxine increases bone turnover and increases resorption in excess of formation; thyrotoxicosis thus leads to bone loss. Excess growth hormone leads to gigantism and acromegaly (according to the age of onset), with enlargement of the bones. Absence of growth hormone will lead to proportional short stature; where there is a wider pituitary failure the reduction in gonadotrophins will cause bone loss.

Biochemical measures of bone turnover

Knowledge of bone physiology allows interpretation of biochemical measures of bone turnover. These include the plasma bone-derived alkaline phosphatase and osteocalcin, and urinary total hydroxyproline and cross-linked, collagen-derived peptides. The first two of these are closely related to osteoblast function, and the second two to bone resorption. As formation and resorption are closely coupled, such measurements are usually closely related to each other, and to overall bone turnover (Epstein 1988).

Plasma alkaline phosphatase (largely derived from osteoblasts) provides a crude but readily accessible index of bone formation, being increased during periods of rapid growth and particularly where bone turnover is greatly increased, as in Paget's disease. Where more skeletal specificity is required measurement of bone-derived alkaline phosphatase can be useful. Early measurements of serum osteocalcin (bone Gla protein) were widely variable and depended on the origin, sensitivity, and stability of the antibodies used. Total urinary hydroxyproline excretion is influenced by dietary collagen (gelatin) and reflects both resorption and new collagen synthesis. The recent development of methods for the measurement of urinary collagen-derived pyridinium cross-links promises to give a reliable indication of bone resorption, unrelated to new collagen formation and uninfluenced by diet (Black *et al*. 1988; Eyre 1992; Seibel *et al*. 1992). The nomenclature of these degradation products has become unnecessarily complex. The cross-links of collagen fibres are formed via aldehydes derived from lysyl and hydroxylysyl residues. When collagen is degraded fragments containing these cross-links, referred to as pyridium cross-links, are excreted in the urine. The cross-linked peptides are not further metabolized and the amount excreted is unaffected by dietary collagen (often in the form of gelatin). Pyridinoline (formed from hydroxylysyl cross-links) and

deoxypyridinoline (from lysyl cross-links) derived from bone collagen are excreted in the same ratio as they exist in bone. Pyridinoline also comes from other tissues, especially tendon and cartilage. Since bone is metabolically more active than cartilage the main source of urinary pyridium cross-linked, collagen-derived peptides is the skeleton. There is a diurnal variation in the amount excreted and according to the method used but significant increases are described in preadolescent growth and also in high turnover bone disease, particularly Paget's disease (Editorial 1992b).

New immunoassay methods are being developed which are likely to supersede the original HPLC (high-performance liquid chromatography) procedure. Methods which measure other aspects of collagen turnover, such as the circulating concentration of carboxyl- and amino-terminal fragments of the molecule, have yet to establish themselves in practice.

Bone in disease

The causes of many bone diseases may now be explained by our increased understanding of skeletal physiology. These diseases may be acquired or inherited, and provide examples of disorders of bone formation and resorption, of mineralization, of matrix synthesis, of enzyme function, and of abnormal cell biology. Many conditions previously thought of as skeletal curiosities are now known to have a biochemical basis (Whyte 1990; Royce and Steinmann 1993; Smith 1993). Some are dealt with in more detail elsewhere.

Osteoporosis

In osteoporosis, imbalance between resorption and formation leads to a reduction in the amount of bone per unit volume without a change in its composition (Riggs and Melton 1988; Smith 1990a; Riggs and Melton 1992; Dempster and Lindsay 1993). This reduction leads to microarchitectural deterioration, increased fragility and increased fracture risk. Although osteoporosis is most frequent with increasing years, especially in postmenopausal women, there are a number of other recognizable causes that reduce the density of the younger skeleton (Table 7). These include anorexia nervosa, and excessive exercise in females associated with premenopausal oestrogen lack (Johnston and Longcope 1990; Prior *et al*. 1990). The latter provides a striking example of the opposing effects of mechanical stress and hormone deficiency on the skeleton. A topical cause of osteoporosis is cardiac transplantation (Sambrook *et al*. 1994).

Common
Increasing age
Menopause
Immobility
Less common
Cushing's syndrome
Hypogonadism
Coeliac disease
Rare
Generalized mastocytosis
Osteogenesis imperfecta
Turner's syndrome
Topical
Oestrogen deficient
Anorexia nervosa
Obsessional exercise
Satellite travel
Unknown
Idiopathic osteoporosis
Juvenile
Pregnancy associated
In young people

Table 7 Causes of osteoporosis

Immobilization causes both osteoblastic failure and osteoclast excess. Cushing's syndrome and excessive therapeutic corticosteroids suppress osteoblastic activity.

The amount of bone in the adult depends on the peak bone mass and the subsequent rate of loss. As already stated, peak bone mass depends on genetic, nutritional, mechanical, and endocrine interactions. The rate of bone loss varies considerably between individuals and may also be genetically determined. It is also accelerated by certain so-called risk factors, such as immobility, excessive thinness, alcoholism, cigarette smoking, low calcium intake, prolonged periods of amenorrhoea, and early menopause. There is a relationship between reduced bone mass and fractures, but it is not a close one, as fractures depend as much on injury, including falls (especially in later life), as on loss of bone. How much excessive cytokine activity, especially of interleukin 6, contributes to postmenopausal bone loss is controversial (Manolagas and Jilka 1995). No differences have been found in circulating levels of this cytokine between normal and osteoporotic women (Khosla *et al*. 1994) but this is not unexpected since the action of interleukin 6 is likely to be very localized.

Osteomalacia and rickets

These conditions are distinguishable only by age. There is defective mineralization of the organic bone matrix due to vitamin D deficiency or a disturbance of its metabolism. The causes (Table 8) are best understood against the known metabolic pathways of vitamin D (Fig. 15). They include vitamin D deficiency, often in immigrants to the United Kingdom from the India subcontinent (Smith 1990b) where an increased utilization of vitamin D (related to chapatti consumption) combines with lack of sunlight and lack of vitamin D in the diet, malabsorption (particularly coeliac disease), and renal disease. The causes of rickets from renal glomerular failure (renal glomerular osteodystrophy) are quite different from those due to renal tubular disease (Fig. 16), and may be modified by aluminium intoxication in those on dialysis. The most frequent cause of renal tubular rickets is inherited hypophosphataemia, but there are many others (Thakker and O'Riordan 1988; Thakker 1992; Rowe 1994). Investigation of one of the rarer forms of rickets, so-called vitamin D-dependent rickets, has shown that it may be due to defective 1a-hydroxylation of 25(OH)D_3 (type I) or to end-organ resistance to 1,25-dihydroxycholecalciferol (type II). The molecular basis of the latter has been shown to be a number of mutations at either end of the 1,25-receptor (Fig. 17) (Editorial 1989; Ritchie *et al*. 1989; Thakker 1992; Hewison and O'Riordan 1994; Rut *et al*. 1994).

'Nutritional'
Immigrants from India to the United Kingdom
Housebound elderly people
Malabsorption
Coeliac disease
Small bowel resection
Renal
Renal glomerular failure
Renal tubular disorders:
Inherited X-linked hypophosphataemia
Multiple renal tubular defects
Rare
Vitamin D dependent (type I, type II)
Oncogenous (hypophosphataemic)

Table 8 Some causes of rickets and osteomalacia

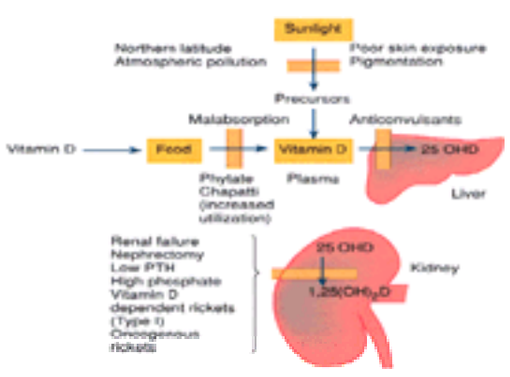


Fig. 15 The sources and metabolism of vitamin D and causes of rickets and osteomalacia.

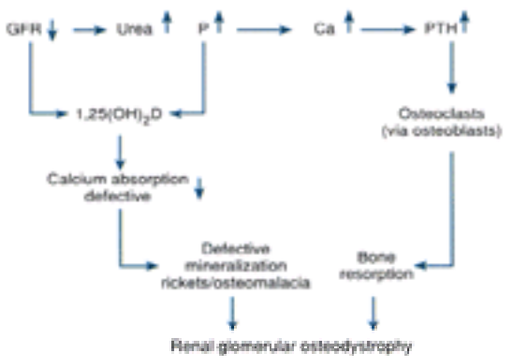


Fig. 16 The effects of renal glomerular failure on the skeleton.

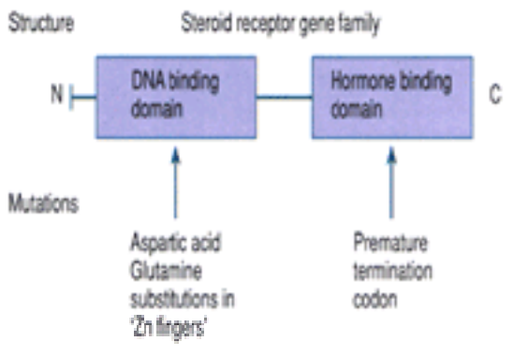


Fig. 17 Some molecular changes identified in type II vitamin D-dependent rickets.

Parathyroid bone disease

Most patients with primary hyperparathyroidism do not have clinically manifest bone disease (osteitis fibrosa cystica), although in the majority the histological appearances are abnormal and improve after removal of the parathyroid tumour (Christiansen *et al*. 1990). Osteitis fibrosa cystica is characterized by an excess of fibrous tissue and osteoclasts within lesions that may become cystic (brown tumours). The most common cause of primary hyperparathyroidism is a parathyroid adenoma, and it is now recognized that a number of these are monoclonal in origin (Friedman *et al*. 1990) with a chromosomal situation similar to multiple endocrine neoplasia type II (Thakker *et al*. 1989). Parathyroid overactivity also occurs in multiple endocrine neoplasia type II associated with mutations in the RET proto-oncogene (Gardner *et al*. 1994).

Parathyroid overactivity can also result from prolonged hypocalcaemia (secondary hyperparathyroidism). Rarely, the parathyroids may become autonomous with hypercalcaemia (tertiary hyperparathyroidism). Finally familial hypocalciuric hypercalcaemia and familial hypocalcaemia can result from mutations in the parathyroid calcium-sensing receptor (Brown *et al*. 1993; Pollock *et al*. 1994). Parathyroid insufficiency can result from inadvertent surgical removal of the parathyroids or may be idiopathic. The condition of pseudohypoparathyroidism is characterized by the presence of short third and fourth metacarpals, ectopic calcification in the skin and in the basal ganglia, mental simplicity, and cataracts (Breslau 1989). It has now been shown that in a number of such patients there are mutations in the genes controlling the G-protein system that links parathyroid hormone to its effector tissue (Spiegel 1990; Gordeladze and Johansen 1994; Schwindinger *et al*. 1994; Wilson *et al*. 1994; Lefkowitz 1995). This provides an explanation of the resistance to parathyroid hormone in this condition (Fig. 18). Interestingly there are similar G-protein defects in polyostotic fibrous dysplasia (Levine 1991).

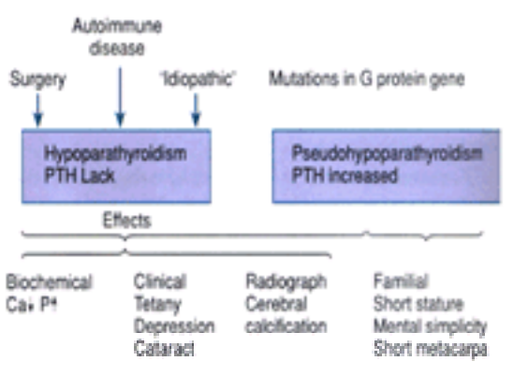


Fig. 18 The causes and effects of parathyroid hormone insufficiency or resistance. Mutations in the Ca $^{2+}$-sensing receptor may also mimic parathyroid resistance.

Paget's disease

In this disorder the main skeletal abnormality appears to be overactivity of large multinucleated osteoclasts. At first this leads to excessive bone resorption but, because of the linkage between osteoclastic and osteoblastic activity, excessive new bone formation follows. Normally this linkage is sufficiently close to prevent hypercalcaemia, however rapid the bone turnover, but this may occur if patients with Paget's disease are immobilized, in which case the osteoblasts become less active whilst resorption continues. It is currently considered that Paget's disease may be related to a viral infection of the osteoclasts (Gordon *et al*. 1994), possibly with the respiratory syncytial virus (Kahn 1990), but current attempts to identify the viral nucleic acid continues to give conflicting results (Smith 1992).

Disorders of bone matrix

Classically these disorders have not been regarded as metabolic disorders of bone but it is now clear that mutations, especially in the collagen genes, can produce a wide variety of skeletal disorders (Table 9 and Chapter 2.2).

Collagen I	Osteogenesis imperfecta
Collagen II	Achondrogenesis type II
	Hypochondrogenesis
	Spondyloepiphyseal dysplasia congenita and others
	Ocular Stickler's syndrome
Collagen IX	Multiple epiphyseal dysplasia
	Diastrophic dwarfism
Collagen X	Metaphyseal dysplasia (type Schmid)
FGF receptor 2	Crouzon's syndrome
FGF receptor 3	Achondroplasia
PTH/PTHrP receptor	Metaphyseal dysplasia (type Jansen)

For references see text and Spranger *et al.* (1994).
FGF, fibroblast growth factor; PTH, parathyroid hormone; PTHrP, parathyroid hormone-related peptide.

Table 9 Gene mutations in the chrondrodysplasias

Osteogenesis imperfecta

Four clinical types are recognized, although many patients with this disorder cannot be accurately classified (Smith 1986; Cole 1988; Byers 1993). In the first, type I, the disorder is dominantly inherited and often appears to be due to a non-functional allele for type I collagen. The main result is that osteoblasts (and fibroblasts) make half as much type I collagen as they ought to, causing an inherited form of osteoporosis and affecting other tissues containing type I collagen (teeth, sclera, heart valves). Type II osteogenesis imperfecta is the lethal form of the disease, this is most often due to a single-base mutation in one of the type I collagen genes leading to an amino-acid substitution for one of the glycine residues normally present in every third position in the amino-acid chain. The importance of glycine is that it has no side-chain and can therefore fit into the interior of the helix (Sykes 1990). Replacement of such a 'core' glycine by an amino acid with a side-chain disturbs helix formation. The effect depends on the size of the side-chain and on the position of the replacement amino acid relative to the carboxyl-terminal end of the collagen molecule, where helix formation begins. The most lethal effects appear to be produced by large substitutions near to the carboxyl-terminal end of the molecule (Fig. 19).

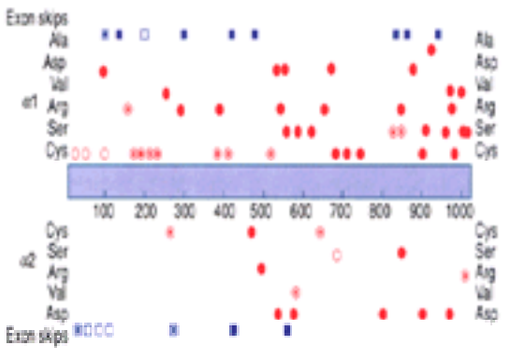


Fig. 19 Mutations in the collagen genes known to cause osteogenesis imperfecta (from Sykes (1990) with permission). Axial glycine (circles) and exon skip (squares) mutations in the genes for the a_1- (upper) and a_2- (lower) subunits of type I collagen. The scale indicates the osteogenesis imperfecta phenotype from lethal (solid) through severe (dot) to mild (open). Note that the collagen triple helix winds up from the carboxyl (right) to the amino (left) terminal and that the least severe phenotypes are often associated with mutations near the amino terminus. Ala, Asp, etc., refer to the amino acid replacing the axial glycine.

Type III osteogenesis imperfecta is a progressively disabling disease. Infants survive but the skeleton becomes more and more deformed. The chemistry of this condition is not so well known as in other forms but again it is probably frequently related to a disturbance of helix formation by glycine substitutions. Type IV is rare and is a dominantly inherited disease. Extensive linkage studies have demonstrated that dominantly inherited (type I and IV) osteogenesis imperfecta is linked to mutations in either the a_1- or a_2-chain of type I collagen (Sykes *et al*. 1990). A detailed study on the variable expression of the mutant osteogenesis imperfecta gene demonstrates the importance of somatic mosaicism (Wallis *et al*. 1990). The relationship between genotype and phenotype is still complex (Smith 1994).

Other forms of inherited disorders of connective tissue

These include Marfan's syndrome, Ehlers–Danlos syndrome, and the skeletal dysplasias. It is clear that in most cases Marfan's syndrome is due to mutations in the fibrillin genes (Pyeritz 1990; Francke and Furthmayr 1994; Pereira *et al*. 1994; Pyeritz 1993) on chromosome 15, although linkage to chromosome 3 has also been described (Collod *et al*. 1994). Since fibrillin is the major component of the suspensory ligament of the lens and of the elastin-associated microfibrillar system, mutations in its genes provide a neat explanation for the combination of lens dislocation and aortic dissection found in this disorder (Kielty *et al*. 1995).

There are many different types of the Ehlers–Danlos syndrome. The cause of the most common forms, types 1, 2, and 3, has yet to be elucidated. In type 4 the major complication is rupture of the middle-sized blood vessels resulting from a variety of mutations in the gene for type III collagen. In type 6 Ehlers–Danlos syndrome, whose main features are ocular fragility and scoliosis, there is evidence of hydroxylysine deficiency. Another type of Ehlers–Danlos syndrome (type VII, with excessive joint mobility) results from defective removal of the procollagen extensions from the precursor molecule due to mutations at the peptidase site.

Skeletal dysplasias

These conditions (Wynne-Davies *et al*. 1985) which include achondroplasia, spondyloepiphyseal dysplasia, and achondrogenesis, were long thought of as orthopaedic curiosities. However, numerous mutations have now been described. To begin with, exon deletions were described in the type II collagen gene in patients with spondyloepiphyseal dysplasia (Lee *et al*. 1989), and single-base mutations in the same gene in infants with achondrogenesis type II (Vissing *et al*. 1989). There was also evidence of genetic linkage of spondyloepiphyseal dysplasia congenita to type II collagen (Anderson *et al*. 1990). It is possible that mutations in type II collagen, which is the major collagen of cartilage, will be increasingly described as a cause of those disorders characterized by abnormal cartilage formation, although a number have been excluded (Spranger *et al*. 1994; Winterpacht *et al*. 1994). Indeed recent work (Table 9) has linked multiple epiphyseal dysplasia to mutations in type IX collagen (a minor fibre-associated collagen) (Briggs *et al*. 1994) and abnormal type IX collagen has been described in diastrophic dysplasia (Diab *et al*. 1994). Mutations in type X collagen have also been described in the Schmid type of metaphyseal dysplasia (type X collagen is specifically produced by hypertrophic chondrocytes) (McIntosh *et al*. 1994; Wallis *et al*. 1994). Unexpectedly, skeletal development is normal in type X collagen-null mice (Rosati *et al*. 1994). Most interesting of all mutations are those in the type 3 fibroblast growth-factor receptor (FGFR3) which have been identified as the cause of achondroplasia (Editorial 1994; Le Merrer *et al*. 1994; Rousseau *et al*. 1994; Shiang *et al*. 1994; Velinov *et al*. 1994). The reason why such mutations should lead to specific changes in the skeleton is not yet understood. Finally mutations in a similar receptor, FGFR2, appear to cause Crouzon's syndrome (Reardon *et al*. 1994). Another (activating) mutation, this time in the PTH/PTHrP receptor, has recently been identified as a cause of the Jansen type of metaphyseal dysplasia (Schipani *et al*. 1995).

Enzyme disorders

Some disorders that affect the skeleton result from abnormal enzyme function. These include hypophosphatasia, homocystinuria, and alkaptonuria.

Hypophosphatasia

In hypophosphatasia there are variable abnormalities in the skeleton associated with differing phenotypes. A common biochemical feature is an abnormally low plasma alkaline phosphatase associated with an excess of phosphoethanolamine in the urine. The disorder may be lethal, or it may cause hypercalcaemia in infants and early fusion of the cranial sutures (Whyte 1990). By contrast the disease can be mild in adults, although pathological fractures of the long bones may occur.

Various mutation in the gene for the tissue non-specific alkaline phosphatase have been described in some patients (Weiss *et al*. 1988; Henthorn and Whyte 1992; Orimo *et al*. 1994) but not in others. Family studies have now linked the defect to chromosome 1 (Greenberg *et al*. 1990).

Homocystinuria

This condition, which mimics Marfan's syndrome, is due to a defect in cystathionine synthase, a pyridoxine-dependent enzyme. Accumulation of homocysteine behind the metabolic block leads to homocystinuria. Skeletal features of this disease are very similar to those of Marfan's syndrome with arachnodactyly and scoliosis. There is also dislocation of the lenses.

Alkaptonuria

In this disorder there is a deficiency of homogentisic acid oxidase; the accumulation of homogentisic acid is associated with the features of ochronosis, with pigmentation of articular cartilage, early osteoarthritis, and calcification of the intervertebral discs.

Disorders of cell biology

In two rare conditions, osteopetrosis and myositis ossificans progressiva, there is evidence that the disorder is primarily due to abnormalities in the bone cells.

Osteopetrosis

There are a number of inherited disorders in which there is an excessive amount of bone within the skeleton. The most well recognized is marble bone disease (Albers–Schönberg disease). This condition may be severe and recessively inherited, occurring in infancy, or it may be mild with dominant inheritance. There are intermediate forms and a rare form due to deficiency of carbonic anhydrase II (Sly *et al*. 1983; Fathallah *et al*. 1994). In animals and in man, varying osteoclast defects are described and there is evidence that the infantile form may be cured by successful transplantation of normal osteoclasts (Fischer *et al*. 1986). Interestingly, osteopetrosis in the micro-ophthalmic *op*/*op* mouse is associated with a defect in the gene for macrophage colony-stimulating factor (Yoshida *et al*. 1990) and mice lacking the c-*fos* oncogene develop osteopetrosis, emphasizing the importance of this system in bone physiology (Wang *et al*. 1992). Jackson *et al*. (1994) suggest that osteoblasts may also be defective in mammalian osteopetrosis.

Fibrodysplasia (myositis) ossificans progressiva

In this disorder progressive ossification of the major striated muscles is associated with characteristic skeletal abnormalities from birth (Connor and Evans 1984; Cohen *et al*. 1993). The cause is unknown but it seems that the fibroblasts or mesenchymal cells within muscle behave as if they are osteoblasts to produce cartilage and true bone within muscles. This may be due to an inherited cellular abnormality; alternatively it could represent, for instance, an excess of one of the morphogenetic proteins (Smith and Triffitt 1986). The combination of skeletal abnormalities and ectopic ossification, which often occurs in a specific order, would seem to implicate proteins involved in skeletal patterning as well as bone morphogenesis (Storm *et al*. 1994).

The skeleton in malignant disease

The skeleton is often affected in malignant disease, which is the most frequent cause of hypercalcaemia in hospital patients. This hypercalcaemia results from excessive bone resorption and increased renal tubular reabsorption of calcium. The relative importance of these depends on the type and distribution of the cancer. Bone resorption in cancer is due to an increase in osteoclastic activity (either as a direct response to hormonal factors or indirectly via other cells), and to direct resorption of bone by the malignant cells themselves. The complexity of the mechanisms involved illustrates the multiple effects of cytokines and hormones on the skeleton (Mundy 1989; Orloff *et al*. 1989; Gutierrez *et al*. 1990; Martin 1990).

Neoplastic hypercalcaemia may occur in three clinical settings—haematological malignancy, solid tumour with metastases, and solid tumours without. In the first of these the hypercalcaemia is usually due to the release of osteoclast-activating factors by the malignant cells (occasionally, as in lymphoma, the plasma calcium is increased because of excessive formation of 1,25-dihydroxycholecalciferol by the lymphomatous tissue; Cox and Haddad 1994); in the case of myeloma this activating factor is a lymphokine called lymphotoxin. Osteoclast-activating factors have been divided into those from T cells, B cells, and monocytes.

In solid tumours with lytic metastases, predominantly from the breast, lung and kidney, a major contribution to hypercalcaemia is direct destruction of bone. However, where there are few or no detectable metastases, other mechanisms must be involved. Chief amongst these is the production of parathyroid hormone-related peptide by the tumour itself (Martin 1990). This hormone may also be relevant in metastatic and haematological malignancies. Parathyroid hormone-related peptide has been detected mainly in lung tumours, but also occurs in those from the breast (the mRNA for parathyroid hormone-related peptide is present in lactating mammary tissue) (Heath *et al*. 1990). The main biochemical effect of parathyroid hormone-related peptide is to produce changes similar to those of parathyroid hormone. These are hypercalcaemia and hypophosphataemia, which has led in the past to terms like 'pseudohyperparathyroidism' and 'ectopic parathyroid hormone syndrome'. It is now clear that the concentrations of parathyroid hormone are low, although nephrogenous cAMP is increased. Increases in circulating parathyroid hormone-like protein have been recorded in such patients (Budayr *et al*. 1989). Finally, although the increase in parathyroid hormone-related peptide provides a neat explanation for hypercalcaemia in the absence of metastases, it may be only part of the story because the peptide may act in concert with other factors locally on bone. These have already been dealt with and include transforming growth factors-a and -b, prostaglandins, and interleukins (Table 10).

Factor	Actions
Parathyroid hormone-related peptide (PTHrP)	See Table 6
Transforming growth factor-α (TGF-α)	Stimulates osteoclastic bone resorption
	Produced by some tumours as PTHrP
Tumour necrosis factor (TNF)	Powerful bone-resorbing cytokine; synthesized by normal activated macrophages
	Stimulates formation and activity of osteoclasts
	Probably produced by host cells under action of tumour-produced granulocyte-macrophage colony-stimulating factor
Interleukin 1	Effect similar to TNF; associated with solid tumours
	Acts synergistically on bone resorption with TNF, parathyroid hormone, PTHrP, and TGF-α
Prostaglandins	Unlikely as a mediator for the humoral hypercalcaemia of

Table 10 Factors involved in the humoral hypercalcaemia of malignancy

Conclusions

The skeleton fulfils many functions whose requirements are often in opposition. Bone is a very complex structure, which is continually being dismantled and replaced. This complexity leads to its fascination for investigators and provides many opportunities for a wide variety of bone diseases. Recent examples of rare diseases are provided by the discovery of mutations in the G-protein and Ca^{2+}-sensing receptor systems, in the minor collagens, and in the fibroblast growth factor receptors. Much more work is required before we fully understand the causes of the common polygenic disorders such as osteoporosis.

Chapter References

Anderson, I.J., Goldberg, R.B., Marion, R.W., Upholt, W.B., and Tsipouras, P. (1990). Spondyloepiphyseal dysplasia congenita: genetic linkage to type II collagen (COL2A1). *American Journal of Human Genetics*, **46**, 896–901.

Avioli, L.V. and Krane, S.M. (1990). *Metabolic bone disease and related disorders*, (2nd edn). W.B. Saunders, Philadelphia, PA.

Azria, M. (1989). *The calcitonins: physiology and pharmacology*. Karger, London.

Baron, R. *et al*. (1993). Cellular and molecular biology of the osteoclast. In *Cellular and molecular biology of bone*, (1st edn) (ed. M. Noda), pp. 445–95. Academic Press, New York.

Bentz, H. *et al*. (1989). Purification and characterisation of a unique osteoinductive factor from bovine bone. *Journal of Biological Chemistry*, **246**, 20805–10.

Bernards, A. and Gusella, J.F. (1994). The importance of genetic mosaicism in human disease. *New England Journal of Medicine*, **331**, 1447–9.

Berridge, M.J. (1993). Inositol trisphophate and calcium signalling. *Nature*, **361**, 315–25.

Black, D., Duncan, A., and Robins, S.P. (1988). Quantitative analysis of the pyridinium cross links of collagen in urine using ion-paired reversed-phase high-performance liquid chromatography. *Analytical Biochemistry*, **169**, 197–203.

Boskey, A. (1994). Bone and cartilage mineralization. In *Bone formation and repair*, (1st edn) (ed. C.T. Brighton, G. Friedlaender, and J.M. Lane), pp. 23–48. American Academy of Orthopedic Surgeons, Rosemount, IL.

Boyde, A. (1972). Scanning electron microscope studies of bone. In *The biochemistry and physiology of bone*, Vol. 1 (2nd edn) (ed. G. H. Bourne), pp. 259–310. Academic Press, New York.

Breslau, N.A. (1989). Pseudohypoparathyroidism: current concepts. *American Journal of the Medical Sciences*, **298**, 131–40.

Briggs, M.D. *et al*. (1994). Genetic mapping of a locus for multiple epiphyseal dysplasia (EDM2) to a region of chromosome 1 containing a type IX collagen gene. *American Journal of Human Genetics*, **55**, 678–84.

Brighton, C.T., Friedlaender, G., and Lane, J.M. (1994). *Bone formation and repair* (1st edn). American Academy of Orthopedic Surgeons, Rosemount, IL.

Brown, E.M. *et al*. (1993). Cloning and characterization of an extracellular Ca^{2+}-sensing receptor from bovine parathyroid. *Nature*, **366**, 575–80.

Budayr, A.A. *et al*. (1989). Increased serum levels of a parathyroid hormone-like protein in malignancy-associated hypercalcaemia. *Annals of Internal Medicine*, **111**, 807–12.

Burgeson, R.E. (1988). New collagens, new concepts. *Annual Review of Cell Biology*, **4**, 551–77.

Byers, P.H. (1993). Osteogenesis imperfecta. In *Connective tissue and its heritable disorders*, (1st edn) (ed. P.M. Royce and B. Steinmann), pp. 317–50. Wiley-Liss, New York.

Canalis, E., McCarthy, T.L., and Centrella, M. (1989). The role of growth factors in skeletal remodelling. *Endocrinology and Metabolism Clinics of North America*, **18**, 903–18.

Caulfield, M.P. and Rosenblatt, M. (1990). Parathyroid-hormone receptor interactions. *Trends in Endocrinology and Metabolism*, **1**, 164–8.

Centrella, M., Horowitz, M.C., Wozney, J.M., and McCarthy, T.L. (1994). Transforming growth factor-b gene family members and bone. *Endocrine Reviews*, **15**, 27–39.

Chambers, T.J. (1988). The regulation of osteoclastic development and function. *Ciba Foundation Symposium*, (ed. D. Evered and S. Harnett), **i**, 92–107.

Christiansen, P., Steiniche, T., Mosekilde, L., Hesson, I., and Melsen, F. (1990). Primary hyperparathyroidism: changes in trabecular bone remodelling following surgical treatment—evaluated by histomorphometric methods. *Bone*, **11**, 75–9.

Coe, F.L. and Favus, M.J. (1991). *Disorders of bone and mineral metabolism*, (1st edn). Raven Press, New York.

Cohen, R.B. *et al*. (1993). The natural history of ectopic ossification in patients who have fibrodysplasia ossificans progressiva. A study of forty-four patients. *Journal of Bone and Joint Surgery*, **75A**, 215–19.

Cole, W.G. (1988). Osteogenesis imperfecta. *Baillière's Clinical Endocrinology and Metabolism*, **2**, 243–65.

Collod, C. *et al*. (1994). A second locus for Marfan syndrome maps to chromosome 3 p24.2–p25. *Nature Genetics*, **8**, 264–8.

Connor, J.M. and Evans, D.A. (1984). Fibrodysplasia ossificans progressiva. The clinical features and natural history of 34 patients. *Journal of Bone and Joint Surgery*, **64B**, 76–83.

Cox, M. and Haddad, J.G. (1994). Lymphoma, hypercalcaemia and the sunshine vitamin. *Annals of Internal Medicine*, **21**, 709–12.

DeLuca, H.F., Krisinger, J., and Darwish, H. (1990). The vitamin D system: 1990. *Kidney International*, **38** (Suppl. 29), S2–S8.

Demay, M. *et al*. (1993). Parathyroid hormone biosynthesis and action. In *Cellular and molecular biology of bone*, (1st edn) (ed. M. Noda), pp. 321–41. Academic Press, New York.

Dempster, D.W. and Lindsay, R. (1993). Pathogenesis of osteoporosis. *Lancet*, **341**, 797–801.

Diab, M., Wu, J-J., Shapiro, F., and Eyre, D. (1994). Abnormality of type IX collagen in a patient with diastrophic dysplasia. *American Journal of Medical Genetics*, **49**, 402–9.

Editorial (1989). Zinc fingers and vitamin D resistance. *Lancet*, **i**, 487.

Editorial (1992*a*). New bone? *Lancet*, **339**, 463–4.

Editorial (1992*b*). Pyridinium crosslinks as markers of bone resorption. *Lancet*, **340**, 278–9.

Editorial (1994). Fingering fibroblast growth factor receptors. *Nature Genetics*, **8**, 1–2.

Epstein, S. (1988). Serum and urinary markers of bone remodelling: assessment of bone turnover. *Endocrine Reviews*, **9**, 437–49.

Eriksen, E.F., Steiniche, T., Mosekilde, L., and Melsen, F. (1989). Histomorphometric analysis of bone in metabolic bone disease. *Endocrinology and Metabolism Clinics of North America*, **18**, 919–54.

Evered, D. and Harnett, S. (ed.) (1988). Cell and molecular biology of vertebrate hard tissue. *Ciba Foundation Symposium*, **136**.

Eyre, D. (1992). New biomarkers of bone resorption. *Journal of Clinical Endocrinology and Metabolism*, **74**, 470A–C.

Fathallah, D.M., Bejaoui, M., Sly, W.S., Lakhoua, R., and Dellagi, K. (1994). A unique mutation underlying carbonic anhydrase II deficiency syndrome in patients of Arab descent. *Human Genetics*, **94**, 581–2.

Fischer, A. *et al*. (1986). Bone marrow transplantation for immunodeficiencies and osteopetrosis: European survey 1968–1985. *Lancet*, **ii**, 1080–3.

Francke, U. and Furthmayr, H. (1994). Marfan's syndrome and other disorders of fibrillin. *New England Journal of Medicine*, **330**, 1384–5.

Fraser, D.R. (1995). Vitamin D. *Lancet*, **345**, 104–7.

Friedman, E. *et al*. (1990). Genetic abnormalities in sporadic parathyroid adenomas. *Journal of Clinical Endocrinology and Metabolism*, **71**, 293–7.

Gardner, E. *et al*. (1994). Haplotype analysis of MEN 2 mutations. *Human Molecular Genetics*, **3**, 1771–4.

Glass, C.K. and Brown, M.A. (1993). Molecular mechanisms of estrogen and thyroid hormone action. *Cellular and molecular biology of bone*, (1st edn) (ed. M. Noda), pp. 257–85. Academic Press, New York.

Gordeladze, J.O. and Johansen Per, W. (1994). G-proteins: implications for pathophysiology and disease. *European Journal of Endocrinology*, **131**, 557–74.

Gordon, M.T., Mee, A.P., and Sharpe, P.T. (1994). Paramyxoviruses in Paget's disease. *Seminars in Arthritis and Rheumatism*, **23**, 232–4.

Greenberg, C.R. *et al*. (1990). Infantile hypophosphatasia: localisation within chromosome region 1p36.1–34 and prenatal diagnosis using linked DNA markers. *American Journal of Human Genetics*, **46**, 286–92.

Gutierrez, G.E., Poser, J.W., Katz, M.S., Yates, A.J.P., Henry, H.L., and Mundy, G.R. (1990). Mechanism of hypercalcaemia of malignancy. *Baillière's Clinical Endocrinology and Metabolism*, **4**, 119–38.

Hall, J.G. (1988). Review and hypothesis. Somatic mosaicism: observations related to clinical genetics. *American Journal of Human Genetics*, **43**, 355–63.

Hardingham, T. (1981). Proteoglycans: their structure, interactions and molecular organisation in cartilage. *Biochemical Society Transactions*, **9**, 489–97.

Hardingham, T. and Fosang, A.J. (1992). Proteoglycans: many forms and many functions. *FASEB Journal*, **6**, 861–70.

Hauschka, P.V., Chen, T.L., and Mavrakos, A.E. (1988). Polypeptide growth factors in bone matrix. *Ciba Foundation* Symposium, **136**, 207–25.

Heaney, R.P. (1986). Calcium, bone health and osteoporosis. *Bone and Mineral Research*, **4**, 255–301.

Heaney, R.P. (1993). Thinking straight about calcium. *New England Journal of Medicine*, **328**, 503–5.

Heath, D.A., Senior, P.V., Varley, J.M., and Beck, F. (1990). Parathyroid hormone-related protein in tumours associated with hypercalcaemia. *Lancet*, **335**, 66–9.

Heinonen, A., Pekka, O., Kannus, P., Sievanen, H., Mantari, A., and Vuori, I. (1993). Bone mineral density of female athletes in different sports. *Bone and Mineral*, **23**, 1–14.

Henthorn, P.S and Whyte, M.P. (1992). Missense mutations of the tissue non-specific alkaline phosphatase gene in hypophosphatasia. *Clinical Chemistry*, **38**, 2501–5.

Hewison, M. and O'Riordan, J.L.H. (1994). Vitamin D resistance. *Baillière's Clinical Endocrinology and Metabolism*, **8**, 305–15.

Holick, M.F. (1988). Skin: site of synthesis of vitamin D and a target for the active form 1,25-dihydroxy vitamin D_3. *Annals of The New York Academy of Sciences*, **548**, 14–26.

Hruska, K.A., Rolnick, F., Duncan, R.L., Medhora, M., and Yarnakawa, K. (1993). Signal transduction in osteoblasts and osteoclasts. *Cellular and molecular biology of bone*, (1st edn) (ed. M. Noda), pp. 413–44. Academic Press, New York.

Hulmes, D.S. (1992). The collagen superfamily—diverse structures and assemblies. *Biochemistry*, **27**, 49–67.

Hustmyer, F.G., Peacock, M., Hui, S., Johnston, C.C., and Christian, J. (1994). Bone mineral density in relation to polymorphism at the vitamin D receptor gene locus. *Journal of Clinical Investigation*, **94**, 2130–4.

Jackson, M.E., Shalloub, V., Lian, J.B., Stein, G.S., and Marks, S.C., Jr. (1994). Abberant gene expression in cultured mammalian bone cells demonstrates an osteoblast defect in osteopetrosis. *Journal of Cellular Biochemistry*, **55**, 366–72.

Johnston, C.C. and Longcope, C. (1990). Premenopausal bone loss—a risk factor for osteoporosis. *New England Journal of Medicine*, **323**, 1271–3.

Johnston, C.C. *et al*. (1992). Calcium supplementation and increases in bone mineral density in children. *New England Journal of Medicine*, **327**, 82–7.

Kahn, A.J. (1990). The viral etiology of Paget's disease of bone: a new perspective. *Calcified Tissue International*, **47**, 127–9.

Kanis, J.A. and Passmore, R. (1989). Calcium supplementation of the diet. Not justified by present evidence. *British Medical Journal*, **298**, 131–40, 205–8.

Karaplis, A.C. *et al*. (1994). Lethal skeletal dysplasia from targeted disruption of the parathyroid hormone-related peptide gene. *Genes and Development*, **8**, 277–89.

Khosla, S., Peterson, J.M., Egan, K., Jones, J.D., and Riggs, B.L. (1994). Circulating cytokine levels in osteoporotic and normal women. *Journal of Endocrinology and Metabolism*, **79**, 707–11.

Kielty, C.M. *et al*. (1995). Marfan syndrome; fibrillin expression and microfibrillar abnormalities in a family with predominant ocular defects. *Journal of Medical Genetics*, **32**, 1–6.

Kronenberg, H.M. *et al*. (1986). Structure and expression of the human parathyroid hormone gene. *Recent Progress in Hormone Research*, **42**, 641–62.

Lee, B., Vissing, H., Ramirez, F., Rogers, D., and Rimoin, D. (1989). Indentification of the molecular defect in a family with spondyloepiphyseal dysplasia. *Science*, **244**, 978–80.

Lefkowitz, R.J. (1995). G proteins in medicine. *New England Journal of Medicine*, **332**, 186–7.

Le Merrer, M. *et al*. (1994). A gene for achondroplasia–hypochondroplasia maps to chromosome 4p. *Nature Genetics*, **6**, 318–21.

Levine, M.A. (1991). The McCune-Albright syndrome. The ways and wherefores of abnormal signal transduction. *New England Journal of Medicine*, **325**, 1738–40.

MacIntyre, I. (1986). The hormonal regulation of extracellular calcium. *British Medical Bulletin*, **42**, 343–52.

Manolagas, S.C. and Jilka, R.H. (1995). Bone marrow, cytokines, and bone remodelling. *New England Journal of Medicine*, **332**, 305–11.

Martin, T.J. (1990). Properties of parathyroid hormone-related protein and its role in malignant hypercalcaemia. *Quarterly Journal of Medicine*, **76**, 771–86.

Martin, T.J., Ng, K.W., and Nicholson, G.C. (1988). Cell biology of bone. *Baillière's Clinical Endocrinology and Metabolism*, **2**, 1–29.

Martin, T.J., Ng, K.W., and Suda, T. (1989). Bone cell physiology. *Endocrinology and Metabolism Clinics of North America*, **18**, 833–58.

Mason, D.J., Suva, L.J., Stueckle, S., Hillam, R., and Skerry, T.M. (1995). Candidate genes in the osteocyte response to mechanical loading in vivo. *Bone and Tooth Society Meeting*, March (Abstract).

McIntosh, I., Abbott, M.H., Warman, M.L., Olsen, B.R., and Francomano, C.A. (1994). Additional mutations of type X collagen confirm COL10A1 as the Schmid metaphyseal chondrodysplasia locus. *Human Molecular Genetics*, **3**, 303–7.

Morrison, N.A. *et al*. (1994). Prediction of bone density from vitamin D receptor alleles. *Nature*, **367**, 284–7.

Mundy, G.R. (1989). Hypercalcaemia of malignancy. In *Metabolic bone disease: cellular and tissue mechanisms* (ed. C.S. Tam, J.N.M. Heersche, and T.M. Murray), pp. 146–56. CRC Press, Boca Raton, FL.

Nijweide, P.J., Van Der Plas, A., and Olthof, A.A. (1988). Osteoblastic differentiation. *Ciba Foundation Symposium*, **136**, 61–77.

Noda, M. (1993). *Cellular and molecular biology of bone*, (1st edn). Academic Press, New York.

Nordin, B.E.C. (1990). Calcium homeostasis. *Clinical Biochemistry*, **23**, 3–10.

Nordin, B.E.C. and Heaney, R.P. (1990). Calcium supplementation of the diet: justified by present evidence. *British Medical Journal*, **300**, 1056–60.

Orimo, H., Hayashi, Z., Watanabe, A., Hirayama, T., Hirayama, T., and Shimada, T. (1994). Novel missense and frameshift mutations in the tissue-nonspecific alkaline phosphatase gene in a Japanese patient with hypophosphatasia. *Human Molecular Genetics*, **3**, 1683–4.

Orloff, J.J., Wu, T.L., and Stewart, A.F. (1989). Parathyroid-like proteins: biochemical responses and receptor interactions. *Endocrine Reviews*, **10**, 476–95.

Owen, M. and Friedenstein, A.J. (1988). Stromal stem cells: marrow derived osteogenic precursors. *Ciba Foundation Symposium*, **136**, 42–60.

Pereira, L. *et al*. (1994). A molecular approach to the stratification of cardiovascular risk in families with Marfan's syndrome. *New England Journal of Medicine*, **331**, 148–53.

Pocock, N.A., Elsman, J.A., and Hopper, J.L. (1987). Genetic determinants of bone mass in adults: a twin study. *Journal of Clinical Investigation*, **80**, 706–10.

Pollock, M.R. *et al*. (1994). Autosomal dominant hypocalcaemia caused by a Ca^{2+}-sensing receptor gene mutation. *Nature Genetics*, **8**, 303–7.

Prior, J.C., Vigna, Y.M., Scheichter, M.T., and Burgess, A.E. (1990). Spinal bone loss and ovulatory disturbances. *New England Journal of Medicine*, **323**, 1221–7.

Prockop, D.J. and Kivirikko, K.I. (1984). Heritable diseases of collagen. *New England Journal of Medicine*, **311**, 376–86.

Pyeritz, R.E. (1990). Marfan syndrome. *New England Journal of Medicine*, **323**, 987–9.

Pyeritz, R.E. (1993). The Marfan syndrome. In *Connective tissue and its heritable disorders*, (1st edn) (ed. P.M. Royce and B. Steinmann), pp. 437–68. Wiley-Liss, New York.

Reardon, W., Winter, R.M., Rutland, P., Pulleyn, L.J., Jones, B.M., and Malcolm, S. (1994). Mutations in the fibroblast growth factor receptor 2 gene cause Crouzon syndrome. *Nature Genetics*, **8**, 98–103.

Reichel, H., Koeffler, H.P., and Norman, A.W. (1989). The role of the vitamin D endocrine system in health and disease. *New England Journal of Medicine*, **320**, 980–91.

Riggs, B.L. and Melton, L.J. (1988). *Osteoporosis. Etiology, diagnosis and management*. Raven Press, New York.

Riggs, B.L. and Melton, L.J. (1992). The prevention of osteoporosis. *New England Journal of Medicine*, **327**, 620–7.

Ritchie, H.H. *et al*. (1989). An ochre mutation in the vitamin D receptor gene causes hereditary 1,25-dihydroxy vitamin D_3 resistant rickets in three families. *Proceedings of the National Academy of Sciences (USA)*, **86**, 9783–7.

Robey, P.G. (1989). The biochemistry of bone. *Endocrinology and Metabolism Clinics of North America*, **18**, 859–902.

Robey, P.G. (1994). Normal bone formation In *Bone formation and repair*, (1st edn.) (ed. C.T. Brighton, G. Friedlaender, and J.M. Lane), pp. 3–12. American Academy of Orthopedic Surgeons, Rosemount, IL.

Rosati, R. *et al*. (1994). Normal long bone growth and development in type X collagen-null mice. *Nature Genetics*, **8**, 129–35.

Rousseau, F. *et al*. (1994). Mutations in the gene encoding fibroblast growth factor receptor-3 in achondroplasia. *Nature*, **371**, 252–4.

Rowe, P. (1994). Molecular biology of hypophosphataemic rickets and oncogenic osteomalacia. *Human Genetics*, **94**, 457–67.

Royce, P.M. and Steinmann, B. (1993). *Connective tissue and its heritable disorders*, (1st edn). Wiley-Liss, New York.

Russell, R.G.G. (1990). Bone cell biology: the role of cytokines and other mediators. In *Osteoporosis 1990*, (ed. R. Smith), pp. 9–33. Royal College of Physicians Conference Publications, London.

Rut, A.R., Hewison, M., Kristjansson, K., Luisi, B., Hughes, M.R., and O'Riordan, J.L.H. (1994). Two mutations causing vitamin D resistant rickets: modelling on the basis of steroid hormone receptor DNA-binding domain crystal structure. *Clinical Endocrinology*, **41**, 581–90.

Sambrook, P.N. *et al*. (1994). Mechanisms of rapid bone loss following cardiac transplantation. *Osteoporosis International*, **4**, 273–6.

Sampath, T.K., Muthukumaran, N., and Reddi, A.H. (1987). Isolation of osteogenin, an extra-cellular associated, bone inductive protein, by heparin affinity chromatography. *Proceedings of the National Academy of Sciences (USA)*, **84**, 7109–13.

Sandy, J.R., Meghi, S., Farndale, R.W., and Meikle, M.C. (1989). Dual elevation of cyclic AMP and inositol phosphates in response to mechanical deformation of murine osteoblasts. *Biochimica et Biophysica Acta*, **1010**, 265–9.

Schipani, E., Kruse, K., and Juppner, H. (1995). A constitutively activated PTH/PTHrP receptor (PTH-r) causes Jansen type metaphyseal chondrodysplasia (JMC). *Bone*, **16** (No. 1 supplement), 1015 (Abstract).

Schwindinger, W.F., Miric, A., Zimmerman, D., and Levine, M.A. (1994). A novel $G_s a$ mutant in a patient with Albright hereditary osteodystrophy uncouples cell surface receptors from adenylyl cyclase. *The Journal of Biological Chemistry*, **269**, 25387–91.

Seeman, E., Hopper, J.L., and Bach, L.A. (1989). Reduced bone mass in daughters of women with osteoporosis. *New England Journal of Medicine*, **320**, 554–8.

Seibel, M.J., Robins, S.P., and Bilezikian, J.P. (1992). Urinary pyridinium cross links of collagen. Specific markers of bone resorption in metabolic bone disease. *Trends in Endocrinology and Metabolism*, **3**, 263–70.

Shiang, R. *et al*. (1994). Mutations in the transmembrane domain of FGFR3 cause the most common genetic form of dwarfism, achondroplasia. *Cell*, **78**, 335–43.

Sly, W.S., Hewett-Emmett, D., Whyte, M.P., Yu, Y.L., and Tashian, R.E. (1983). Carbonic anhydrase II deficiency identified as the primary defect in the autosomal recessive syndrome of osteopetrosis with renal-tubular acidosis and cerebral calcification. *Proceedings of the National Academy of Sciences (USA)*, **80**, 2752–6.

Smith, R. (1984). Recent advances in the metabolism and physiology of bone. In *Recent advances in physiology*, **10** (ed. P.F. Baker), pp. 317–48. Churchill Livingstone, Edinburgh.

Smith, R. (1986). Osteogenesis imperfecta. *Clinics in the Rheumatic Diseases*, **12**, 655–89.

Smith, R. (1990*a*). *Osteoporosis 1990*. Royal College of Physicians Conference Publications, London.

Smith, R. (1990*b*). Asian rickets and osteomalacia. *Quarterly Journal of Medicine*, **76**, 899–901.

Smith, R. (1992). Paget's disease of bone. Advances and controversy. *British Medical Journal*, **305**, 1379–80.

Smith, R. (1993). Heritable metabolic bone diseases, chondrodysplasia and skeletal poisons. In *Metabolic bone and stone disease*, (3rd edn) (ed. B.E.C. Nordin, A.G. Need, and H.A. Morris), pp. 213–48. Churchill Livingstone, Edinburgh.

Smith, R. (1994). Osteogenesis imperfecta: from phenotype to genotype and back again. *International Journal of Experimental Pathology*, **75**, 233–41.

Smith, R. (1995). Disorders of the skeleton. In *Oxford textbook of medicine*, (3rd edn) (ed. D.J. Weatherall, J.G.G. Ledingham, and D.A. Warrell), pp. 3055–97. Oxford University Press, Oxford.

Smith, R. and Triffitt, J.T. (1986). Bone in muscles: the problems of soft tissue ossification. *Quarterly Journal of Medicine*, **76**, 899–901.

Spiegel, A.M. (1990). Albright's hereditary osteodystrophy and defective G proteins. *New England Journal of Medicine*, **322**, 1461–2.

Spranger, J., Winterpacht, A., and Zabel, B. (1994). The type II collagenopathies: a spectrum of chondrodysplasias. *European Journal of Paediatrics*, **153**, 56–65.

Storm, E.E., Huynth, T.V., Copeland, N.G., Jenkins, N.A., Kingsley, D.M., and Lee, S-J. (1994). Limb alteration in brachypodism mice due to mutations in a new member of the TGFb-superfamily. *Nature*, **368**, 639–42.

Sykes, B. (1990). Bone disease cracks genetics. *Nature*, **348**, 18–20.

Sykes, B. *et al*. (1990). Consistent linkage of dominantly inherited osteogenesis imperfecta to the type I collage loci COL1A1 and COL1A2. *American Journal of Human Genetics*, **46**, 293–307.

Tabas, J.A. *et al*. (1991). Bone morphogenetic protein: chromosomal localisation of human genes for BMP1, BMP2A and BMP3. *Genomics*, **9**, 283–9.

Tam, C.S., Heersche, J.N.M., and Murray, T.M. (1989). *Metabolic bone disease: cellular and tissue mechanisms.* CRC Press, Boca Raton, FL.

Termine, J.D. (1988). Non collagen proteins of bone. *Ciba Foundation Symposium*, **136**, 178–202.

Thakker, R.V. (1992). Molecular genetics of mineral metabolic disorders. *Journal of Inherited Metabolic Disorders*, **15**, 592–609.

Thakker, R.V. and O'Riordan, J.L.H. (1988). Inherited forms of rickets and osteomalacia. *Baillière's Clinical Endocrinology and Metabolism*, **2**, 157–91.

Thakker, R.V. *et al*. (1989). Association of parathyroid tumours in multiple endocrine neoplasia. Type I with loss of alleles on chromosome 11. *New England Journal of Medicine*, **321**, 218–24.

Tickle, C. (1994). On making a skeleton. *Nature*, **368**, 587–8.

Triffitt, J.T. (1987). The special proteins of bone tissue. *Cllinical Science*, **72**, 399–408.

Van der Rest, M. and Garrone, R. (1991). The collagen family of proteins. *FASEB Journal*, **5**, 2814–23.

Vaughan, J. (1981). *The physiology of bone*, (3rd edn). Clarendon Press, Oxford.

Velinov, M., Slaugenhaupt, S.A., Stoilov, I., Scott, C.I., Jr., Gusella, J.F., and Tsipouras, P. (1994). The gene for achondroplasia maps to the telomeric region of chromosome 4p. *Nature Genetics*, **6**, 314–17.

Vissing, H., D'Alessio, M., Lee, B., Ramirez, F., Godfrey, M., and Hollister, D.W. (1989). Glycine to serine substitution in the triple helical domain of Pro a1(II) collagen results in a lethal perinatal form of short-limbed dwarfism. *Journal of Biological Chemistry*, **264**, 18265–7.

Wallis, G.A., Starman, B.J., Zinn, A.B., and Byers, P.H. (1990). Variable expression of osteogenesis imperfecta in a nuclear family is explained by soamtic mosaicism for a lethal point mutation in the a1(I) gene (COL1A1) of type I collagen in a parent. *American Journal of Human Genetics*, **46**, 1034–40.

Wallis, G.A. *et al*. (1994). Amino acid substitutions of conserved residues in the carboxyl-terminal domain of the a1(X) chain of type X collagen occur in two unrelated families with metaphyseal chondrodysplasia type Schmid. *American Journal of Human Genetics*, **54**, 169–78.

Wang, Z-Q. *et al*. (1992). Bone and haematopoietic defects in mice lacking *c-fos. Nature*, **360**, 741–5.

Weiss, J.J. *et al*. (1988). A missense mutation in the human liver/bone/kidney alkaline phosphatase gene causing a lethal form of hypophosphatasia. *Proceedings of the National Academy of Sciences (USA)*, **85**, 8666–9.

Whyte, M.P. (1990). Heritable metabolic and dysplastic bone disease. *Endocrinology and Metabolism Clinics of North America*, **19**, 133–72.

Wilson, L.C., Oude Luttikhuis, M.E.M., Clayton, P.T., Fraser, W.D., and Trembath, R.C. (1994). Parental original of Gsa gene mutations in Albright's hereditary osteodystrophy. *Journal of Medical Genetics*, **31**, 835–9.

Winterpacht, A., Hilber, M., Schwarze, U., Mundlos, S., Spranger, J., and Zabel, B. (1994). Autosomal dominant spondylarthropathy due to a type II procollagen gene (COL2A1) point mutation. *Human Mutation*, **4**, 257–62.

Wozney, J.M. (1993). Bone morphogenetic proteins and their gene expression. In *Cellular and molecular biology of bone*, (1st edn) (ed. M. Noda), pp. 131–67. Academic Press, New York.

Wozney, J.M. *et al*. (1988). Novel regulators of bone formation: molecular clones and activities. *Science*, **242**, 1528–34.

Wynne-Davies, R., Hall, C.M., and Apley, A.G. (1985). *Atlas of skeletal dysplasias*, (1st edn). Churchill Livingstone, Edinburgh.

Yoshida, H. *et al*. (1990). The murine mutation osteopetrosis is in the coding region of the macrophage colony stimulating factor gene. *Nature*, **345**, 442–3.

Young, M.F., Ibaraki, K., Kerr, M.J., and Heegaard, A.-M. (1993). Molecular and cellular biology of the major non-collagenous proteins in bone. In *Cellular and molecular biology of bone*, (1st edn) (ed. M. Noda), pp. 191–234. Academic Press, New York.

2.5 The physiology of the joint and its disturbance in inflammation

Paul Mapp, Cliff. R. Stevens and David R. Blake

No part of the physiology of bones abounds more in hypotheses and less in discoveries than the history of the synovial system. Many discussions and few facts; a long series of assumed principles; a brief assemblage of proofs; this is the analysis of all the hitherto known works on the subject. The notions already received throw but little light on those yet to be acquired.

Marie-Francois-Xavier Bichat (1771–1802).

This chapter covers selected aspects of the physiology of the joint in health and in disease. The normal synovium and synovial fluid will be described. Other areas of the joint such as the cartilage have been covered in previous chapters. Further, we will characterize the structural changes to the synovium and chemical changes in the fluid that occur in chronic inflammation and address the physiological implications of these changes. Finally we will evaluate possible mechanisms by which these changes may lead to joint damage.

The synovium

The normal human body contains 187 synovial joints. In this article the term synovium will be used to describe the connective tissue bounded by the fibrous joint capsule on one side and by the joint space on the other.

One of the most striking histological features of the normal synovium is a lack of cellularity, apart from a discontinuous layer of cells, the synovial intimal cells, on the internal surfaces of the joint. Although variously described as synoviocytes, lining cells, or surface cells, here they will be referred to as intimal cells.

Intimal cells have been described by many authors but a consensus has yet to be reached as to their cellular phenotype, relative proportions, and functions. They form a thin discontinuous layer lining the joint space and separating it from the fibro-fatty tissue and fibrous capsule that surrounds the whole joint. The intimal cells line all surfaces of the joint other than the cartilage and menisci. Synovial intimal cells in common with other connective tissues are derived in the embryo from the mesenchyme either directly or indirectly via the blastema (O'Rhailly and Gardner 1978). What are apparently intimal cells may also be generated under certain other circumstances in connective tissue, as in the case of a pseudoarthrosis, and as regenerated synovium covering the connective tissue of the joint capsule following synovectomy. Similar cells can be created in animal models when air is injected into loose connective tissue. The whole synovium is bathed in synovial fluid since there is no true basement membrane beneath the intimal cells. There is neither a continuous layer of cells nor any basement underneath them and, therefore, the term synovial membrane is strictly speaking a misnomer.

Histology and ultrastructure of the normal synovium

Cellular morphology

By routine light microscopy the intimal cell layer appears as a homogeneous population of varying thickness. In some areas the cells may lie three or four deep, whilst in others places there are gaps where no intimal cells are seen and the underlying stroma is apparently in direct contact with the joint fluid. The underlying tissue, however, may have markedly different appearances and this has led some authors to classify synovia on this basis. Divisions into areolar, fibrous, and fatty synovia have been applied to the synovium, although other researchers have recognized intermediate types, such as fibroareolar synovium. In general, the normal synovium appears as a quiescent connective tissue containing some macrophages and fibroblasts with a thin layer of cells opposing the joint space. The particular nature of the underlying connective tissue is unlikely to be of significance.

Ultrastructural examination of the synovium has yielded much data but little agreement with regard to the predominant intimal cell type or to their function. Intimal cells were first described by Barland *et al*. (1962) and divided into types A and B on a purely morphological basis. Other authors have made similar distinctions but have included an intermediate cell type, postulating that this is a precursor of the two intimal cell types. The type A cell is characterized morphologically by a prominent Golgi apparatus, numerous vesicles and vacuoles, and many filopodia and mitochondria. This description is typical of a tissue macrophage. Type B cells, in contrast, contain large amounts of endoplasmic reticulum and a few large vacuoles, correlating well with ultrastructural descriptions of fibroblasts at other sites. The weight of scientific evidence suggests that the type A cell is a bone marrow-derived macrophage and the type B cell is a mesenchymal fibroblast. The reader is particularly directed to the work of Edwards and Willoughby (1982) which demonstrates the arrival in the synovium of bone marrow-derived macrophages using intracellular granules as a genetic marker.

Vascular system

Capillaries occur within the synovium and also the underlying tissues (fat, skeletal muscle, etc.). The synovium is considerably more vascular than the structures supporting the joint, such as the capsule, ligaments, and tendons.

The first description of the vasculature within joints was by Hunter in 1743, who described the 'circulus articuli vasculosus'. He stated: 'All around the neck of the bone there is a greater number of arteries and veins which ramify into smaller branches and communicate with one another by frequent anastomoses like those of the mesentery. This might be called the circulus articuli vasculosus, the vascular border of the joint.'

At the margins of articular cartilage the synovium forms villi and folds, into which the plexus sends arcades of capillaries. The circulus articuli vasculosus anastomoses with deeper perichondrial vessels that communicate with the bone marrow vasculature. There are anastomosing plexuses in the deeper layers of areolar lining tissue, with decreasing vascular calibre towards the surface. A plexus of arterioles and venules also forms a quadrilateral array close to the surface of the synovium. This plexus yields capillary loops which pass through or immediately below the layer of synovial intimal cells. The surface of the tissue carries a much denser capillary bed than the deeper tissue layers. The dimensions of the mesh of the arteriolar/venous plexus are between 0.9 and 1.5 mm.

Arteriovenous anastomoses have been demonstrated to occur commonly in the joints of humans and animals. The pattern of the arteriovenous anastomoses suggests an important role in the regulation of the articular and, perhaps, the epiphysial blood flow, possibly by a shunt mechanism redistributing the blood flow between certain vascular beds. The arteriovenous anastomoses between muscle, capsule, and bone help to explain the production of endosseous blood stasis during rest or muscular

inactivity and suggests that the anastomoses are related to regional haemodynamic regulation.

The vascularity of normal synovium appears to depend considerably on the type of underlying tissue. Capillary density is greatest where the substructure is areolar or adipose tissue and is virtually zero over fibrous areas. Away from the margins of articular cartilage the synovium is flatter, and the synovial capillaries form a flat, loose, polyhedral network supplied by several articular arteries. The capillaries are sinuous, which may reduce wall stress during motion, and are occasionally so coiled as to resemble glomerular capillaries.

Morphometry of normal human knee synovia has provided data on the spatial distribution of these vessels. The density of the capillaries is 240/mm ^{2} of synovial lining and their modal depth is 35 μm from the intimal cell/joint space boundary. The mean intercapillary distance is 17 μm (Stevens *et al*. 1991); similar data have been obtained in rabbit synovia (Knight and Levick 1983). From these data the functionally optimal range of vascular parameters for the synovium can be extrapolated. The bulk of the constituents of the synovial fluid are derived from this blood supply.

Nerve supply

Joints are supplied by both primary and accessory nerves. Primary nerves are branches of peripheral nerves passing near to the joint, whilst accessory nerves are branches of intramuscular nerves crossing the joint capsule. Some joints, such as the knee and ankle, also receive a nerve supply from cutaneous nerves in the overlying skin. The nerves to any one particular joint always arise from more than one level in the spinal cord. About 50 per cent of the axons that comprise articular nerves are less than 5 μm in diameter, those that are less than 2 μm in diameter are unmyelinated. These fibres carry nociceptive information with a slow conduction velocity. The nerve supply to the synovium was, until relatively recently, thought to be sparse. However, it is now known that the synovium contains a good supply of unmyelinated nerve fibres (Mapp *et al*. 1990a) These are of two types: the postganglionic sympathetic adrenergic fibres located around the larger blood vessels are responsible for the control of articular blood flow; the unmyelinated C fibres, in contrast, are responsible for pain transmission. The latter group of fibres are not normally active and are thought only to fire during tissue damage, either mechanical or chemical.

The unmyelinated fibres arise from cell bodies which are located in the dorsal root ganglion, close to the spinal cord. In addition to its peripheral projection the cell body also has a central projection to the spinal cord. The fine diameter nerve fibres terminate in the superficial layers of the spinal cord, laminae I and II of the dorsal horn, from where they synapse with ascending fibres in the spinal cord. The cell bodies are the site of manufacture of neuropeptides which are transported both centrally and peripherally along the nerve axons. The functions of these peptides are not clearly understood in the central nervous system but some, for instance substance P, are thought to be involved in pain transmission. Peripherally, where the effects of these peptides are much easier to investigate, they are clearly involved in the inflammatory process, both its induction and modulation. They have the capacity individually or in synergism with other neuropeptides or mediators to modulate, mediate, or prime for an inflammatory response. In addition, they have an effect on vascular tone.

Having examined the structure of the synovium, we turn to the fluid that is derived from it and bathes the whole joint cavity.

Synovial fluid

The study of normal synovial fluid is difficult since the aspiratable volume of fluid is very small, in the range 0 to 4 ml. The fluid is regarded as an ultrafiltrate of the plasma, but in addition components secreted by the synovial cells are also present. The most important of these is hyaluronate. Hyaluronate is a linear repeating disaccharide, b-D-glucuronyl-b-D-*N*-acetyl-glucosamine, of high molecular mass (upwards of 10 000 kDa). Hyaluronate forms the central axis of the proteoglycan aggregates necessary for the functional integrity of articular cartilage and other extracellular matrices. It comprises the major macromolecular species of the synovial fluid and is responsible for the unique visco-elastic properties of the synovial fluid. More recently another component of human synovial fluid has been identified, lubricin, which as its name implies is thought to be involved in the lubrication of the joint (Swann *et al*. 1985). Lubricin is a protein with oligosaccharide side-chains and has a molecular mass of 166 kDa. *In vitro* its lubricating activity is concentration dependent, and at concentrations thought to be present in the joint it acts as a highly efficient lubricant. A similar molecule has been isolated from bovine synovial fluid. As well as acting as a lubricant, synovial fluid is also the source of nutrition for the articular cartilage.

The mechanism of formation of the synovial fluid is thought to be similar to that of interstitial fluid in any other body cavity. The flow across the capillary wall is driven by the difference between the pressure in the capillary and the joint fluid. This is opposed by both the osmotic and colloid pressure gradients, the Starling hypothesis (Starling 1896). Small physiological molecules of molecular mass less than 10 kDa are in full equilibrium with the plasma. A graphic plot of the diffusion coefficient of a molecule (which takes into account mass and size) against the synovial permeability yields a broadly linear curve. Larger molecules such as proteins have only restricted access to the normal synovial fluid. The total protein content of the synovial fluid is 13 mg/l (compared with the serum concentration of 65 to 80 mg/l); most of which is albumin. Proteins of higher molecular weight, such as fibrinogen, are excluded. Whilst proteins arrive in the joint at rates inversely proportional to their molecular size, their rates of egress are similar. This is a reflection of the fact that protein is cleared from the synovial fluid by bulk flow through the lymphatic system, which is not dependent on the size of the molecule.

Intra-articular pressure

The overall pressure within the normal joint is subatmospheric, pressures of –30 to –60 mmH_2O are typically recorded at rest and during movement. This is true both in relaxation and during weight bearing. The 'joint space' is therefore only an apparent space, since atmospheric pressure pressing on the soft tissues around the joint will cause the synovium to be in close apposition to the articular cartilage surface. This has the effect of optimizing the transfer of nutrients from the superficial blood vessels in the synovium to the avascular cartilage.

The intra-articular pressure is to be distinguished from the pressure that arises between the two cartilage surfaces when under load.

Inflammatory disease

Histology of the inflamed synovium

Examination of the synovium in chronic inflammation reveals a wide variety of morphological changes, many of which may alter the physiology of the joint. The intimal cell layer becomes thickened, both by an increase in cell number and an increase in cell size; hyperplasia and hypertrophy, respectively. The intimal cell layer is six to eight cells deep. Underlying the intimal layer, changes typical of chronic inflammation are seen. There is a large increase in the number of cells, comprising mainly macrophages and lymphocytes. The lymphocytes accumulate perivascularly around the postcapillary venules and may organize into foci resembling lymphoid follicles; plasma cells are seen at the margins of the follicles. Large numbers of macrophages are also seen scattered widely throughout the tissue, these are thought to be migrating to the intimal cell layer replacing the turnover in type A cells.

Vasculature

The dominating vascular feature of chronic inflammatory disease is a vasculitis localized to the venules and capillaries, although it has been demonstrated that arterioles also become involved to a lesser degree. Using vital microscopy to study the rheumatoid synovium it has been found that venular dilatation of varying calibre, resulting in uneven outline and slow, almost stagnated, corpuscular flow, is also a feature. In severe cases, focal and segmental venular-capillary necrosis and thrombosis are found in association with connective tissue necrosis. However, in less severe disease, vascular derangement is largely reversible and manifested by excessive venular-capillary dilatation, leakage, and occasionally by focal, partial, or temporary microvasculature obstruction.

Of particular relevance to the physiology of the joint are the changes seen in the synovial vasculature. Assuming that the data on capillary density and depth determined in normal joints represent the optimal situation, there is a significant reduction in vascularity in the functionally important superficial region of the synovium. The capillary density in rheumatoid arthitis is reduced to about one-third of the normal value, from 240/mm ^{2} to 80/mm^{2}; mean intercapillary distance is increased from 17 μm to 37 μm, reflecting not only an increase in spatial distribution of the vasculature but also a thickened synovial lining. The average capillary distance from the joint cavity is increased from 32.5 μm to 93.3 μm (Stevens *et al*. 1991). This apparent burial of the vasculature is due to the thickening of the synovial lining layer which would appear to proceed at a greater pace than angiogenesis (Fig. 1). In addition to these changes multilamination of the basement membranes (Matsubara *et al*. 1983) surrounding the blood vessel is also seen, due to stimulation of endothelial cells (by unknown mediators) to produce excessive cellular components of basement membrane; this represents a further barrier to nutrient exchange. This is compounded by the deposition of fibrin which is often seen

on the surface of the inflamed synovium.

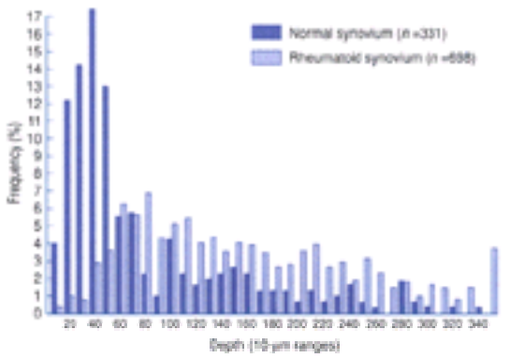


Fig. 1 Histogram showing the distribution of blood vessels in normal and rheumatoid synovia. The figures referred to relate to the distance of the vessels from the intimal cell surface.

The height of endothelium in small vessels is increased in chronic synovitis; an increase in the number of vessels with histochemical features of so-called high endothelial venules is seen in tissue from patients with rheumatoid arthritis. These vessels are thought to be the main site of lymphocyte traffic in lymph fluid and could represent a facilitated pathway for lymphocyte entry in pathological circumstances. One of the most frequent histological findings affecting small vessels in chronic inflammation is that of perivascular cuffing by lymphocytes.

Nerve supply

In common with the vasculature, the nerve supply which normally extends to the synovial/joint space interface, is buried deeper within the rheumatoid synovium. It is probable that this is due to proliferation of the synovium, but it may be due to release of factors which are inhibitory to nerve proliferation or the inactivation of peptides which promote neuronal outgrowth, such as protease nexin 1 (Meier *et al*. 1990). The blood vessels that are in the superficial layers of the synovium, up to 150 μm from the joint space, appear not to be innervated by postganglionic sympathetic nerve fibres. This would imply that there is poor control over the superficial vascular supply. Free nerve fibres, assumed to be sensory in nature, are also absent from this area.

However, nerves fibres, both free and perivascular, are present in the deeper inflamed tissues (Fig. 2). These fibres appear to contain reduced amounts of neuropeptide, as determined by immunocytochemistry. The probable explanation for this is that the stored peptide is being released into the joint cavity. The justification for this view, and its consequences, are discussed in a later section of this chapter.

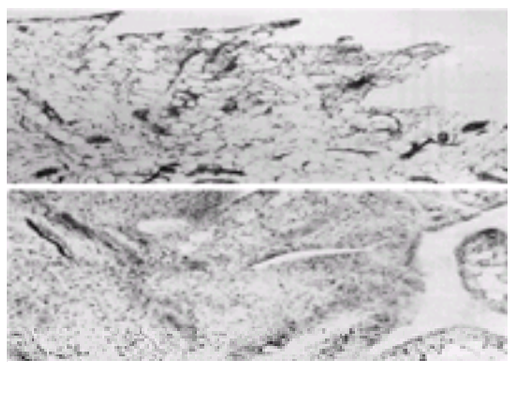


Fig. 2 Photomicrographs of normal (a) and rheumatoid (b) synovia. The specimens have been stained by immunocytochemistry to show all nerve fibres. Note the rich innervation of the normal specimen compared with the paucity of nerve fibres in the disease state.

In summary the inflamed joint shows the following characteristics:

1. infiltration with numerous inflammatory cells;
2. the intimal layer becomes hypertrophic and hyperplastic;
3. the vasculature is buried deeper within the synovium rendering the tissue relatively hypoxic;
4. the nervous supply is also buried deeper in the synovium, perhaps releasing cells from normal neuronal regulation;
5. the synovial fluid markedly increases in volume and changes in composition.

We now discuss how some of these basic changes induced by inflammation in turn influence the inflammatory process.

Ischaemia and hypoxia

The presence of a large cellular infiltrate in the inflamed synovium would lead one to think that there is an increased utilization of oxygen, often referred to as increased oxygen demand. This increased demand is not readily quantifiable. A working definition of ischaemia is that the steady-state rate of energy production, i.e. ATP synthesis, cannot be supported by the existing blood flow. Since the necessary levels of energy production cannot be maintained this stimulates anaerobic metabolism which in turn gives rise to lactic acid production and a fall in venous oxygen tension.

In order to determine whether the joint is ischaemic, attempts have been made to measure the blood flow in the synovium. Some studies, using the xenon-153 clearance technique have reported an increased perfusion in chronically inflamed synovia (Dick *et al*. 1970). The validity of such studies is in question because they fail to take account of the possibility of arteriovenous shunting, which would give rise to the impression of perfusion but would result in ischaemic, non-perfused synovial pockets. Other studies using iodine-123 or tritiated water appear to show that perfusion in chronic synovitis is markedly reduced (Simkin and Pizzorno 1979.) The stage of the disease at which the perfusion is measured may be the all important factor here. In the early stages of rheumatoid arthritis, the temperature of the joint is raised relative to normal whereas in late-stage disease the joint is cooler than normal. The implication of this is that in the early stages of the disease blood flow increases but in the later stages it is decreased, the most affected joints being the ones with the least blood flow. Microvessel plugging by inflammatory cells also serves to reroute blood from the peripheral capillary beds. This is associated to a varying degree with microthrombi and extravascular aggregates of platelets.

Synovial fluid oxygen tension is more easily quantifiable than synovial ischaemia. The oxygen tension (PO_2) within the inflamed joint cavity (Lund-Olsen 1970) has been shown to be lower (27 mmHg) than in traumatic effusions (63 mmHg) or osteoarthritic knee fluids (43 mmHg). These results are confirmed by independent studies. In addition to a low oxygen tension, many studies have shown alterations in rheumatoid synovial fluid physiology which would be anticipated in a hypoxic state: namely, a raised carbon dioxide tension (PCO_2 up to 150 mmHg), raised lactate (up to 10 mmol), lowered glucose, and acidosis (pH as low as 6.6). The most comprehensive study summarizing these changes is by Falchuk *et al*. (1970), who showed that joints with the lowest oxygen tension (as low as 9 mmHg) also exhibited large increases in carbon dioxide tension and lactate; the same joints showed severe microvascular obliteration in the synovial membrane.

In summary the inflamed joints are hypoxic because of increased metabolic demand of the inflamed synovium, and the inadequate oxygen delivery due to poor perfusion of the inflamed joint. It is important to note that physiological measurements in synovial fluid can vary with the patient's mobility. This is evident from a recent

study which showed that movement of an inflamed joint can induce significant further decrease in the oxygen tension from a resting baseline (Tsoulfa *et al*. 1989). This observation has particular relevance to the concept of joint 'ischaemia-reperfusion' injury (Blake *et al*. 1989).

Synovial metabolism

The oxygen consumption of the rheumatoid synovial membrane per gram of excised tissue is approximately 20 times that of normal (Page-Thomas and Dingle 1955; Dingle and Page-Thomas 1956). This work also demonstrated that the activity of the glycolytic (Embden–Meyerhof) pathway for ATP production was markedly increased in rheumatoid compared with normal synovium. The raised metabolic rate of the rheumatoid synovium concomitantly raises the demand for ATP. The intracellular production of this molecule can be achieved by either the aerobic or anaerobic oxidation of glucose via the tricarboxylic acid (**TCA**) cycle or the glycolytic pathway respectively. The oxygen-dependent TCA cycle is a much more efficient producer of ATP than the anaerobic system and so is generally favoured in normoxic tissues. The fact that the rheumatoid synovium favours the anaerobic glycolytic pathway indicates its hypoxic nature. In support of this, synovial intimal cells in rheumatoid arthritis contain significantly more glyceraldehyde-3-phosphate and lactate dehydrogenase activity than those of normal tissue. These are major enzymes of the glycolytic pathway. Their increase is more reasonably explained by a response to tissue hypoxia than by elevated metabolic activity, as mitochondrial oxidation in the synovium is not similarly enhanced.

The terminal end product of the anaerobic oxidation of glucose is lactate. The ratio of lactate to glucose, therefore, should give an indication of oxygen status in the synovium. This is indeed the case, it has been shown that lowered glucose and raised lactate levels correlate well with falls in oxygen tension and pH. Proton nuclear magnetic resonance (**NMR**) spectroscopy confirms the peculiar anaerobic environment in the inflamed joint. Comparison of paired synovial fluids and sera show that the concentrations of lactate, 3-D-hydroxybutarate and acetoacetate are all much higher in the synovial fluid than the serum pair. Additionally, high levels of ketone bodies are also found. The hypoxic conditions prevailing in the inflamed joint lead to an increase in fatty acid oxidation. However, the unavailability of NAD+ results in an inability of the citric acid cycle to oxidize fully acetyl coenzyme A. This process leads to a build-up of unoxidizable metabolites, ketone bodies, recently demonstrated by NMR spectroscopy and illustrated in Fig. 3.

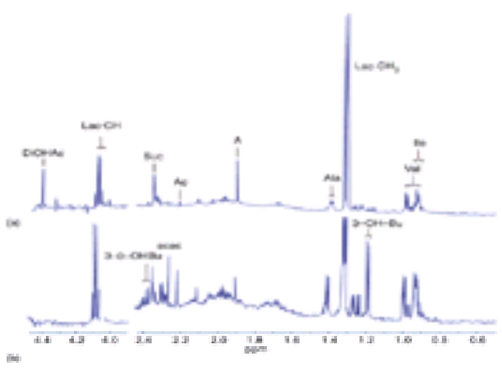


Fig. 3 600 MHz proton nuclear magnetic resonance spectra of low-molecular-mass, non-protein bound components of isolates from osteoarthritic (a) and rheumatoid (b) synovial fluids. DiOHAc, dihydroxyacetone; Lac, lactate; Suc, succinate; Ac, acetone; A, acetate; Ala, alanine; Val, valine; Ile, isoleucine; 3-D-OH Bu, 3-D-hydroxybutarate; Acac, acetoacetate. (Samples were run in a Varian 600 MHz machine operating at 600 MHz at ambient temperature.)

Acute exacerbation of chronic hypoxia

A distinguishing property of the joint is its ability to move. The inflamed joint of a mobile patient is unavoidably subjected to movement. This situation has profound significance in joint pathogenesis through the effects of hypoxic–reperfusion injury (Woodruff *et al*. 1986).

The environment for hypoxia and reperfusion in rheumatoid synovitis is created by the unique topography of the component parts of a joint which predisposes the synovium to pressure-induced fluctuations in blood supply. The synovium is the innermost confine of the joint space and its fluid. This moveable and compressible stroma is backed by the more rigid confines of the musculature and ligaments. It might be anticipated, therefore, that centripetal thickening and increasing fluid volume generates a pressurized system. This is indeed the case. We have confirmed (Blake *et al*. 1989) that the resting intra-articular pressure in chronically inflamed joints is slightly above atmospheric pressure (+5 to 10 mmHg) compared with that of normal joints which is subatmospheric (–5 mmHg). When normal joints (knee) are exercised, the pressure falls further, whereas in chronically inflamed joints intra-articular pressure rises, sometimes as high as 300 mmHg and always above the capillary perfusion pressure in inflammation of 30 to 60 mmHg; this is illustrated in Fig. 4. It is therefore suggested that in chronically inflamed joints, the intra-articular pressure rises during exercise above a critical level sufficient to occlude parts of the capillary bed, thus inducing acute ischaemia in an already hypoxic environment. Recent independent observations have supported the hypothesis of ischaemia induced by intra-articular pressure. James *et al*. (1990) have shown in patients with rheumatoid arthritis with knee effusions that intra-articular pressure elevations in the range encountered during daily activity can compromise blood flow. This effect is associated with increased synovial fluid lactate, elevated carbon dioxide tension and decreased pH. In rabbit knees intra-articular pressures as low as 19 mmHg cause synovial capillaries to assume a more flattened, elliptical profile (Levick 1990).

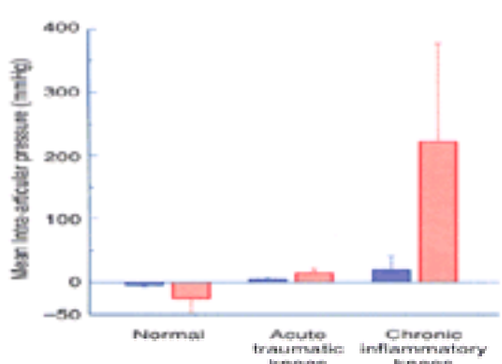


Fig. 4 Pressure changes on exercise. Solid bars, rest; hatched bars, exercise.

In mobile patients who may be subjecting their joints to pressure-induced ischaemia, subsequent rest allows reperfusion of blood and reoxygenation of the synovium, albeit to an inadequate degree. This eventuality, has pathological repercussions consequent on the generation of reactive oxygen metabolites which accompanies postischaemic reperfusion.

Biochemical consequences of synovial hypoxia

There are a number of biochemical consequences of hypoxia; calcium imbalance is arguably the most influential of these events. Intracellular homeostasis, with respect to Ca^{2+}, is dependent upon energy in the form of ATP. When ATP availability is compromised, as in the hypoxic synovium, cytosolic Ca^{2+} levels rise. Mitochondria are avid accumulators of Ca^{2+}, and when overloaded with Ca^{2+} their function is impaired (Nayler *et al*. 1979), resulting in the further depletion of ATP levels and the maintenance of a calcium influx. Many important cellular functions and control mechanisms are Ca^{2+} dependent. The consequences of hypoxia-derived Ca^{2+} imbalance must be considered as possible contributing factors in the pathogenesis of rheumatoid arthritis (Fig. 5). The mechanism for increased vascular permeability for instance, which is associated with inter-endothelial gap formation, is thought to involve changes in the actin cytoskeleton of the endothelial cell (Shasby *et al*. 1985). These changes can be elicited by calcium imbalance since actin binding proteins which regulate the endothelial cytoskeleton are calcium dependent (Hinshaw *et al*. 1988). The failure of cellular phosphorylation mechanisms also leads to accumulation of adenosine and of its breakdown products,

including hypoxanthine and xanthine which are the substrates for the xanthine oxidase enzyme system (see below). The outcome of these events in the hypoxic synovium confers the facility for the inappropriate generation of reactive oxygen metabolites upon reperfusion of blood. There is much evidence to show that postischaemic reperfusion of the synovium evokes the generation of reactive oxygen metabolites, despite the underlying hypoxic environment which is not conducive to most conventional radical generating systems. One proposed mechanism that can achieve this is dependent on the enzyme xanthine oxidase (McCord 1987), which is present within the capillaries of the synovium (Allen *et al*. 1987). The non-pathological dehydrogenase form of this enzyme oxidizes hypoxanthine and xanthine to uric acid using NAD+ as an electron acceptor. Under ischaemic conditions, however, this enzyme can be converted to an NAD+-independent form which catalyses the same reaction using molecular oxygen as an electron acceptor and resulting in superoxide anion generation. The inappropriate production of superoxide in a biological system can have catastrophic repercussions, particularly when the system facilitates the redox environment required for consequent hydroxyl radical (·OH) formation. The synovium is such a system, enabling the spontaneous or enzymatic conversion of superoxide to hydrogen peroxide, which can then fuel a Fenton-type reaction catalysed by decompartmentalized iron (Morris *et al*. 1986) forming the cytotoxic ·OH radical thus:

$$H_2O_2 + Fe^{2+} \rightarrow {\times}OH + OH^- + Fe^{3+}$$

Iron is decompartmentalized as a consequence of hypoxia and can be detected bound to citrate in the synovial fluid. In such a form it may promote hydroxyl radical formation.

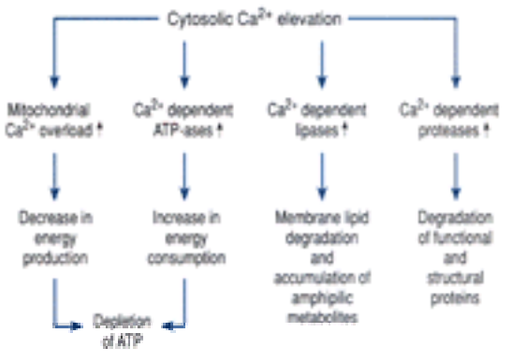


Fig. 5 Role of elevated cytosolic calcium in tissue injury.

Biological consequences of reactive oxygen metabolites

Biochemical analysis of synovial fluid reveals the effects of radical damage to endogenous biomolecules in rheumatoid arthritis.

Lipids

The cell membrane, composed primarily of polyunsaturated fatty acids, is a well-studied target for radical attack, and the resultant process of lipid peroxidation causes cell membrane damage. Polyunsaturated lipids undergoing the process of peroxidation give rise to the formation of lipid breakdown products such as lipid hydroperoxides, endoperoxides, aldehydes, and alkanes. These can be measured and correlate well with the clinical severity of the disease.

Proteins

Exposure of proteins to free radicals will lead to denaturation, cross-linking, aggregation, and fragmentation. There is much evidence of radical-mediated protein damage occurring in inflammatory synovitis. Immunoglobulin G (IgG) when exposed to free radicals *in vitro* develops a characteristic autofluorescence (excitation 360 nm, emission 454 nm) and forms both monomeric and aggregated complexes (Lunec *et al*. 1985). IgG with these characteristics is present in rheumatoid synovial fluid.

Additionally, the methionine residues on a_1-antitrypsin (the primary inhibitor of leucocyte elastase) are oxidized to sulphoxide adducts by free radicals, rendering the molecule biologically inactive. Such radical damaged a_1-antitrypsin is present in inflammatory synovial fluid.

Carbohydrates

Free radicals can also react with polysaccharides and induce fragmentation. The major macromolecule of the synovial fluid, which accounts for its viscosity, is the glycosaminoglycan hyaluronate. Loss of viscosity of the synovial fluid is thought to be due to the depolymerization of hyaluronate and an increase in the synovial concentration of dialysable hyaluronan fragments and saccharide monomers. The alternative explanation for the reduction in the molecular weight of hyaluronan is defective synthesis with premature termination of the polysaccharide chain. Pulse–chase labelling experiments have supported the view that the presence of short-chain molecules of hyaluronan in the synovial fluid is due to post-synthetic degradation rather than defective synthesis. As neither normal nor inflammatory synovial fluids contain any measurable hyaluronidase activity, it is inferred that the depolymerization of hyaluronan is the result of free radical attack. The degradation products have different biological properties from that of the native molecule. Low-molecular-mass fragments of hyaluronan are proangiogenic, whereas the high-molecular-mass hyaluronan has inhibitory effects on endothelial cell proliferation. Proliferation of both fibroblasts and mitogen-stimulated lymphocytes is similarly affected.

Effects on cells and cell viability

There are important effects of free radicals upon cells when considered as a whole. T lymphocytes in general have been shown to be susceptible to free radical attack. Free radicals are more toxic to T-cytotoxic/suppressor than to T-helper cells *in vitro* (Allan *et al*. 1986). These results support the speculation that free radical reactions, as well as inducing the denaturation of proteins, can prevent immune suppressor/cytotoxic cells from controlling autoimmune reactions with subsequent immunopathological consequences. As well as the cytotoxic aspects of free radical attack on whole cells, which occurs at high concentrations, the data (Murrell *et al*. 1990) also suggest that at low concentrations free radicals can stimulate proliferation of cells. Under *in vitro* conditions fibroblasts were stimulated to proliferate by low concentrations of free radicals (the fibroblasts themselves release their own free radicals), and fibroblasts were inhibited from proliferation but were still viable when the endogenous free radicals were inhibited.

Many other examples of oxidative damage to the joint have been reported. Apart from these damaging or cytotoxic properties of free radicals, they also have the capacity to interfere subtly with finely controlled physiological mechanisms. One example of this occurs in the control of vascular tone by vasoactive substances.

The activity of endothelium-derived relaxing factor (**EDRF**) has been ascribed to the nitrogen-centred radical nitric oxide, which is synthesized in endothelial cells by the Ca^{2+}- dependent enzymatic conversion of L-arginine. It activates the cytosolic guanylyl cyclase of the adjacent smooth muscle cells increasing the levels of cGMP and resulting in relaxation of the cells and hence vasodilation. Nitric oxide can interact with superoxide via peroxynitrite to produce the highly cytotoxic hydroxyl radical. *In vitro* this interaction abrogates the EDRF activity of nitric oxide. This is demonstrated by the ability of superoxide dismutase to enhance and prolong the relaxant effect of nitric oxide. Nitric oxide is unstable and rapidly degrades to nitrate and nitrite, which are relatively stable in biological fluids. In inflammatory joints the synovial fluid levels of nitrite exceed those of the paired serum samples, thus indicating local production. The implication of this is that nitric oxide is also produced locally in the synovium. Vascular tone is controlled by the opposing actions of relaxing and contracting factors released from the endothelium in response to neurogenic stimuli or platelet products. Superoxide has itself been implicated as an endothelium-derived contracting factor along with endothelin. An obvious vascular response to tissue ischaemia would be to lower its resistance to blood flow by vascular relaxation. This has been shown to be the case in some systems where hypoxia-induced release of prostacyclin (PGI_2) is sufficient to elicit endothelium-dependent dilation. Hypoxia also stimulates the release of EDRF from rabbit femoral artery and aorta. In the pulmonary circulation, however, it has been known for many years that hypoxia causes vascular constriction. The benefit of this is in rerouting the majority of the blood flow to the best oxygenated parts of the lung. Other pulmonary studies have shown that hypoxia blocks EDRF production and induces the

release of a constricting factor from the endothelium. These conflicting reports highlight the heterogeneity of endothelial cell responses to the same stimulus in different situations and environments, allowing the possibility of an apparently appropriate response to eventually become pathological.

Hypoxia has other quite pronounced effects on endothelial cells, for example on von Willebrand factor release (discussed later), reducing plasminogen activator activity, and increasing superoxide dismutase activity. The endothelium also produces other very powerful vasoactive substances which may be mediated by hypoxia. These include platelet-activating factor and the highly potent vasoconstrictor endothelin. In addition, endothelial cells play a pivotal role in the initiation, regulation, and maintenance of the inflammatory process via the induction of changes in their interactions with blood constituents. These factors, when considered in parallel with the susceptibility to perturbation by hypoxia of the endothelial cell, confer its prominent role in the pathogenesis of ischaemic states.

Leucocyte chemotaxis and adhesion

The initiating factors of inflammatory cell infiltration are adherence to endothelium and chemotaxis. Synovial inflammatory cell infiltration is a major characteristic of rheumatoid arthritis as is the expression of membrane-associated adhesion molecules. The contribution of hypoxia to this situation may have considerable significance, particularly where oxygen free radicals are a consequence. Superoxide, hydrogen peroxide, or both can generate a chemoattractant for neutrophils from extracellular fluid (Perez *et al*. 1980; Petrone *et al*. 1980). The identity of the neutrophil chemotactic factor(s) generated in ischaemia is unknown. However, possible candidates are cyclo-oxygenase and lipoxygenase metabolites of arachidonic acid, such as 12-hydroxyeicosatetraenoic acid which is found to be raised in hypoxic states and is a neutrophil chemoattractant. In support of this, cultured endothelial cells are capable of responding to hypoxia by producing a neutrophil chemoattractant which is suppressible by lipoxygenase inhibitors (Farber *et al*. 1987).

The adhesion component of leucocyte infiltration can be facilitated by oxygen radicals (Suzuki *et al*. 1989) and by hypoxia (Mullane and Pinto 1987) alone. Interestingly, the adhesion of neutrophils to ischaemic arterial vasculature is associated with a loss of endothelium-dependent vasorelaxant activity.

Microvascular barrier function

In any inflammatory condition the manifestation of the inflammatory response is controlled by the permeability of the microvascular blood/tissue barrier. Increased transvascular exchange of macromolecules, emigration of leucocytes, and oedema formation are indicative of modified barrier function and are characteristic of synovial inflammation. The dynamics of barrier function, particularly solute exchange, are modulated physiologically by a complex assortment of mediators. During hypoxia, however, homeostasis is disrupted by, as yet, poorly understood mechanisms involving mediators of altered barrier function, such as platelet activating factor, complement, leukotrienes, and prostaglandins. Increased permeability during ischaemia or hypoxia has been reported in many vascular systems *in vitro* and *in vivo*, for example heart (Armiger and Gavin 1975), lung (Lockhart and Saiag 1981), brain (Oleson 1986), and cultured bovine endothelium (Ogawa *et al*. 1990).

In addition to the effects of hypoxia alone, rheumatoid synovial microvascular function and permeability can be altered by the membrane damaging effects of oxygen free radicals. Measurement of von Willebrand factor illustrates this. von Willebrand factor, otherwise known as factor VIII-related antigen, is a large adhesive glycoprotein which is synthesized in endothelial cells and megakaryocytes and plays an important role in the adhesion of platelets to damaged vessel walls. Small subunits of von Willebrand factor are released constitutively by endothelial cells, accounting for normal circulating plasma levels. Larger multimers of von Willebrand factor are stored in specific organelles called Weibel–Palade bodies within endothelial cells. Interestingly, these bodies are common in the synovium and have been implicated as a marker of angiogenesis in rheumatoid arthritis (Kumar *et al*. 1985). A variety of physiological stimuli can induce acute release of these multimers from the Weibel–Palade bodies *in vitro* including hypoxia, by a Ca^{2+}-dependent process. Elevated levels of plasma von Willebrand factor have been observed in various disease states including rheumatoid arthritis (Woolf *et al*. 1987; Pottinger *et al*. 1989) where, in contrast to normal, levels increase with simple 'physiologic' exercise. This phenomenon has been taken to indicate vascular damage (Greaves *et al*. 1987).

The attempts of the joint to limit the hypoxic damage

The immediate local response to tissue hypoxia is elicited by the vasculature, utilizing vasoactive substances which serve to enhance perfusion. It is clear that the facility for this response is impaired in rheumatoid synovitis. A more long-term strategy for the re-establishment of adequate perfusion is neovascularization of hypoxic tissue. The facility for this, termed angiogenesis, most certainly exists in the rheumatoid synovium, to an extent which some believe to be pathological.

Under conditions of physiological stress, such as hypoxia, cell survival may become dependent on the expression of a specialized set of proteins, termed stress proteins.

Angiogenesis

The formation and growth of new blood vessels is under rigid control, so much so that in health angiogenesis only occurs in wound healing and during endometrial regeneration. Functional aberrations of the physiological regulation of new capillary growth, such as seen in ischaemia and reperfusion, are implicated in the pathogenesis (if not aetiology) of many neoplastic and non-neoplastic diseases.

The neovascularization of a wound progresses, logically, towards the centre where hypoxic macrophages are stimulated to produce angiogenic factors. The persistent synovitis characteristic of rheumatoid arthritis equates with persistent, though perhaps not fully effective, angiogenesis. The resulting fibroproliferative, inflammatory pannus invades and elicits the destruction of articular cartilage leading to joint deformation and loss of function. Clearly there is angiogenesis in the rheumatoid synovium, although morphometry suggests that this does not keep pace with synovial proliferation. This situation, as has been discussed earlier, promotes hypoxia in the peripheral region of the synovium which is rich in macrophages. The evidence that rheumatoid synovial tissue macrophages and hypoxic macrophages releases a substance capable of inducing angiogenesis invite the hypothesis that hypoxia drives synovial angiogenesis. The macrophage-derived angiogenic factor is thought to be tumour necrosis factor-a. This factor has been localized histochemically to the synovial lining cells in rheumatoid arthritis but no prominent staining is seen in osteoarthritis non-inflammatory samples.

Stress proteins

Often inaccurately given the general term 'heat-shock proteins', these ubiquitous intracellular molecules form two groups. True heat-shock proteins contain a specific consensus region within the promotor region of the gene while the second group, the glucose-regulated proteins, do not have this DNA sequence. Although the genes are normally quiescent or only partially active, stress protein mRNAs are actively transcribed following physiological stress, the proteins induced being dependent on the particular stressor; they are classified into families according to molecular weight. An excellent review of the subject is available (Lindquist 1986).

Rheumatological interest in the heat-shock proteins followed the implication of the 60 kDa family in the development of adjuvant-induced arthritis in rats (Van Eden *et al*. 1988). These observations have been extended to include other animal models of arthritis. Additionally, patients with rheumatoid arthritis have circulating isotype-restricted antibodies to stress proteins and these same proteins are abundantly present in inflamed but not normal synovia; the high levels of the stress-inducible 72 kDa protein give a good indication of the physiological trauma of synoviocytes within the inflamed joint (Winrow *et al*. 1990a).

No other information is available with respect to stress protein synthesis in joints. It has been shown *in vitro* that glucose-regulated proteins are synthesized during hypoxia and heat-stress proteins on reperfusion (Sciandra *et al*. 1984). Similarly, animal experiments show that *de novo* intracellular synthesis occurs during exercise. Measurement of the 78 kDa glucose-regulated proteins synthesized during glucose deprivation is difficult to quantify *in situ* since the available antisera which recognize this protein also bind with other members of the 70 kDa family. The usual method of cell radiolabelling is impractical *in vivo*.

Natural defence mechanisms exist to limit damage from reactive oxygen metabolites. The 32 kDa heat-shock protein, now characterized as the enzyme haem oxygenase, is specifically induced by hydrogen peroxide, by heavy metals including iron, and by its substrate haem. Interestingly, it is also induced by the antirheumatic compound auranofin. Haem oxygenase, in the presence of NADPH and oxygen, cleaves haem to form biliverdin which is then rapidly reduced to bilirubin, a radical scavenger. It has been suggested that increased production of this enzyme may be necessary to facilitate the turnover of haem-containing proteins, like the mitochondrial cytochromes which contribute to oxidative phosphorylation by respiration-linked electron transfer (Welch 1990). We have postulated that in the inflammatory joint, production of this enzyme would be beneficial, decreasing the levels of haem whilst additionally increasing the levels of the antioxidant bilirubin and thus limiting damage by the haem iron catalysed formation of hydroxyl radicals; the low oxygen tension within the joint may limit activation of this enzyme (Winrow *et al*. 1990b). Auranofin and thiol reactive agents may act by helping to stimulate production of the 32 kDa protein and thus affect vascular changes; it is known that

D-penicillamine suppresses neovascularization by generating low levels of hydrogen peroxide (Matsubara *et al*. 1989).

Reflex muscle inhibition

Exercise of the inflamed knee joints gives rise to elevated intra-articular pressures which is central to the mechanism of ischaemic reperfusion injury. However, this is not the case in all joints with effusions, there is a dramatic difference between a joint with an acute traumatic effusion and one with a chronic effusion. Measurement of the intra-articular pressure in the knee after quadriceps setting reveals that the patient with acute trauma is barely able to raise the intra-articular pressure at all (13 mmHg) whilst the patient with chronic inflammation is able to achieve pressures in excess of 250 mmHg. This phenomenon is explained by the fact that the patient with acute trauma has reflex muscular inhibition of the quadriceps whereas the patient with a chronic condition does not. This may be a defensive mechanism in the patient with acute trauma attempting to limit hypoxic–reperfusion injury. The mechanism behind this inhibition is not understood, neither the manner by which it is overcome in the chronic inflammatory condition. However, it is reasonable to speculate that the depletion of the sensory nerve fibres in chronic arthritis (see earlier) is the cause. The transition from the acute to the chronic condition obviously involves the nervous system but again how this is regulated is not known. The nervous system also has other influences on the chronic inflammatory condition, some of which are only now being elucidated.

Nervous system

A number of clinical observations point to a neurogenic mechanism operative in chronic arthritis. The clearest example of this is provided by the synovitis that often accompanies reflexed sympathetic dystophy. Sympathetic activity has also been implicated in the aetiology of the frozen shoulder (Maini *et al*. 1989). More controversial are the observations relating to hemiplegia and rheumatoid arthritis. In cases in which the rheumatoid arthritis develops after the hemiplegia the joints in the hemiplegic limbs are spared, but partially affected limbs are not always spared (Thompson and Bywaters 1962). In cases of patients paralysed by polio before the onset of the arthritis there is almost total sparing of the paralysed limbs. While it is tempting to consider this as clear evidence for a neural component in rheumatoid arthritis, other causal factors such a movement-induced hypoxic–reperfusion injury (see earlier) need to be considered.

In animal model systems the available evidence is clearer. Experimental evidence for a neurogenic effect in acute arthritis is provided by the fact that antidromic (reversed) stimulation of the articular C fibres results in vasodilation and plasma extravasation within the joint. The intra-articular injection of capsaicin, a compound known specifically to stimulate C fibres, gives similar results. Intra-articular infusion of 6-hydoxydopamine, a compound which stimulates post-ganglionic sympathetic nerves to release the contents of their peripheral terminals, can also produce a prolonged increase in synovial plasma extravasation. The acute vascular response to intra-articular injection of inflammatory compounds, such as carrageenan, can be significantly inhibited by prior joint denervation.

In chronic arthritis, substance P levels (see later) have been shown to be greater in those joints which develop more severe adjuvant arthritis (Levine *et al*. 1984), and injection of substance P into minimally involved joints increases the severity of the adjuvant disease in the injected joint. The destruction of the ankle joint in the same model is ameliorated by sciatic nerve section as compared with the contralateral non-operated side.

In human disease we have demonstrated that there is a change in the morphology of the nerve supply to the joint and have speculated on the release of neuropeptides in the inflamed joint.

Pain

One of the outstanding features of the inflamed joint is undoubtedly the pain which arises from it. As described earlier, the synovium contains unmyelinated C fibres which usually remain silent. However, in an inflammatory reaction the threshold of these fibres becomes lowered. This effect can be induced experimentally by agents such as prostaglandin E_1 and bradykinin (reviewed by Konttinen *et al*. 1994). Both substances are found in inflammatory joint fluid. Not only can the nociceptive fibres alter their threshold but they may alter their receptive field so as to widen it. Thus movement in the normal range which was once painless now becomes painful.

Neurogenic inflammation

In addition to a nociceptive function, it is now widely accepted that neuropeptides contained within these nerve fibres play a role in the inflammatory response. The activation of the sensory nerves results not only in an impulse transmission to the spinal cord but also a reversed (antidromic) transmission through the extensive branches in the peripheral C fibres. This is known as an axon reflex. This reversed transmission results in the release of neuropeptides, such as substance P, and stimulates an acute reaction—neurogenic inflammation (Foreman 1987). Radioimmunoassays point to the release of neuropeptides into the inflammatory joint fluid. There are numerous neuropeptides present in the synovium and a comprehensive description of all their actions is not possible. Here we have selected a few for closer examination.

Substance P

In 1931 von Euler and Gaddum extracted from equine brain and intestine a material which was hypotensive and smooth muscle stimulating (von Euler and Gaddum 1931). They designated this activity preparation P (for powderable) and in later publications referred to it as substance P. However, it was not until 1971 that the peptide was purified (Chang *et al*. 1971), sequenced, and synthesized.

Substance P is a small polypeptide of 11 amino acids and is one of a family of peptides, the tachkykinins, which share 4 common amino acids at the carboxyl terminal. Two precursor messenger RNAs have been isolated, one containing the substance P sequence (a-preprotachykinin) and the other containing substance P and the sequence for another tachykinin, neurokinin A. There appears, however, to be only one preprotachykinin gene, with alternative RNA splicing resulting in the production of the two messages.

Substance P in joints

C fibres containing substance P are mostly seen with their terminals at the boundary between the synovium and the joint space. The location of these fibres makes them ideally suited to play a part in monitoring the intra-articular environment or to act as sensory receptors. Since both the peptides associated with these nerves, substance P and calcitonin gene-related peptide, are found in the sensory nerves, a nociceptive role seems a likely hypothesis. Now that the presence of nerves containing substance P in human joints has been established, what evidence do we have that they contribute to arthritis?

In experimental animals antidromic stimulation of articular C fibres results in vasodilation and plasma extravasation within the joint. Intra-articular injection of capsaicin, the hot component of chili peppers which is known to stimulate C fibres, yields similar results. Direct intra-articular injection of substance P also produces vasodilation and plasma extravasation. This effect is probably mediated by the release of endothelial cell-derived relaxing factor (nitric oxide) The induction of an acute inflammatory response within the joint with an agent such as carrageenan may also be mediated by substance P. In knee perfusion experiments, pretreatment of the test knee with the substance P antagonist d-Pro^4, d-$Trp^{7,9,10}$ -SP (4 to 11) before challenge with carrageenan produced a 93 per cent reduction in the inflammatory response compared with knees treated with carrageenan alone (Lam and Ferrell 1989). Substance P levels have been shown to be greater in those joints that develop the most severe arthritis, and the injection of substance P into minimally involved joints increases the severity of the adjuvant disease in the injected joint. This and other work adds up to a substantial body of evidence that substance P can cause joint inflammation in animal models, is there evidence for an involvement in human disease?

Detection of substance P in synovial fluid

One obvious way to establish a link with chronic inflammation in human disease is to demonstrate the presence of the peptide in the inflammatory synovial fluid. This problem has been addressed by several investigators with differing results. The studies to date have utilized immunoassay methods and all have employed different polyclonal antisera, which makes comparison of results difficult. One initial report (Devillier *et al*. 1986) suggested that inflammatory fluids contained more tachykinin-like immunoreactivity than non-inflammatory fluids. Larsson *et al*. (1989) found no tachykinins (substance P or neurokinin A) in their inflammatory fluids but neurokinin A was present in the traumatic injury control group. In four types of arthritis (rheumatoid, osteoarthritis, Reiter's syndrome, and traumatic arthritis) the synovial fluid levels of substance P were higher than those seen in paired plasma samples (Marshall *et al*. 1990). Since the levels are higher in the synovial fluid than the plasma, the source of the substance P is probably intra-articular. The only known source are the C fibres, although there has been one report that a subpopulation of endothelial cells may also contain substance P (Loesch and Burnstock 1988). Whether these cells acquire substance P by uptake or by synthesis remains to be determined. The significance of absolute levels of substance P, or indeed any parameter measured in joint fluid, is difficult to assess since the dynamics of the system

in terms of production, degradation, and clearance rates are unknown. Furthermore, there is no method yet available to show whether the substance P detected is bioactive or simply a smaller non-active fragment, as the potency of the molecule might lead one to suspect.

These reservations are substantiated by recent data (Joyce 1993) demonstrating that authentic levels of substance P in joint fluids are less than 4.7 pg/ml, which was the detection limit of the assay. The study showed that the inhibition of degrading enzymes and correct extraction procedures were all important in obtaining reliable results.

In vitro investigations

The effects of substance P on a wide variety of cell types *in vitro* have been investigated. These data show release of inflammatory mediators from various cell types, including the degranulation of mast cells (Mazarek *et al*. 1981), the stimulation of prostaglandin E_2 and collagenase from synoviocytes (Lotz *et al*. 1987), secretion of prostaglandins and thromboxane from macrophages (Hartung and Tokya 1983; Hartung *et al*. 1986), and the modulation of immunoglobulin production from lymphoid tissue (Sanitz *et al*. 1986). Substance P is also reported to induce the release of interleukin 1 from cultured mouse macrophages (Kimball *et al*. 1985). Thus tachykinins may contribute to the maintenance of chronic arthritis. These activities of substance P require a dose of the peptide in the nanomolar range, which is consistent with published figures for the dissociation constant of substance P with its receptor (0.5 to 2.0 nmol). Other cells, such as neutrophils, only appear to be activated directly by much larger concentrations of the peptide (greater than 10 µmol) suggesting a lack of physiological relevance. However, lower concentrations of substance P, in the nanomolar range, appear to be able to prime neutrophils to respond to other mediators (Perianin *et al*. 1989). In particular, neutrophils have been shown to act in a synergistic manner with substance P to other chemotactic peptides, such as complement-derived C5a or bacterial fMLP. Low doses of substance P have also been shown to modulate the action of other peptides, such as calcitoin gene-relate peptide. Collectively these results suggest that in addition to a direct effect of substance P on several cell types a second modulatory mechanism is also operative. The nervous system may therefore be capable of priming cells to respond to lower doses of certain agents, than might otherwise be the case, leading to an apparent hypersensitivity. This mechanism may be particularly important in allergic responses, such as that seen in asthma.

Calcitonin gene-related peptide

Calcitonin gene-related peptide (**CGRP**) is a 37 amino acid peptide produced by alternative splicing of the primary transcript of the calcitonin gene. It is particularly abundant in the sensory nervous system, being present in approximately 50 per cent of primary sensory neurones. Behaviourial and electrophysiological studies have shown that CGRP may play a part in pain perception and reports have indicated that it can potentiate the hyperalgesia caused by intradermal injection of substance P. In addition to its presence in primary sensory nerves, CGRP has also been found in motor end plates and is stored and synthesized in motor neurones (Gibson *et al*. 1988). This is consistent with reports describing motor-related actions for this peptide, including a role as a muscle trophic factor. In both human and animal studies, CGRP has been shown to be a potent vasodilator, capable of inducing a protracted increase in microvascular blood flow when injected extravascularly. Although relatively inactive on its own, CGRP potentiates oedema induced by other mediators of increased microvascular permeability, including substance P.

The majority of substance P-containing nerve fibres also contain CGRP, and electron immunocytochemical studies have revealed that secretory granules in sensory nerves contain immunoreactivity for both peptides (Merighi *et al*. 1987). It is interesting to note that co-injection of CGRP with substance P into human skin converts the long-lasting vasodilation induced by CGRP into a transient response. Experiments in animals reveal that this phenomenon is dependent on the action of proteases from mast cells stimulated by substance P (Brain and Williams 1988). Thus if both CGRP and substance P are released simultaneously from nerve terminals, CGRP will exert a relaxant effect on vascular smooth muscle. The increased blood flow can then be controlled by proteolysis of CGRP by enzymes released from the mast cells in response to substance P. If this inactivation mechanism is overruled by an excess of CGRP, generated locally by depletion of substance P or by loss of mast cell proteases, protracted erythema, as occurs following local infection or injury, could result.

Since CGRP and substance P coexist, it might be expected that CGRP has a similar distribution to substance P in articular tissues. However, substance P-immunoreactive nerves are less abundant than CGRP neurones, while CGRP-immunoreactivity is found in motor nerves of muscles surrounding the motor nerves and long bones.

Neuropeptide Y

Neuropeptide Y is a 36 amino acid peptide. Neuropeptide Y is associated with catecholaminergic nerves because of its co-localization with tyrosine hydroxlase and dopamine b-hydroxylase in sympathetic peripheral ganglion cells. Potent vasoconstrictor effects on the vasculature and presynaptic inhibition of transmitters released from non-adrenergic, cholinergic nerves and non-adrenergic, non-cholinergic nerves are among the reported actions (Allen and Bloom 1986). Nerve fibres containing neuropeptide Y have been localized almost exclusively to blood vessels in synovial tissues and periosteum (Mapp *et al*. 1990b), unlike the peptides substance P and CGRP which also occur as free nerve fibres.

Degradation of peptides

The local activities of regulatory peptides depend on a combination of release and clearance from the vicinity of their receptors. Many peptides have a short half-lives, rapid clearance being due to membrane-bound peptidases. These enzymes have characteristic regional distributions, and their activities may vary during inflammation. Understanding the local topography of membrane-bound peptidases and their relation to the sites of release and action of those peptide that may act as substrates is therefore essential to understanding peptidergic regulatory systems in the normal and diseased synovium.

Neutral endopeptidase (NEP; EC 3.4.24.11) is responsible for the majority of the degradation of substance P in the human synovium. NEP is capable of the hydrolysis of many peptides including substance P (Skid *et al*. 1984) and CGRP (Katyama *et al*. 1991) This enzyme has been shown to be identical to the common acute lymphoblastic leukaemia antigen (CALLA) (Letarte *et al*. 1988) and is the CD10 determinant.

Its activity has been shown in the synovium of patients with chronic arthritis (Sreedharan *et al*. 1990). The activity of NEP was found to be higher in all patients with rheumatoid arthritis and some patients with degenerative joint disease, when compared with controls with traumatic arthritis. Our own data show the localization of this enzyme in human synovium; fibroblasts are its source, a restricted population of cells surrounding blood vessels being responsible for the majority of the activity (Mapp *et al*. 1992) (Fig. 6).

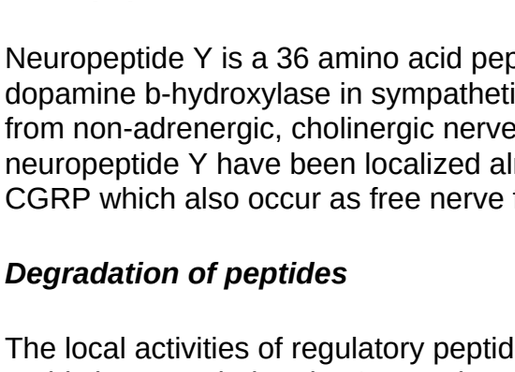


Fig. 6 Photomicrographs of rheumatoid synovia showing (a) the localization (black reaction product) of neutral endopeptidase, an enzyme that degrades neuropeptides, surrounding blood vessels and (b) at higher power, to a restricted population of cells surrounding those blood vessels.

These may represent a specialist subpopulation of fibroblasts. Since the function of NEP is probably the degradation of locally released regulatory peptides, its presence around the blood vessels makes it ideally located to inactivate vasoactive peptides, such as substance P, which are released from perivascular nerve fibres.

In addition to NEP, we (Walsh *et al*. 1993) have also localized several other peptide degrading enzymes to the synovium, including angiotensin converting enzyme,

dipeptidyl peptidase IV, and aminopeptidase M. Angiotensin converting enzyme is localized to the endothelia of all vessels, and aminopeptidase M to the lining layer of the synovium and scattered cells in the stroma. The localization of these peptidases in the human synovium suggests not only a role limiting the duration of action of regulatory peptides but also in localizing activities to the vicinity of their release. This functional compartmentalization of vascular and stromal regions in synovium is essential to the local regulatory function of peptides and is likely to influence reponses to exogenous peptides administered into non-physiological compartments during experimental investigations (Fig. 7).

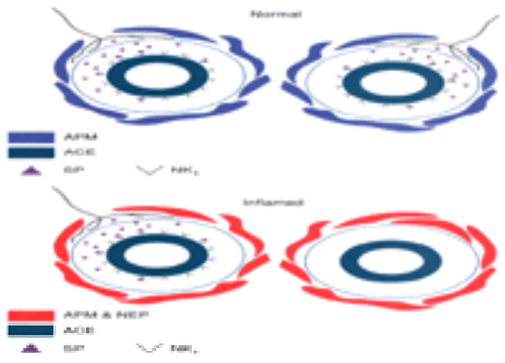


Fig. 7 Schematic diagram showing the distribution of enzymes in and around blood vessels, capable of degrading neuropeptides, in the normal and inflamed joint. Note how the neuropeptides are compartmentalized by the degrading enzymes. In disease nerve fibres may be absent and receptors downregulated. Additionally, degrading enzymes are induced. (Diagram by courtesy of Dr D.A. Walsh, London Hospital Medical College.)

Summary

In this chapter we have shown how we believe the altered states of the synovium and synovial fluid arise in terms of the physiology of the joint. This is not intended to be a comprehensive review, since we have not attempted to address the involvement of the many other important mediator systems, to which separate chapters are devoted. The mechanisms that we have described are surely going on in parallel with several others. The concepts that are advanced are widely accepted in other areas of biology but are not often applied to the field of rheumatology. As these ideas gain acceptance it is to be hoped that our understanding of the underlying causes of why the inflammation in a joint can be so extraordinarily persistent is enhanced and novel therapeutic strategies developed.

The researches of many commentators have already thrown much darkness on the subject, and it is probable that, if they continue, we shall soon know nothing at all about it.

Mark Twain.

Chapter References

Allan, I.M., Lunec, J., Salmon, M., and Bacon, P.A. (1986). Reactive oxygen species (ROS) selectively deplete T-lymphocytes via a hydroxyl radical dependant mechanism. *Agents Actions*, **19**, 351–2.

Allen, J.M. and Bloom, S.R. (1986). Neuropeptide Y: a putative neurotransmitter. *Neurochemistry Internationa*l, **8**, 1–8.

Allen, R.E., Outhwaite, J., Morris, C.J., and Blake, D.R. (1987). Xanthine oxido-reductase is present in human synovium. *Annals of the Rheumatic Diseases*, **46**, 843–5.

Armiger, C. and Gavin, J.B. (1975). Changes in the microvasculature of ischaemic and infarcted myocardium. *Laboratory Investigations*, **33**, 51–6.

Barland, P., Novikoff, A.B., and Hamerman, D. (1962). Electron microscopy of the human synovial membrane. *Journal of Cell Biology*, **14**, 207–20.

Blake, D.R. *et al*. (1989). Hypoxic-reperfusion injury in the inflamed human joint. *Lancet*, **i**, 289–93.

Brain, S.D. and Williams, S.D. (1988). Substance P regulates the vasodilator activity of calcitonin gene-related peptide. *Nature*, **335**, 73–5.

Chang, M.M., Leeman, S.E., and Niall, H.D. (1971). The amino acid sequence of substance P. *Nature New Biology*, **232**, 86–7.

Devillier, P., Weill, B., Renoux, M., Menkes, C., and Pradelles, P. (1986). Elevated levels of tachykinin-like immunoreactivity in joint fluids from patients with inflammatory rheumatic diseases. *New England Journal of Medicine*, **314**, 1323.

Dick, W.C. *et al*. (1970). Derivation of knee joint synovial perfusion using the xenon (^{133}Xe) clearance technique. *Annals of the Rheumatic Diseases*, **29**, 131–4.

Dingle, J.T.M. and Page-Thomas, D.P. (1956). *In vitro* studies on human synovial membrane. A metabolic comparison of normal and rheumatoid tissue. *British Journal of Experimental Pathology*, **37**, 318–36.

Edwards, J.C.W. and Willoughby, D.A. (1982). Demonstration of bone marrow-derived cells in the synovial lining by means of giant intracellular granules as genetic markers. *Annals of theRheumatic Diseases*, **41**, 177–82.

Falchuk, K.H., Goetzl, E.J., and Kulka, J.P. (1970). Respiratory gases of synovial fluids. An approach to synovial tissue circulatory–metabolic imbalance in rheumatoid arthritis. *Americal Journal of Medicine*, **49**, 223–31.

Farber, H.W., Center, D.M., and Rounds, S. (1987). Effect of ambient oxygen on cultured endothelial cells from different vascular beds. *American Journal of Physiology*, **253**, H875–83.

Foreman, J.C. (1987). Peptides and neurogenic inflammation. *British Medical Bulletin*, **43**, 386–400.

Gibson, S.J. *et al*. (1988). Calcitonin gene-related peptide messenger RNA is expressed in the sensory neurones of dorsal root ganglia and also in the spinal motorneurones in man and rat. *Neuroscience Letters*, **91**, 283–8.

Greaves, M., Pickering, C., Knight, G., Boulton, A.J.M., Ball, J., and Preston, F.E. (1987). Changes in the factor VIII complex in diabetic ketoacidosis: evidence of endothelial damage? *Diabetologica*, **30**, 160–5.

Hartung, H.-P. and Toyka, K.V. (1983). Activation of the macrophage by substance P: induction of oxidative burst and thromboxane release. *European Journal of Pharmacology*, **89**, 301–6.

Hartung, H.-P., Woiters, K., and Toyka, K.V. (1986). Substance P binding properties and studies on cellular responses in guinea pig macrophages. *Journal of Immunology*, **136**, 3856–61.

Hinshaw, D.B., Armstrong, B.C., Burger, J.M., Beals, T.F., and Hyslop, P.A. (1988). ATP and microfilaments in cellular oxidant injury. *American Journal of Pathology*, **132**, 479–88.

James, M.J., Cleland, L.G., Rofe, A.M., and Leslie, A.L. (1990). Intra-articular pressure and the relationship between synovial perfusion and metabolic demand. *Journal of Rheumatology*, **17**, 521–7.

Joyce, T.J., Yood, R.A., and Carraway, R.E. (1993). Quantitation of substance P and its metabolites in plasma and synovial fluid from patients with arthritis. *Journal of Clinical Endocrinology and Metabolism*, **77**, 632–7.

Katyama, M., Nadel, J.A., Bunnet, N.W., DiMaria, G.U., Haxhiu, M., and Borson, D.B. (1991). Catabolism of calcitonin gene-related peptide and substance P by neutral endopeptidase. *Peptides*, **12**, 563–7.

Kimball, E.S., Persico, F.J., and Vaught, J.L. (1985). Substance P, neurokinin A, and neurokinin B induce generation of IL1-like activity in P388D1 cells: possible relevance to arthritic diseases. *Journal of Immunology*, **141**, 3564–9.

Knight, A.D. and Levick, J.R. (1983). The density and distribution of capillaries around a synovial cavity. *Quarterly Journal of Experimantal Physiology*, **63**, 629–44.

Konttinen, Y.T. *et al*. (1994). Peripheral and spinal mechanisms in arthritis, with particular reference to treatment of pain and inflammation. *Arthritis and Rheumatism*, **37**, 965–82.

Kumar, P., Erroi, A., Sattar, A., and Kumar, S. (1985). Weibel–Palade bodies as a marker for neovascularization induced by tumour and rheumatoid angiogenesis factors. *Cancer Research*, **45**, 4339–48.

Lam, F.Y. and Ferrell, W.R. (1989). Inhibition of carrageenan induced inflammation in the rat knee joint by substance P antagonist. *Annals of the Rheumatic Diseases*, **48**, 928–32.

Larsson, J., Ekblom, A., Henriksson, K., Lundeberg, T., and Theodorsson, E. (1989). Immunoreactive tachykinins, calcitonin gene-related peptide and neuropeptide Y in human synovial fluid from inflamed knee joints. *Neuroscience Letters*, **100**, 326–30.

Letarte, M. *et al*. (1988). Common acute lymphocytic leukaemia antigen is identical to neutral endopeptidase. *Journal of Experimental Medicine*, **168**, 1247–53.

Levick, J.R. (1990). Hypoxia and acidosis in chronic inflammatory arthritis, relation to vascular supply and dynamic effusion pressure. *Journal of Rheumatology*, **17**, 579–82.

Levine, J.D., Clark, R., Devois, M., Helms, C., Moskowitz, M.A., and Basbaum, A.I. (1984). Intraneuronal substance P contributes to the severity of experimental arthritis. *Science*, **226**, 574–9.

Lindquist, S. (1986). The heat-shock response. *Annual Review of Biochemistry*, **55**, 1151–91.

Lockhart, A. and Saiag, B. (1981). Altitude and the human pulmonary circulation. *Clinical Science* (*London*), **60**, 599–605.

Loesch, A. and Burnstock, G. (1988). Ultrastructural localisation of serotonin and substance P in vascular endothelial cells of rat femoral and mesenteric arteries. *Anatomy and Embryology*, **178**, 137–42.

Lotz, M., Carson, D.A., and Vaughan, J.H. (1987). Substance P activation of rheumatoid synoviocytes: neural pathways in the pathogenesis of arthritis. *Science*, **235**, 893–5.

Lund-Olsen, K. (1970). Oxygen tension in synovial fluids. *Arthitis and Rheumatism*, **3**, 769–6.

Lunec, J., Blake, D.R., McCleary, S.J., Brailsford, S., and Bacon, P.A. (1985). Self perpetuating mechanisms of immunoglobulin G aggregation in rheumatoid inflammation. *Journal of Clinical Investigation*, **76**, 2084–90.

McCord, J.M. (1987). Oxygen-derived radicals: a link between reperfusion injury and inflammation. *Federation Proceedings*, **46**, 2402–6.

Maini, R., Cooper, C., Kidd, B.L., Cole, J.D., and Cawley, M.I.D. (1989). Use of laser Doppler flowmetry and transcutaneous oxygen tension electrodes to assess local autonomic dysfunction in patients with frozen shoulder. *Journal of the Royal Society of Medicine*, **82**, 536–8.

Mapp, P.I. *et al*. (1990*a*). Substance P, calcitonin gene-related peptide and C flanking peptide of neuropeptide Y-immunoreactive fibres are present in normal synovium but depleted in patients with rheumatoid arthritis. *Neuroscience*, **37**, 143–53.

Mapp, P.I. *et al*. (1990*b*). Receptors to the neuropeptide substance P in the human synovium. *British Journal of Rheumatology*, **29**, 22.

Mapp, P.I., Walsh, D.A., Kidd, B.L., Cruwys, S.C., Polak, J.M., and Blake, D.R. (1992). Localisation of the enzyme neutral endopeptidase to the human synovium. *Journal of Rheumatology*, **19**, 1838–44.

Marshall, K.W., Chiu, B., and Inman, R. (1990). Substance P and arthritis: analysis of plasma and synovial fluid levels. *Arthritis and Rheumatism*, **30**, 87–90.

Matsubara, T., Velvart, M., Odermatt, B.F., Spycher, M.A., Rüttner, J.R., and Fehr, K. (1983). The thickening of basement membrane in synovial capillaries in rheumatoid arthritis. *Rheumatology International*, **3**, 57–64.

Matsubara, T., Saura, R., Hirohata, K., and Ziff, M. (1989). Inhibition of human endothelial cell proliferation *in vitro* and neovascularisation *in vivo* by D-pennicillamine. *Journal of Clinical Investigation*, **83**, 158–67.

Mazarek, N., Pecht, I., Teichburg, V.I., and Blumberg, S. (1981). The role of the *N*-terminal tetrapeptide in the histamine releasing action of substance P. *Neuropharmacology*, **20**, 1025–7.

Meier, R., Spreyer, P., Ortmann, R., Harel, A., and Monard, D. (1990). Induction of glia-derived nexin following peripheral nerve lesion. *Nature*, **342**, 548–50.

Merighi, A., Polak, J.M., Gibson, S.J., Gulbenkian, S., Valentino, K.L., and Peirone, S.M. (1987). Ultrastructural studies on calcitonin gene-related peptide-, tachykinins-, and somatostatin-immunoreactive nerves in rat dorsal root ganglia: evidence for the colocalisation of different peptides in single secretory granules. *Cell and Tissue Research*, **254**, 101–9.

Morris, C.J., Blake, D.R., Wainwright, A.C., and Steven, M.M. (1986). Relationship between iron deposits and tissue damage in the synovium: an ultrastructural study. *Annals of the Rheumatic Diseases*, **45**, 21–6.

Mullane, F.M. and Pinto, A. (1987). Endothelium, arachidonic acid and coronary vascular tone. *Federal Proceedings*, **46**, 54–62.

Murrell, G.A.C., Francis, M.J.O., and Bromley, L.. (1990). Modulation of fibroblast proliferation by oxygen free radicals. *Biochemistry Journal*, **265**, 659–65.

Nayler, W.G., Poole-Wilson, P.A., and Williams, A. (1979). Hypoxia and calcium. *Journal of Molecular and Cell Cardiology*, **11**, 683–706.

Ogawa, S., Gerlach, H., Esposito, C., Pasagian-Macaulay, A., Brett, J., and Stern, J. (1990). Hypoxia modulates the barrier and coagulant function of cultured bovine endothelium. Increased monolayer permeability and induction of procoagulant properties. *Journal of Clinical Investigation*, **85**, 1090–8.

Oleson, S.-P. (1986). Rapid increase in blood–brain barrier permeability during severe hypoxia and metabolic inhibition. *Brain Research*, **368**, 24–9.

O'Rhailly, R. and Gardner, E. (1978). The embryology of moveable joints. In *Joints and synovial fluid* (ed. L. Sokoloff), pp. 43–103. Academic press, New York.

Page-Thomas, D.P. and Dingle, J.T.M. (1955). *In vitro* studies of rheumatoid synovium. Preliminary metabolic comparison between synovial membrane and villi. *British Journal of Experimental Pathology*, **36**, 195.

Perez, H.D., Weksler, B.B., and Goldstein, I.M. (1980). Generation of a chemotactic lipid from arachidonic acid by exposure to a superoxide-generating system. *Inflammation*, **4**, 313–28.

Perianin, A., Snyderman, R., and Malfoy, B. (1989). Substance P primes neutrophil activation: a mechanism for neurological regulation of inflammation. *Biochemical and Biophysical Research Communications*, **161**, 520–4.

Petrone, W.F., English, D.K., Wong, K., and McCord, J.M. (1980). Free radicals and inflammation: superoxide-dependent activation of a neutrophil chemotactic factor in plasma. *Proceedings of the National Academy of Sciences* (*USA*), **77**, 1159–63.

Pottinger, B.E., Read, R.C., Paleolog, E.M., Higgins, P.G., and Pearson, J.D. (1989). von Willebrand factor is an acute phase reactant in man. *Thrombosis Research*, **53**, 387–94

Sanitz, A.J., Befus, D., and Bienenstock, J. (1986). Differential effects of vasoactive intestinal peptide, substance P, and somatostatin on immunoglobulin synthesis and proliferation of lymphocytes from Peyers patches, mesenteric lymph nodes and spleen. *Journal of Immunology*, **136**, 152–6.

Sciandra, J.J., Subjeck, J.R., and Hughes, C.S. (1984). Induction of glucose-regulated proteins after anaerobic exposure and of heat-shock proteins after reoxygenation. *Proceedins of the National Academy of Sciences* (*USA*), **81**, 4843–7.

Shasby, M., Lind, S.E., Shasby, S.S., Goldsmith, J.C., and Hunninghake, G.W. (1985). Reversible oxidant-induced increase in albumen transfer across cultured endothelium: alterations in cell shape and calcium homeostasis. *Blood*, **65**, 605.

Simkin, P.A. and Pizzorno, J.E. (1979). Synovial permeability in rheumatoid arthritis. *Arthritis and Rheumatism*, **22**, 689–96.

Skid, R.A., Engelbrecht, A., Johnson, A.R., and Erdos, E.G. (1984). Hydrolysis of substance P and neurotensin by converting enzyme and neutral endopeptidase. *Peptides*, **5**, 769–76.

Sreedharan, S.P., Goetzl, E.J., and Malfroy, B. (1990). Elevated tissue concentration of the acute common lymphoblastic leukaemia antigen (CALLA)-associated neutral endopeptidase (3.4.24.11) in human chronic arthritis. *Immunology*, **71**, 142–4.

Starling, E.H. (1896). On the absorbtion of fluids from connective tissue spaces. *Journal of Physiology*, **19**, 312–26.

Stevens, C.R., Williams, R.B., Farrell, A.J., and Blake, D.R. (1991). Hypoxia and inflammatory synovitis. *Annals of the Rheumatic Diseases*, **50**, 124–32.

Suzuki, M. *et al*. (1989). Superoxide mediates reperfusion-induced leukocyte-endothelial cell interactions. *American Journal of Physiology*, **257**, H1740–5.

Swann, D.A., Silver, F.H., Slayter, H.S., Stafford, W., and Shore, E. (1985). The molecular structure and lubricating activity of lubricin isolated from bovine and human synovial fluids. *Biochemistry Journal*, **225**, 195–201.

Thompson, M. and Bywaters, E.G. (1962). Unilateral rheumatoid arthritis following hemiplegia. *Annals of the Rheumatic Diseases*, **21**, 370–7.

Tsoulfa, G. *et al*. (1989). Elevated IgG antibody levels to the mycobacterial 65-kDa heat shock protein are characteristic of patients with rheumatoid arthritis. *Scandinavian Journal of Rheumatology*, **30**, 519–27.

Van Eden, W. *et al*. (1988). Cloning of the mycobacterial epitope recognised by T lymphocytes in adjuvant arthritis. *Nature*, **331**, 171–3.

von Euler, U.S. and Gaddum, J.H. (1931). An unidentified depressor substance in certain tissue extracts. *Journal of Physiology*, **72**, 74.

Walsh, D.A. *et al*. (1992). Localisation and characterisation of substance P binding to human synovial tissue in rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **51**, 313–17.

Walsh, D.A., Mapp, P.I., Wharton, J., Polak, J.M., and Blake, D.R. (1993). Neuropeptide degrading enzymes in normal and inflamed human synovium. *American Journal of Pathology*, **142**, 1–12.

Welch, W.J. (1990). The mammalian stress protein response: cell physiology and biochemistry of stress proteins. In *Stress proteins in biology and medicine* (ed. R.I. Morimoto, A. Tissieres, and C. Georgopoulos), pp. 223–78. Cold Spring Harbor Press, New York.

Winrow, R., Mojdehi, G., Mapp, P.I., Rampton, D.S., and Blake, D.R. (1990 *a*). Immunohistological localisation of stress proteins in inflammatory tissue. In *Stress proteins in inflammation* (ed. R. Burdon, C. Rice-Evans, D. Blake, and V. Winrow), pp. 237–51. Richelieu Press, London.

Winrow, V.R., McLean, L., Morris, C.J., and Blake, D.R. (1990 *b*). The heat shock response and its role in inflammatory disease. *Annals of the Rheumatic Diseases*, **49**, 128–32.

Woodruff, T., Blake, D.R., Freeman, J., Andrews, F.J., Salt, P., and Lunec, J. (1986). Is chronic synovitis an example of reperfusion injury? *Annals of the Rheumatic Diseases*, **45**, 608–11.

Woolf, A.D., Wakerley, G., Wallington, T.B., Scott, D.G.I., and Dieppe, P.A. (1987). Factor VIII related antigen in the assessment of vasculitis. *Annals of the Rheumatic Diseases*, **46**, 441–7.

2.6 Skeletal muscle damage

Joan M. Round and D. A. Jones

Weakness, feelings of fatigue, and myalgia are amongst the most common symptoms complained of by patients with a variety of rheumatic and autoimmune diseases. In evaluating these symptoms it is important to have an understanding of the underlying physiology of normal skeletal muscle and this requires a knowledge of muscle development, the structure of muscle fibres, the organization of fibres into muscles and muscle groups, and the interactions between muscles and motor nerves.

Structure: the contractile apparatus

The major proteins of the contractile apparatus are actin and myosin.

Actin is a protein of great antiquity that occurs in a very similar form in most animal and plant cells and is polymerized into what appear to be double helical strands (F actin). F actin, together with the proteins tropomyosin and troponin form the thin filaments of the contractile apparatus. Tropomyosin blocks myosin-binding sites until caused to move by calcium ions binding to troponin C (Fig. 1). Actin filaments join at one end to form the Z-line structure with a-actinin making the connections between actin filaments (Fig. 2).

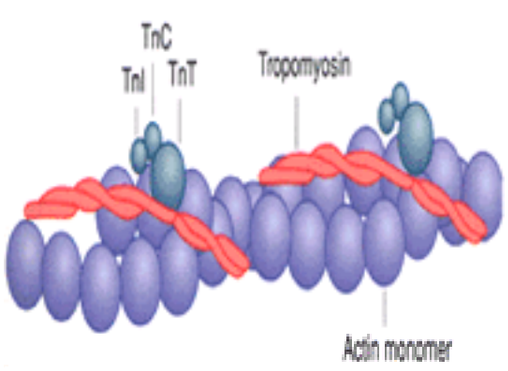


Fig. 1 Part of an actin filament together with tropomyosin and the troponin subunits.

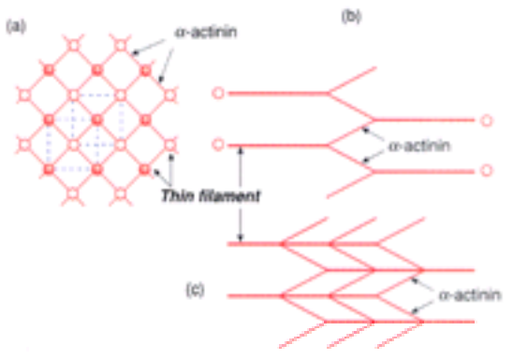


Fig. 2 Actin filaments joining to form the Z line. (a) View of the Z line along the axis of the thin filaments showing the square array. Solid symbols indicate thin filaments coming out of the page (or Z line), open symbols indicate thin filaments going into the page. (b) View from the side; note that only a-actinin connections in the plane of the page are shown, for every thin filament there will be another two connections in a plane at right angles to the page; (b) represents a simple Z-line structure found in fast muscle. (c) More complex Z line found in slow skeletal and cardiac muscle.

Myosin molecules consist of two identical heavy chains combined with four light chains. The myosin tail combines with other tails to bind the myosin molecules together to form a thick filament with the globular enzymic head regions, which combine with actin, projecting from the backbone of the filament (Fig. 3). Thick myosin filaments are arranged so that the thin actin filaments can slide between them. The unit from Z line to Z line is known as a sarcomere (Fig. 4).

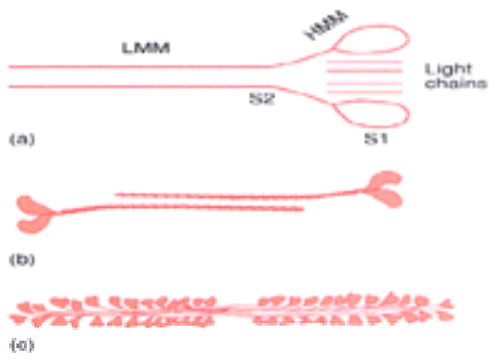


Fig. 3 Myosin structure and assembly into thick filaments. (a) Schematic arrangement of myosin subunits. (b) and (c) The basic double-headed myosin units aggregating to form a thick filament.LMM, light meromyosin; HMM, heavy meromyosin; S1, subunit 1; S2, subunit 2.

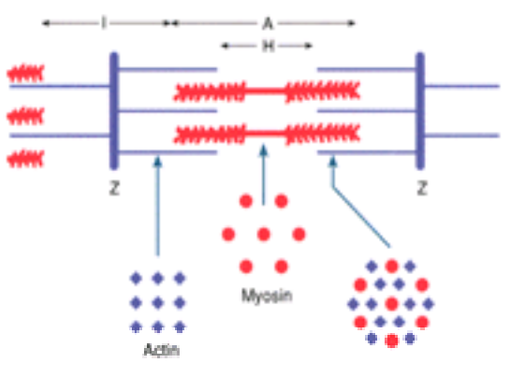


Fig. 4 The arrangement of thick and thin filaments to form a sarcomere. Below: cross-sections of the sarcomere.

A variety of structural proteins including C- and M-line proteins, titin, a-actin, desmin, and dystrophin are found associated with the structural apparatus and probably function to maintain sarcomere structure (Fig. 5).

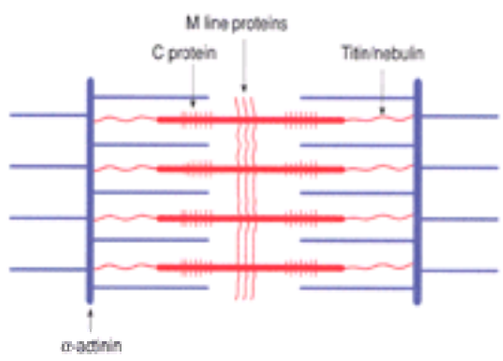


Fig. 5 Structural proteins in the sarcomere.

Bundles of 100 to 400 thick filaments form myofibrils (Fig. 6), which are separated from each other within the muscle fibre by sarcoplasmic reticulum, T tubules, and sometimes mitochondria (Fig. 7). Myofibrils vary in size but average around 1 μm in diameter. Myofibrillar protein constitutes about 80 per cent of the muscle fibre volume and it is primarily this protein which is lost when a muscle fibre atrophies.

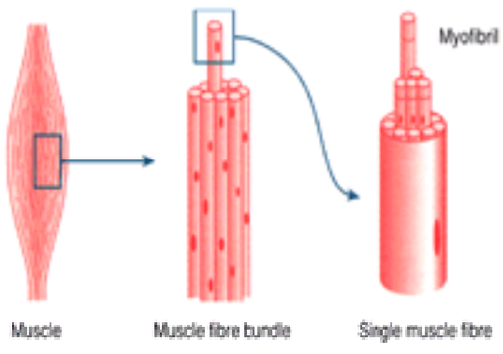


Fig. 6 Diagrammatic representation of the relationship between the muscle, muscle fibres, and myofibrils. In reality, the myofibrils are very much smaller in relation to the muscle fibre than shown here.

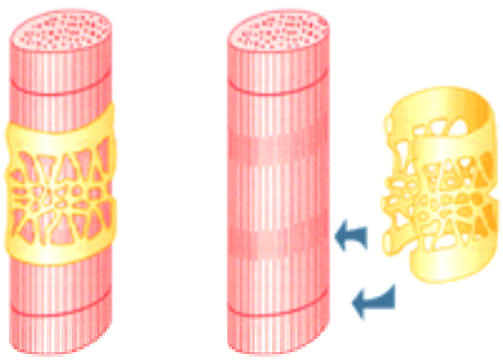


Fig. 7 The sarcoplasmic reticulum envelops a myofibril.

The muscle fibre is a multinucleate structure and, in healthy fibres, the nuclei are situated at the periphery of the cell just under the cell membrane. Muscle fibres grow by the addition of myofibrils, and there can be as few as 50 myofibrils in developing fetal muscle fibres while adult muscle fibres contain about 2000 myofibrils.

The membranous sarcoplasmic reticulum which encases the myofibrils acts as a store for the uptake and release of calcium. In mammalian skeletal muscle, the plasma membrane invaginates twice in every sarcomere forming a complex branching tubular network, the T tubules, which run through the whole fibre contacting every myofibril (Fig. 8). The lumen of the T tubule is continuous with the extracellular space and is separated from the interior of the sarcoplasmic reticulum even though the membranes of the T tubules and the sarcoplasmic reticulum are in close proximity at many points. The portions of the sarcoplasmic reticulum where the T tubules are in closest proximity are known as terminal cysternae (Fig. 9).

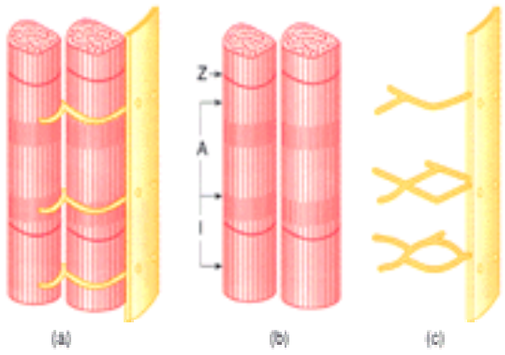


Fig. 8 The T tubules in relation to a sarcomere.

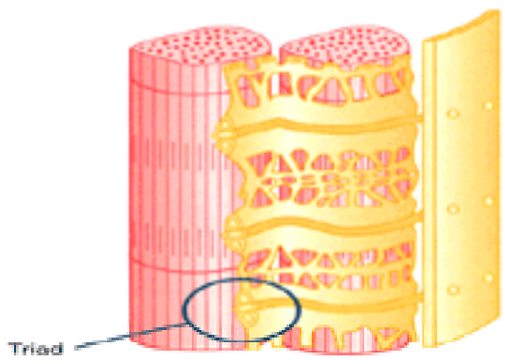


Fig. 9 T tubules, sarcoplasmic reticulum, and the myofibrils. Insert shows the structure of a triad.

Each muscle fibre is bounded by its sarcolemma, the innermost portion of which is the plasma membrane, which is the true boundary of the cell. This surface (plasma) membrane of skeletal muscle has all the properties of selective permeability and transport which are found in other excitable tissues, such as nerve, but it also has a number of distinctive features. Associated with the T-tubular membranes are modified calcium channels, known as DHP receptors as they bind dihydropyridines such as verapamil. These receptors are part of the link between the surface and the sarcoplasmic membranes that couples the passage of an action potential along the surface membrane to the release of calcium from the sarcoplasmic reticulum. The T-tubular membranes also contain a large number of chloride channels that act to stabilize the membrane potential. Where these channels are blocked or absent, as in myotonia congenita, the surface membrane is susceptible to repetitive bursts of firing in response to a single depolarization.

The surface membrane contains a variety of cytoskeletal proteins that may have the additional role of connecting the underlying sarcomere structure to the surface. One of these proteins, dystrophin, is of particular interest as its congenital absence or modification gives rise to the family of muscular dystrophies. How the absence of this one protein causes the fibre damage and degeneration seen in these conditions is not known and remains one of the major challenges in linking molecular genetics to cellular processes (McArdle *et al*. 1995).

Outside the plasma membrane lies the basement membrane, which is not a membrane in the usually accepted sense of the word as it does not have a lipid bilayer structure but consists of a loose glycoprotein and collagen network. The basement membrane is freely permeable and may enclose more than one fibre. Following muscle fibre damage the basement membrane forms a framework within which regeneration may occur.

At the end of a muscle fibre the outer membranes become irregular and indent to form a close link with the connective tissue. The connective tissue elements come together to form the tendons which join muscles to the bony skeleton (Fig. 10).

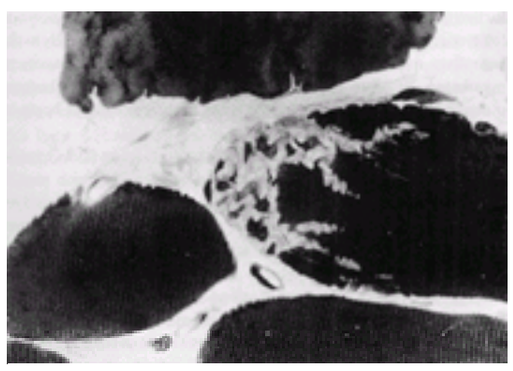


Fig. 10 Myotendinous junction. Low-power electron micrograph.

Situated between muscle fibres are fibroblasts which secrete collagen fibres to form a connective tissue matrix, the endomysium. Groups of 10 to 100 muscle fibres are bound together by a thicker layer of connective tissue, the perimysium, to form fascicles (Fig. 11). Fibroblasts and their major product, collagen, play an important role in maintaining muscle structure. Muscle fibres that suffer irreparable damaged are replaced by collagen fibres secreted by the ubiquitous fibroblasts; massive replacement of muscle tissue by connective tissue can be seen in biopsies from the muscles of patients with advanced muscular dystrophy or steroid-resistant autoimmune muscle disease (Fig. 12).

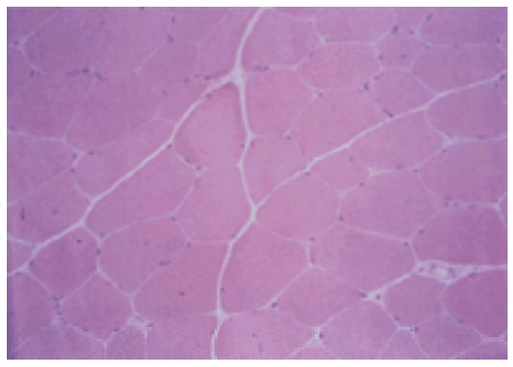


Fig. 11 Transverse section through a portion of human muscle showing delineation of the tissue into fascicles. Light microscopy, haematoxylin and eosin stain.

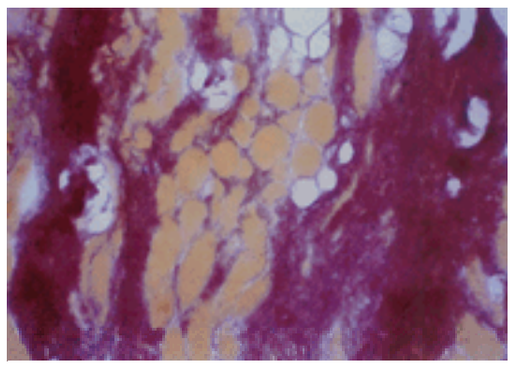


Fig. 12 Transverse section of quadriceps muscle from a 15-year-old boy with Duchenne muscular dystrophy showing replacement of muscle fibres by fibrous connective tissue. Van Geison stain; fibrous tissue stained red, muscle fibres stained yellow.

Small blood vessels and motor axons traverse the perimysial spaces to make contact with the muscle fibres. Muscle spindles, which detect change in muscle length, are also found in the perimysium enclosed in connective tissue envelopes (Fig. 13). Each muscle is covered by a thick outer connective tissue layer, the epimysium.

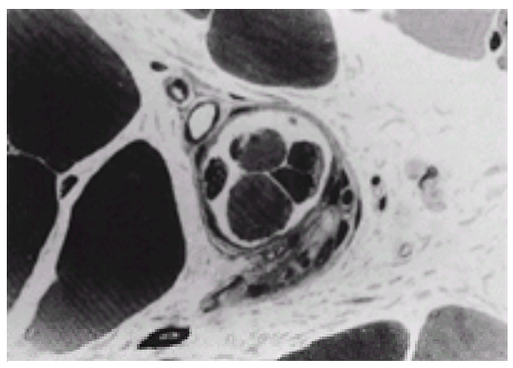


Fig. 13 Low-power electron micrograph of a muscle spindle. The intrafusal spindle fibres are contained within a connective tissue capsule, separating them from the extafusal fibres which make up the bulk of the muscle.

Embryonic origins and fetal development

Skeletal muscle is derived from the embryonic mesoderm. A myogenic fate is imposed on undifferentiated mesodermal cells by the action of a family of myogenic determining factors, the *myf* genes: *myf3* (*MyoD*), *myf4* (myogenin), *myf5*, and *myf6* (*MRF4*) (mouse genes in parentheses) (Buckingham 1992). These helix–loop–helix transcription factors continue to exert a controlling action on the genes encoding many muscle proteins and, together with thyroid hormone, they regulate the sequential development of muscle in the fetus. There is evidence that *myf* genes are also important in the adult in regulating gene expression in damaged and denervated muscle (Buckingham 1994).

At about the sixth week of gestation, mesodermal cells begin to divide and differentiate to form myoblasts. Some myoblasts remain as single cells with mitotic potential, but the majority aggregate and fuse to form primary myotubes attached at their ends to the tendons and the developing skeleton. Within a developing myotube a central chain of nuclei forms (Fig. 14), surrounded by basophilic cytoplasm rich in polyribosomes.

Fig. 14 Cultured myoblasts fuse to form myotubes. Note the chain of central nuclei.

Midway along the primary myotubes further myoblasts aggregate and fuse to form secondary myotubes. At first the primary and secondary myotubes share a common basement membrane (Fig. 15), but eventually the secondary myotubes develop a separate basement membrane, make contact with the tendon, and become independent of the primaries. In the human fetus the transition from myoblasts to primary myotube takes place at around the seventh to ninth week of gestation and by the end of this period the primordia of most muscle groups are well defined. At this time the synthesis of the contractile proteins, actin and myosin, occurs, the first signs of cross-striation are visible (Fig. 16), and contractile activity begins.

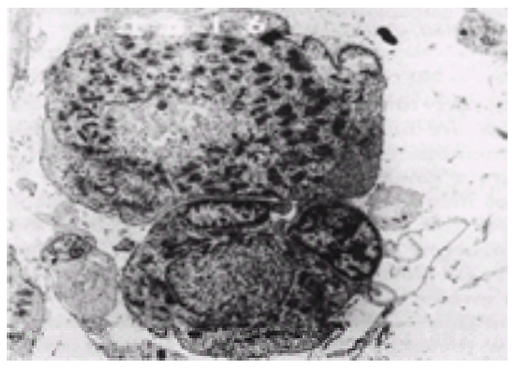


Fig. 15 Developing myotubes. Primary myotube with two attached myoblasts forming secondary myotubes. Electron micrograph of tissue at 9 weeks of gestation.

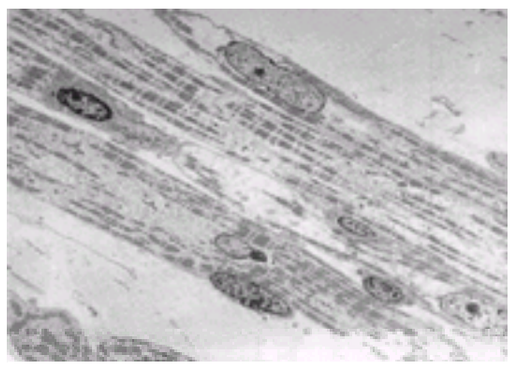


Fig. 16 Longitudinal section of developing myotubes showing the formation of myofibrils. Electron micrograph of tissue at 12 weeks of gestation.

From 11 weeks of gestation onwards there is a proliferation of myofibrils leading to hypertrophy of the muscle fibres, which also grow in length by the addition of sarcomeres at the ends. At 16 to 17 weeks a further population of myotubes becomes apparent and are known as the tertiary myotubes. These myotubes are small, adhere closely to the secondary myotubes, and are enclosed within the same basement membrane. By 18 to 23 weeks the tertiary myotubes have become independent and the nuclei of the more mature myotubes begin to move to the periphery of the fibre.

Under the influence of the *myf* genes the developing fibres express a number of different myosins which can be identified with monoclonal antibodies; these myosins include an embryonic form and an intermediate fast type that confers the properties of the type 2c fibres (often seen in regenerating adult muscle) when stained for myosin ATPase. The primary myotubes can be identified throughout embryonic development as they alone express adult slow myosin from about 9 weeks of gestation (Drager *et al*. 1987). At around 10 weeks the nervous system makes contact with the developing muscle fibres and, in response to the contractile activity imposed by the motor nerve, the fibres differentiate so that, eventually, fetal myosins are no longer expressed and about 50 per cent of fibres contain adult slow myosin and 50 per cent fast myosin. This process, which is apparent by about 32 weeks of gestation, is not fully completed in human muscle until a few months after birth.

Satellite cells

Integral to the growth of skeletal muscle and its recovery after damage are the satellite cells, first described by Mauro (1961).

During development of skeletal muscle some of the mesodermal cells fail to fuse but remain closely apposed to the walls of the developing myotubes. These undifferentiated myoblasts are eventually enclosed by the basement membrane of the mature muscle fibres but remain outside the sarcolemma and are known as satellite cells. Under the light microscope it is very difficult to distinguish muscle fibre nuclei from the 5 to 10 per cent of nuclei that belong to satellite cells, but with the electron microscope the position of the satellite cells beneath the basement membrane, but outside the plasma membrane, of the muscle can be seen (Fig. 17). If the muscle fibre suffers damage, the satellite cells are activated to divide, form new myonuclei, and begin the process of regeneration.

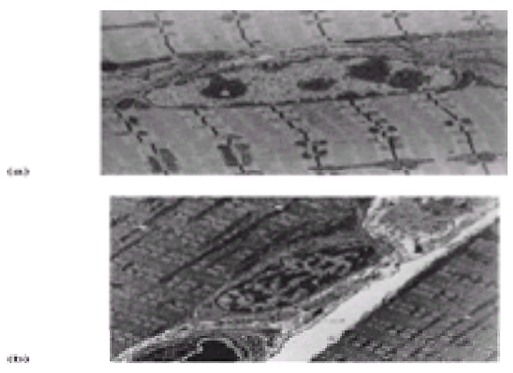


Fig. 17 Muscle fibre nucleus and satellite cell. (a) Muscle fibre nucleus lying just below the plasma membrane of the muscle cell. (b) Satellite cell lying outside the plasma membrane with a layer of cytoplasm around its nucleus.

Satellite cells isolated from both regenerating and non-regenerating muscle can be distinguished from embryonic myoblasts on the basis of their myosin heavy-chain expression, they also differ in their expression of desmin, an integral membrane protein, and enolase, a glycolytic enzyme. These differences suggest that mature satellite cells are a distinct class of myogenic cells. There is also some differentiation of function within the satellite cell population, some cells functioning as stem cells while others seem readily available to fuse and form new myotubes (Schultz and McCormick 1994).

In the new-born child about 30 per cent of nuclei are satellite cells, but this number decreases to between 5 and 10 per cent in mature muscle. Satellite cells are more plentiful in oxidative fibres in all mammals examined and more common in the region of the motor endplates (neuromuscular junctions) where the motor axon makes contact with the muscle fibre (Fig. 18).

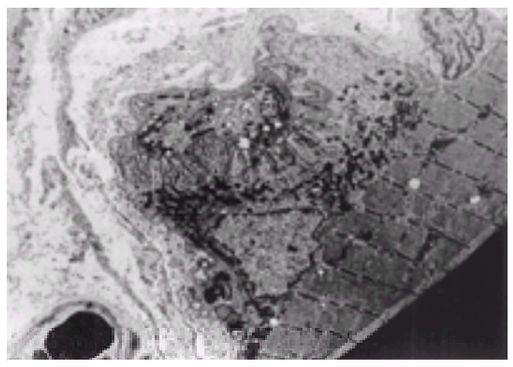


Fig. 18 Neuromuscular junction under low-power electron microscopy. There is only one neuromuscular junction per muscle fibre, the appearance is of very regular infoldings of the muscle plasma membrane.

Satellite cells are fusiform in shape, about 25 µm in length in most mammals, and using electron microscopy, it can be seen that there are multiple branches emanating from their poles. It appears that these branches are connected with the ability of satellite cells to migrate within the basement membrane (Schultz 1976). No evidence has been found that satellite cells communicate in any way with the fibres to which they are adjacent, they appear to be separate entities. Rare reports of cytoplasmic continuity appear to be associated with the proliferation of new myotubes rather than with communication between satellite cells and mature muscle fibres

Vital stains, and antibodies to membrane proteins are now available to aid recognition of satellite cells. Human satellite cells express the neural cell adhesion molecule (NCAM) cell-surface glycoprotein but not its embryonic form (ECAM) when muscle is growing. In regenerating muscle both NCAM and ECAM are expressed, while another marker of activation is upregulation of the *myf* genes. The use of tritiated thymidine and bromodeoxyuridine as nuclear labels in experimentally damaged animal preparations has enabled dividing satellite cells to be identified and, using this technique, the numbers of cells activated after a particular injury has been calculated (Schultz 1989).

In tissue culture, satellite cells respond to fibroblast growth factor, insulin-like growth factor 1, transforming growth factor, platelet-derived growth factor, adreonocorticotropin hormone, and melanocyte-stimulating hormone.

The response of satellite cells to these factors is variable among species and between different cultures of cells. Some of these factors, such as insulin-like growth factor 1 and fibroblast growth factor, are present in muscle *in vivo*, and receptors for insulin-like growth factor 1 have been demonstrated on satellite cells. It is probable that they play some role in satellite cell proliferation and fusion during growth.

To date the only known function of satellite cells in mature muscle is the production of new myonuclei and the initiation of new myotube formation, although recent work has suggested that, in early development, they may play a role in the production of the extracellular matrix proteins, fibronectin and laminin (Schultz and McCormick 1994).

The appearance of activated satellite cells is a good indication of perturbation occurring in a muscle, such as damage or the stimulus for growth. In rats the response to injury begins after several hours with the expression of myoD and myogenin. Proliferation of nuclei occurs after about 24 h and new myotubes begin to develop. Successive cycles of damage lead eventually to a reduction in the numbers of satellite cells and in the efficiency of the remaining population. This fact may provide an explanation for the decline in regenerative ability seen in the muscles of patients with advanced Duchenne muscular dystrophy (Webster and Blaud 1990) or chronic inflammatory muscle disease that is not responsive to steroids. The number of satellite cells which survives each cycle of damage also depends on the way in which the myofibres were killed; experiments with animals have shown that satellite cells do not survive ischaemic damage any better than other cells, and repair of the muscle fibres is initiated by satellite cells which migrate from adjacent undamaged areas of the muscle. Migration occurs extensively along the longitudinal axis of a muscle while only a small amount of transverse migration can occur. When muscle is injured its entire satellite cell population can potentially be recruited to aid in repair. Migration between muscles is very limited, so that the ability of a muscle to regenerate depends on its intrinsic satellite cell population.

The size of adult muscle

The number of fibres in each human muscle is probably set by 24 weeks of gestation. In the rat, fibre numbers do not change during life while the mean fibre cross-sectional area increases nearly 10-fold from the new-born to adult animal (Rowe and Goldspink 1969). There are considerable practical and ethical problems involved in making measurements of fibre size and number in children and adults, but the limited data available suggest that there is an increase in size without a change in fibre numbers (hypertrophy without hyperplasia) as the muscles grow in size and strength. The final cross-sectional areas of adult muscle fibre are reached shortly after puberty (Fig. 19). In an adult man about 40 to 45 per cent of the body weight is muscle and this figure is slightly lower for women. The mean cross-sectional area of fibres in a biopsy from the quadriceps muscle in a normal man is about 3500 to 7500 µm ^{2} and in the normal woman from 2000 to 5000 µm^{2}.

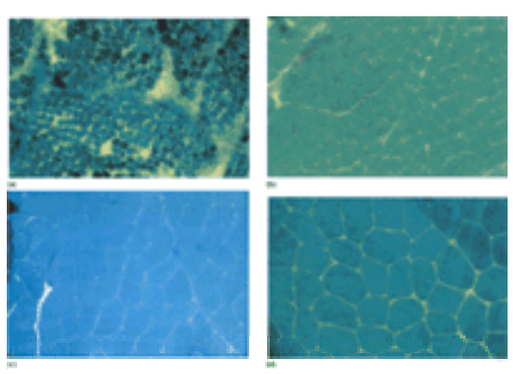


Fig. 19 Growth of muscle fibre size in the human quadriceps. (a) In a baby aged 8 months. (b) In a child of 5 years. (c) In a boy aged 14 years. (d) In an adult male. All at the same magnification.

Innervation of muscle fibres

In a mature muscle, one motor neurone will, through axonal branches, supply a number of fibres scattered throughout the muscle. In a healthy muscle the innervation is almost entirely random and adjacent fibres are likely to be supplied by branches from different motor neurones. All the muscle fibres supplied by one motor neurone form a motor unit (Fig. 20). The number of fibres constituting a motor unit varies from muscle to muscle and may be as few as 10 fibres in a small muscle of the hand to several thousand in a large muscle like the quadriceps. The finer the control of movement, the smaller are the motor units.

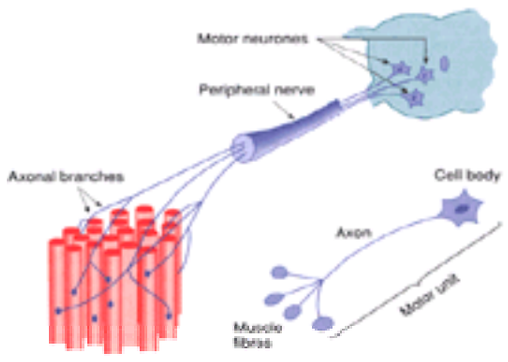


Fig. 20 The concept of the motor unit. A motor unit consists of the motor neurone and all the scattered muscle fibres which it innervates.

Loss of innervation by trauma or by death of the motor neurone, as in motor neurone disease, causes severe atrophy and ultimately necrosis of the deprived fibre if the muscle fibres are not reinnervated by another healthy axon.

Muscle fibre types

Muscle fibres fall into groups determined by their biochemical and physiological properties which are in turn determined by the type of motor neurone controlling the motor unit to which they belong. Various histochemical stains can be used on sections of unfixed frozen muscle to differentiate different fibre types, the most commonly used are the myosin ATPase stain with preincubation at either acid or alkaline pH and the NADH tetrazolium reductase or succinate dehydrogenase stains for mitochondrial activity. An indicator of glycolytic activity, such as myophosphorylase, is commonly assessed (Fig. 21). Three main fibre types can be identified. Type 1 (slow oxidative) fibres with low myosin ATPase activity at alkaline pH, high mitochondrial content, and relatively low glycolytic enzyme activity. Type 2 fibres (fast glycolytic) which stain darkly with myosin ATPase at pH 9.4 and have high glycolytic but lower mitochondrial content than type 1 fibres. The type 2 fibres can be subdivided into 2a and 2b, the former having a higher mitochondrial content and being more fatigue resistant. These two fibre subgroups can be differentiated by the myosin ATPase stain at pH 4.3 and 4.6 (Table 1).

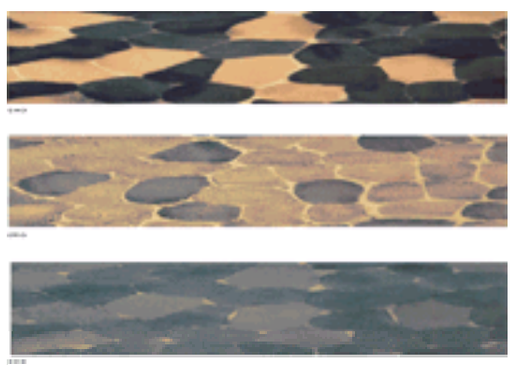


Fig. 21 Serial transverse sections of human quadriceps muscle stained with (a) ATPase at pH. 9.4, (b) NADH tetrazolium reductase for mitochondrial activity, and (c) for phosphorylase. Fibres high in phosphorylase are low in mitochondrial activity and stain darkly (type 2) with ATPase at pH 9.4.

	Muscle fibre type			
Stain	**1**	**2a**	**2b**	**2c**
ATPase (pH 9.4)	1+	3+	3+	3+
ATPase (pH 4.6)	3+	0	3+	3+
ATPase (pH 4.3)	3+	0	0	2+
NADH-TR	3+	2+	1+	2+
Phosphorylase	1+	3+	3+	3+

NADH-TR, NADH tetrazolium reductase is a stain for mitochondrial activity. Numbers indicate intensity of staining.

Table 1 Histochemical staining of human muscle fibre types

The broad classification of skeletal muscle fibres into three main types is adequate for most purposes when considering normal and pathological situations, but recent investigations show that there are complex combinations between the *myl* and other genes which control the expression of contractile proteins and, especially with the fast fibres, there can be variable amounts of mixed gene expression in single fibres.

Contractile properties

Fast muscles rapidly develop force and also relax more rapidly than do the slow muscles. The fast extensor digitorum longus muscle of a mouse, when stimulated at a frequency of 10 to 20 Hz, reacts quickly enough for the tension to fall back to the baseline before the next impulse, but in the slower soleus muscle the next impulse comes before relaxation is complete and the contraction is superimposed on the tension remaining from the previous stimulus. In this way the individual twitches are said to summate or fuse. When stimulated at a sufficiently high frequency the muscle will produce a smooth plateau of force. The frequency required to achieve this plateau is known as the fusion frequency and this frequency is higher for fast compared with slow muscles (Fig. 22).

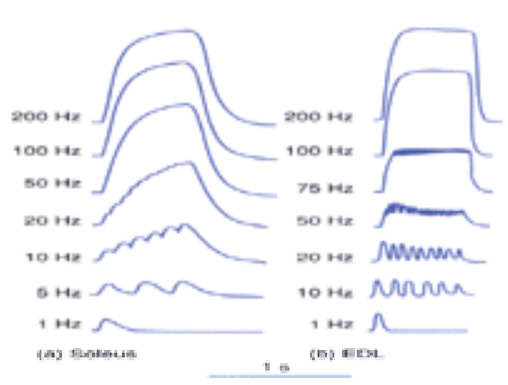


Fig. 22 Force generated during a 500-ms tetanus at different frequencies (isolated muscles, 25°C). (a) Mouse soleus; (b) mouse extensor digitorum longus.

As first recorded by Ranvier (1873), there is a clear connection between the appearance and the contractile properties of certain skeletal muscles, such as the red and white muscles of a chicken or the soleus and extensor digitorum longus in a mouse. These muscles are, however, somewhat unusual in consisting predominantly of one fibre type while the majority of skeletal muscles are a mixture of different types.

In animal preparations it is possible to characterize the functional properties of individual motor units in a muscle and, on the basis of size, speed, and fatiguability, they are found to fall between two extremes: large, fast, and fatiguable; or small, slow, and fatigue resistant (Fig. 23) (Table 2) (e.g. Burke *et al*. 1971) and all the evidence indicates that a similar type of organization applies to human muscle.

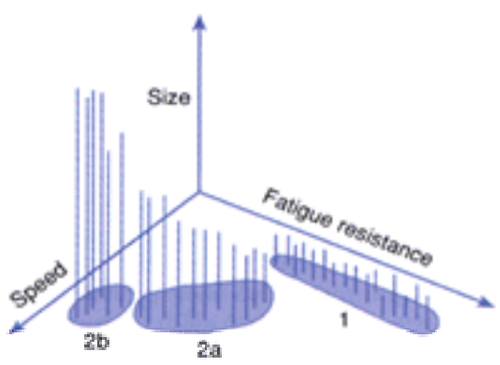


Fig. 23 Contractile properties of motor units. The size, speed, and fatigue resistance of different types of motor unit, as defined by their histochemical properties.

Type	Guinea pig and rabbit	Cat
1	**SO** Slow twitch Oxidative	**S** Slow twitch
2a	**FOG** Fast twitch Oxidative/glycolytic	**FR** Fast twitch Fatigue resistant
2b	**FG** Fast twitch Glycolytic	**FF** Fast twitch Fatiguable

After Burke *et al*. 1971; Peter *et al*. 1972.

Table 2 The relationships between histochemical fibre type classification and physiological properties; two ways are given of classifying the physiological properties

Examining the histochemical properties of the motor units shows that large units tend to be made up of type 2b fibres while the small slow units are predominantly composed of type 1 fibres. Type 2a motor units span a range of size and fatigue resistance which is reflected in the broad spectrum of their mitochondrial enzyme activities.

After severe traumatic damage, muscle fibres regenerating from active satellite cells within the basement membrane will express early myosins as they develop, the final myosin expression being determined through the neuromuscular junction by the motor nerve branch. If reinnervation follows the death of the motor neurone and loss of the muscle fibres constituting its motor unit, regenerating fibres may be innervated in a nonrandom fashion. This type of regeneration leads to muscle fibre grouping which can be detected with histochemical stains (Fig. 24(a)) and by giant action potentials on an electomyographaph (Fig. 24(b)).

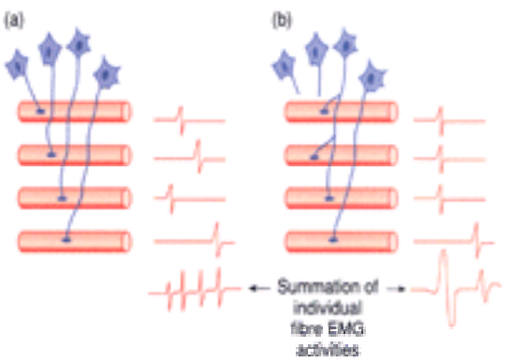


Fig. 24 (a) Normal muscle with fibres innervated by different motor neurones. The activity recorded is the summation of asynchronous activity in the fibres. (b) Muscle after denervation and reinnervation. A single motor neurone supplies a group of fibres giving a giant action potential.

Length–tension relationship

There are two components of the total force that can be measured: the passive tension which is due to stretching the connective tissue elements of the muscle and, possibly, the structural protein, titin; and the active tension which is superimposed on the passive tension when the muscle is stimulated (Fig. 25). In the discussion that follows we are concerned only with the active tension, the passive component being subtracted from the total force.

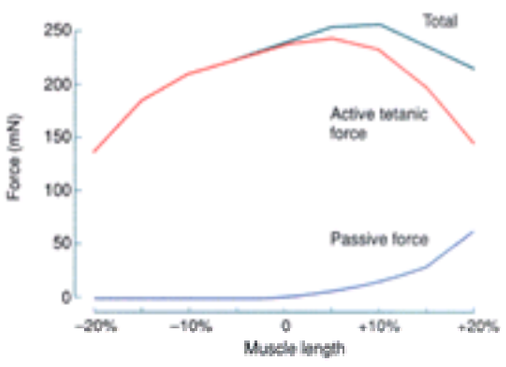


Fig. 25 The two components of force, passive and active tension. Total force is the sum of the two components. Mouse soleus muscle, 26°C, stimulated at 100 Hz.

The main feature of the length–tension relationship is that force declines on either side of an optimum length and, by extrapolating the line at longer lengths, a value can be predicted at which no tension would be generated. This value corresponds to the point where the thick and thin filaments no longer overlap.

Force–velocity characteristics

As the velocity of shortening increases so the force sustained by a muscle rapidly diminishes, eventually reaching a velocity at which force can no longer be sustained at all; this is the maximum velocity of unloaded shortening (v_{max}).

The force at zero velocity of shortening (isometric force) is often referred to as P_0. Muscles of different sizes and therefore different isometric strengths can be compared by expressing the force (P) at a particular velocity as a fraction of P_0.

Force during muscle stretch

In the body, muscles work in pairs, while one muscle shortens its antagonist is stretched, and many movements, such as walking down stairs or lowering weights, involve the extension of active muscles (Fig. 26(a)). The force generated during this type of movement is considerably greater than the isometric force (Fig. 26(b)) and varies with velocity. With increasing velocity of stretch, force begins to plateau, reaching a value of about 1.8 times the isometric force (Fig. 27).

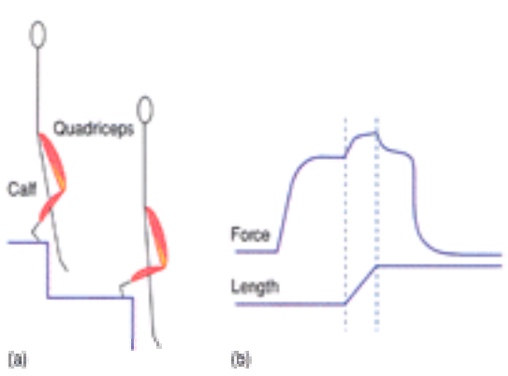


Fig. 26 Stretching muscles. (a) The quadriceps and calf muscles are stretched when lowering the body weight down a step. (b) Mouse soleus muscle stimulated to develop maximum isometric force and then stretched at a constant rate.

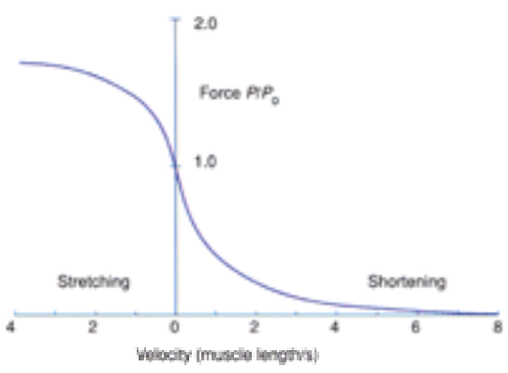


Fig. 27 Force of mouse soleus muscle during stretch and shortening at 26°C.

The overall shape of the curve shown in Fig. 27, with force increasing with stretch and decreasing with shortening, may have important consequences for the stability of sarcomeres acting in series. If one sarcomere is stronger than the next, it might be imagined that the stronger would pull out and extend the weaker, eventually destroying the fibre. However, this is not the case, as when the sarcomeres begin to move, the force generated by the stronger will fall as it shortens (moving down the right-hand portion of the curve in Fig. 27) while the force of the weaker sarcomere will increase as it is stretched (moving up the left-hand side of Fig. 27). Consequently, sarcomeres of unequal strength (due to differences in length, activation, or even damage) can coexist and function to transmit force along the length of the fibre. This holds true when the muscle is working on the left-hand side of the length/tension curve. When a muscle fibre is held at longer lengths, differences in sarcomere lengths along the length of the fibre may be such that this mechanism cannot stabilize the fibre, and shorter and stronger sarcomeres towards the ends of the fibre will pull out and possibly damage the longer weaker sarcomeres in the middle.

ATP splitting during contraction

During an isometric contraction there is no movement, no external work is done by the muscle, and all the energy liberated appears in the form of heat. During shortening, heat is still produced but the muscle also performs work which is the product of the force and distance moved. It was first observed by Fenn in 1923 that during shortening the total energy liberated, in the form of heat plus work, is greater than that occurring during an isometric contraction (Fenn effect) (Fenn 1923). However, heat production falls to low values when muscle is stretched, a fact first noticed by Fenn. More recently, Curtin and Davies (1973) showed that ATP splitting was likewise very low during stretching (Fig. 28).

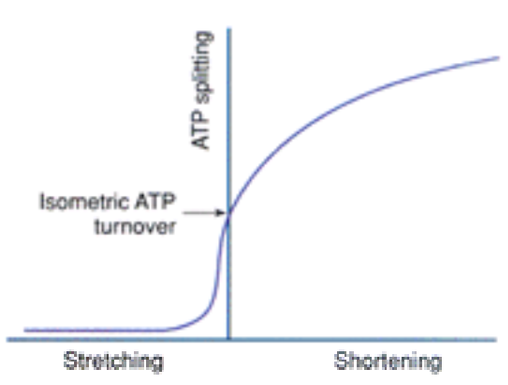


Fig. 28 ATP turnover in muscle allowed to shorten or while being stretched.

Ageing and muscle function

During and after the fifth decade there is a marked loss of muscle mass and a decrease in strength which is particularly notable in women when their hormonal

balance is changed after the menopause. The loss is less severe in men below the age of about 70 years. The more rapid loss of muscle mass in women is reflected in a similar change in bone, with women being much more susceptible to osteoporosis than men.

The loss of muscle in elderly individuals appears to be progressive and by the age of 90 years muscle mass can be reduced by 30 per cent. The loss of muscle bulk is reported to be greater than can be accounted for by atrophy of the muscle fibres, indicating that fibre numbers may be reduced (Grimby and Saltin 1983). It has been suggested that just as neuronal loss from the brain causes memory impairment in the elderly, so loss of motor neurones from the anterior horn causes a loss of motor units. In elderly rats there is a preferential loss of the fast motor units, suggesting that the fast motor neurones may be more susceptible to the processes of ageing. The consequence of this loss is that the slow motor units become larger as a result of reinnervation of the previously fast fibres by axon branches from slow motor neurones, and the slow fibres may become larger as a result of greater use (Kanda and Hashizume 1989). It is likely that a similar process occurs in ageing human muscle.

Histochemical examination of muscle specimens taken at autopsy from elderly subjects with no reported neurological symptoms frequently show areas of fibre type grouping that suggest some neurogenic disturbance has occurred (Fig. 29). These appearances could also be due to peripheral damage to motor nerves. It is of interest that the small muscles in the foot can show clear evidence of neuropathic change with increasing age, which is thought to be a consequence of wearing shoes. Increased collagen cross-linking and lipofuchsin deposits in the muscle are also common findings. Whatever its cause, the loss of muscle may eventually become disabling so that subjects cannot rise from a chair or visit the toilet unaided. A decrease in mobility is an important cause of loss of independence and postponing its onset will become increasingly important as the number of elderly people surviving in the population increases. There is no evidence that habitual exercise can help to prevent the changes associated with ageing although, by preserving cardiac and respiratory function, regular exercise will help the elderly make the best use of the muscle that remains.

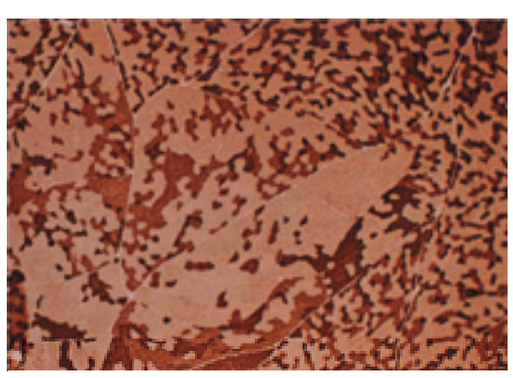


Fig. 29 Adductor pollicis muscle of the hand from an 80-year-old woman with no clinical or neurological signs. Transverse section of an autopsy sample, several fascicles consist largely of only type 1 fibres indicating reinnervation. ATPase stain at pH 9.4.

When examining muscle biopsy specimens from elderly patients with suspected muscle pathology, the fact that reduced fibre areas and some evidence of muscle fibre grouping are common findings in ageing muscle and do not point to any specific pathology should be borne in mind.

Methods of measuring damage

The most common presentation of muscle damage is weakness, sometimes accompanied by myalgia, often of specific muscle groups. If weakness is found it is valuable to quantify the degree in order to have a measure of severity and of progress and recovery with treatment. Simple functional tests such as the ability to rise from a chair or squatting position or stand on tiptoe are useful guides to strength but more objective measures are desirable. There are a variety of commercially available machines that allow measurements at different angles and speeds of contraction, but in clinical practice it is probably most useful to assess the isometric strength (isometric means the same length, i.e. with the muscle unable to move) in a relatively simple testing chair (e.g. Edwards *et al*. 1977; Fig. 30). These techniques are relatively easy to apply and reliable measurements can be made even on children after the age of 5 or 6 years.

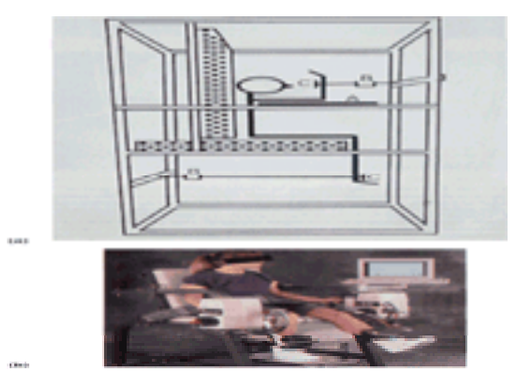


Fig. 30 (a) Diagrammatic representation of apparatus for measuring isometric strength of the quadriceps and elbow flexors. The subject is firmly held in the chair with a strap around the hips. (b) Cybex isokinetic strength testing chair.

Just as is the case with damaged cardiac muscle, damaged skeletal muscle fibres will leak soluble constituents into the circulation and the measurement of enzymes such as creatine kinase or lactate dehydrogenase in the blood can give valuable backup to clinical examination. Both creatine kinase and lactate dehydrogenase have specific muscle isoenzyme fractions which, if measured, allow effluxes from skeletal muscle to be distinguished from enzyme coming from cardiac muscle or other tissues. Myoglobin can also be used as a marker. It appears in the circulation somewhat sooner than creatine kinase but it is an expensive assay and provides little or no clinical advantage over the measurement of creatine kinase.

Histological examination of a muscle biopsy specimen from large muscle groups such as the quadriceps or gastrocnemius, obtained either by the percutaneous needle or choncotome technique, can often provide a diagnosis (Fig. 31). For smaller muscle groups it may be necessary to perform an open biopsy.

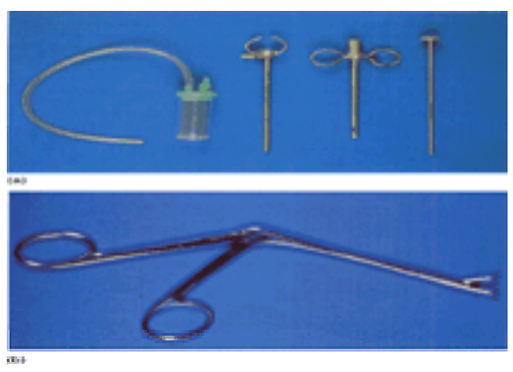


Fig. 31 (a) The University College Hospital (UCH) modification of Bergstrom's original percutaneous muscle biopsy needle. The cutting edge of the inner hollow cylinder is introduced within the outer pointed casing which contains a window at its lower end through which the muscle biopsy is taken. The sterile trap is attached

to a suction mechanism which assists in pulling the muscle tissue into the outer window. The central solid portion of the instrument is used only to remove the specimen after the needle has been withdrawn. (b) Conchotome forceps for percutaneous muscle biopsy.

Electromyograms from weak or myalgic muscle groups can also provide valuable information in confirming a diagnosis. Recording electomyogram activity from an affected muscle group is particularly useful in differentiating myopathic from neuropathic conditions. Polyphasic potentials tend to be seen in inflammatory myopathies while giant potentials are characteristic of neuropathies where muscle fibre grouping has occurred.

Mechanisms of damage

Damage to skeletal muscle can occur in healthy subjects after exercise or accident, as a secondary complication of disease, or as the result of pathology affecting either motor nerves or the muscle fibres themselves. For those whose primary interest is in rheumatic disease the two latter types of damage will be of greatest interest, but details of exercise-induced and ischaemic muscle damage will be helpful in later discussions of the muscle pathologies which occur in polymyositis, polyarteritis nodosa, systemic lupus erythematosus, and other related conditions.

Although the primary causes of muscle damage are varied, there is general agreement that the final stages of muscle cell death are brought about by a combination of relatively few pathways. These pathways are loss of ATP energy supply, loss of intracellular calcium homeostasis, and overactivity of reactions mediated by oxygen free radicals. Microscopically, these processes are revealed as muscle fibre necrosis with increases in lysosomal enzyme concentrations and invasion of the tissue with mononuclear cells.

A fall in intracellular ATP is associated with cell death, probably due to interference with the functions of the muscle cell membrane which leads to an efflux of soluble proteins from the muscle accompanied by an influx of calcium ions. Calcium accumulation will lead to mitochondrial overload and failure of oxidative metabolism which can further exacerbate the situation. Raised intracellular calcium will also stimulate calcium-activated proteases and phospholipases which cause further membrane disruption and a vicious circle of damage leading to final cell death. Abnormal mitochondrial function with poor metabolism of fatty acids can lead to the production of free radical species which will also damage cellular membranes by lipid peroxidation (Fig. 32).

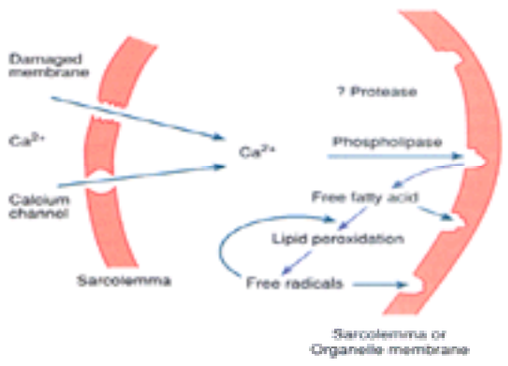


Fig. 32 Suggested mechanisms of muscle damage. Calcium entry may activate phospholipases and, possibly, protease enzymes. The free fatty acids liberated will, in turn, have a detergent effect on cell membranes and can be the substrate for free radical attack.

It is now important to examine the different ways in which this final common pathway of damage can be activated.

Damage due to physical trauma

Skeletal muscles are frequently subjected to physical trauma during everyday life and this risk is increased with participation in contact sports. There is little information about the extent of muscle fibre damage occurring as the result of bumps and bruises, but it is probable that most of the painful sensations experienced are due to inflammation of the skin, subcutaneous tissues, and the muscle fascia, with relatively little involvement of the muscle fibres.

After severe crush injuries or in cases where a patient has been found lying unconscious on a hard surface for some hours, severe muscle fibre necrosis (rhabdomyolysis) can occur with the release of intracellular proteins including myoglobin into the circulation, and a serious risk of subsequent acute renal failure due to precipitation of this muscle protein in the proximal renal tubules. Part of the mechanism which causes rhabdomyolysis in this situation may be that large areas of muscle are rendered ischaemic while the patient is lying for hours in one position and subsequently become reperfused (see below). The measurement of myoglobin in the laboratory is time consuming and expensive and the more commonly estimated soluble muscle protein is creatine kinase, although lactate dehydrogenase and aspartate transaminase are also indicators of muscle damage and are raised in parallel with creatine kinase.

Metabolic depletion

Large metabolic fluxes occur during intense exercise and it is natural to wonder whether excessive metabolic depletion leads to long-lasting changes in muscle which might be classified as damage.

Work with isolated muscle preparations subjected to prolonged damaging stimulation has shown that a build-up of calcium occurs within the muscle fibres, followed by an efflux of soluble enzymes and structural damage to the cells (Jones *et al*. 1984). The raised internal calcium leads to a stimulation of phospholipases which damage the cell membranes. The liberated free fatty acids are then oxidized leading to the production of free radicals, highly reactive species which can initiate a chain reaction leading to further membrane damage (Fig. 32).

Damage during fatiguing exercise

There is little evidence that conventional exercise testing such as work on a cycle ergometer or treadmill leads to significant muscle damage. Heavy exercise of this type maintained for several hours can lead to a rise of two or three times normal in circulating creatine kinase, reaching a peak about 24 h after exercise with a rapid return to normal within 48 h. In the late 1970s the possibility was explored that patients with various muscle disorders might have their condition exacerbated in response to a standard exercise protocol (Brooke *et al*. 1979) and, although some abnormalities were found, it was concluded that significant damage was not incurred in patients after this standard protocol. In healthy subjects, conventional hard work such as running uphill or working against a high load on a cycle ergometer, although metabolically depleting, produced very little muscle damage.

Damage due to eccentric exercise

Although a marked transient (two to three times the upper limit of normal) rise in plasma creatine kinase may occur after any heavy exercise, the most interesting and the most damaging exercise is that in which the muscle is extended while lowering the body weight; many movements, such as walking down stairs, lowering weights, when a person runs downhill, or in an experimental situation when a subject walks backwards down a treadmill, involve the extension of active muscles. This type of exercise is often termed 'eccentric' or plyometric.

The force generated during this type of movement is considerably greater than the isometric force and varies with velocity (see above). After this type of exercise the muscle is weak and some reduction in force can persist for 2 or 3 weeks in severely damaged muscle, subsequently there is a steady return to normal strength.

There are two major consequences of eccentric exercise which remain unexplained despite the considerable attention devoted to its investigation during the last 10 years. The first is the origin of the characteristic pain and tenderness experienced 1 or 2 days after an unaccustomed bout of heavy exercise which contains an eccentric component and with which most of us will be familiar. The second is the nature of the mechanism that causes the delayed rise in plasma creatine kinase which peaks at about 5 or 6 days after the exercise and subsequently falls to normal but precedes the muscle cell damage which is not apparent histologically until about 10 days after the exercise.

Muscle pain after eccentric (plyometric) exercise

During eccentric exercise the subject is aware of high forces generated in the muscle and tendons, but there is no burning sensation like that produced in muscles by metabolically demanding high-intensity isometric or concentric activity such as running up a long flight of stairs. After about 20 min of eccentric activity the subject usually notices a loss of force and increased tremor accompanied by an inability to fully flex or extend the affected limb. These sensations are generally described as unusual but not as painful. By 6 to 12 h after exercise, discomfort develops in the exercised muscles and typically the subject notices pain the morning after the exercise. The major sensation is not of continuous pain but of discomfort elicited by pressure on the affected muscle. Some degree of muscle oedema is also present. The degree of discomfort may be quantified by measuring the pressure required to elicit a painful response (Fig. 33).

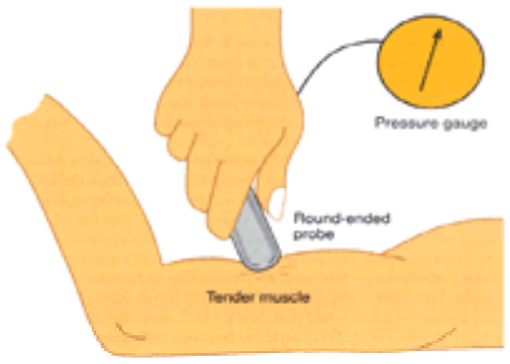


Fig. 33 Quantitative assessment of muscle tenderness. A round-ended probe attached to a pressure transducer is pressed into the muscle and the pressure recorded at which pain is first felt.

Associated with the tenderness is a feeling of muscle stiffness which limits the range of movement of the limb. Electromyography of the affected limb has not demonstrated any unusual electrical activity.

Damage after eccentric (plyometric) exercise

The most intriguing feature of muscle damage induced by eccentric exercise is the delay which occurs in many subjects between the end of the precipitating exercise and the evidence of damage in terms of both the leakage of soluble constituents from the muscle fibres and the necrosis and regeneration visible microscopically in a biopsy of tissue from the affected muscle.

A delay of some hours might be explicable in terms of slow diffusion of large proteins into the general circulation, as seen with the appearance of cardiac damage after infarction. However, following eccentric exercise there is usually a delay of about 2 days before any change is seen and the peak response is seen 4 to 6 days after the exercise (Fig. 34). Observations on patients undergoing orthopaedic surgery during which large amounts of muscle are either cut or at least roughly handled have shown relatively small increases in plasma creatine kinase, which reached peak levels 1 or 2 days after the operation (Fig. 35) (Jones *et al*. 1991), so there is a clear difference between the time course of damage caused by eccentric exercise and that provided by direct physical trauma to muscle fibres. The delay between exercise and damage could be explained if the exercise were to cause an initial change in the muscle which was amplified over the next few days until it resulted in muscle necrosis. There are a number of changes in the structure and contractile characteristics of the muscle which may be indicators of this initial damage. Immediately after the exercise, changes such as Z-line streaming (Fig. 36) and sarcomeres that have the appearance of being pulled apart can be seen in the ultrastructure of the muscle. These changes may be limited to isolated sarcomeres or wider areas of disruption may be seen (Friden *et al*. 1983). Immediately after eccentric exercise there is a loss of maximum voluntary isometric force which is not connected with pain or loss of central drive as similar force loss is seen with high frequencies of stimulation. The loss of strength and contractile properties persists for hours and sometimes for days (Newham *et al*. 1983a). Long-lasting alterations in function are generally ascribed to 'damage' rather than 'fatigue' since the latter carries with it the implication that recovery will occur within a period of minutes rather than hours. These long-lasting changes have been taken as evidence of 'microdamage' to the muscle fibres which, if of sufficient magnitude, could have a cumulative effect overwhelming the capacity for repair and leading to necrosis of a segment or the whole of the muscle fibre.

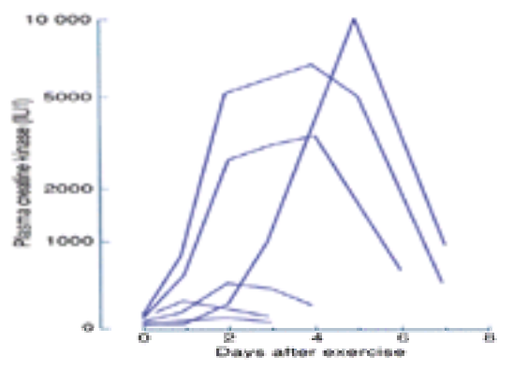


Fig. 34 Plasma creatine kinase levels following 30-min stepping exercise. Subjects stepped on and off a stool adjusted to just above knee height at a frequency of 15 cycles per minute using, every time, the same leg to step up and the opposite to step down.

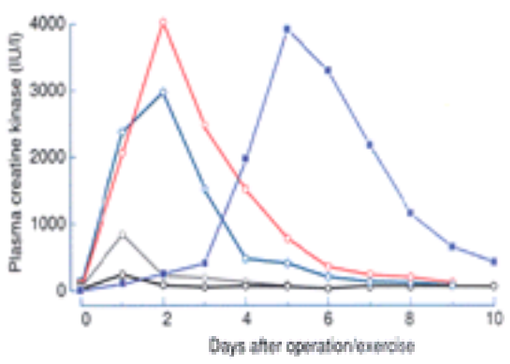


Fig. 35 Plasma creatine kinase (CK) measured in patients undergoing a range of orthopaedic operations. The values are compared with those obtained from one subject (■) who had undertaken eccentric exercise of the biceps.

Fig. 36 Electron micrograph of skeletal muscle after eccentric exercise to show Z-line streaming.

One of the first descriptions of Z-line streaming and sarcomere disruption as a consequence of eccentric exercise was made from observations on biopsies from the quadriceps muscles of subjects who had undertaken stepping exercise (Newham *et al*. 1983b). It was noted at the time, although not commented on, that there was no correlation between the extent of damage seen with electron microscopy and the subsequent creatine kinase release. In a further series of experiments using radioactive technetium pyrophosphate to identify damaged muscles, it was shown that as a result of stepping exercise the gluteus and one of the adductor muscles was affected, but not the quadriceps muscle (Newham *et al*. 1986).

Thus there is evidence of microscopic damage to the quadriceps which does not progress to full-scale degeneration of the muscle fibres (although it is still possible that gross muscle damage only ensues if some threshold is exceeded). The Z-line streaming and disruption of sarcomere structure probably includes damage to the sarcoplasmic reticulum and T-tubule systems and it is reasonable to assume that the immediate loss of force and change in the force/frequency relationship are reflections of the extent of this damage. Assuming that changes in contractile properties are indicators of 'microdamage', it has been noted that these changes do not correlate well with the extent of subsequent fibre degeneration, as indicated by creatine kinase release. This inconsistency is illustrated well by considering the effects of training. Figure 37 shows the loss of strength after a series of bouts of exercise at 2-week intervals (Newham *et al*. 1987). On each occasion the initial loss of force approached 50 per cent with recovery taking a little over a week and a very similar pattern of change was seen with the force/frequency relationship. Figure 38 shows the very different pattern of creatine kinase release from the exercised muscle between the first and the two subsequent bouts of exercise. After the first exercise session there was a large release of creatine kinase into the circulation with individual values ranging from 1570 to 10 904 IU/l (upper limit of normal 200 IU/l), reaching a peak at 5 days and returning to base levels within 10 days. Following the second and third exercise sessions there was no significant increase in circulating creatine kinase, indicating that the muscle had become resistant to the dramatic destructive process. This finding is in contrast to the changes that occurred in the contractile properties of the muscle with training and shows quite clearly that if microdamage is responsible for the changes in contractile properties it is unlikely to be a precursor of the delayed-onset muscle degeneration, especially as the extent of the initial force loss was so similar on the three successive occasions.

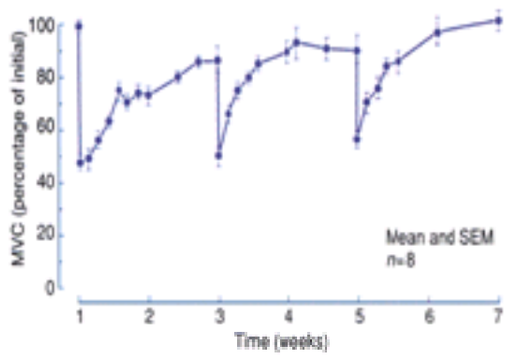


Fig. 37 Changes in strength with repeated exercise. Eight subjects exercised their forearm flexors with maximal eccentric contractions, repeating the exercise every 2 weeks. Maximum isometric strength was measured at regular intervals and the results are expressed as a percentage of the strength before the first period of exercise.

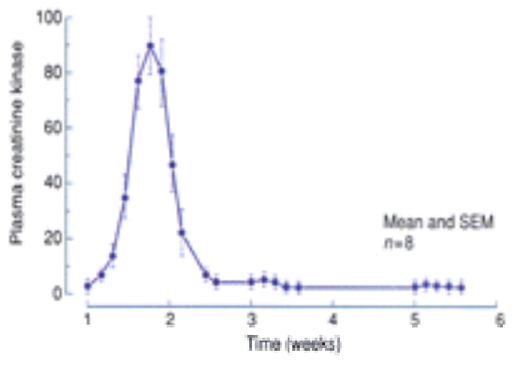


Fig. 38 Changes in creatine kinase (CK) response to repeated exercise. Plasma CK measurements were made on the subjects described in Fig. 37 and the results are expressed as a percentage of the highest value obtained for each individual.

Muscle biopsy after eccentric exercise

If a muscle biopsy is taken a few days after the peak of the creatine kinase rise, fibre necrosis and regeneration with invading mononuclear cells are seen, and in many cases the microscopic picture transiently resembles that seen in the muscle of patients suffering from inflammatory muscle disease (Round *et al*. 1987) (Fig. 39). Unlike the situation in inflammatory muscle disease where regeneration may be slow even after steroid therapy, further biopsies of this exercise-damaged muscle show that regeneration is virtually complete 3 to 4 weeks later (Fig. 39).

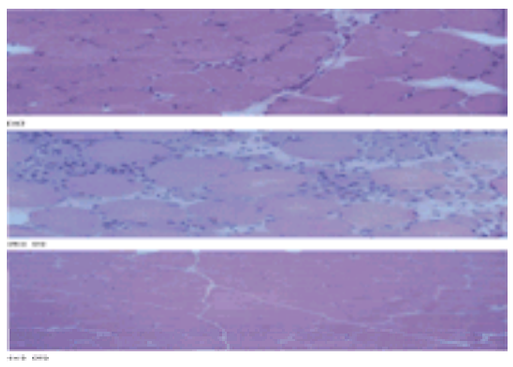


Fig. 39 Haematoxylin and eosin stains of transverse sections of: (a) biceps brachii, 10 days after eccentric exercise; (b) quadriceps muscle, 12 days and (c) 6 weeks after eccentric exercise.

Muscle damage after reperfusion of an ischaemic limb

The techniques of vascular surgery are now so well advanced that it is possible to salvage limbs which a few years ago would have been amputated. One of the side-effects of these advances has been the occurrence of occasional cases where, following successful revascularization, the patient experiences severe rhabdomyolysis often accompanied by acute renal failure. In this situation the huge rise in circulating creatine kinase and myoglobin is not seen to occur immediately after revascularization but 24 to 48 h later, suggesting that damage to the muscle has occurred not as a direct result of ischaemia but due to some subsequent damaging effect that occurred after oxygenated blood flow has been restored to the limb.

Muscle specimens taken immediately after surgery show very little evidence of fibre damage, but biopsies obtained several days later show the characteristic picture of fibre necrosis, regeneration, and infiltration of the tissue by mononuclear cells (Fig. 40). This chain of events, although occurring over a shorter time span, closely resembles that seen after damaging eccentric exercise. If adequate vascularization can be maintained postoperatively, recovery of the limb musculature is generally satisfactory although full strength may not always be restored (Adiseshiah *et al*. 1992).

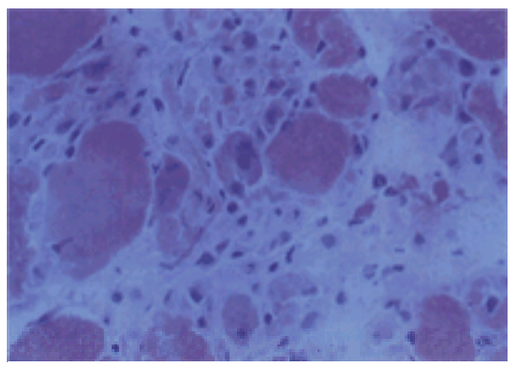


Fig. 40 Ischaemic muscle damage. Histological appearance of a biopsy from anterior tibialis muscle 10 days after surgical revascularization of the limb.

The mechanism of muscle damage following reperfusion is unclear but there are suggestions that processes mediated by free radicals are involved, radicals being generated when oxygenated blood returns to the metabolically depleted tissue. In a number of tissues the conversion of xanthine dehydrogenase to xanthine oxidase has been implicated in the process of reperfusion injury. A similar mechanism might be postulated for skeletal muscle but for the fact that xanthine oxidase is not found in skeletal muscle fibres (McArdle and Jackson 1994; McCord 1985). However, the enzyme is found in the capillary endothelium and it is possible that this endothelium is the site of primary damage in reperfusion injury. The rhabdomyolysis would then be a secondary phenomenon consequent upon the increased permeability of the damaged capillaries.

Muscle pathology in rheumatic disease

Weakness, particularly of proximal skeletal muscles, is frequently reported in patients with rheumatic disease. There are two major causes of such weakness. First, loss of muscle fibres by some destructive process leading to a biopsy picture of necrosis and regeneration, usually accompanied by invasion of the muscle by mononuclear cell infiltrates; such a picture is commonly seen in biopsies from patients with polymyositis and dermatomyositis. Second, atrophy of the muscle cells with loss of contractile material may occur due to disuse of the muscles through joint pain, as in rheumatoid arthritis, or to the prolonged negative nitrogen balance produced by the use of anti-inflammatory steroid therapy, a common finding in many patients who have received treatment for an autoimmune rheumatic disease. In the latter case it is the type 2 fibres which show the most severe atrophy. In some cases a mixed picture of both inflammation and atrophy is seen, as when polymyositis has been controlled by long-term steroid therapy but where the muscle tissue still shows focal inflammatory changes.

Where patients respond well to steroid therapy, muscle strength will show good recovery, but because of the nitrogen losses caused by the high steroid doses, recovery may take many months. Treatment in these cases must be aimed at balancing the therapeutic effects of the steroids against their catabolic side-effects, which may themselves influence the eventual degree of muscle regeneration (Edwards *et al*. 1979).

When examining muscle biopsy specimens it is important to differentiate between the atrophy which affects predominantly type 2 (fast) fibres, as seen for example in steroid myopathies (Fig. 41), and the scattered atrophic fibres of both types 1 and 2 which occur in certain acquired neuropathic disorders such as motor neurone disease, where the motor neurones are sequentially destroyed (Fig. 42). The cause of this relentless destruction is not known.

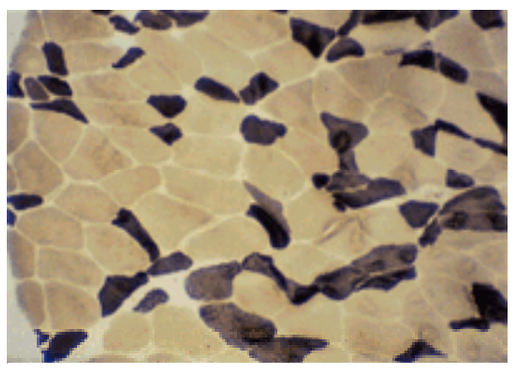


Fig. 41 Type 2 fibre atrophy in a 30-year-old woman on long-term steroid therapy. ATPase stain at pH 9.4

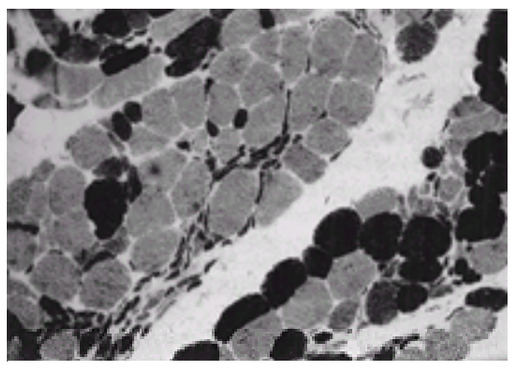


Fig. 42 Motor neurone disease. Transverse section quadriceps muscle from a 35-year-old man. Note the presence of atrophy in some type 1 and type 2 fibres, denoting loss of motor units. ATPase stain at pH 9.4.

The mechanism of muscle damage in rheumatic disease

Skeletal muscle damage is apparent in many autoimmune rheumatic diseases. There may be an apparently direct attack by the immune system on muscle fibres which have undergone some change in cell surface antigens, perhaps due to a virus infection, rendering them 'foreign' to the patient's immune system. It is notable that the expression of certain HLA determinants may predispose patients to such an abnormal responses after viral infections. Alternatively, muscle damage may be secondary to changes in the blood vessels and capillaries which are closely associated with muscle cells. Immune reactions against the endothelial cells of small blood vessels cause inflammation and necrosis. As a biproduct of this inflammatory process, muscle fibres may become damaged either by mechanisms mediated by free radicals or simply by a loss of blood supply, which reduces oxyen supply to the fibres. As discussed above, a loss of muscle energy will lead to fibre damage and necrotic changes followed by invasion of the tissue by mononuclear cells. The nature of these immune reactions is outside the range of this chapter; they are discussed elsewhere in this textbook.

In polymyositis and dermatomyositis there is widespread, patchy, muscle fibre breakdown accompanied by a lymphocytic invasion of the tissue, particularly in perivascular and perimysial areas (Fig. 43); subsequently, many macrophages are also seen associated with necrotic muscle fibres, and regenerating muscle cells are present (Fig. 44). In severe acute polymyositis, massive necrosis of muscle fibres with severe myoglobinuria and a consequent threat to renal function has been reported. The inflammatory appearances seen in a muscle biopsy usually improve with steroid therapy, but atrophy of type 2 fibres may delay the recovery of full muscle strength.

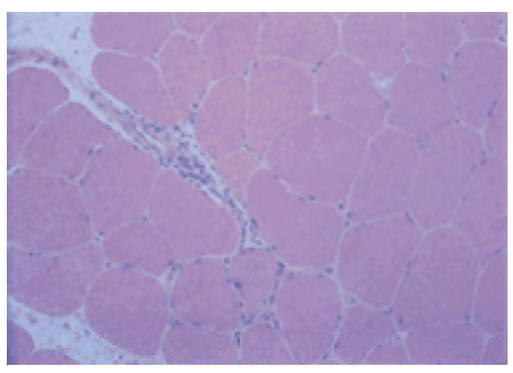


Fig. 43 Early polymyositis in a 38-year-old man. Inflammatory cells are seen in perimysial areas. Haematoxylin and eosin stain.

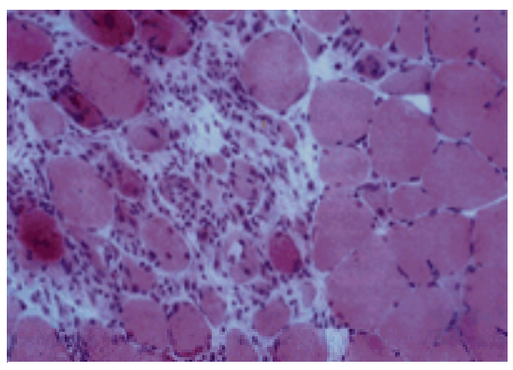


Fig. 44 Severe polymyositis in a 34-year-old woman. Necrotic and regenerating fibres are surrounded by inflammatory cells. One fascicle remains relatively normal showing the 'patchy' nature of the muscle damage. Haematoxylin and eosin stain.

In some cases the patient does not respond to steroids and the damage to muscle fibres becomes relentless, resembling that seen in muscular dystrophy with vacuolated fibres and replacement of contractile material with fibrous connective tissue and fat (Fig. 45). Some of these cases have been identified as inclusion body myositis. In this condition there is insidious muscle weakness accompanied by selective atrophy with an unusual distribution that affects, in particular, the quadriceps in the lower limbs, and forearm and finger flexors in the upper limbs. Most cases are sporadic but familial forms with variable genotypes also occur, in the latter cases inflammatory infiltrates of T cells and macrophages are not usually seen and the creatine kinase may be normal.

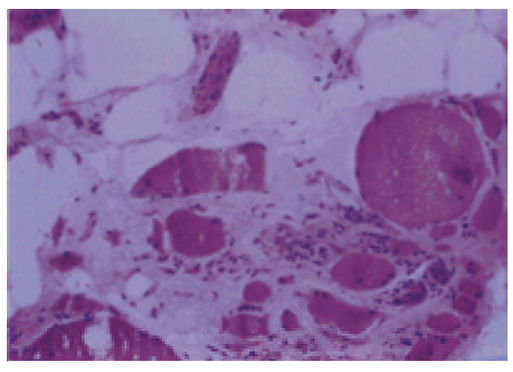


Fig. 45 Steroid-resistant inflammatory muscle disease. Much of the quadriceps muscle is replaced by fat and fibrous tissue. Haematoxylin and eosin stain.

Muscle biopsies reveal the presence of fibres containing characteristic rimmed vacuoles filled with hyaline eosinophilic inclusions. Recent immunohistochemical and electron microscopic studies have shown that these inclusions consist of cytoplasmic twisted tubulofilaments, b-amyloid, ubiquitin, and an array of other proteins including prion proteins similar to those seen in the cerebral plaques in Alzheimer's disease. The latter finding has led to the intriguing suggestion that the degenerative process in inclusion body myositis may have something in common with the degenerative brain changes in Alzheimer's disease (Askanas *et al*. 1993).

Inclusion body myositis is usually considered to be an autoimmune disease and prognosis in these cases is poor, but it is unclear whether the proposed autoimmune process leads to the formation of inclusion-containing vacuoles or the presence of vacuolar inclusions provokes the immune response. Genetic evidence for a possible immune origin in inclusion body myositis is the fact that 90 per cent of patients possess the DR3 allele of the MHC complex, usually on a haplotype marked by HLA B8 and have a deletion at the C4A (complement component) locus, or on a haplotype marked by HLA B18 BfF1 with a deletion of C4B. Both these extended haplotypes have been associated with other autoimmune diseases such as myasthenia gravis, systemic lupus erythematosus, and autoimmune thyroid disease (Garlepp and Mastaglia 1996). Despite this genetic evidence, the variable extent of inflammatory infiltrates and the poor response to steroids where inflammatory changes are seen, lead to the conclusion that the autoimmune status of this disease is far from proven. To date, the mechanisms by which amyloid, prion, and other proteins are produced in excess are not known. No genetic variations have been found in the genes encoding the production of these proteins in patients with inclusion body myositis. The existence of a 'master switch' gene that controls the production of all these proteins remains at present in the realms of speculation.

If multisystem vasculitis occurs in association with autoimmune rheumatic diseases, the skeletal muscles may be affected and fibre damage may subsequently occur. Inflammatory vascular changes are associated with myositis in polyarteritis nodosa (Fig. 46), Sjögren's syndrome, and rheumatoid arthritis. Although vasculitis can occur in rheumatoid arthritis, a more common cause of weakness is the disuse atrophy of both type 1 and type 2 fibres, associated with loss of mobility owing to joint pain. If there has been prolonged steroid therapy, atrophy of type 2 fibres predominates.

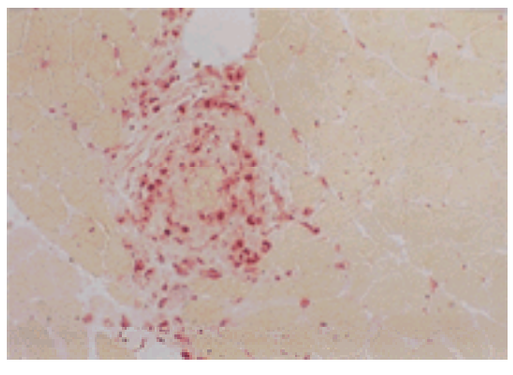


Fig. 46 Polyarteritis nodosa (PAN) inflammatory cells surround a blood vessel. Transverse section of quadriceps muscle. Acid phosphatase stain.

Muscle biopsy in systemic lupus erythematosus is seldom helpful as, although a broad spectrum of vascular abnormalities has been described (Weisman and Zvaifler 1980), the muscle usually appears relatively normal.

In scleroderma the disordered collagen metabolism leads to excess collagen deposits which produce swelling and partial obstruction of the capillaries and small blood vessels in the muscles, which in turn leads to ischaemic changes and necrosis of the muscle cells, sometimes with the presence of inflammatory infiltrates. If the ischaemic damage is widespread, interstitial and perivascular fibrosis is also seen.

Muscle pathology in patients with rare genetic defects

Malignant hyperpyrexia (hyperthermia)

There are some rare pathological conditions in which inherited metabolic defects can lead to severe damage after heavy exercise or unusual metabolic stress. One such condition is malignant hyperpyrexia in which skeletal muscle responds during anaesthesia to the muscle relaxant halothane with a prolonged contracture causing metabolic depletion of the muscle. The first signs of danger are muscle rigidity, a rise in the patient's temperature, and the release of potassium into the circulation which brings with it the risk of cardiac arrest. If the patient survives this acute emergency, the skeletal muscles subsequently undergo necrosis (rhabdomyolysis) with the release of massive amounts of myoglobin and other soluble proteins into the circulation, causing the risk of acute renal failure and another life-threatening crisis. Rapid treatment with dantrolene and cooling has proved effective in the therapy of this serious emergency. The gene defect in this condition has been identified in pigs but, to date, the human gene defect has not been characterized.

The glycogenoses

Other genetically determined conditions give rise to defects in muscle energy metabolism. These include defects in the enzymes of the glycolytic pathway, mitochondrial enzymes concerned with both pyruvate and fatty acid metabolism, and the cytochrome component of the electron transport chain. All of these pathways are important in the secondary supply of energy to the muscles as the primary reserves of phosphocreatine become depleted. In general, patients with metabolic defects are of normal or near normal strength when rested but are limited in their exercise endurance.

When a defect occurs in the glycolytic pathway as, for example, in myophosphorylase deficiency (McArdle's disease) or phosphofructokinase deficiency (Tauri's disease), if prolonged vigorous exercise is undertaken, the exercised muscle tends to go into a painful contracture (which is electrically silent, unlike muscle cramp in normal subjects which is electrically active). However, after a short period of rest, patients with myophosphorylase deficiency can often restart exercise for a short time as bloodborne glucose reaches the muscle cells (second wind phenomenon). This is not the case in phosphofructokinase deficiency, where the gene mutation affects an enzyme further up the glycolytic pathway. Patients vary in the severity of their symptoms, but if severely affected patients are forced to continue exercising, extensive rhabdomyolysis of the exercised muscle can occur. Fortunately these patients are generally well aware of their limitations and take care to reduce the exercise level before any crisis can develop.

Just as patients differ in the severity of their symptoms, they also differ in their ability to express myophosphorylase protein. Some patients produce mRNA but no protein, others produce no mRNA or protein, and one patient has been described who produced mRNA at normal levels but reduced enzyme protein. This molecular heterogenicity implies the involvement of multiple mutations. The muscle glycogen phosphorylase gene is approximately 14 kb in size and has been assigned to the long arm of chromosome 11 (Lebo *et al*. 1984). A C to T mutation in codon 50 has been identified on at least one allele in all patients studied. However, heterozygotes do occur although McArdle's disease is usually considered as an autosomal recessive condition, indicating that these heterozygotic patients must possess an additional mutation(s). Several new mutations have recently been characterized (Bartram *et al*. 1994), but there appears to be no obvious clinical correlation between the severity of the disease and the different mutation patterns, suggesting that other factors must be involved.

McArdle's disease is now categorized as type 5 in a group of inherited glycogenoses. Not all of these conditions affect skeletal muscle. A summary of those associated with muscle damage is given below.

Von Gierke's disease. Here there is a deficiency of the enzyme glucose-6-phosphatase leading to hepatomegaly, growth retardation, hypoglycaemia, and lactic acidosis. The liver, kidney, and skeletal muscles are affected.

Pompe's disease (acid maltase deficiency). The lysosomal enzymes a-1,4- and a-1,6-glucosidase are deficient. In children all tissues are affected with cardiomyopathy, severe weakness and large accumulations of glycogen in the muscle (Fig. 47). Death is usually within the first year of life. In adults the disease may present with a proximal myopathy of variable severity, and weakness of the respiratory muscles.

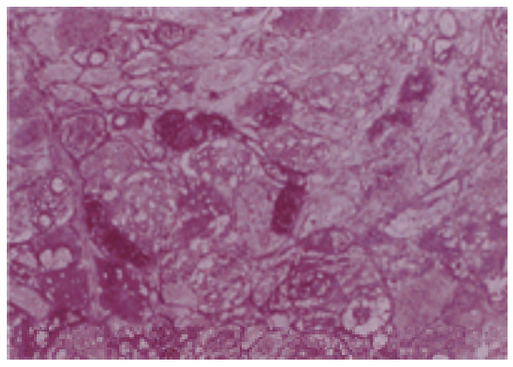


Fig. 47 Acid maltase deficiency. Severe vacuolar myopathy, the fibres are packed with glycogen. Periodic acid–Schiff stain.

Cori–Forbes disease (debrancher deficiency). The absence of debrancher enzyme leads to the production of limited dextrin as the myophosphorylase enzyme removes from glycogen the glycosyl units attached by 1,4-linkages but cannot attack any branch points. This abnormal glycogen accumulates in the muscle. Clinical symptoms vary from severe muscle weakness in childhood to asymptomatic adult forms.

McArdle's disease. In this condition the enzyme myophosphorylase is deficient. The disease is discussed above.

Tauri's disease. Here the enzyme phosphofructokinase is deficient in skeletal muscle and red cells. The clinical symptoms and exercise impairment are very similar to those of McArdle's disease. Because the deficiency is also present in the red cells, a haemolytic anaemia may occur, often detectable by a raised reticulocyte count in the peripheral blood.

Mitochondrial myopathies

Patients with defective mitochondrial function might be expected to have similar problems to those with defects in the glycolytic pathway, as they have severely limited exercise capacity and mild exercise is associated with breathlessness, acidosis, and a high level of blood lactate, such as would be seen in a normal subject exercising under hypoxic conditions. In practice they do not present with muscular damage following exercise. Although often severely limited in their exercise capacity, the normal glycolytic function in these patients must provide sufficient energy to prevent the development of muscle contractures. Histologically, these patients show abnormal 'ragged red' fibres in biopsy sections treated with trichrome stain (Fig. 48), the periphery of the fibres being packed with numerous red-staining mitochondria. Scattered necrotic fibres are also seen. Biochemical analysis of muscle biopsy specimens can be used to identify the location of the genetic defect in the electron transport chain.

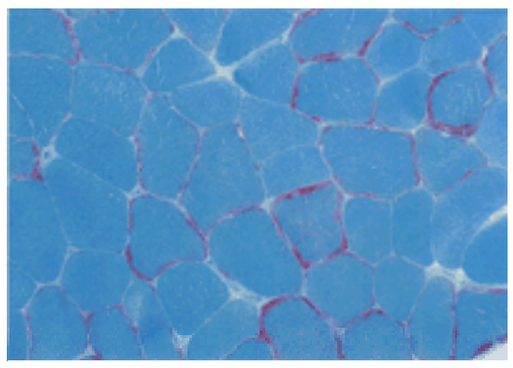


Fig. 48 'Ragged red fibres' in a transverse section of quadriceps muscle from a 19-year-old boy with a mitochondrial myopathy. Dark red staining at the periphery of the fibres is caused by aggregations of mitochondria. Trichrome stain.

Disorders of fat metabolism

These disorders should properly be classified under mitochondrial myopathies, as fat metabolism occurs within the mitochondria, but because defects in fatty acid metabolism do not have such a drastic effect on energy supply as defects in the electron transport chain, they are considered separately. Patients with disorders of fatty acid metabolism have symptoms of weakness, exercise intolerance, muscle stiffness, and pain (sometimes accompanied by myoglobinuria). Symptoms are most evident at times when free fatty acids are the main substrates for energy metabolism, such as during prolonged submaximal exercise, particularly in the fasting state. The b-oxidation of fatty acids depends on the transport of free fatty acids into the mitochondria by a shuttle mechanism which involves carnitine and the carnitine palmitoyl transferase (**CPT**) enzymes. Low plasma carnitine and deficiencies of CPT have been described. Patients lacking carnitine are weak, whereas those lacking the CPT enzymes are of fairly normal strength at rest but after fasting or exercise they may show evidence of muscle damage, with a raised creatine kinase and occasionally myoglobinuria. Carnitine is synthesized only in the liver and is taken up by muscle from plasma. Muscle biopsy may show large fat droplets in the muscle fibres. Treatment with oral carnitine has proved successful in children with a carnitine deficieny.

Patients with disorders of fatty acid metabolism often report episodes of muscle pain and myoglobinuria. While exercise, particularly in the fasting state, may sometimes be the precipitating factor, virus infections or alcohol are often implicated. It is unlikely that metabolic depletion is the cause of muscle damage in these patients and it is possible that high levels of free fatty acids may be the immediate cause. Exercise mobilizes free fatty acids from body fat depots and muscle triglycerides but, because fatty acid transport into the mitochondria is defective in these patients, the free fatty acids (mostly stearic and palmytic) accumulate within the muscle fibres and due to their detergent effect, damage the muscle fibre membranes.

Chapter References

Adiseshiah, M., Round, J.M., and Jones, D.A.. (1992). Reperfusion injury in skeletal muscle: a prospective study in patients with acute limb ischaemia and claudicants treated with revascularisation. *British Journal of Surgery*, **79**, 1026–9.

Askanas, V., Masako, B., Engel, W.K., Alvarez, R.B., Tome, F., and Leclerc, A. (1993). Prion protein is accumulated in inclusion body myositis. *Clinical Neuroscience and Neuropathology: Neuroreport*, **5**, 25–8.

Bartram, C., Edwards, R.H.T., Clague, J., and Benyon, R.J. (1994). McArdle's disease, a rare frameshift mutation in exon 1 of the muscle glycogen phosphorylase gene. *Biochimica et Biophysica Acta*, **1226**, 341–3.

Brooke, M.H., Carroll, J.E., Davis, J.E., and Hagberg, J.M. (1979). The prolonged exercise test. *Neurology*, **29**, 636–43.

Buckingham, M. (1992). Making muscle in mammals. *Trends in Genetics*, **8**, 144–9.

Buckingham, M. (1994). Which myogenic factors make muscle? *Current Biology*, **4**, 61–3.

Burke, R.E., Levine, D.M., Zajac, F.E., Tsairis, P., and Engel, W.K. (1971). Mammalian motor units: physiological–histochemical correlation in three types in cat gastrocnemius. *Science*, **174**, 709–12.

Curtin, N.A. and Davies, R.E. (1973). Chemical and mechanical changes during stretching of activated frog skeletal muscle. *Symposia on Quantitative Biology* XXVII, 619–26.

Drager, A., Weeds, A.G., and Fitzsimmons, R.B. (1987). Primary, secondary and tertiary myotubes in developing muscle: a new approach to the analysis of human myogenesis. *Journal of the Neurological Sciences*, **81**, 19–43.

Edwards, R.H.T, Young, A., Hosking, G.P., and Jones, D.A. (1977). Human skeletal muscle function: description of tests and normal values. *Clinical Science and Molecular Medicine*, **52**, 283–90.

Edwards, R.H.T, Wiles, C.M., Round, J.M., Jackson, M.J., and Young, A. (1979). Muscle breakdown and repair in polymyositis: a case study. *Muscle and Nerve*, **2**, 223–8.

Fenn, W.O. (1923). A quantitative comparison between the energy liberated and the work performed by the isolated sartorius muscle of the frog. *Journal of Physiology*, **58**, 175–203.

Friden, J., Sjostrom, M., and Ekblom, B. (1983). Myofibrillar damage following intense eccentric exercise in man. *International Journal of Sports Medicine*, **4**, 170–6.

Garlepp, M.J. and Mastaglia, F.L. (1996). Inclusion body myositis. *Journal of Neurology, Neurosurgery and Psychiatry*, **60**, 251–5.

Grimby, G. and Saltin, B. (1983). The ageing muscle: a mini review. *Clinical Physiology*, **3**, 209–18.

Jones, D.A., Jackson, M.J., McPhail, G., and Edwards, R.H.T. (1984). Experimental mouse muscle damage: the importance of external calcium. *Clinical Science*, **66**, 317–22.

Jones, D.A., Round, J.M., and Carli, F. (1991). Plasma creatine kinases in patients following routine surgery: a comparison with experimental muscle damage. *Journal of Physiology*, **438**, 173.

Kande, K. and Hashizume, K. (1989). Changes in properties of the medial gasrocnemius motor units in ageing rats. *Journal of Neurophysiology*, **4**, 737–46.

Lebo, R.V. *et al*. (1984). *Science*, **225**, 57–9.

Mauro, A. (1961). Satellite cells of skeletal muscle fibres. *Journal of Biophysics, Biochemistry and Cytology*, **9**, 493–8.

McArdle, A. and Jackson, M. (1994). Intrcellular mechanisms involved in damage to skeletal muscles. *Basic and Applied Myology*, **4**, 43–51.

McArdle, A., Edwards, R.H.T., and Jackson, M.J. (1995). How does dystrophin deficiency lead to muscle degeneration?—evidence from the MDX mouse. *Neuromuscular Disorders*, **5**, 445–56.

McCord, J.M. (1985). Oxygen derived free radicals in post ischaemic tissue injury. *New England Journal of Medicine*, **312**, 159–63.

Newham, D.J., Mills, K.R., Quigley, B.M., and Edwards, R.H.T. (1983 *a*). Pain and fatigue following concentric and eccentric muscle contractions. *Clinical Science*, **64**, 55–62.

Newham, D.J., McPhail, G., Mills, K.R., and Edwards, R.H.T. (1983 *b*). Ultrastructural changes after concentric and eccentric contractions *Journal of the Neurological Sciences*, **61**, 109–22.

Newham, D.J., Jones, D.A., Tolfree, S.E.J., and Edwards, R.H.T. (1986). Skeletal muscle damage: a study of isotope uptake, enzyme efflux and pain after stepping. *European Journal of Applied Physiology*, **55**, 106–12.

Newham, D.J., Jones, D.A., and Clarkson, P.M.N. (1987). Repeated high force eccentric exercise: effects on muscle pain and damage. *Journal of Applied Physiology*, **63**, 1381–6.

Peter, J.B., Barnard, V.R., Edgerton, V.R., Gillespie, C.A., and Stempel, K.E. (1972). Metabolic profiles of three fibre types of skeletal muscles in guinea pigs and rabbits. *Biochemistry*, **11**, 2627–33.

Ranvier, M.L. (1873). Proprietes et structures differentes des muscles rouges et des muscles blancs chez les lapins et chez les raies. *Comptes Rendus de l'Academie de Sciences*, **77**, 1030–4.

Round, J.M., Jones, D.A., and Cambridge, G. (1987). Cellular infiltrates in human skeletal muscle: exercise induced damage as a model for inflammatory muscle disease? *Journal of the Neurological Sciences*, **82**, 1–11.

Rowe, R.W.D. and Goldspink, G. (1969). Muscle fibre growth in five different muscles in both sexes of mice: 1. Normal mice. *Journal of Anatomy*, **104**, 519–30.

Schultz, E. (1976). Fine structure of satellite cells in growing skeletal muscle. *American Journal of Anatomy*, **147**, 49–70.

Schultz, E. (1989). Satellite cell behaviour during skeletal muscle growth and regeneration. *Medicine and Science in Sports and Exercise*, **21**, S181–S186.

Schultz, E. and McCormick, K.M. (1994). Skeletal muscle satellite cells. *Reviews of Physiology, Biochemistry and Pharmacology*, **123**, 213–57.

Webster, C. and Blau, H.M. (1990). Accelerated age-related decline in replicative life span of Duchenne muscular dystrophy myoblasts: implications for cell and gene therapy. *Somatic Cell Molecular Genetics*, **16**, 557–65.

Weisman, M.H. and Zvaifler, N.J. (1980). Vasculitis in connective tissue diseases. *Clinics in the Rheumatic Diseases*, **6**, 351–72.

2.7 Biomechanics of articulations and derangements in disease

Anthony Unsworth and Andrew A. Amis

Introduction

Synovial joints are nature's bearings which means that large loads can be transmitted through the joints while the frictional forces created by the accompanying motion are kept to a minimum. This is achieved by a combination of the properties of cartilage and synovial fluid which together maintain coefficients of friction which are less than 0.02. This, at first sight, is surprising since cartilage surfaces are rough compared with engineering bearings, the transmitted loads are high, and the large range of sliding speeds found in human joints means that a whole range of lubrication mechanisms need to be enlisted in order to maintain efficient articulation.

Typically, in the hip joint, loads of 3000 N are common, but an important feature of these loads is that they are 'dynamic' and hence only have a duration of the order of 0.1 s. However, a lower limb might undergo up to 200 million impacts of between 2000 and 5000 N in a lifetime and this sort of demand often leads to some degree of failure. Osteoarthritis is a common disease which bears many of the typical signs of mechanical failure of the material. The first sign of failure is the 'softening' of the cartilage which reflects the change in the way it responds to loads. Fibrillation, or a breakdown in collagen structures of the surface layer then leads to excessive wear of the cartilage which sometimes progresses to total exposure of the underlying bone.

It is clear that mechanical factors alone cannot fully explain the aetiology of osteoarthritis but in this chapter we will discuss the synovial joint as a bearing and describe the mechanical factors which may contribute to that failure. This includes the nature of the loading and how the cartilage responds by distributing it over the joint. The role of the menisci in the knee will be discussed and the consequences of damage to ligaments. The lubrication of joints and the mechanisms which have been identified as protecting joint surfaces from wear will lead on to a review of the breakdown of cartilage.

Mechanics of joints

The science of mechanics, whether applied to engineering artefacts or to human joints, is generally studied under one of three main categories: statics, kinematics, or dynamics, and these may be defined as:

1. statics is the study of forces without reference to motion;
2. kinematics is the study of motion without reference to forces;
3. dynamics is the study of forces and motion.

It is convenient to examine the biomechanics of articulations under these headings, each of which produce data relevant to different situations.

Static force analysis of joints

Statics is normally concerned with the study of the equilibrium of a body. For a human joint, the mechanical influences acting can be classified as either forces, which tend to cause a linear motion (or translation), or moments, which tend to cause rotational effects. These may be about an axis of flexion, for example. The general principle, arising directly from Newton's laws, is that for a static analysis, all forces and moments are assumed to be in equilibrium. If this were not the case, then the body being analysed would be subjected to a resultant load or moment which would cause it to accelerate. This would then mean that a dynamic analysis would be needed.

Static equilibrium uses the resolution of forces to account for linear translation effects, and the summation of the moments acting about a point in a body to balance rotations. Although the body is, of course, a three-dimensional structure, it is often acceptable to simplify a situation into two dimensions, in order to obtain an approximation for forces or moments acting on a joint. This applies to joints which have almost uniplanar motion, such as the interphalangeal joints of the fingers, where out of plane effects are small. In this case, the general situation where forces can act in three mutually perpendicular directions, and where moments can act to cause rotational effects about the three axes (i.e. involving six degrees of freedom), can be simplified to perhaps one equation that examines rotation about one axis, and then to two further components of force in order to calculate the joint load.

Before illustrating the equilibrium analysis of a simplified case, it should be noted that all such analyses involve several stages of work prior to the equilibrium analysis itself. This arises because it is not normally possible to perform invasive procedures, so assumptions must be made as to the state of load or position of internal structures.

The first step is to measure the external load acting on the body. This is sometimes simple, such as when a weight is being supported in the hand. Often, however, the biomechanical researcher must use a purpose-made load transducer, that can produce (usually) an electrical output for analysis. In the case of an interphalangeal joint, this may well be a relatively simple device such as a pinch strength meter, which gives a force-related output alone. Here, the analyst must assume that the force displayed acts along the axis of the finger tip pads. It is more complex if walking is involved, since the load can pass through a large area of foot to floor contact. Also, the loading may not be perpendicular to the floor, since there could be shearing actions (i.e. a tendency towards skidding), such as when the heel strikes the floor. In order to analyse complex situations, purpose-made load transducers are used. In the case of foot to floor contact, this is known generically as a 'force plate', which can give simultaneous output of all six degrees of freedom, i.e. three forces and three moments. Further, by using an array of force sensors (e.g. at each corner of a rectangular plate), it is possible to show the position and orientation of the resultant force, and thus the position where it acts on the sole of the foot. A further refinement to study pressures between the body and objects on which it acts uses sheets of electroconductive rubber with electrodes painted on to each surface. The resistance of this device changes as the rubber is compressed.

Having found the external loads, the researcher must turn to the internal loads. The majority of the loads acting on human joints arise as a reaction to the tensions in the muscles crossing the joints, and not (even in the lower limb) as a result of the weight of the body or other external loads being supported. This is because of leverage effects. The tendons generally pass much closer to the axis of a joint than does the line of action of an external load. This mechanical disadvantage may be considerable. At the elbow, for example, the moment arm of the biceps tendon may be 35 mm, while that of a load in the hand may be 350 mm, in which case the tendon tension must be ten times the external load, if that were the only muscle acting. Given that it is the muscles which cause most of the load, it is obviously of great importance to identify which of them are active, and to obtain data on their physical locations, especially their moment arms about axes of rotation, and the directions in which they act.

Muscle activity is discerned by electromyography, normally using skin-mounted electrodes. Data on muscles lying deep to the surface can be obtained by means of needle electrodes, but this is an invasive procedure. For analysis of a static situation there is much information published on electromyography (e.g. Basmajian 1967). For analysis of a moving subject, when different muscles will be switching on and off (e.g. when rising from a chair), it is necessary to use electromyography equipment, since the patterns of activity vary between subjects. This applies particularly to antagonistic actions which stabilize joints, causing great increases in the joint force.

The positions of the lines of action of muscles and tendons have usually been found by dissecting cadavers. If a range of joint positions is to be analysed, wires can be embedded in the tendons and three-dimensional data found by biplanar radiography (An *et al*. 1979), or by direct digitization of the structures *in vitro* at a range of joint positions (Amis *et al*. 1979). If the joint is surrounded by the fibre 'bellies' of muscles, such as at the hip, it is necessary to estimate the line of action within the muscle cross-section. In this case, the tissue can be frozen, sliced and the centroid of the cross-section found by digitizing a photograph. As medical imaging has advanced, it has become possible to obtain cross-sectional data by means of MRI scans of living subjects with physiological muscle tone. As the scanners and their software improve, so it is becoming possible to obtain multiple slices at a range of postures as a joint is flexed during a scanning procedure.

It is unusual for a joint to be acted on by a single muscle, and so the analyst must use the electromyography data to prepare a list of those that are active and then decide how they share the effort. There has been a range of optimization criteria put forward for this. One such technique is minimizing the energy expended during a movement. These assumptions are necessary in order to reduce the number of unknown variables and produce a solvable set of equations. Some workers have assigned the share of forces according to the physiological cross-sectional areas of the muscles (which allows for muscle fibre pennation geometry), and this may be appropriate for strenuous isometric actions, when the muscles are coming to the limit of stress that they can create (Amis *et al*. 1980b).

With all the muscle, geometrical, and external force data collected, the analyst can produce a mathematical model of the joint, placing all the lines of action of the forces with the correct sharing of the muscle tensions, and then examine the equilibrium of the joint:

Consider the finger tip pinch shown in Fig. 1:

For moment equilibrium about the axis of flexion,

$$\mathbf{70\ N \times 15\ mm = \mathit{T} \times 5\ mm,\ so\ \mathit{T} = 210\ N.}$$

For force equilibrium along the distal phalanx,

$$\mathbf{70\cos 65° + \mathit{T}\cos 50° = \mathit{F}_{axial},\ so\ \mathit{F}_{axial} = 165\ N.}$$

For force equilibrium across the distal phalanx,

$$\mathbf{\mathit{F}_{shear} + 70\sin 65° = \mathit{T}\sin 50°,\ so\ \mathit{F}_{shear} = 98\ N.}$$

By the theorem of Pythagoras, the (resultant force)$^2 = F^2_{axia} + F^2_{sheal}$, so the resultant force = 192 N and the angle of this force relative to the axis of the phalanx = $\tan^{-1}(F_{shear}/F_{axia}) = 31°$.

As noted above, it is the tendon tension, and not the external force, that causes the majority of the force transmitted across the joint itself.

Sometimes the analysis predicts a resultant joint force which is directed outside of the area of contact between the bones. This represents a potential instability, and ligament tensions may be invoked to regain equilibrium. For example, an analysis of pinching action showed that the axial force along the proximal phalanx, which would stabilize the metacarpopharyngeal joint, was overcome by the shearing action of the flexor tendon tensions when the metacarpopharyngeal joint was flexed. This would cause a palmar subluxation. Stability thus entailed tension in the collateral ligaments. In the rheumatoid hand, when the ligaments are slack because of loss of the joint surfaces, the shearing action causes palmar subluxation (Weightman and Amis 1982). This situation is at an extreme in the knee, where the tibiofemoral joint depends on the cruciate ligaments to resist the anterior–posterior actions of the muscles when walking (Morrison 1968).

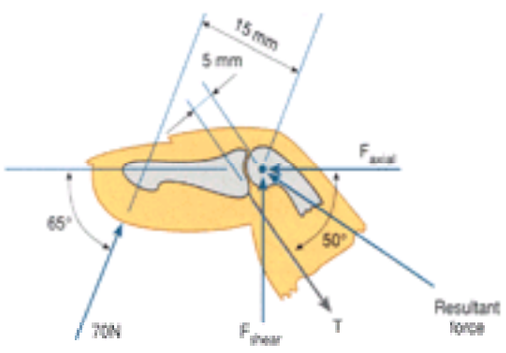


Fig. 1 Two-dimensional force analysis of finger tip pinch action. The flexor tendon tension *T* = 210 N, and the resultant joint force = 192 N at an angle of 31° to the axis of the distal phalanx.

Kinematics of joints

The most common aims of kinematic studies are to provide data on the patterns of movements of the body, or else to provide detailed information about the axes of rotation of individual joints, that can be used when making force analyses as described above.

The study of patterns of motion started approximately a century ago, when Braune and Fischer examined the limb movements when walking by means of stroboscopic pictures (Paul 1978). This is a method used more recently by Murray *et al*. (1964) where the flashes of light reflect from polished rods attached to the limbs. For three-dimensional analysis of movement, it is necessary to use multiple viewing directions. This produces large amounts of data (such as from high-speed cine-films) that must be synchronized, but allows details such as transverse rotations of the segments of the lower limb to be discerned (Levens *et al*. 1948) and a full three-dimensional force analysis to be carried out (Paul 1967: the hip; Morrison 1968: the knee). More recently, spatial data collection has used video cameras linked to computers, which allows reflective markers on joints to be tracked and their co-ordinates calculated automatically using the specialist software.

As well as photographic, non-contacting methods it is also possible to attach measuring instruments directly to the limb. Electrogoniometers can give outputs relating to rotations about more than one axis at a joint. For example, Kettlekamp *et al*. (1970) produced graphs of knee flexion–extension, abduction–adduction, and internal–external rotation during walking. Some studies of joint function have used such devices powered by batteries and combined with data loggers in order to obtain information on the numbers and magnitudes of joint movements in daily activities. The drawbacks of such devices are that they can inhibit subject actions and also move relative to the skeleton since they are usually held in place simply by tapes.

Detailed kinematic studies of individual joints often use multiple exposure radiographs or tracking of small lights or other markers. At any time, there will be an 'instant centre of rotation', about which all other points on a body are rotating. Since, by definition, these points are all on arcs of circles centred at the axis, it is a simple geometrical calculation to identify the axis from the paths of several points (Fig. 2). This methodolgy has shown that the complexity of the carpus, for example, can be reduced to a simple uniaxial rotation about the proximal pole of the capitate during wrist flexion–extension and abduction–adduction (Youm *et al*. 1978); a finding

which has been used to justify the design of simple spherical geometries for the articular surfaces of wrist prostheses.

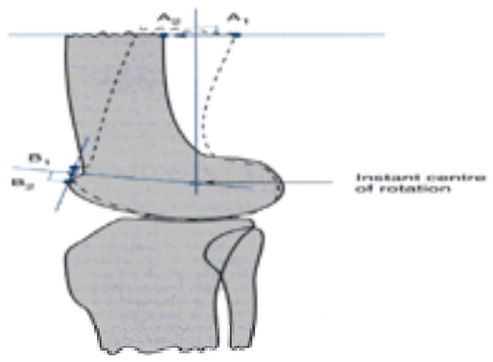


Fig. 2 Method to find 'instant centre of rotation' of a joint from double-exposure radiographs.

Dynamic analysis of joints

This approach is necessary when movements occur fast enough for inertial effects to cause significant errors if they are ignored. As the speed or direction of motion of limb segments changes, so their mass must be accelerated in order to cause the change in velocity, i.e. forces are required to overcome inertia. This applies to both linear translation and rotation. In the extreme, inertial effects limit the speed with which a joint can be moved, since active muscle tension capability diminishes as the speed of shortening increases (Elftman 1966), and large accelerations occur as the limb segment reaches the limits of motion. Amis *et al*. (1980*a*) showed that high-speed elbow flexion movements could cause accelerations of 20 *g*, while Paul (1974) reported accelerations of 4 *g* in the lower limb during walking.

In order to allow for the inertial effects of accelerations, data must be available on the masses, locations of the centre of mass and inertial properties of the limb segments. Fortunately, much of this is available in the literature (e.g. Contini and Drillis 1966). However, such data are based on a small number of cadavers, which may not have the same mass distribution as a particular experimental subject. It would be a major undertaking to attempt to gain such information from live subjects, and thus it is not normally attempted.

Joint forces during locomotion

The approaches to joint analysis described in the three sections above must all come together if a complex activity such as human locomotion is to be analysed, with simultaneous collection of spatial, electromyographic and force plate data. When the foot to floor forces are known, it is possible to work back into the body to estimate the joint forces. The external force action causes moments about the joints, which can have their equilibrium examined by means of 'static' methods that add in the inertial effects, which must themselves be calculated after analysing the accelerations of the limb segments. This is obviously a large undertaking, and while the data manipulation has been reduced by powerful computers, the data collection still requires a well set-up and calibrated gait analysis facility.

Typical force predictions for walking have been approximately three, four, and five times body weight for the ankle, knee and hip, respectively, with much higher loads anticipated in sporting activities.

Although it has been shown that such a gait analysis is the only way to find out the force distribution across a prosthetic knee when in use postoperatively, for example, (Johnson *et al*. 1980) the amount of work required has meant that such analyses have usually been done only as part of a specific research project, and not as part of the general aim to monitor progressive changes in motion or force patterns in patients.

Direct measurement of joint forces

It is obvious that the many stages and assumptions needed to perform a dynamic analysis leading to joint force predictions will lead to a large band of potential error in the results. It would, therefore, be very attractive to be able to measure the joint forces directly. Unfortunately, this is difficult to accomplish ethically, but there have been several reports of such measurements, all at the hip.

The method used has been to implant a hip prosthesis in which the neck of the femoral component has been instrumented by strain gauges that respond to the deflections caused by loads applied to the joint. The early work by Rydell (1966) was hampered by the need to use transcutaneous wires, but Bergmann *et al*. (1993) transmitted a signal containing all components of load on the joint from within a sealed cavity in the prosthesis. This has allowed the activities of patients to be followed for up to 2 years postoperatively, and thus to monitor a return to relatively normal gait. It is reassuring that this work has largely confirmed the earlier force plate predictions.

Lubrication of human joints

Having examined the loads passing through joints and seen how to measure these in clinical situations, we need to determine how the joints can respond to these with the minimum of damage. This leads us to look at the lubrication mechanisms in joints. Lubrication of human joints is important for two main reasons. A well lubricated joint has low friction and hence low energy loss, and if the lubrication mechanism actually separates the cartilage surfaces with a film of synovial fluid, then this will minimize wear and may help slow the process of osteoarthritic change (Unsworth 1984). There are a number of lubrication mechanisms which have been identified (Dowson 1967), some of which are better than others. Best of all is full-fluid-film lubrication. Here, the two surfaces which are loaded together, are completely separated by a film of lubricant, albeit a very thin film (about 0.5 m). This reduces both friction and wear. For the surfaces to be held apart by a film of fluid, the fluid must be pressurized in some way. One of the earliest mechanisms to be identified was 'hydrodynamic' lubrication (Reynolds 1886). Here the rolling and sliding motion between the surfaces generates pressure within the fluid. Thus motion and fluid viscosity are important. Another way is to pump the fluid between the surfaces using an external pump. This is called 'externally pressurized' or 'hydrostatic' lubrication.

If the load placed on the joint is not steady but varies, such as when we jump in the air and land on the ground, pressure can be generated by 'squeeze film' mechanisms. This can be explained by imagining two surfaces separated by a film of fluid suddenly being squeezed together. This puts a high pressure on the fluid which therefore resists the applied force. Obviously with time this fluid film disappears unless it is recharged by more fluid, but for dynamic loads as found in human joints, it is a most important mechanism (Unsworth *et al*. 1975).

If these pressures are very high compared with the elastic properties of the solid boundaries (about 2 MPa in the case of cartilage), then significant elastic deformation or 'squashing' of the surface will occur and this changes the geometry of the lubricating film. This in turn changes the pressures generated. When this happens, the mechanism is called 'elastohydrodynamic lubrication' (EHL or EHD lubrication). Under fluid-film conditions, and all the above mechanisms are classified as fluid-film lubrication, the coefficient of friction is of the order of 0.01 or lower. However, under some conditions a fluid film cannot be generated to separate the two surfaces. Here, intimate contact between the surfaces can be prevented by a boundary lubricant. A good boundary lubricant attaches itself to the solid surfaces by molecular forces and modifies the surfaces. This consequently reduces friction but not as much as fluid-film lubrication does (e.g. coefficient of friction in the range 0.1 to 0.5). The coefficient of friction is also independent of speed of sliding or load in boundary lubrication, whereas it is strongly dependent on both in fluid-film lubrication.

In reality many practical devices spend much of their working life running under a combination of fluid film and boundary lubrication known as 'mixed lubrication'. Here part of the load is carried by the fluid film and part by the solid to solid contact. Consequently coefficients of friction tend to be between 0.1 and 0.05.

Most surfaces, cartilage included, are rough on the microscopic scale (Jones and Walker 1968; Sayles *et al*. 1979) so separation of the surfaces only occurs when the fluid film is so thick that the roughnesses on the surfaces do not touch each other (Fig. 3). Thus for complete separation and hence fluid-film lubrication, the film

thickness must be greater than the combined roughnesses of the two surfaces.

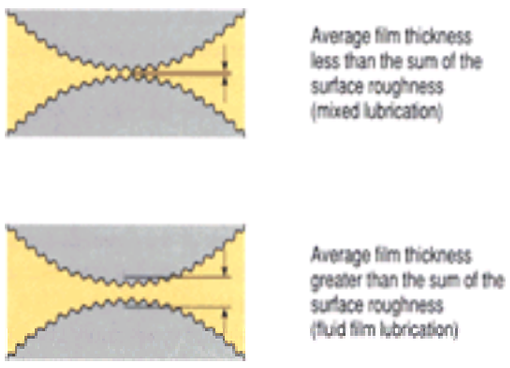


Fig. 3 Separation of surfaces.

A brief history of studies of human joint lubrication

The subject of human joint lubrication has only recently become clear in scientific terms, and therefore we will review briefly the history of the studies of human joint lubrication, followed by a summary of the current thoughts on the subject.

The early texts suggested that human joints were fluid-film lubricated (Reynolds 1886; MacConaill 1932). Comparisons of anatomical features of the hip and knee with features of engineering bearings led to the belief that hydrodynamic lubrication was the principal mechanism. In order to have hydrodynamic lubrication, an entraining velocity and a physical wedge of fluid between the cartilage surfaces are necessary. This, according to MacConaill, came from the shape of the articulating surfaces. He even claimed that the menisci of the knee act as 'tilting pads' to help optimize the load capacity of such joints. This was a reference to the work of Michell (1905). MacConaill also said that if these menisci were damaged, the resulting increase in friction was the cause of the higher incidence of osteoarthritis found to accompany this damage. This proved not to be the case, as Seedhom (1976) showed considerably increased contact stresses in knee articular cartilage following damage to menisci and this was more likely to be the cause of early osteoarthritic changes.

Jones (1934) was the first to measure the coefficient of friction in a joint. A horse's stifle joint produced a coefficient of friction of 0.27 when dry and 0.02 when lubricated with synovial fluid. These experiments were carried out under unphysiological conditions of high constant load and low sliding speed, therefore 2 years later, he used a proximal interphalangeal joint of the finger as the fulcrum of a pendulum (Jones 1936). His results were consistent with fluid-film lubrication.

Charnley (1959; Charnley 1960*a*; Charnley 1960*b*) repeated Jones's experiments using a similar apparatus but found that the friction was the same whether synovial fluid was present or not, and using a pendulum for human ankles found that the friction was independent of velocity of sliding and therefore boundary lubricated not fluid film. Charnley's measured values of friction were, however, very low for boundary lubrication (0.005 to 0.024).

Such low values of coefficient of friction at very low speeds of movement led McCutchen (1959) and Lewis and McCutchen (1959) to postulate a self-pressurizing hydrostatic lubrication mechanism which relied on pockets of liquid in the soft cartilage surface. Also, because cartilage was porous, they believed that as the surface was loaded, the cartilage matrix compressed, pushing pressurized synovial fluid from its pores into the interface of the joint. This they called 'weeping lubrication'. Evidence of the pockets of liquid was supported by Unsworth *et al*. (1975) but the 'weeping' mechanism was shown by Maroudas (1967) to be controversial. She showed that the permeability of cartilage was so low as to be unrealistic as a method of replenishing fluid films between cartilage surfaces within the time scale of normal gait. However, this permeability was important for nutrition (Maroudas 1968).

Dintenfass (1963) in his review of human joint lubrication favoured elastohydrodynamic lubrication as a mechanism. This was supported by Tanner (1966) and much later by Higginson and Unsworth (1981). However, squeeze film lubrication can act together with elastohydrodynamic lubrication, and Fein (1966–67) showed that in human joint applications this combination can be an important feature of effective lubrication.

The importance of the type of loading in joints can be seen from a series of experiments carried out by Linn (1967; Linn 1968) and Linn and Radin (1968). They used an 'arthrotripsometer' to measure the coefficient of friction in dog's ankle joints under constant load, which is unphysiological and predisposes against fluid films, and showed that mixed lubrication was present in the joints. In other words both hydrodynamic and boundary lubrication were acting together. The effects of boundary lubrication were examined by digesting synovial fluid with hyaluronidase and trypsin in turn. Trypsin did not alter the synovial fluid's viscosity but doubled friction, while hyaluronidase reduced viscosity but did not affect friction. Clearly this pointed to a protein component of synovial fluid acting as a boundary lubricant. However, it is important to emphasize that these tests were carried out with a constantly loaded ankle joint whereas in real life such joints are dynamically loaded (loads from almost zero to five times body weight every second). O'Kelly *et al*. (1978) and Roberts *et al*. (1982) repeated Linn and Radin's work but using the dynamic walking cycles of Paul (1967) and English and Kilvington (1979), respectively, and found fluid-film lubrication throughout. An obvious conclusion is that under physiological loads, (dynamic), fluid-film lubrication exists in human joints, while under static loads this reverts to 'mixed' lubrication, where boundary friction dominates. So what happens when we stand still under steady load for long periods? McCutchen (1959) showed that a type of self-acting hydrostatic lubrication (weeping lubrication) could force synovial fluid out of cartilage pores into the joint space to lubricate the joints by replenishing 'pools' of synovial fluid trapped in the surface roughnesses. The opposite of this was 'boosted lubrication' (Walker *et al*. 1968), a theory that fluid flowed from the trapped pools into the cartilage thereby concentrating the synovial fluid in the pools and increasing its viscosity to keep the surfaces apart. The same experimental evidence was often used by both sets of workers to justify the different conclusions. In the event Ling (1974) produced a theoretical model of elastic porous discs on impervious substrata and showed that weeping and boosted were not mutually exclusive.

The other important feature of human joints is their surface elasticity. Soft elastic layers are beneficial to lubrication and friction in pure-sliding (Bennett and Higginson 1970), especially in squeeze film situations (Gaman *et al*. 1974). A combination of elasticity and porosity was not significantly better than elasticity alone (Higginson and Norman 1974) since the permeability of cartilage was far too low to affect lubrication.

A combination of dynamic loading (physiological) and soft elastic surfaces had a very great effect (Unsworth *et al*. 1987). So much so that the friction in a Charnley hip prosthesis could be reduced by a factor of 10 by adding a soft elastic layer. Thus the important features of healthy behaviour of human joints are high viscosity synovial fluid (greater than 0.010 Pa s), dynamic loads, and a soft elastic surface (hardness 4 to 8 N/mm ^{2}).

Theoretical analyses

Using the experimental evidence already referred to, a number of analytical and numerical studies have been carried out to help understand the lubrication mechanisms. An elegant combined theoretical and numerical analysis of the movement of fluid in and out of cartilage concluded that 'The natural lubrication process is neither the weeping mechanism nor the boosted mechanism' (Mansour and Mow 1977). In any analysis, simplifying assumptions are necessary and several workers have followed the advice of Higginson and Norman (1974) and neglected cartilage permeability. This is extremely convenient since elastohydrodynamic lubrication of soft impermeable layers is quite well documented in rolling/sliding situations for cylindrical contacts (Hooke and O'Donoghue 1972). This is probably sufficient for knee or ankle analysis, though not for the hip. They give an equation for minimum fluid-film thickness of hmin – $U^{0.6}\,W^{-0.2}\,R$, where U = /ER and W = w/ER. R is the effective radius, the viscosity of the lubricant, U is the rolling/sliding speed, E is the equivalent elastic modulus, and w is the load/unit length on a cylindrical contact. For a typical knee in normal walking this gives a film thickness of 1.4×10^{-7} m. Since cartilage is rough (10^{-6} m) this film thickness is unlikely to separate the cartilage surfaces unless the soft compliant surfaces are more tolerant of surface roughnesses than are metals. Higginson and Unsworth (1981) show this to be the case and when account of this is taken, the film thickness is increased to about 6×10^{-7} m. This is still insufficient to separate the surfaces of cartilage. However, Dowson and Jin (1986) showed that using the ankle joint as a model the effect of high pressures generated on the roughnesses of the surfaces was to squash the high spots and produce a smoother surface. Adding squeeze film to the analysis increased film thicknesses to between 0.6 and 1.2 m, while the original surfaces had been smoothed to between 0.05 and 0.9 m depending on the wavelength of the surface roughness. Thus for the first time theory had supported experiments on human joints.

This theory relates only to line contacts and so hips are not truly represented. However, Dowson and Yao (1990) have analysed elliptical contacts and although the film thicknesses are smaller than for knees and ankles, Yao (1990) suggests that when combined with the microelastohydrodynamic lubrication of Dowson and Jin (1986) then fluid-film lubrication will be predicted theoretically.

The work of Dowson and Jin (1986) was taken further by Yao and Unsworth (1993) and while the basic mechanism of microelastohydrodynamic lubrication was still seen to be active, the predicted film thicknesses for realistic cartilage surfaces were seen to be very much smaller than the earlier predictions.

Summary of human joint lubrication

The important aspects of human joint lubrication are given in Fig. 4, which shows a hip joint at various stages in a walking cycle. During the lightly loaded swing phase, a relatively thick film of fluid can be entrained between the cartilage surfaces (Higginson and Unsworth 1981). When the heel strikes the ground, the load increases rapidly as the entraining velocity approaches zero. Here the thick film of fluid, generated during the swing phase, begins to squeeze out and the fluid-film thickness reduces. However, since the load is only applied for a short time (about 0.1 s, Fig. 4), the film does not reduce by too much at this stage. During the next phase of the walking cycle, the 'stance phase', the load reduces and the entraining velocity increases (both of which help to generate elastohydrodynamic lubrication), and so the fluid film separating the joint surfaces is maintained. Finally at 'toe off', the load is near to maximum and the entrainment velocity is low, so 'squeeze film' lubrication is called upon to maintain the fluid film and prevent surface to surface contact. With normal healthy tissue under physiological loads and motion, both theory and experiments show this to be true. During the time that the cartilage surfaces are completely separated by synovial fluid, little or no wear can take place. However, after resting, or standing motionless for long periods, we might expect our joints to squeeze all the fluid out from between the surfaces and therefore to have cartilage rubbing on cartilage (McCutchen 1959; Walker *et al*. 1968; Unsworth *et al*. 1975). When we start to move the joint again the potential for wear is high, except that a 'boundary lubricant' in the form of a protein (Swann *et al*. 1974) is present in the synovial fluid to keep friction low and help reduce wear. Once the motion has begun again, fluid-film lubrication takes over.

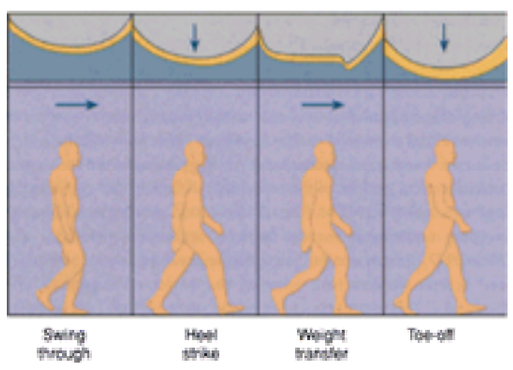


Fig. 4 A hip joint at various stages in a walking cycle.

Mechanical factors in the pathogensis of osteoarthrosis

Although the initiation and progression of articular cartilage breakdown is not well understood, the process seems likely to include both 'mechanical' and 'biochemical' factors, which may interact. From the viewpoint of the bioengineer, cartilage is seen as a bearing material which is subjected to repetitive impulsive loads and movements. For a normal engineering structure, such a situation would mean that it might be liable to accumulate microdamage that might eventually lead to fatigue failure. Such a mechanism is also attractive for explaining the breakdown of articular cartilage, since it is a tissue which has limited regenerative capability. With people typically taking 2 to 3 million steps per annum, this means that cartilage must probably resist in excess of a 100 million load cycles during a lifetime, which is effectively an infinite life in engineering design terms. In such a situation, a very small change in conditions can have a very large effect on the life of the structure. At 10^8 load cycles, a fatigue graph of cycles to failure versus stress is virtually flat even on a logarithmic scale, so a 5 per cent increase in stress can reduce the life to failure by a factor of 100.

Musculoskeletal tissues generally obey Wolff's law in that their form and structure are adapted to withstand the mechanical demands placed on them. Although this law was originally aimed at osseous tissue, it is accepted that articular cartilage also obeys it, and measurements have shown that the strength and stiffness of the articular surfaces increase as the stresses imposed locally increase. For example, Kempson *et al*. (1971) showed that the cartilage of the load-bearing area on the superior part of the femoral head had higher indentation stiffness than that from the inferior aspect, which was rarely loaded. This was later expanded to other joints of the lower limb by Swann and Seedhom (1985), who found a correlation between compressive modulus and predicted compressive stress in use. Kempson (1979) also found that there was a highly significant correlation with glycosaminoglycan content, and hypothesized that the loss of stiffness which follows early cartilage damage can be linked to loss of integrity of the collagenous framework that contains the glycosaminoglycans.

In an experiment on the cartilage of the femoral condyles, Kempson (1979) showed a significant loss of cartilage tensile strength with advancing age, and Weightman (1976*a*; Weightman 1976*b*) found a highly significant loss of cartilage fatigue strength with age. Thus, it seems that articular cartilage is more easily damaged with advancing age, and the exact mechanism for this is not known. However, since body weight does not fall greatly in most people with increasing age, the cartilage stresses will remain constant in normal activities, so one can foresee a convergence with the diminishing cartilage tensile and fatigue strengths which would lead to failure as a result of repetitive load cycles that eventually exceed a damage threshold. One hypothesis related to this is that an increase in stress can result from changes in the underlying bone, either a localized irregularity in the contour of the subchondral plate or a more general stiffening of the bone (Radin and Paul 1970). Repetitive loading experiments on rabbit joints (Radin *et al*. 1984) have shown that impacts caused trabecular microfractures, and that the healing response, associated with localized thickening of the damaged trabeculae, led to increased stiffness. These changes preceded cartilage damage. Trabecular hypertrophy is seen frequently as cartilage is lost from an osteoarthritic joint, and this is also part of a functional adaptation to the loss of the ability of the softened cartilage to support loads, which leads to localized areas of higher pressure. The growth of peripheral osteophytes is a further part of this adaptive process.

The studies above were careful to avoid testing damaged areas. With time, the early degenerative changes of softening and fibrillation (vertical fissures in the cartilage) seem to be found almost universally, but this is not the same as saying that fibrillation inevitably progresses to symptomatic osteoarthritis. In younger joints, fibrillation is found primarily in the non-loaded areas. This has been reported for the hip (Byers *et al*. 1970), the knee (Waugh *et al*. 1980) and the elbow (Goodfellow and Bullough 1967). These areas were found to be quite different from those which had 'progressive' lesions that led to full-thickness cartilage loss and which were associated with high stresses in use, such as the superior pole of the femoral head or the centre of the end-face of the radial head. It is interesting that there are differing frequencies with which the joints are affected by clinical degenerative changes. Beyond 70 years of age, most people have patellofemoral osteoarthritis, yet degeneration is rarely seen in the ankle. Part of the explanation comes from the pattern of movements, with 'linear' motion, such as in the ankle, appearing to be less damaging than a mixture of sliding and rotation, such as at the hip. This difference was emonstrated clearly by Goodfellow and Bullough (1967) between the humeroulnar and humeroradial articulations of the elbow. From a mechanical viewpoint, this appears to be related to the ability of the cartilage structure to adapt to resist the dominant loads. This can be visualized clearly through the demonstration of surface 'split lines', which display the orientation of the collagen fibres in the articular surfaces. They are able to align themselves with the sliding of a hinge joint, but cannot do so for a sliding and rotating joint, when the shear forces would be able to act across their orientation and hence to 'separate' the fibre bundles.

It is usually accepted that the areas of fibrillation, noted above, which are found in the majority of joints, do not lead directly to osteoarthritis. However, since these are areas of softening, it must be a possibility that cartilage damage could spread from them into the areas that usually carry significant forces. If there is a local area of cartilage softening, it will not carry its normal share of the joint force when the contact area covers it, thus, it is logical to expect elevated contact pressures in the adjacent cartilage. This phenomenon has been reported for chondromalacic lesions and for cartilage repairs that did not carry normal pressures when loaded (Shahgaldi *et al*. 1991). It is possible that large movements of joints that bring the loaded contact area over the softened zones at the normally unloaded aspect, can lead to expansion of the softened area until it is acted upon regularly, leading to breakdown in the heavily-loaded area. This is similar to the hypothesis of Swann and Seedhom (1985) that damage relates to occasional overload of cartilage that is not adapted to high stresses, and may be the cause of progression of patellar changes from localized degeneration of the extreme 'odd' medial articular facet, which only makes contact with the femur in full flexion (135°), across the full width of

the patella.

The discussion above has concentrated on 'primary' osteoarthritis, in which the cartilage changes have an insidious onset throughout a lifetime, and for which there is no obvious cause. There are, however, many cases where there is an obvious mechanical factor, often referred to as 'secondary' arthritis. The underlying principle is that there is a range of physiological stresses to which the articular surface is adapted, and that any sudden increase in loading, to a 'pathological' level, will initiate damage. This can arise from a wide range of orthopaedic problems, such as deformity of the joint surfaces due to malunion or malreduction of articular fractures, depression of the subchondral plate due to trabecular crushing, or to deformity of the limb due to angulation of an extra-articular fracture. There is also a spectrum of disturbances arising from abnormal growth, muscle actions, etc. A typical situation would be an inaccuracy of alignment following a femoral supracondylar fracture, which might leave a valgus angulation. In such a case, the line of action of the body weight would be shifted laterally at the knee, and this would cause elevated stresses on the lateral tibiofemoral articulation. This can then be the cause of a 'vicious circle', since lateral compartment cartilage loss will increase the valgus angulation on weight bearing, which will increase the stresses on the damaged articulation (that may now be devoid of cartilage). Total knee replacement can then act to restore the leg alignment, and balance the loads acting across the joint. Prior to the widespread acceptance of joint replacement, the procedure of choice was a realignment osteotomy, usually with a touch of overcorrection, that would allow the damaged area a respite, when reconstitution of a fibrocartilagenous surface and restoration of a healthy-looking joint space were sometimes observed (Maquet 1976). This approach is still used in young patients, for whom joint replacement is not advisable.

As well as bony damage, osteoarthritic changes can be initiated by soft tissue damage. At the knee, which has an inherently unstable articular geometry, the joint kinematics are controlled largely by the cruciate ligaments, which stabilize the tibiofemoral articulation against the anterior–posterior actions of the muscles. Rupture of the anterior cruciate ligament is common in sport, and this allows abnormal excursions of the femur across the tibial plateau. The resulting contact forces cause rapid articular damage in animals (Fig. 5), and this has been a widespread animal model for osteoarthritis (Muir and Carney 1987). The changes in humans are less drastic, and are often caused secondarily to meniscal damage. Loss of stability allows abnormal bone excursions, which then causes the menisci to be overridden by the femoral condyles, leading to damage. It is well known that the menisci transmit much of the knee joint force (Seedhom and Hargreaves 1979), so meniscal damage in a ligament-deficient knee combines abnormal motion with abnormal stresses, and can lead to degeneration. Meniscal damage alone has been suggested as a means of inducing arthritis experimentally (Shapiro and Glimcher 1980). With a slower rate of onset of cartilage damage, this was suggested as being more realistic than cruciate sectioning when inducing osteoarthritic changes experimentally.

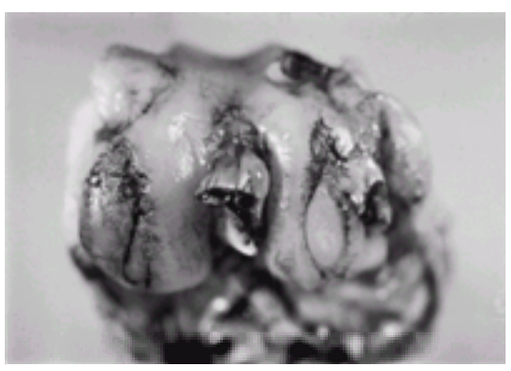


Fig. 5 Distal surfaces of rabbit femur, 6 months after division of the cranial cruciate ligament, showing areas of full-thickness cartilage erosion on the condyles. Cartilage damage has been highlighed by Indian ink staining (Meachim 1972).

Failure of lubrication mechanism

Another contributory factor to the development of osteoarthritis is a failure of the lubrication mechanism. This most likely acts in conjunction with one of the other failure mechanisms already discussed.

Under normal conditions, human joints function by having the two articular cartilage surfaces separated by a full film of synovial fluid. This does two things to help limit the damage to articular cartilage. First, by covering the rough surfaces of cartilage with a film of viscous fluid under pressure, the compressive stress transmitted to the cartilage is lower overall than if the two rough surfaces were in direct contact. We have seen that rough surfaces only contact at high spots, giving rise to very high stresses. The second method is by reducing the shear stress due to friction acting on the surfaces of the articular cartilage. Unsworth (1984) showed that these two factors could make all the difference to the failure of articular cartilage from fatigue stresses. In other words, if a full fluid-film mechanism exists between the cartilage surfaces, then they are likely to withstand a lifetime of use, whereas if the two surfaces rub directly on each other, premature failure is to be expected. The effects of this can be dramatic, reducing the life of 50-year-old cartilage to 4 years if lubrication fails (Unsworth 1984). This then raises another important feature of the way that the properties of synovial fluid and cartilage are affected by age, trauma, and disease. Anything that affects the joint's ability to produce a film of fluid can potentially cause damage to the cartilage surfaces and hence increase the effects of osteoarthritis.

Summary

It has been argued that cartilage is subjected to dynamic loading which can lead to a fatigue process causing damage to the surface layers of the joint. In addition, changes in the subchondral bone properties or in the stiffness of cartilage due to mechanical damage can also lead to accelerated mechanical failure of the articular surfaces. Changed kinematics in the joint, which for example might be due to the rupture of the anterior cruciate ligament of the knee, can also increase the resultant contact force and cause premature failure; as can damage to menisci or a failure of the lubrication mechanisms which normally promote fluid-film lubrication. All of these mechanisms effectively increase the resultant stress acting on the cartilage and this in turn will accelerate failure due to fatigue.

Chapter References

Amis, A.A., Dowson, D., and Wright, V.(1979). Muscle strengths and musculoskeletal geometry of the upper limb. *Engineering in Medicine*, **8**, 41–8.

Amis, A.A., Dowson, D., and Wright, V.(1980*a*). Analysis of elbow joint forces due to high speed forearm movements. *Journal of Biomechanics*, **13**, 825–31.

Amis, A.A., Dowson, D., and Wright, V.(1980*b*). Elbow joint force predictions for some strenuous isometric actions. *Journal of Biomechanics*, **13**, 765–75.

An, K.-N., Chao, E.Y., Cooney, W.P., and Linscheid, R.L. (1979). Normative model of human hand for biomechanical analysis. *Journal of Biomechanics*, **12**, 775–88.

Basmajian, J.V. (1967). *Muscles alive*, (2nd edn). Williams and Wilkins, Baltimore.

Bennett, A. and Higginson, G.R. (1970). Hydrodynamic lubrication of soft solids. *Journal of Mechanical Engineering Science*, **12**, 218–22.

Bergman, G., Graichen, F., and Rohlmann, A. (1993). Hip joint measured in two patients. *Journal of Biomechanics*, **26**, 969–90.

Byers, P.D., Contempomi, C.A., and Farkas, T.A. (1970). A post mortem study of the hip joint. *Annals of the Rheumatic Diseases*, **29**, 15–31.

Charnley, J. (1959). The lubrication of animal joints. Institution of Mechanical Engineers Symposium on Biomechanics, pp. 12–22. London, Institute of Mechanical Engineers.

Charnley, J. (1960*a*). The lubrication of animal joints in relation to surgical reconstruction by arthroplasty. *Annals of the Rheumatic Diseases*, **19**, 10–19.

Charnley, J. (1960*b*). How our joints are lubricated. *Triangle*, **4**, 175–80.

Contini, R. and Drillis, R.J. (1966). *Body segment mass properties*. Technical report No. 1166.03, New York University.

Dintenfass, L. (1963). Lubrication in synovial joints: a theoretical analysis—a rheological approach to the problems of joint movements and joint lubrication. *Journal of Bone and Joint Surgery*, **45A**,

1241.

Dowson, D. (1967). Modes of lubrication in human joints. *Proceedings of the Institution of Mechanical Engineers*, **181**, 45–54.

Dowson, D. and Jin, Z.-M. (1986). Micro-elastohydrodynamic lubrication of synovial joints. *Engineering in Medicine*, **15**, 63–5.

Dowson, D. and Yao, J.Q. (1990). A full solution to the problem of film thickness prediction in natural synovial joints. In *Mechanics of coatings*, Proceedings of 16th Leeds/Lyon Symposium on Tribology, Lyon 1989, pp. 91–102. Elsevier, Amsterdam.

Elftman, H. (1966). Biomechanics of muscle. *Journal of Bone and Joint Surgery*, **48A**, 363–77.

English, T.A. and Kilvington, M. (1979). *In vivo* records of hip loads using a femoral implant with telemetric output (a preliminary report). *Biomedical Engineering*, **1**, 111–15.

Fein, R.S. (1966–67). Are synovial joints squeeze film lubricated? *Proceedings of the Institution of Mechanical Engineers*, **181**, 125–8.

Gaman, I.D.C., Higginson, G.R., and Norman, R. (1974). Fluid entrapment by a soft surface layer. *Wear*, **28**, 345–52.

Goodfellow, J.W. and Bullough, P.G. (1967). The pattern of ageing of the articular cartilage of the elbow joint. *Journal of Bone and Joint Surgery*, **49B**, 175–81.

Higginson, G.R. and Norman, R. (1974). The lubrication of porous elastic solids with reference to the functioning of human joints. *Journal of Mechanical Engineering Science*, **16**, 250–57.

Higginson, G.R. and Unsworth, A. (1981). The lubrication of natural joints. In *Tribology of natural and artificial joints* (ed. J.H. Dumbleton), pp. 47–72. Elsevier, Amsterdam.

Johnson, F., Leitl, S., and Waugh, W. (1980). The distribution of load across the knee: a comparison of static and dynamic measurements. *Journal of Bone and Joint Surgery*, **62B**, 346–9.

Jones, F.S. (1934). Joint lubrication. *Lancet*, **i**, 1426–7.

Jones, F.S. (1936). Joint lubrication. *Lancet*, **i**, 1043.

Jones, H.P. and Walker, P.S. (1968). Casting techniques applied to the study of human joints. *Journal of the Institute of Science and Technology*, **14**, 57–68.

Kempson, G.E. (1979). Mechanical properties of articular cartilage. In *Adult articular cartilage*, (2nd edn) (ed. M.A.R. Freeman), pp. 333–414. Pitman Medical, Tunbridge Wells.

Kempson, G.E., Freeman, M.A.R., and Swanson, S.A.V. (1971). The determination of a creep modulus for articular cartilage from indentation tests on the human femoral head. *Journal of Biomechanics*, **4**, 239–50.

Kettlekamp, D.B., Johnson, R.J., Smidt, G.L., Chao, E.Y.S., and Walker, M. (1970). An electrogoniometric study of knee motion in normal gait. *Journal of Bone and Joint Surgery*, **52A**, 775–90.

Levens, A.S., Inman, V.T., and Blosser, J.A. (1948). Transverse rotation of the segments of the lower extremity in locomotion. *Journal of Bone and Joint Surgery*, **30A**, 859–72.

Lewis, P.R. and McCutchen, C.W. (1959). Experimental evidence for weeping lubrication in animal joints. *Nature*, **184**, 1285.

Ling, F.F. (1974). A new model of articular cartilage in human joints. *Transactions of the American Society of Mechanical Engineers, Journal of Lubrication Technology*, **96**, 449–507.

Linn, F.C. (1967). Lubrication of animal joints: I—the arthrotripsometer. *Journal of Bone and Joint Surgery*, **49A**, 1079–97.

Linn, F.C. (1968). Lubrication of animal joints: II—the mechanism. *Journal of Biomechanics*, **1**, 193.

Linn, F.C. and Radin, E.L. (1968). Lubrication in animal joints III—the effect of certain chemical alteration of the cartilage and lubricant. *Arthritis and Rheumatism*, **11**, 674–82.

MacConaill, M.A. (1932). Function of intra-articular fibrocartilages with special reference to the knee and inferior-radio-ulnar joints. *Journal of Anatomy*, **66**, 210–27.

McCutchen, C.W. (1959). Sponge-hydrostatic and weeping bearings. *Nature*, **184**, 1284–5.

Mansour, J.M. and Mow, V.C. (1977). On the natural lubrication of synovial joints: normal and degenerative. *Transactions of the American Society of Mechanical Engineers, Journal of Lubrication Technology*, **F99**, 163–73.

Maquet, P. (1976). *Biomechanics of the knee*. Springer-Verlag, Berlin.

Maroudas, A. (1967). Hyaluronic acid films. *Proceedings of the Institution of Mechanical Engineers*, **181**, 122–4

Maroudas, A. (1968). Physiochemical properties of cartilage in the light of ion exchange theory. *Biophysics Journal*, **8**, 575–95.

Meachim, G. (1972). Light microscopy of Indian ink preparation of fibrillated cartilage. *Annals of the Rheumatic Diseases*, **31**, 457–64.

Michell, A.G.M. (1905). The lubrication of plane surface. *Zeitschrift, Mathematik und Physik*, **52**, 123–37.

Morrison, J.B. (1968). Bioengineering analysis of force actions transmitted by the knee joint. *Biomedical Engineering*, **3**, 164–70.

Muir, H. and Carney, S.L. (1987). Pathological and biochemical changes in cartilage and other tissues of the canine knee resulting from induced joint instability. In *Joint loading: biology and health of articular structures* (eds H.J. Helminen *et al*.), pp. 47–63. Wright, Bristol.

Murray, M.P., Drought, A.B., and Kory, R.C. (1964). Walking patterns of normal men. *Journal of Bone and Joint Surgery*, **46A**, 335–60.

O'Kelly, J., Unsworth, A., Dowson, D., Hall, D.A., and Wright, V. (1978). A study of the role of synovial fluid and its constituents in the friction and lubrication of human hip joints. *Engineering in Medicine*, **7**, 78–83.

Paul, J.P. (1967). Forces transmitted by joints in the human body. *Proceedings of the Institution of Mechanical Engineer, Part 3J*, **181**, 1–15.

Paul, J.P. (1974). Force actions transmitted in the knee of normal subjects and by prosthetic joint replacements. In *Knee joint replacement*, pp. 126–131. Institution of Mechanical Engineers Press, C271/74.

Paul, J.P. (1978). Kinetic measurement systems and biomechanical data. *Journal of Human Movement Studies*, **4**, 179–90.

Radin, E.L. and Paul, I.L. (1970). Does cartilage compliance reduce skeletal impact loads? *Arthritis and Rheumatism*, **13**, 139.

Radin, E.L., Martin, R.B., Burr, D.B., Caterson, B., Boyd, R.D., and Goodwin, C. (1984). Effects of mechanical loading on the tissues of the rabbit knee. *Journal of Orthopaedic Research*, **2**, 221–34.

Reynolds, O. (1886). On the theory of lubrication and its application to Mr Beauchamp Towers' experiments. *Philosophical Transactions of the Royal Society of London*, **177**, 157–235.

Roberts, B.J., Unsworth, A., and Mian, N. (1982). Modes of lubrication in human hip joints. *Annals of the Rheumatic Diseases*, **41**, 217–24.

Rydell, N.W. (1966). Forces acting on the femoral head prosthesis. *Acta Orthopaedica Scandinavica*, Suppl. No. 88.

Sayles, R.S., Thomas, T.R., Anderson, J., Haslock, I., and Unsworth, A. (1979). Measurement of the surface microgeometry of articular cartilage. *Journal of Biomechanics*, **12**, 257–67.

Seedhom, B.B. (1976). Load bearing function of the menisci. *Physiotherapy*, **62**, 223–6.

Seedhom, B.B. and Hargreaves, D.J. (1979). Transmission of load in the knee joint with special reference to the role of the menisci. *Engineering in Medicine*, **8**, 220–28.

Shahgaldi, B.F., Amis, A.A., Heatley, F.W., McDowell, J., and Bentley, G. (1991). Repair of cartilage lesions using biological implants: a comparative histological and biomechanical study in goats. *Journal of Bone and Joint Surgery*, **73B**, 57–64.

Shapiro, F. and Glimcher, M.J. (1980). Induction of osteoarthritis in the rabbit knee joint: histological changes following meniscectomy and meniscal lesions. *Clinical Orthopaedics*, **147**, 287–95.

Swann, A.C. and Seedhom, B.B. (1985). The surveying of the stiffness of articular cartilage across total joint surfaces with reference to early stages of osteoarthritis. *British Journal of Rheumatology*, **24**, 219.

Swann, D.A., Radin, E.I., Nazimac, N., Weisse, P.A., Curran, N., and Lawinnek, G. (1974). Role of hyaluronic acid in joint lubrication. *Annals of the Rheumatic Diseases*, **33**, 318.

Tanner, R.I. (1966). An alternative mechanism for the lubrication of synovial joints. *Physics in Medicine and Biology*, **11**, 119.

Unsworth, A. (1984). Some biomechanical factors in osteoarthrosis. *British Journal of Rheumatology*, **23**, 173–6.

Unsworth, A., Dowson, D., and Wright, V. (1975). The frictional behaviour of human synovial joints. *Transactions of the American Society of Mechanical Engineers, Journal of Lubrication Technology*, July, 369–76.

Unsworth, A., Pearcy, M.J., and White, E.F.T. (1987). Soft layer lubrication of artificial hip joints. In *Proceedings of the Institution of Mechanical Engineers International Conference Volume on Tribology*, pp. 715–24. MEP, Bury St Edmunds.

Walker, P.S., Dowson, D., Longfield, M.D., and Wright, V. (1968). Boosted lubrication in synovial joints by fluid entrapment and enrichment. *Annals of the Rheumatic Diseases*, **27**, 512–20.

Waugh, W., Newton, G., and Tew, M. (1980). Articular changes associated with a flexion deformity in rheumatoid and osteoarthritic knees. *Journal of Bone and Joint Surgery*, **62B**, 180–83.

Weightman, B.O. (1976 *a*). Variations in the tensile fatigue resistance of human articular cartilage. *Proceedings of the 22nd Annual Meeting of the Orthopaedic Research Society*, New Orleans.

Weightman, B.O. (1976 *b*). Tensile fatigue of human articular cartilage. *Journal of Biomechanics*, **9**, 193–200.

Weightman, B.O. and Amis, A.A. (1982). Finger joint force predictions related to the design of joint replacements. *Journal of Biomedical Engineering*, **4**, 197–205.

Yao, J.Q. (1990). A study of the deformation and lubrication of synovial joints. Ph.D thesis. Leeds University.

Yao, J.Q. and Unsworth, A. (1993). Asperity lubrication in human joints. *Proceedings of the Institution of Mechanical Engineers, Part H*, **207**, 245–54.

Youm, Y., McMurthy, R.Y., Flatt, A.E., and Gillespie, T.E. (1978). Kinematics of the wrist. *Journal of Bone and Joint Surgery*, **60A**, 423–31.

2.8 The neurophysiology of pain

Hans-Georg Schaible

General pain physiology (Table 1)

Table 1 Definitions (I)

Pain and nociception

Pain is the sensation which is evoked by potential or actual tissue-damaging stimuli (noxious stimuli) or by tissue injury. Pain has a sensory discriminative aspect consisting of the detection and identification of the noxious event. The affective (emotional) aspect of pain describes that a pain sensation is usually experienced as unpleasant. Furthermore noxious stimuli may elicit reactions of the motor system (e.g. withdrawal) and of the autonomic nervous system (e.g. changes in blood pressure, heart rate), and these reactions are defined as the motor aspect and the autonomic (vegetative) aspect of pain. Finally, pain sensations are compared and evaluated in regard to experiences of pain in the past and these mental processes form the cognitive aspect of pain. The occurrence of these components indicates that large parts of the central nervous system are involved in the responses to noxious stimuli (Perl 1984; Willis 1985; Besson and Chaouch 1987; Schmidt 1989; Willis and Coggeshall 1991).

The term nociception defines the neuronal events which occur in the peripheral and central nervous system when noxious stimuli are applied (Perl 1984; Willis 1985; Besson and Chaouch 1987; Schmidt 1989; Willis and Coggeshall 1991). Usually pain is a consequence of the activation of the nociceptive system but situations are known in which tissue damage and the resulting nociceptive processes do not evoke pain (e.g. during general anaesthesia) or in which pain occurs in the absence of noxious stimuli, i.e. without overt activation of the peripheral nociceptive system. It is of clinical importance that pain does not only depend on the processes of nociception but also on psychological and social factors.

The nociceptive system (Fig. 1)

A schematic illustration of the anatomy of the nociceptive system is shown in Fig. 1. The nociceptive system consists of neurones in the peripheral and central nervous system which are activated by the application of noxious stimuli to the innervated tissue. A proportion of neurones in the nociceptive system is specifically and exclusively activated by noxious stimuli (nociceptive specific neurones). In the central nervous system a large proportion of neurones shows convergence of non-nociceptive and nociceptive inputs, i.e. the neurones and pathways are not only involved in nociception but also in the processing of non-nociceptive sensory information (for references see Dubner and Bennett 1983; Perl 1984; Willis 1985; Besson and Chaouch 1987; Schmidt 1989; Willis and Coggeshall 1991 ; Mense 1993; Schaible and Grubb 1993; Wall and Melzack 1994).

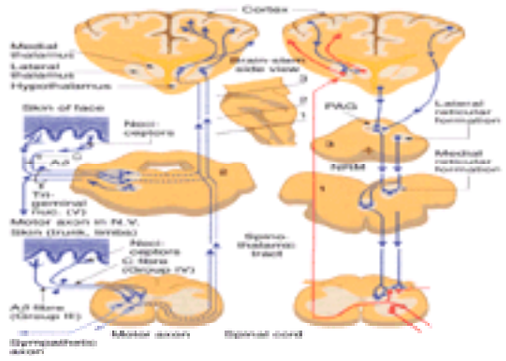


Fig. 1 The nociceptive system. The graph on the left shows a section through the spinal cord (bottom), of which the sensory neurones receive the afferent input from nociceptors supplying trunk and limbs. Sensory spinal cord neurones project to motor neurones and sympathetic neurones of the spinal cord (their axons leave the spinal cord in the ventral root) and to the brainstem and thalamus (top) via ascending tracts. The afferent input from the head is processed in the brainstem (above). Thalamic neurones project to the cerebral cortex. The graph on the right displays descending inhibitory systems which originate in the brainstem (two sections in the middle) and descend to the spinal cord. The brainstem is also influenced by pathways descending from the cortex. The central inset shows a side view of the brainstem with the three levels of the sections numbered 1–3. PAG, periaqueductal grey; NRM, nucleus raphe magnus. (Reproduced with permission from Schmidt 1989.)

Nociceptive afferent fibres of the peripheral nervous system (nociceptors) convey nociceptive sensory information from the innervated tissue to the spinal cord or to the brainstem. Functionally speaking, most of them are polymodal, i.e. they respond to mechanical, thermal, and chemical noxious stimuli. Nociceptive afferent units have unmyelinated axons with conduction velocities of less than 2.5 m/s (group IV or C fibres), a smaller proportion has thinly myelinated axons (group III or A d-fibres, conduction velocities 2.5–30 m/s). In contrast to non-nociceptive mechanoreceptors with thick myelinated axons (group II or A b-fibres, conduction velocities greater than 30 m/s), the sensory endings of nociceptive units are not encapsulated by or linked to specific corpuscular (receptive) structures but are free nerve endings. Numerous nociceptive afferent fibres not only have afferent sensory functions but also efferent functions by releasing peptides into the innervated tissue thereby exerting effects on the vasculature (Holzer 1988; see below). It should be noted, however, that not all slowly conducting afferent units with free nerve endings are nociceptors; some of them are involved in thermo- and mechanoreceptive functions.

The afferent sensory information is processed in the spinal cord (and in the brainstem for inputs from the head). Neurones encode the intensity and duration of noxious stimuli. Many spinal cord neurones show considerable convergence of inputs, i.e. they are activated from the skin and the deep tissue and/or viscera. Ascending axons convey sensory information to supraspinal targets such as the brainstem and the thalamus. The most important tracts are the spinoreticular and spinothalamic pathways, both of which ascend in the anterolateral quarter of the spinal cord. Interneuronal connections within the spinal cord constitute segmental pathways to the spinal cord output systems, the motor neurones in the ventral horn, and the sympathetic neurones in the lateral part of the grey matter. These pathways generate co-ordinated reflex responses which are, however, modified and controlled by descending systems from the supraspinal sites.

In the brainstem, several nuclei (e.g. the periaqueductal grey, PAG, the nucleus raphe magnus, NRM) are interconnected and form pathways which descend to the spinal cord. These pathways (mainly located in the dorsolateral funiculus) mediate a descending inhibition of neurones in the spinal cord, i.e. they partially suppress the processing of nociceptive information in spinal cord neurones (Fig. 1, right side). The brainstem nuclei are thought to be activated by nociceptive inputs via ascending tracts (thus forming a negative feed back system between spinal cord and brainstem) but they are also influenced from supraspinal sites (Gebhart 1986; Besson and Chaouch 1987; Willis 1988).

Activation of neurones in the thalamus and cortex is necessary to evoke the sensory, emotional, and cognitive components of the pain reaction (conscious sensation with unpleasant feeling, evaluation of the pain). Neurones in the ventrobasal complex of the lateral thalamus and in the cortical areas SI and SII are thought to be involved in the discriminative sensory aspects of pain since these neurones encode the parameters of a noxious stimulus such as intensity, localization, onset, and offset. Other neurones in the supraspinal structures such as neurones of the posterior thalamus and the intralaminar nuclei of the thalamus, and connections to the hypothalamus and the limbic system are thought to be involved in the generation of affective aspects of pain, in arousal reactions, and in pain memory (Perl 1984; Willis 1985 ; Guilbaud 1988; Kenshalo and Willis 1991).

The nervous system should not be understood just as a system of wires and switches which respond in a stereotyped manner to sensory stimuli. Considerable plasticity in the nociceptive system has been found in recent years (Dubner and Ruda 1992; Willis 1992; Coderre *et al*. 1993; McMahon *et al*. 1993; Schaible and Grubb 1993; Wall and Melzack 1994). Functional plasticity is an activity-dependent modification of neuronal processing and is observed during activation of the nervous system by long-lasting noxious stimuli such as inflammation. Structural plasticity describes anatomical changes in the nervous system such as sprouting of neurones, which is often seen after damage to the nervous system itself. Under both conditions a noxious stimulus will produce neuronal responses in the nervous system which are different to those seen in the nervous system under normal conditions. Changes in the nociceptive processing are thought to underly changes in the pain sensations such as allodynia and hyperalgesia (often present during inflammation, see below) or abnormal pain (often present after damage to the nervous system).

Neuronal basis of nociception in the joint

The anatomy of the innervation of joints

Articular nerves contain sensory afferent fibres and sympathetic efferent fibres. A proportion of about one-quarter of nerve fibres in the joint nerve consists of thick myelinated fibres (group II units) and thinly myelinated afferent fibres (group III units). Unmyelinated afferent fibres (group IV fibres) and the unmyelinated efferent units form the vast majority of all nerve fibres in the joint nerve. Non-nociceptive group II afferents (see below) are equipped with corpuscular endings of the Ruffini, Golgi, and Pacini type and these are located in the fibrous structures of the joint. By contrast afferent group III and IV fibres, many of which are nociceptive (see below), terminate as non-corpuscular or 'free nerve endings' in the fibrous capsule, the adipose tissue, the ligaments, the menisci, and the periosteum. The presence of non-corpuscular sensory nerve endings in the synovial layer (a major site of inflammatory foci in arthritic diseases) is still disputed since electron microscopy studies failed to identify nerve endings at these sites, whereas peptidergic structures in this layer have been assumed to represent afferent fibres (see Johansson *et al*. 1991; Schaible and Grubb 1993; Schmidt *et al*. 1994).

In their ultrastructure the non-corpuscular free nerve endings consist of a series of spindle-shaped thick segments connected by waist-like thin segments ('string-of-beads-appearance'). These beads and the end bulb contain mitochondria, glycogen particles, vesicles and 'bare' areas of axon sheaths which are not covered by Schwann cell processes. The beads are assumed to represent multiple receptive sites (Heppelmann *et al*. 1990). Proportions of the joint afferents (mainly unmyelinated fibres) are peptidergic, i.e. they express one or several neuropeptides such as substance P, neurokinin A, calcitonin gene-related peptide, and somatostatin, which are transported from the perikarya in the dorsal root ganglia to the peripheral endings in the tissue and also to the terminals in the spinal cord (Schaible and Grubb 1993; Schmidt *et al*. 1994). In the periphery, neuropeptides are thought to produce a neurogenic inflammation after their release upon stimulation of the afferent fibres since they lead to vasodilatation and plasma extravasation in the innervated tissue (Holzer 1988). Neuropeptides released in the terminals in the spinal cord modify the process of signal transmission (Duggan and Weihe 1991 ; Hökfelt *et al*. 1994; Urban 1994).

Nociception in the joint—peripheral mechanisms

Most joint afferents are mechanosensitive, i.e. they are activated by pressure applied to the joint and by movements of the joint. The receptive fields of the afferents are the areas that cause the fibre to fire on palpation. Figure 2 displays the receptive fields of two joint afferents in the capsule (the black spots). The receptive fields correspond to the location of the sensory endings of the afferent fibres in the tissue. According to their mechanosensitivity, afferent fibres are classified as non-nociceptive or nociceptive units. Non-nociceptive units are those which have sensory endings activated by innocuous mechanical stimuli, such as gentle pressure applied to the joint and movements in the working range of the joint. Most of them are group II and group III fibres and some are group IV units (see section I). Figure 2(a) shows a non-nociceptive group III unit which is activated by extension (ext.) and inward rotation (pronation, IR) in the working range of the joint (both are innocuous, normally not painful movements). Typically these units exhibit higher discharge rates during noxious movements such as noxious inward rotation (n.IR) but in the whole range of reactions these units do not clearly discriminate innocuous from noxious stimuli. The non-nociceptive units are considered to serve proprioceptive functions; their activation, however, does not necessarily cause conscious sensations which are related to the joint (the major sensation that is ascribed to the joint is pain).

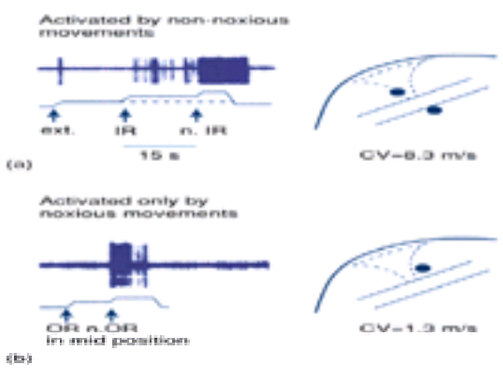


Fig. 2 (a) Response properties of a non-nociceptive unit with a low mechanical threshold in the innocuous range. The graph on the left shows the action potentials recorded from the axon during the application of the movements of the knee indicated by the traces underneath. Phasic response to extension of the knee, ext., and tonic response to inward rotation of the knee, IR, in extension (both movements in the working range of the knee joint); stronger response to noxious inward rotation of the knee, n.IR., in extension. The dots in the inset on the right show the location of two receptive fields of the unit in the capsule. (b) Nociceptive group IV unit with a high threshold. Same type of display as in (a). Strong response to noxious outward rotation, n.OR (and other noxious movements) but no response to outward rotation

in the working range, OR, in mid-position (and any other movement in the working range). The dot shows the location of the receptive field in the capsule. CV, conduction velocity of the axon.

By contrast, nociceptive units are those which have sensory endings that are only activated by noxious stimuli, such as strong pressure applied to the joint and movements exceeding the working range of the joint (thus these units are high-threshold units). Nociceptive afferents are group III and IV units. Figure 2(b) displays a nociceptive group IV unit which does not respond to innocuous movements such as outward rotation in the working range (OR) but responds to noxious outward rotation against the resistance of the tissue (n.OR). These units clearly signal noxious events. Most of them also respond to chemical stimulation, e.g. to algesic substances such as bradykinin. These afferents are polymodal nociceptors. In conscious humans, sensations of pain have been evoked when noxious mechanical, thermal, and chemical stimuli such as acids and hypertonic salt solutions were applied to the fibrous structures of the joint, the ligaments, and the fibrous capsule. By contrast, the application of mechanical and thermal stimuli to the cartilage did not evoke sensations and, interestingly, not much evidence was provided for sensations evoked by stimulation of the synovial layer (see Schaible and Grubb 1993). Muscle nerves contain nociceptive units which have response properties similar to those of joint afferents (Mense 1993).

A further group of afferents does not respond to innocuous or noxious mechanical stimuli; they are mechanoinsensitive afferents (Schaible and Schmidt 1988). These units do not have an apparent sensory function under normal conditions but they are important during inflammation of the joint (Schaible and Grubb 1993; see below).

Nociception in the joint—mechanisms in the central nervous system

In the spinal cord the nociceptive afferent input from joints is processed in neurones of the dorsal and ventral horn of the grey matter. With regard to the 'sensory channels' activated by joint input there are two subsets of neurones; one that has only afferent inputs from deep structures such as joint and muscle and another one which has additional cutaneous input. Thus the afferent input from the joint is further processed in spinal cord neurones which are not exclusively activated by joint afferents but also by afferents from other tissues. Since numerous afferent fibres project to one and the same spinal cord neurone (convergence) the receptive fields of the neurones are much larger than those of single afferent units. Figure 5(b) shows a spinal cord neurone which had a receptive field in the knee joint and the adjacent muscles.

Fig. 3 Overview of neuronal events in the course of an inflammation in the joint. The graph displays the spinal cord with the afferent input from the joint via sensory fibres and the motor and sympathetic output pathways. (Reproduced from Schaible and Grubb 1993, with permission.)

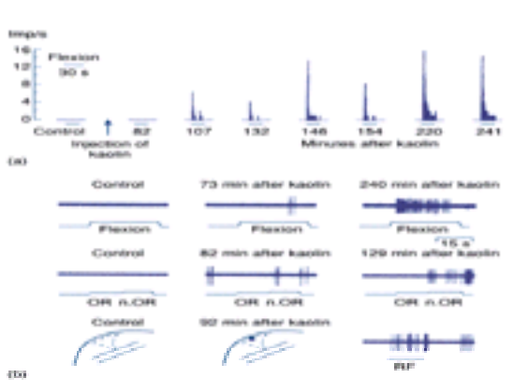


Fig. 4 (a). Sensitization of a nociceptive group III unit to mechanical stimulation of the joint during development of inflammation in the joint. The graph shows peristimulus time histograms indicating the number of action potentials elicted in this afferent fibre by flexion of the knee. No response to flexion of the normal joint. After the induction of inflammation by kaolin and carrageenan the previously high-threshold unit developed responses to flexion of the joint. (b). Sensitization of a mechanoinsensitive group IV unit during development of inflammation of the knee joint. No responses to mechanical stimulation before inflammation (control, traces on the left), i.e. no response to flexion, outward rotation, OR, noxious outward rotation, n.OR, and palpation of the joint (no receptive field identified). After induction of inflammation by kaolin and carrageenan there was induction of mechanosensitivity (traces in the middle and on the right). At this stage responses were elicted by movements and by palpation of the joint (with identification of the receptive field, dot). RF, pressure applied to the receptive field. (Modified from Schaible and Schmidt 1988.)

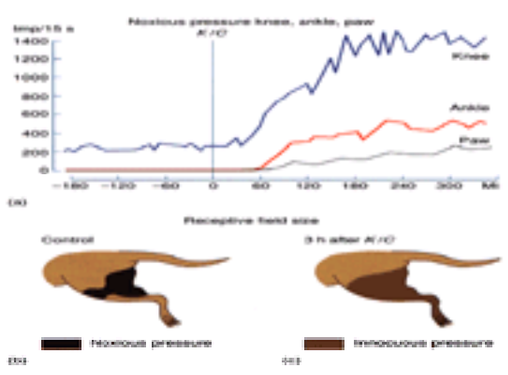


Fig. 5 Generation of hyperexcitability in a spinal cord neurone of the rat during development of inflammation in the knee joint. (a) Responses of the neurone (impulses/15 s) to noxious pressure applied to the knee, ankle, and paw before and after induction of inflammation in the knee joint (k/c). During development of inflammation in the knee the responses to noxious pressure applied to the knee were progressively enhanced and responses were also elicited by noxious pressure applied to the non-inflamed ankle and paw. (b) Receptive field of the neurone before inflammation. The neurone was only activated by noxious pressure applied to the knee and the adjacent muscles. (c) Expanded receptive field of the neurone 3 h after induction of inflammation. At this stage the neurone could be activated by gentle pressure applied to the whole leg. (Reproduced from Neugebauer *et al*. 1993 with permission.)

With regard to their activation threshold, the neurones with joint input have been classified as nociceptive specific and wide dynamic range. Nociceptive-specific

neurones with articular input respond only to noxious compression of the joint and other structures of the hind limb(s) and/or to noxious movements of the knee (movements against the resistance of the joint structures). Wide dynamic range neurones show substantial responses to innocuous pressure applied to the knee and other structures and to movements in the working range of the knee but show more pronounced responses to stimuli of noxious intensity. The convergence of nociceptive afferents to the spinal cord neurones provides some explanation for the diffuse and badly localized nature of joint pain.

The response characteristics of a spinal cord neurone do not only depend on its afferent inputs (non-nociceptive and/or nociceptive fibres). Numerous spinal cord neurones with joint input are tonically inhibited by descending inhibitory systems (see above). Indeed, the interruption of descending inhibitory influences leads to enhanced sensitivity and enlargements of receptive fields of many spinal cord neurones (Gebhart 1986; Willis 1988; Schaible and Grubb 1993). It is possible that intrinsic properties of the neurone (e.g. biophysical parameters) also contribute to its response behaviour.

A proportion of the nociceptive-specific and wide dynamic range neurones project to supraspinal sites (see Fig. 1). They activate supraspinal pathways and in awake humans and animals they evoke pain sensations (Willis 1985; Willis and Coggeshall 1991). Most nociceptive-specific and wide dynamic range neurones with joint input belong to intraspinal pathways mediating and controlling motor events. Reflexes evoked by joint afferents are important in eliciting protective muscle movements, i.e. movements counteracting hyperflexion, hyperextension, and hyperrotation. These noxious movements have marked effects on a-motor neurones which activate skeletal muscle fibres and initiate movements. Under normal (innocuous) conditions joint afferents only exert weak effects on the a-motor neurones. By contrast, the application of innocuous stimuli to the joint leads to considerable reflex discharges in g-motor neurones (g-motor neurones adjust the length of muscle spindles thereby optimizing the gain of the reflexes which control the length of the muscle, for details see Schmidt 1989). Through this pathway joint afferents may participate in the regulation of joint stiffness and joint stability during normal movement (Johansson *et al*. 1991).

The neurobiological response of the nociceptive system to joint inflammation

Studies using models of experimental inflammation (injection of crystals and/or carrageenan for induction of acute inflammation and injection of Freund's complete adjuvant (FCA) for induction of chronic inflammation) have provided evidence for complex changes in the peripheral and central nervous system during inflammatory diseases, indicating that the nervous system undergoes functional plasticity during inflammation (Fig. 3). Thus in the course of inflammatory lesions in the joint, the nociceptive system is not only activated but also modified to some extent. The changes observed are thought to be responsible for the allodynia (pain during application of innocuous stimulation such as gentle pressure and movements within the working range of the joint), hyperalgesia (enhanced pain during noxious stimulation), and persistent pain in the joint which are typical symptoms of joint inflammation. The activation of the nociceptive system may also have an influence on the inflammatory lesion due to the efferent function of the nervous system (Schaible and Grubb 1993).

As indicated in Fig. 3, there is an increased afferent inflow into the central nervous system by activation and sensitization of afferent fibres supplying the inflamed joint, thereby constituting the peripheral basis of pain in the inflamed joint (see below). Spinal cord neurones with input from the inflamed joint develop a state of hyperexcitability which significantly increases the gain of the processing of the afferent input. These changes in the spinal cord are thought to be produced by the increased or additional release of several transmitters/modulators and by the activation of receptors on postsynaptic neurones which are less or not activated under physiological conditions (see below). Changes in the release of transmitters and in the activation of postsynaptic receptors may be associated with an upregulation of the synthesis of transmitters in the long-term range. Output channels of the spinal cord (ascending axons, motoric and sympathetic pathways) are modified accordingly (see below).

Activation and sensitization of joint afferents—the peripheral neuronal basis of inflammatory pain in the joint

Numerous afferent units supplying an inflamed joint exhibit ongoing discharges and show increased responsiveness to mechanical stimuli applied to the joint, i.e. they are sensitized. Many non-nociceptive group III and IV afferents show increased responses to movements in the working range of the joints. This form of sensitization enhances the afferent input to the spinal cord quantitatively. Importantly, in most nociceptive group III and IV afferents the activation threshold is lowered: they start to respond to normally innocuous stimuli such as movements within the working range of the joint and gentle pressure applied to the joint. Figure 4(a) shows the sensitization of a high-threshold group III unit to mechanical stimuli. Before inflammation the unit responded only to noxious rotation of the joint (similar to the unit displayed in Fig. 2(b) but, during development of an acute inflammation, responses were also elicited by movements in the working range: after induction of inflammation this nociceptor started to respond to flexion in the working range of the joint (see Fig. 4(a). This form of sensitization represents a quantitative and a qualitative change since sensitized nociceptors transmit their message 'noxious movement' to the spinal cord in response to an ordinarily innocuous stimulus (Schaible and Grubb 1993; Schmidt *et al*. 1994) and thus the nociceptive system in the central nervous system is already activated by normally non-painful stimuli.

In addition to the sensitization of mechanosensitive non-nociceptive and nociceptive afferent fibres, a proportion of mechanoinsensitive afferent fibres are rendered mechanosensitive during inflammation. Figure 4(b) displays a unit which was not excited even by noxious stimulation before inflammation (traces on the left: no response to pressure applied to the joint, i.e. no identification of the receptive field, and no response to movements). During development of inflammation the fibre could be activated even by innocuous movements and gentle pressure (see action potentials in the traces on the right and the dot on the knee joint showing the location of the receptive field identified during inflammation). It was proposed that these units are 'silent nociceptors' which play a role particularly under inflammatory conditions (see above). By their activation these newly recruited nociceptive units increase the afferent drive of spinal neurones providing, by their recruitment, spatial and temporal facilitation in the spinal cord (Schaible and Grubb 1993).

Mechanisms underlying the activation and sensitization of joint afferents

The mechanosensitivity of joint afferents during inflammation may be altered by physical changes in joint tissues such as increase of the intra-articular pressure by an effusion, which is also accompanied by a decrease in joint compliance (see Fig. 3). The effect of this is that, during movements, articular pressure increases more in diseased joints than in normal joints for the same volume of fluid present. In addition to changes in the gross pathology of the joint in arthritis there are also changes in the microvasculature. Both factors could activate joint afferents but experimental evidence is sparse. Of great importance for the activation and sensitization of joint afferents is their chemosensitivity to inflammatory mediators which occur in inflamed tissue (Fig. 3). Inflammatory mediators do not only produce the inflammatory response (Moncada *et al*. 1979; Salmon and Higgs 1987; Sedgwick and Willoughby 1989) but also exert significant effects on afferent fibres (Mense 1993; Schaible and Grubb 1993; Schmidt *et al*. 1994). Whilst these extraneuronal changes represent stimuli to the receptive endings, there may be also (secondary) changes in the terminal sites themselves that contribute to the increased mechanosensitivity.

The application of inflammatory mediators to the joint has shown chemosensitivity of joint afferents for several inflammatory compounds, namely prostaglandins (in particular for PGE_2 and PGI_2), bradykinin, and serotonin. Chemosensitivity of joint afferents to other inflammatory mediators (e.g. leukotrienes) has not been experimentally tested. In significant proportions of afferent group III and IV fibres these inflammatory mediators induce patterns of firing which are similar to those seen under inflammatory conditions, i.e. excitation and/or induction of ongoing activity and an increase of the responses to mechanical stimuli (sensitization towards mechanical stimuli) and thus these mediators are thought to play an important role in the generation of inflammation-evoked activity.

In addition, individual mediators such as prostaglandins and bradykinin have been shown to interact. Both PGE_2 and PGI_2 enhance the excitatory effects of bradykinin in a proportion of neurones, and the combination of both mediators may sensitize afferent units more to mechanical stimuli than the individual mediators. The combined action of mediators may be more close to pathophysiological conditions in a joint than the action of just one mediator since several inflammatory mediators are present in inflammatory exudates at the same time. Interestingly, neuropeptides which are peripherally released from afferent fibres (producing neurogenic inflammation) failed to evoke marked effects in afferent units as far as this has been tested (Schaible and Grubb 1993; Schmidt *et al*. 1994).

Recordings from single units have shown two further important aspects of the chemosensitivity. First, each of the mediators tested excited and sensitized only proportions of the articular afferents; thus joint afferents are not homogeneous with regard to their chemosensitivity. Second, mediators excited and sensitized proportions of non-nociceptive, nociceptive, and initially mechanoinsensitive afferent units; thus there is no specific relationship between the mechanical threshold and the chemosensitivity of the afferent units. Inflammatory mediators usually do not, however, excite and sensitize non-nociceptive group II afferents (Schaible and Grubb 1993).

Activation of spinal cord neurones and generation of spinal hyperexcitability—a complex response of the central nervous system to peripheral inflammation

Figure 3 highlights the importance of the spinal cord as an integrative site for sensory and reflex functions. During the development of an acute inflammation in the joint, nociceptive-specific and wide dynamic range neurones in the spinal cord show enhanced responsiveness to noxious stimuli applied to the inflamed joint.

Nociceptive-specific neurones exhibit a reduction in their mechanical threshold such that the application of innocuous stimuli to the inflamed joint is sufficient to excite the neurones (Schaible and Grubb 1993). Figure 5 shows the changes of the response properties of a spinal cord neurone during development of inflammation in the knee joint of the rat. Before inflammation the neurone had a high-threshold receptive field in the knee and the deep tissue of the thigh and lower leg (black area in Fig. 5(b)), i.e. the neurone responded to noxious pressure on to the knee and the adjacent structures but not to noxious pressure applied to the ankle and paw (Fig. 5(a)). After induction of inflammation (k/c), the neurone showed increases of the responses to noxious pressure applied to the knee and it developed also responses to noxious pressure applied to ankle and paw (Fig. 5(a)), i.e. the neurone exhibited an expansion of the receptive field. At this stage innocuous pressure to all parts of the receptive field was sufficient to activate the neurone, i.e. the excitation threshold had changed from high to low (Fig. 5(b), stippled area).

The increased responses to stimuli applied to the inflamed joint are most likely induced by the increase of the afferent inflow due to activation and sensitization of afferent units (see Fig. 3 and above). The changes of the response properties of spinal cord neurones during inflammation do not, however, simply reflect the afferent impulses from the joint. The enhanced responses to mechanical stimuli applied to non-inflamed regions adjacent and remote from the joint and the expansion of the receptive fields (Fig. 5), as well as increased responses to electrical stimulation of peripheral nerves (not shown), indicate that the spinal cord neurones with articular input develop a state of hyperexcitability in which the responsiveness of the neurones to inputs from inflamed and non-inflamed areas is increased (the putative mechanisms of the spinal sensitivity changes are described in the next section). Numerous observations suggest that similar changes occur during chronic inflammation. It is thought that these neuronal changes underly the pain, allodynia, and hyperalgesia in the inflamed joint, adjacent and even remote non-inflamed tissue (Schaible and Grubb 1993).

Other mechanisms, however, have been found to counteract the increase in the excitability of spinal cord neurones to some extent, namely an increase in heterotopic inhibitory influences on spinal neurones with input from inflamed areas (i.e. the neurones may be inhibited by noxious stimuli applied to areas outside the receptive field of the neurone) and an increase in the effectiveness of the descending inhibition of spinal cord neurones (see Fig. 1).

Changes of the responsiveness of spinal cord neurones have effects on the motor system. In humans a painful joint always influences motor performance in that extensive movements are avoided and the limb is kept in a position where pain is minimized. In animals, there is an increase in the flexion reflex during long-lasting noxious stimulation of the knee, which is consistent with the nociceptive flexion reflex concept (stating that withdrawal is the adequate reaction to a noxious stimulus), but increases in inhibitory influences have also been demonstrated. Probably there is a new balance between excitatory and inhibitory drives from spinal and supraspinal neurones to motor neurones which makes one keep an injured or inflamed joints in mid-position where joint afferents are less active than at flexed and extended positions (Schaible and Grubb 1993).

Mechanisms underlying the synaptic activation of spinal cord neurones and the development of hyperexcitability (Table 2)

The communication between neurones is provided by chemical transmission in synapses. Upon activation (propagation of action potentials into the axonal terminal) the presynaptic neurone releases excitatory or inhibitory transmitters which cross the synpatic cleft, bind to receptors, and then open ion channels on the postsynaptic neurones (Fig. 6). The sum of all excitatory and inhibitory synaptic potentials evoked in the neurone determines whether the postsynaptic neurone will fire. Importantly, many neurones do not only contain a classical transmitter (small molecules such as L-glutamate or biogenic amines) but rather exhibit a coexistence of classical transmitters and one or several neuropeptides such as substance P and calcitonin gene-related peptide in their presynaptic endings. Postsynaptic neurones contain a variety of receptors for classical transmitters and neuropeptides (Duggan and Weihe 1991; Urban 1994).

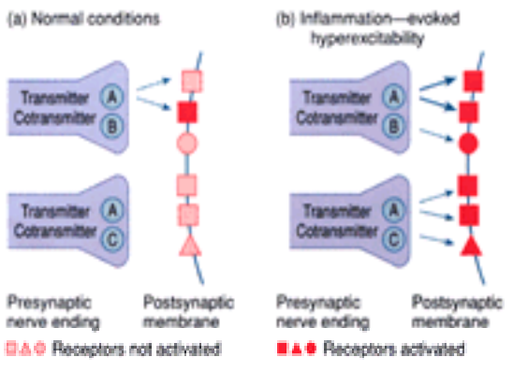


Fig. 6 Proposed mechanisms involved in the generation and maintenance of inflammation-evoked hyperexcitability of spinal cord neurones. The graphs in (a) and (b) show the membrane of a postsynaptic neurone with different receptors for transmitters. The ligands for these receptors, the transmitters, and cotransmitters, are stored in and released from the presynaptic nerve terminals connected to the postsynaptic neurone. (a) Synaptic function under normal conditions. Only the upper presynaptic ending is active and only one transmitter is released. In this situation many receptors are not activated. (b) Synaptic function during inflammation. Both presynaptic endings are activated and they release a number of transmitters which are bound to the receptors on the postsynaptic side.

γ-motor neurone	Nerve cell in the spinal cord which forms neuromuscular endings in the skeletal muscle. Activation of γ-motor neurones causes contraction of the muscle
Excitatory amino acids	Amino acids which cause depolarization (synaptic activation) of neurones. They serve as transmitters at synapses
γ-motor neurone	Nerve cell in the spinal cord of which the axon terminates on muscle spindles in the skeletal muscle. The activation of γ-motor neurones enhances the sensitivity of muscle spindles which measure the length of the skeletal muscle
Postganglionic sympathetic neurone	Efferent neurone of the sympathetic nervous system which is synaptically activated by preganglionic neurones in sympathetic ganglia. The postganglionic neurone mediates the effect of the sympathetic nervous system on the tissue
Postsynaptic receptor	Ligand-gated ion channel in postsynaptic membrane which is opened by a transmitter
Presynaptic ending	Presynaptic nerve terminal at which transmitter(s) is (are) released
Synapse	Microstructure by which nerve cells are connected. The presynaptic nerve terminal releases transmitter which diffuses to receptors on the membranes of the postsynaptic neurone and opens there ligand-gated ion channels
Transmitter	Endogenous compound that is released from a presynaptic ending and acts on a receptor in the postsynaptic membrane
Vasoconstrictor neurone	Neurone of the sympathetic nervous system which causes vasoconstriction of vessels

Table 2 Definitions (II)

Synaptic transmission is not only the basis of the normal communication between neurones; due to the complexity of the pre- and postsynaptic processes, changes in the synaptic transmission are also thought to be involved in the plasticity of the nervous system. Figure 6 illustrates some properties of the synaptic function proposed for the normal state and for the hyperexcitable state of spinal cord neurones, which is an example of the functional plasticity in the nociceptive system. Under normal conditions only a proportion of the afferent fibres are activated by innocuous and noxious stimuli and only some of several transmitters are released from presynaptic endings. Consequently only some receptors on the postsynaptic side are activated. During inflammation numerous afferent neurones are sensitized such that they are activated by innocuous and noxious stimuli (see above). Therefore more presynaptic endings in the spinal cord are active, and a 'cocktail' of transmitters and/or modulators is released which is quantitatively and qualitatively different to that released during noxious stimulation of normal tissue. Hence more and different receptors on the postsynaptic neurones are activated by the ligands. This results in a stronger activation on the postsynaptic neurone as shown in Fig. 5. In addition, sensitivity of postsynaptic neurones may be altered by intracellulalar processes which are induced by the synaptic activation (Schaible and Grubb 1993).

Excitatory amino acids such as L-glutamate are transmitters with a wide distribution in the central nervous system. Table 3 shows that L-glutamate acts on three classes of receptors at postsynaptic neurones (Monaghan *et al*. 1989; Headley and Grillner 1990; Urban 1994). Electrophysiological and behavioural experiments have shown that excitatory amino acids and their receptors are important in nociception (Headley and Grillner 1990; Schaible and Grubb 1993; Urban 1994). Figure 7(a) displays the involvement of non-NMDA and **NMDA** (*N*-methyl-D-aspartate) receptors in the synaptic activation of a spinal cord neurone which responds to innocuous and noxious pressure applied to the knee joint. The graph shows the action potentials of the spinal cord neurone which are elicited by application of innocuous and noxious pressure to the knee joint. Whilst the spinal ionophoretic administration of the antagonist, 6-cyano-7-nitroquinoxaline-2,3-dione (**CNQX**), at the non-NMDA receptor reduced the responses to innocuous and noxious pressure applied to the normal knee, the administration of the antagonist, ketamine, at the NMDA receptor, only reduced the responses to noxious pressure (Fig. 7(a)). Thus, under normal conditions, NMDA receptors seem to be mainly involved in the

nociceptive processing in these neurones but not in the processing of non-nociceptive inputs.

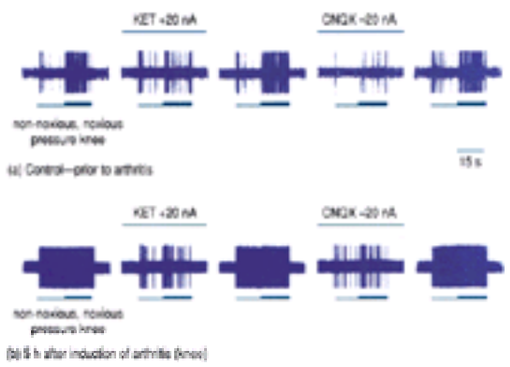


Fig. 7 Responses (action potentials) of a spinal cord neurone to innocuous and noxious pressure applied to the knee joint and reduction of the responses by antagonists at the NMDA receptor (ketamin) and at the non-NMDA receptor (CNQX, see text). In (a) the knee joint is normal and the NMDA antagonist reduces only the response to noxious pressure. In (b) the knee joint is inflamed (the mechanical stimuli evoke stronger responses) and in this situation the NMDA antagonist reduces the responses to innocuous and noxious pressure. CNQX reduces the responses to innocuous and noxious pressure under both conditions. The antagonists were administered ionophoretically to the neurone using multibarrel electrodes for simultaneous recording and drug application.

Receptor	Specific ligands	Functional properties
NMDA	N-methyl-D-aspartate	Ion channel for Na^+, K^+, Ca^{2+} only opened after strong depolarization
Non-NMDA	(RS)-2-amino-3-hydroxy-5-methylisoxazole-4-propionic acid (AMPA); kainate	Ion channel for Na^+, K^+
Metabotropic	E.g. Trans-(+)-1-amino-(1S,3R)-cyclopentane-dicarboxylic acid (t-ACPD)	Coupled to G-proteins Activation of intracellular messenger systems

Table 3 Receptors for excitatory amino acids

The non-NMDA and NMDA receptors are not only involved in the synaptic transmission of nociceptive information; they are also involved in the generation of hyperexcitability (see Fig. 5). Increases in the responsiveness (hyperexcitability) of spinal cord neurones during inflammation and other noxious stimuli are prevented as long as either non-NMDA or NMDA receptors are blocked (see references in Neugebauer *et al*. 1993). Evidence has also been provided that blockade of the G-protein coupled metabotropic glutamate receptors prevents hyperexcitability (Neugebauer *et al*. 1994). Thus all three receptor types for excitatory amino acids must be activated to generate hyperexcitability of the neurone, and the changes of the responses displayed in Fig. 5 do not occur when the receptors for excitatory amino acids are blocked. Excitatory amino acids and their receptors are also involved in the maintenance of hyperexcitability. This is displayed in Fig. 7(b) which shows the same neurone as Fig. 7(a) but now the knee joint is inflamed and the neurone exhibits stronger responses to innocuous and noxious pressure applied to the knee. Both the increased responses to innocuous and noxious pressure applied to the inflamed knee are reduced by the antagonists at the NMDA and non-NMDA receptor (Neugebauer *et al*. 1993).

Some neuropeptides are also of particular interest for synaptic transmission in the spinal cord and its plasticity. They are contained in many afferent units and intrinsic neurones of the spinal cord (Holzer 1988; Duggan and Weihe 1991). The intraspinal release of the tachykinins, substance P and neurokinin A, and of calcitonin gene-related peptide can be evoked by application of noxious stimuli such as noxious pressure applied to the joint. During inflammation of the joint, innocuous stimulation of inflamed tissue may be sufficient to elicit release of the peptide, whilst such stimuli usually do not evoke release when the joint is in normal condition (Schaible and Grubb 1993; Schaible *et al*. 1994). Using antagonists it has been shown that substance P is involved in the synaptic transmission of nociceptive information (Henry 1994) and some evidence has been provided that substance P may also play a role in the generation and maintenance of hyperexcitability (see Neugebauer *et al*. 1995). Inhibitory neuropeptides may also play an important role. The synthesis of dynorphin, an endogenous agonist, at the k-opioid receptor has been found to be increased during long-lasting noxious stimulation such as inflammation. Inhibitory neuropeptides may set a new balance between excitatory drives and inhibitory mechanisms and thus control the degree of pain experienced (Dickenson 1994).

A number of issues have not yet been clarified; for example it is not known to what extent the changes in the central nervous system require the permanent drive from sensitized afferent fibres from the inflamed tissue and to what extent alterations in the central nervous system become independent of the afferent input. Furthermore it is not yet clear whether chronic pain lasting for weeks, months, and years involves the same mechanisms which are at work in acute and persisting acute pain.

Activation of neurones in the supraspinal nervous system during joint inflammation

Since spinal cord neurones change their response pattern during inflammation, neurones in the brainstem and the thalamus receive altered inputs via the ascending tracts. As a consequence, in the course of acute unilateral inflammation and of chronic polyarthritis, neurones in the ventrobasal thalamus and the cortical area SI display enhanced responses to mechanical stimulation of inflamed and non-inflamed areas and increases of their receptive fields (Guilbaud 1988). It is not known whether these alterations in the discharges are entirely due to the inflow from the spinal cord or whether intrinsic mechanisms in the thalamus and cortex contribute to the expression of these changes in response.

The contribution of the nervous system to the inflammatory process in the joint

As outlined above, the inflammatory process is mediated by a number of mediators which are released and/or synthesized in inflamed tissue. There is experimental and clinical evidence that the nervous system may also contribute to or influence the inflammatory process, by efferent functions of afferent fibres and by efferent sympathetic fibres.

Efferent functions of afferent fibres—mechanisms of neurogenic inflammation

The neuropeptides substance P and calcitonin gene-related peptide have been found to be released from afferent fibres supplying numerous organs including the joints upon activation. These peptides produce a neurogenic inflammation characterized by vasodilation and plasma extravasation (Holzer 1988). Other actions of neuropeptides include stimulation of white blood cells and synoviocytes to release and/or produce inflammatory compounds (Levine *et al*. 1987). There is experimental evidence that neuropeptides contribute to the development of the inflammation of the tissue since the expression of acute and chronic inflammatory lesions has been found to be attenuated after impairment of the function of afferent fibres (e.g. by pretreatment with the neurotoxin capsaicin which destroys or inactivates unmyelinated afferent fibres or by immunization against calcitonin gene-related peptide). A role in the contribution to the inflammatory lesion has been ascribed to substance P and calcitonin gene-related peptide. Observations in human joint diseases seem to support the experimental evidence, although the picture from clinical studies is not totally clear (Schaible and Grubb 1993). In this context, findings could be important which showed an upregulation of the synthesis of substance P and calcitonin gene-related peptide in the dorsal root ganglia in segments supplying inflamed tissue (Schaible and Grubb 1993). Neuropeptides, however, do not only support inflammatory reactions; they may also support wound healing (Holzer 1988).

Contribution of efferent sympathetic fibres to the expression of inflammatory disease

The joints are supplied by efferent sympathetic nerve fibres which exhibit ongoing and reflex discharges. From their efferent effects they are vasoconstrictor neurones. As other sympathetic subsystems, such as the cardiac postganglionic sympathetic neurones and the sympathetic adrenal system, the sympathetic efferents to the joint show reflex discharges when noxious stimuli are applied to the joint (Jänig 1985; Schaible and Grubb 1993). Some sympathetic subsystems (e.g. the cardiac system and the adrenal medulla) show more pronounced reflex responses to stimulation of the inflamed joint than to stimulation of the normal joint (Schaible and Grubb 1993).

In the polyarthritic rat, the sympathetic nervous system seems to contribute to the expression of the inflammatory lesions since the reduction of sympathetic activity and/or the blockade of the postsynaptic effects by antagonists at adrenergic receptors in the tissue partially reduced the severity of the lesions. In acute models of experimental inflammation the sympathetic nervous system has not been found to contribute to the expression of the inflammatory lesion (see Schaible and Grubb 1993).

Reduction of activity in afferent fibres and neurones of the central nervous system by antinociceptive compounds

With increasing knowledge of the nociceptive pathways and the reactions of the nervous system to inflammatory lesions in the periphery, the antinociceptive action of drugs can be better related to targets and neuronal events in the nervous system. The spectrum and sites of effects of antinociceptive drugs are more complex than previously thought.

Reduction of afferent discharges

Several classes of compounds have been found to reduce the inflammation-evoked activity in articular afferents. Non-steroidal anti-inflammatory drugs (**NSAIDs**) such as acetylsalicylic acid and indomethacin reduce ongoing and stimulus-evoked discharges. Since NSAIDs inhibit the synthesis of prostaglandins these findings show that elevated levels of prostaglandins are involved in the generation of inflammation-evoked discharges. Not all joint afferents, however, have been found to be sensitive to the effects of prostaglandins (see above). Whilst the reduction of activity in afferent fibres unequivocally shows the originally proposed peripheral action of NSAIDs, there is growing evidence that the antinociceptive action of NSAIDs is not only due to actions in the peripheral inflamed tissue, an additional action in the central nervous system is likely (McCormack and Brune 1991).

In behavioral experiments the application of bradykinin antagonists (of the B_1 and B_2 type) has been shown to reduce inflammation-evoked hyperalgesia suggesting an important role of endogenous bradykinin (Dray and Perkins 1993). A reduction of the sensitivity of joint afferents to mechanical and chemical stimuli has been described for capsaicin, a naturally occurring constituent of hot red peppers. The mechanisms underlying these desensitizing effects are not yet clear. Evidence is accumulating that opioids do not only have actions in the central nervous system (see below) but also in the peripheral nervous system. After peripheral administration of opioids a reduction of the afferent discharges, a reduction of nociceptive behaviours, and a reduction of the release of substance P from joint afferents have been shown. All three opioid receptor types, μ-, d-, and k-receptors, seem to contribute to these effects (Schaible and Grubb 1993). It remains to be determined whether bradykinin antagonists, capsaicin, or peripherally acting opioids can be successfully used for treatment of arthritic pain.

Reduction of activity in neurones of the central nervous system

The reduction of discharges in afferent fibres by peripherally acting compounds will reduce the activity in the central nervous system and thus a reduction of inflammation-evoked activity may be of peripheral origin. It should be noted, however, that it has not been shown to what extent the activity in the central nervous system is dependent on the afferent inputs under pathophysiological conditions (see above). Some potent analgesics such as opioids, e.g. morphine (acting on μ-receptors), have their main targets within the central nervous system. In spine cord transected animals a spinal site of the antinociceptive action of morphine has been demonstrated. It has also been suggested that morphine may increase the descending inhibition of spinal cord neurones by acting on brainstem nuclei (see Fig. 1). There is, however, still no agreement on the importance of the latter mechanism (Dickenson 1994). Both the analgesic effect of morphine and the tolerance for opioids have been found to be enhanced in polyarthritic rats (Kayser 1994).

As mentioned above, NSAIDs may also have additional targets in the central nervous system. It has yet to be shown whether actions of NSAIDs in the central nervous system are also due to the inhibition of prostaglandin synthesis or whether other mechanisms are involved. When the mechanisms of signal transmission in the central nervous system are better understood, a more specific blockade of receptors (e.g. receptors for excitatory amino acids and neuropeptides) may be applicable, provided that the overlap of transmitter/receptors at different sites and nerve functions does not create too many side-effects.

Chapter References

Besson, J.-M. and Chaouch, A. (1987). Peripheral and spinal mechanisms of nociception. *Physiological Reviews*, **67**, 67–186.

Coderre, T.J., Katz, J., Vaccarino, A.L., and Melzack, R. (1993). Contribution of central neuroplasticity to pathological pain: review of clinical and experimental evidence. *Pain*, **52**, 259–85.

Dickenson, A.H. (1994). Where and how do opioids act? In *Proceedings of the 7th World Congress on Pain, Progress in Pain Research and Management*, Vol. 2 (ed. G.F. Gebhart, D.L. Hammond, and T.S. Jensen), pp. 525–52. IASP Press, Seattle.

Dray, A. and Perkins, M. (1993). Bradykinin and inflammatory pain. *Trends in Neurosciences*, **16**, 99–104.

Dubner, R. and Bennett, G.J. (1983). Spinal and trigeminal mechanisms of nociception. *Annual Reviews of Neuroscience*, **6**, 381–418.

Dubner, R. and Ruda, M.A. (1992). Activity-dependent neuronal plasticity following tissue injury and inflammation. *Trends in Neurosciences*, **15**, 96–103.

Duggan, A.W. and Weihe, E. (1991). Central transmission of impulses in nociceptors: events in the superficial dorsal horn. In *Towards a new pharmacotherapy of pain* (ed. A.I. Basbaum and J.-M. Besson), pp. 35–67. John Wiley and Sons, Chichester.

Gebhart, G.F. (1986). Modulatory effects of descending systems on spinal dorsal horn neurones. In *Spinal afferent processing* (ed. T.J. Yaksh), pp. 391–416. Plenum, New York.

Guilbaud, G. (1988). Peripheral and central electrophysiological mechanisms of joint and muscle pain. In *Proceedings of the 5th World Congress on Pain* (ed. R. Dubner, G.F. Gebhart, and M.R. Bond), pp. 201–15. Elsevier, Amsterdam.

Headley, P.M. and Grillner, S. (1990). Excitatory amino acids and synaptic transmission: the evidence for a physiological function. *Trends in Pharmacological Sciences*, **11**, 205–11.

Henry, J.L. (1994). Role of substance P in spinal nociceptive pathways. In *Neuropeptides, nociception and pain* (ed. T. Hökfelt, H.-G. Schaible, and R.F. Schmidt), pp. 209–20. Chapman & Hall, Weinheim.

Heppelmann, B., Meálinger, K., Neiss, W.F., and Schmidt, R.F. (1990). Ultrastructural three-dimensional reconstruction of group III and group IV sensory nerve endings ('free nerve endings') in the knee joint capsule of the cat: evidence for multiple receptive sites. *Journal of Comparative Neurology*, **292**, 103–16.

Hökfelt, T., Schaible, H.-G., and Schmidt, R.F. (eds) (1994). *Neuropeptides, nociception and pain*. Chapman & Hall, Weinheim.

Holzer, P. (1988). Local effector functions of capsaicin-sensitive sensory nerve endings: involvement of tachykinins, calcitonin gene-related peptide and other neuropeptides. *Neuroscience*, **24**, 739–68.

Jänig, W. (1985). Organization of the lumbar sympathetic outflow to skeletal muscle and skin of the cat hindlimb and tail. *Reviews of Physiology, Biochemistry, and Pharmacology*, **102**, 119–213.

Johansson, H., Sjölander, P., and Sojka, P. (1991). Receptors in the knee joint ligaments and their role in the biomechanics of the joint. *CRC Critical Reviews in Biomedical Engineering*, **18**, 341–68.

Kayser, V. (1994). Endogenous opioid systems in the modulation of pain: behavioral studies in rat models of persistent hyperalgesia. In *Proceedings of the 7th World Congress on Pain, Progress in Pain Research and Management*, Vol. 2 (ed. G.F. Gebhart, D.L. Hammond, and T.S. Jensen), pp. 553–68. IASP Press, Seattle.

Kenshalo, D.R., Jr. and Willis W.D. (1991). The role of the cerebral cortex in pain sensation. In *Cerebral cortex*, Vol. 9 (ed. A. Peters), pp. 153–212. Plenum Corporation, New York.

Levine, J.D., Goetzl, E.J., and Basbaum, A.I. (1987). Contribution of the nervous system to the pathophysiology of rheumatoid arthritis and other polyarthritides. *Rheumatic Disease Clinics of North America*, **13**, 369–83.

McCormack, K. and Brune, K. (1991). Dissociation between the antinociceptive and anti-inflammatory effects of the nonsteroidal anti-inflammatory drugs. A survey of their analgesic efficacy. *Drugs*, **41**, 533–47.

McMahon, S.B., Lewin, G.R., and Wall, P.D. (1993). Central hyperexcitability triggered by noxious inputs. *Current Opinion in Neurobiology*, **3**, 602–10.

Mense, S. (1993). Nociception from sceletal muscle in relation to clinical muscle pain. *Pain*, **54**, 241–89.

Monaghan, D.T., Bridges, R.J., and Cotman, C.W. (1989). The excitatory amino acid receptors: their classes, pharmacology and distinct properties in the function of the central nervous system. *Annual Reviews of Pharmacology and Toxicology*, **29**, 365–402.

Moncada, S., Ferreira, S.H., and Vane, J.R. (1979). Pain and inflammatory mediators. In *Handbook of experimental pharmacology*, Vol. 50, Part I: *Inflammation* (ed. J.R. Vane and S.H. Ferreira), pp. 588–616. Springer, Berlin.

Neugebauer, V., Lücke, T., and Schaible, H.-G. (1993). *N*-methyl-D-aspartate (NMDA) and non-NMDA receptor antagonists block the hyperexcitability of dorsal horn neurones during development of acute arthritis in rat's knee joint. *Journal of Neurophysiology*, **70**, 1365–77.

Neugebauer, V., Lücke, T., and Schaible, H.-G. (1994). Requirement of metabotropic glutamate receptors for the generation of inflammation-evoked hyperexcitability in rat spinal cord neurones. *European Journal of Neuroscience*, **6**, 1179–86.

Neugebauer, V., Weiretter, F., and Schaible, H.-G. (1995). The involvement of substance P and neurokinin-1 receptors in the hyperexcitability of dorsal horn neurones during development of acute arthritis in rat's knee joint. *Journal of Neurophysiology*, **73**, 1574–83.

Perl, E.R. (1984). Pain and nociception. In *Handbook of physiology*, Section 1, Vol. III, Part II *The nervous system*, pp. 915–75. American Physiology Society, Bethesda, MD.

Salmon, J.A. and Higgs, G.A. (1987). Prostaglandins and leukotrienes. In *Inflammation—mediators and mechanisms, British Medical Bulletin*, Vol. 2 (ed. D.A. Willoughby), pp. 285–96. Churchill Livingstone, Edinburgh.

Schaible, H.-G. and Grubb, B.D. (1993). Afferent and spinal mechanisms of joint pain. *Pain*, **55**, 5–54.

Schaible, H.-G. and Schmidt, R.F. (1988). Time course of mechanisensitivity changes in articular afferents during a developing experimental arthritis. *Journal of Neurophysiology*, **60**, 2180–95.

Schaible, H.-G., Freudenberger, U., Neugebauer, V., and Stiller, R.U. (1994). Intraspinal release of immunoreactive calcitonin gene-related peptide during development of inflammation in the joint in vivo—a study with antibody microprobes in cat and rat. *Neuroscience*, **62**, 1293–305.

Schmidt, R.F. (1989). Nociception and pain. In *Human physiology* (2nd edn) (ed. R.F. Schmidt and G. Thews), pp. 223–36. Springer, Heidelberg.

Schmidt, R.F., Schaible, H.-G., Meálinger, K., Heppelmann, B., Hanesch, U., and Pawlak, M. (1994). Silent and active nociceptors: structure, functions, and clinical implications. In *Proceedings of the 7th World Congress on Pain, Progress in Pain Research and Management*, Vol. 2 (ed. G.F. Gebhart, D.L. Hammond, and T.S. Jensen), pp. 213–50. IASP Press, Seattle.

Sedgwick, A.D. and Willoughby, D.A. (1989). Initiation of the inflammatory response and its prevention. In *Handbook of inflammation. The pharmocology of inflammation*, Vol. 5 (ed. I.L. Bonta, M.A. Bray, and M.J. Parnham), pp.27–47. Elsevier, Amsterdam.

Urban, L. (ed.) (1994). *Cellular mechanisms of sensory processing. The somatosensory system*. Springer, Berlin.

Wall, P.D. and Melzack, R. (eds) (1994). *Textbook of pain* (3rd edn). Churchill Livingstone, Edinburgh.

Willis, W.D. (1985). *The pain system. The neural basis of nociceptive transmission in the mammalian nervous system*. Karger, Basel.

Willis, W.D. (1988). Anatomy and physiology of descending control of nociceptive responses of dorsal horn neurones: a comprehensive review. *Progress in Brain Research*, **77**, 1–29.

Willis, W.D. (ed.) (1992). *Hyperalgesia and allodynia*. Raven, New York.

Willis, W.D. and Coggeshall, R.E. (1991). *Sensory mechanisms of the spinal cord* (2nd edn). Plenum, New York.

3.1 Cells and mediators

Mark J. Walport and Gordon W. Duff

Introduction

Inflammation is an essential protective process preserving the integrity of organisms against physical, chemical, and infective insults. Inflammation is associated with the production of disease when inappropriately triggered, acutely, as in an attack of gout, or chronically, when the normal mechanisms for the resolution of inflammation do not operate or the stimulus to inflammation cannot be removed.

This chapter will consider the basic physiology and pathophysiology of the innate mechanisms of immunity and their role in inflammation. Emphasis will be placed on the normal regulation of the pathways of inflammation and how these may be disturbed in disease. The major physiological activities of individual components of innate immunity are often redundant, for example there are many receptors capable of recognition of bacteria and immune complexes that co-operatively mediate phagocytosis. An accurate perspective of the function *in vivo* of a molecule or pathway of inflammation is sometimes gained by the analysis of individuals who show inherited total deficiency of a particular protein. Examples of such deficiency states will be considered throughout this chapter.

Variation in host response may play an important part in determining disease susceptibility. It may be that genetic susceptibility to many rheumatic diseases is determined in part by inherited polymorphisms in the function of the molecules and cells involved in inflammatory mechanisms. This will be illustrated by consideration of the role of inherited variations of the complement system in the induction of increased susceptibility to the development of systemic lupus erythematosus.

We will not attempt to provide a comprehensive review of the role of the innate system of immunity in the production of rheumatic disease, but will illustrate the main principles using selected examples. Particular issues that we will consider are: (i) what distinguishes acute from chronic inflammation; (ii) does genetic variation in the mechanisms of innate immunity influence the susceptibility to rheumatic disease; (iii) what are the factors determining the tissue specificity of inflammation in rheumatic diseases; (iv) what inflammatory pathways are implicated in the production of tissue injury; and (v) is there heterogeneity in inflammatory mechanisms between different diseases and between different individuals with the same disease?

Acute versus chronic inflammation

A critical question in understanding the pathogenesis of chronic inflammatory rheumatic diseases is what factors distinguish an acute attack of gout, followed by the restoration of normal joint function, from rheumatoid arthritis where there is prolonged influx of inflammatory cells and chronic synovial injury? There are likely to be many answers to this question, which may be classified under three headings: (i) persistence of the inflammatory stimulus; (ii) failure of normal homeostasis to switch off inflammation after the stimulus is gone; and (iii) escape of a clone of cells releasing proinflammatory stimuli.

There are many examples of the persistence of an inflammatory stimulus causing chronic disease in humans such as: (i) persistence of foreign antigen, due to replication of an infectious agent, for example the inflammatory consequences of bacterial endocarditis; (ii) persistence due to failure of clearance of a foreign substance, typified by asbestos-associated injury; and (iii) autoimmunity, where the antigen is indivisible from the host.

The hypothesis that defects in the homeostasis of inflammation cause chronic inflammation is less well explored, although the diseases associated with inherited immunodeficiency provide some dramatic support for this idea. Chronic granulomatous disease, for example, illustrates the consequences of the failure of electron transport to oxygen in phagocytic vacuoles, causing a precipitous drop in pH in this compartment. Granulomas ensue from the resulting failure to degrade engulfed micro-organisms and autologous debris. The strong association between inherited complement deficiency and systemic lupus erythematosus provides a further example of a host defect in innate immunity leading to chronic inflammation.

Autonomy of a clone of autologous cells leading to the production of inflammatory mediators is a further mechanism for induction of chronic inflammation. One example that illustrates this is the systemic effects of certain benign and malignant tumours, the 'paraneoplastic' syndromes (Butler *et al*. 1987; Sanchez-Guerrero *et al*. 1990). Clonal proliferation of T cells may produce dramatic inflammatory effects, for example some cases of the hypereosinophilic syndrome are associated with the presence of T-cell tumours (Catovsky *et al*. 1980; Fauci 1982; O'Shea *et al*. 1987). Another striking association is of clonal proliferations of large, granular, lymphocytic T cells with a Felty-like syndrome in rheumatoid arthritis (Dhodapkar *et al*. 1994). A third example is provided by the consequences of some 'benign' clonal proliferations of B lymphocytes. Essential mixed cryoglobulinaemia is a disease in which a B-lymphocyte clone produces an IgM rheumatoid factor that complexes with IgG *in vivo*, leading to the production of inflammation mediated by immune complexes and characterized by vasculitis, nephritis, Raynaud's phenomenon, and neuropathy (Brouet *et al*. 1974; Gorevic *et al*. 1980).

Genetics of inflammation

It is likely that inherited susceptibility to the majority of chronic inflammatory diseases is under polygenic control, and in most cases the relative contribution of genes and environment is unknown. The genetic contribution to the aetiology of rheumatic diseases may be considered under two overlapping headings, the genetics of the adaptive immune system and the genetics of innate immunity and of the control of inflammatory responsiveness. Important questions for those interested in the genetics of adaptive immunity are: (i) does the repertoire of immunoglobulin and T-cell receptors vary between different people; and (ii) what is the influence of the thymus and of products of the major histocompatibility complex (**MHC**) in selecting the T-cell receptor repertoire?

An understanding of the genetics of inflammation may have many important outcomes. This is illustrated by consideration of the pathogenesis of infectious disease, which may be influenced by variation in host responsiveness as much as by the particular organism, as follows.

1. Deficiency of a host defence pathway may lead to chronic infection and chronic inflammation, for example chronic infection with lymphochoriomeningitis virus in new-born mice.

2. An over-response to an infectious agent may lead to inflammatory injury of tissue, for example the granulomatous response in tuberculoid leprosy.

Similar heterogeneity of responsiveness, based on genetic polymorphisms, may determine the level of susceptibility to chronic inflammatory rheumatic disease following a particular environmental stimulus and also may influence the precise pattern of inflammatory damage. It is possible, for example, that microscopic polyarteritis nodosa (a leucocytoclastic vasculitis affecting small and medium arteries) and Churg–Strauss vasculitis (characterized by asthma, eosinophilia, and granulomas), both associated with the presence of antineutrophil cytoplasmic antibodies, are caused by the same agent, and that differences in the expression of the two diseases are due to different patterns of host response, which may be determined genetically. The opposite possibility is that they have entirely different causes and that the similarities between them are a result of the relatively limited pathways of expression of vascular inflammation. These possibilities cannot be distinguished on the basis of present evidence.

Cytokine genes

Inherited variation in the expression of cytokines may explain observed differences between individuals in the amplification and maturation of specific and/or innate immune responses. An example that illustrates this is the recent discovery of inherited polymorphisms around the human tumour necrosis factor (**TNF**) genes. Variation in mitogen-stimulated production of TNF-a (but not TNF-b) was correlated with HLA-DR type, suggesting that *TNF*-a genes in different MHC haplotypes may show differences in control mechanisms. Subjects with HLA-DR3 or -DR4 were high producers of TNF-a whereas HLA-DR2-positive individuals were low producers (Jacob *et al*. 1990). A correlation has been reported between the prevalence of lupus nephritis and the presence of HLA-DR2. These observations are compatible with the results of experiments in mice with systemic lupus erythematosus: NZW mice produced low levels of TNF-a and treatment of the (NZB × NZW) F_1 hybrid strain with TNF-a was associated with delay in onset of nephritis and prolonged survival (Jacob and McDevitt 1988).

In human inflammatory diseases, recent attention has focused mainly on the major proinflammatory loci such as the *TNF* gene on the short arm of human chromosome 6 and the interleukin (**IL**)-1 gene cluster on the long arm of human chromosome 2 (Duff 1994).

There is a cluster of lymphokine genes on chromosome 5 (Marsh *et al*. 1994) where several polymorphic variants have also been characterized, but less work has been done in populations of patients with inflammatory or immune diseases.

The increasing evidence that IL-1 and TNF act as important mediators of the inflammatory response led to the rapid characterization of their genes and then to the search for variant forms that may have functional differences making an individual more or less likely to develop a rheumatic disease, or influence the severity of the disease if not directly involved in its initiation. For example, the ability to predict rapid bone erosions in rheumatoid arthritis using markers in the *IL1* and/or *TNF* genes would be of clinical value in the management of new cases of rheumatoid arthritis.

TNF locus

These are three known *TNF* genes in close proximity within the class III region of the MHC. Across this region several DNA polymorphisms have been described in the normal human population. In the promoter region of the *TNF*-a gene (better known as *TNFA*), at 308 bp upstream of the transcription start site, there is an A–G change in nucleotides (Wilson *et al*. 1992) (Fig. 1). The carriage rate of the rarer allele (*TNFA2*) is 16 to 20 per cent in populations from England, Ireland, Germany, Holland, and The Gambia in Africa. The *TNFA2* allele is part of the extended MHC haplotype *A1, B8, DR3, DQ2* (Wilson *et al*. 1993). This particular haplotype of the MHC has been associated with several autoimmune diseases and is sometimes referred to as the 'autoimmune haplotype'. Diseases associated with this haplotype include: insulin-dependent diabetes; autoimmune thyroiditis; systemic lupus erythematosus; myasthenia gravis; and coeliac disease. It seems likely that in most cases *TNF* is not the primary disease locus and its higher frequency in these diseases results from linkage to other genes within the haplotype.

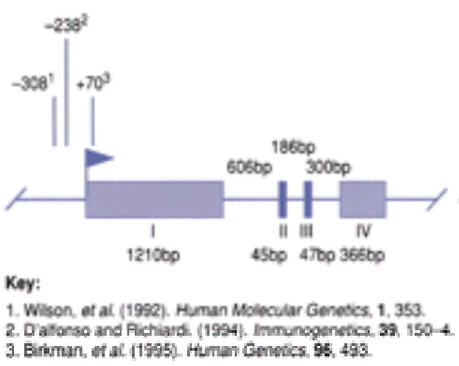


Fig. 1 Polymorphisms within the human *TNFA* locus (6p21).

The 308-bp change in *TNFA* is in the promoter region of the DNA that initiates gene transcription. This polymorphism, therefore, could change the function of the promoter, resulting in altered production of TNF-a. Assays of promoter function where the two allelic promoter fragments of *TNFA* were placed upstream of a bacterial reporter gene indicated that the rarer allele (*TNFA2*) was some six- to eight-fold more efficient at directing transcription than the common allele (*TNFA1*) (G. W. Duff, unpublished observation). However, to date there is little evidence that this results in a reproducible difference in the production of TNF-a protein, although this may be both cell-type and stimulus-specific and cannot at present be excluded.

Interest in *TNF* polymorphisms has extended into infectious diseases and a large study in The Gambia showed that homozygosity for *TNFA2* conferred an eight-fold increased risk of death or major sequelae from cerebral malaria (McGuire *et al*. 1994). Since death in malaria has previously been correlated with high blood levels of TNF-a, this finding suggests that the *TNFA2* allele may be a marker for an exaggerated TNF response that is pathogenic.

The IL1 *gene cluster*

There are three known *IL1* genes designated as *IL1A* (producing IL-1a), *IL1B* (producing IL-1b), and *IL1RN* (producing the IL-1 receptor antagonist). All three genes lie within approx. 500 kb of DNA on the long arm of chromosome 2 (Nicklin *et al*. 1994).

Several base-change and repeat-sequence polymorphisms have been described across this region (Fig. 2) and specific *IL1* cluster genotypes defined by these DNA markers have been found to be significantly over-represented in inflammatory diseases.

Fig. 2 (a) Polymorphisms within the human *IL1A* locus (2q13–14); (b) polymorphisms within the human *IL1B* locus (2q13–14); (c) polymorphisms within the human

IL1RN locus (2q13–14).

IL1A

Several polymorphisms have been described (see Fig. 2(a)). A base variation at position 889 produces two alleles, the rarer of which (*IL1A2*) was found to confer an approximate doubling of relative risk of juvenile chronic arthritis in Norwegian children (McDowell *et al*. 1995). This applied only to the early-onset, pauciarticular variety of juvenile chronic arthritis, and, within this population, *IL1A2* was more significantly over-represented in patients with higher erythrocyte sedimentation rates and those with chronic iridocyclitis, a sight-threatening ocular inflammation that occurs in early-onset, pauciarticular juvenile chronic arthritis.

Further studies involving larger numbers of patients and other ethnic groups need to be done to confirm this finding.

IL1B

A base change in the 5' flanking region (–511) and another within exon 5 (+3953) defines a genotype that occurs significantly more frequently in psoriatic skin disease (Di Giovine *et al*. 1996) (Fig. 2(b)). Studies of this genotype in psoriatic arthritis have not yet been completed, but of great interest is the additional finding that the rarer, +3953 *IL1B* allele is related to the production of IL-1b *in vitro*. Blood mononuclear cells from heterozygous individuals produce about twice as much IL-1b than cells from individuals who are homozygous for the common allele, while those who are homozygous for the rare allele produce approximately four times more IL-1b (Di Giovine *et al*. 1996).

This gene-dose effect on the production of IL-1b seems particularly relevant to the genetic association with psoriasis since IL-1 is thought to be a key mediator of skin inflammation and mice transgenic for an overactive *IL1* gene driven by a skin-specific promoter develop keratinocyte hyperplasia and dermal inflammation with many similarities to human psoriasis (Groves and Kupper 1994).

Preliminary studies in rheumatoid arthritis have suggested that the *IL-1* genotypes may also be relevant to the early occurrence of bone and cartilage erosion, a feature of the disease that is of major clinical significance.

***IL1RN* gene**

Several DNA variations occur within this gene but the most extensively studied marker in clinical populations has been a variable number of tandem repeats of an 86-bp sequence in intron 2 (Tarlow *et al*. 1993) (Fig. 2(c)). At least five alleles occur in most human populations, but the presence of four repeats (allele 1) and two repeats (allele 2) accounts for more than 95 per cent of individuals in all of the populations tested. The frequency of allele 2 (*IL1RN2*) is raised in a large number of inflammatory diseases including systemic lupus (Blakemore *et al*. 1994), ulcerative colitis (Mansfield *et al*. 1994), alopecia areata (Tarlow *et al*. 1994), lichen sclerosis (Clay *et al*. 1994), and some populations of patients with multiple sclerosis (Crusius *et al*. 1995).

A common theme in these studies has been that when patient populations were subdivided by clinical severity scores, the carriage rate of *IL1RN2* was greatest in the groups with the most severe clinical disease, and in some cases these individuals accounted for the disease association overall. This would suggest that the *IL1RN* genotype is a marker for the severity of inflammatory disease. There is some evidence that *IL1RN2* defines individuals with a lower rate of production of the IL-1 receptor antagonist, but this is not a consistent finding and cannot, at present, be accepted as the underlying biological mechanism of the association between the *IL1RN2* gene and disease.

From these and other studies it is becoming clearer that the genotype at the *IL1* region is likely to contribute to the risk of development of chronic diseases that have a prominent inflammatory component. Whether the specific DNA variants that are presently available for typing are themselves the source of differences of the production of IL-1 proteins or are physically linked to other functional polymorphisms in the same chromosomal region (not necessarily in a known *IL1* gene) remains an open question. However, logistic regression analysis indicates that where a composite *IL1* genotype is associated with disease (e.g. psoriasis), the individual markers within the genotype contribute independently of each other. This makes it unlikely that all the known disease associations with the *IL1*gene cluster occur because of a single, linked disease locus in the same region. It seems likely that composite genotypes of the *IL1* cluster and also of the *TNF* cluster will provide clinically useful, predictive information about disease phenotypes in a range of rheumatic diseases.

Cytokines are undoubtedly 'host defence' molecules and, like other molecules involved in innate or specific immunity, their genes show a relatively high degree of polymorphism. This is thought to be the result of selection of specific genotypes by the pressure of infectious diseases. Thus, while specific genotypes for cytokines confer fitness to survive certain infections, it would seem that they may also increase the risk of inappropriate inflammatory responses that can lead to chronic diseases.

Complement deficiency (reviewed in Lachmann and Walport 1987; Morgan and Walport 1991)

One of the clearest examples of an association between inherited variation in the function of host defences and autoimmune rheumatic disease is that between inherited complement deficiency and systemic lupus erythematosus. Only specific complement deficiencies have been associated with systemic lupus, particularly of the classical pathway proteins, C1q, C1r and C1s, C4, C2, and C3. Amongst these proteins there appears to be a hierarchy of association of both disease prevalence and severity with systemic lupus, according to the position of the protein in the activation pathway of complement (Fig. 3). C1q, C1r, and C1s (reviewed in Reid 1989; Bowness *et al*. 1994), and C4 deficiency (reviewed in Hauptmann *et al*. 1986) are associated with a very high prevalence of disease (more than 75 per cent), which tends to be severe. Deficiency of C2, the next protein in the pathway, is associated with a lower prevalence of disease, about 33 per cent, and disease tends to be of a similar severity to that seen amongst cohorts of patients with systemic lupus in the absence of complement deficiency (Ruddy 1986). C3 deficiency, reported in 24 individuals (Botto *et al*. 1993), has only been associated with lupus-like disease in three patients in two families. Such patients only account for a tiny minority of all patients with systemic lupus but may give a very important clue to one of the pathogenetic factors.

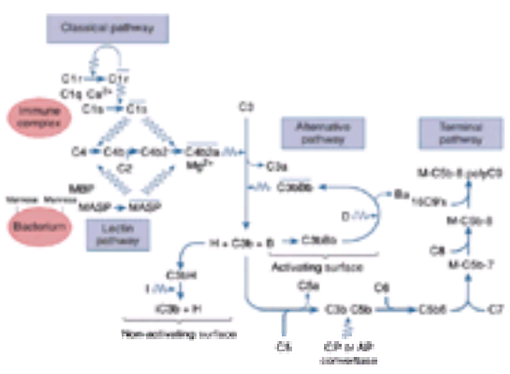


Fig. 3 Simplified diagram of the complement system. Complement may be activated by the classical pathway, the alternative pathway, or the newly characterized 'lectin' pathway, which is initiated by the binding of the collectin, mannose-binding protein, to bacterial surfaces rich in this carbohydrate. The 'amplification loop' of the alternative pathway on an activating surface can amplify C3 activation initiated by the other pathways, or by exogenous proteases from leucocytes, plasma (e.g. plasmin), or bacteria.

It may be that a more subtle inherited deficiency of complement is important as a disease susceptibility gene for the development of systemic lupus erythematosus in a majority of cases. Testing of this hypothesis followed two observations: (i) that there was an increased prevalence of the MHC antigens, HLA-DR3 and -DR2, amongst caucasoid patients with systemic lupus; and (ii) the discovery that C4A and C4B, the two isotypes of C4 encoded in tandem within the MHC, were both highly

polymorphic and that the polymorphism included null alleles (designated Q*0, quantity zero) at both loci. These findings led to examination of the possibility that the disease susceptibility locus within the MHC might be null alleles of one or both loci of C4.

Initial studies showed that there was indeed an increased prevalence of null alleles, mainly of C4A, amongst caucasoid patients with systemic lupus, but the analysis was confounded by strong linkage disequilibrium between C4A null alleles and HLA-DR3. This effectively prevented separation of the relative contribution of these two putative disease-susceptibility genes.

An approach that has been used in attempts to disentangle the relative contributions of HLA-DR3 and C4A Q*0 to disease susceptibility is the examination of the MHC associations with systemic lupus in different racial groups, in which haplotypes containing different combinations of the allotypic variants of the different MHC gene products are found. Using this approach, it has been found that the association of systemic lupus erythematosus with HLA-DR3 breaks down, but the association with C4A Q*0 is sustained. This finding is most striking in Japanese populations amongst whom HLA-DR3 is extremely rare.

The explanation for the association between complement deficiency and increased susceptibility to systemic lupus erythematosus is not certain. It has been known since work by Heidelberger in the 1940s that complement activation by immune complexes interferes with lattice formation and tends to maintain immune complexes in solution. When immune complexes are formed in the presence of complement, precipitation is inhibited, and this is particularly a property of the classical pathway of complement. In contrast, the alternative pathway of complement is implicated in solubilizing preformed immune complexes. The prevalence and severity of systemic lupus associated with inherited complement deficiency is highest for C1q, followed by C4 and then C2. This implies a physiological link between disease induction and an activity of the classical pathway of complement. Inhibition of immune precipitation may be the relevant activity and it has been proposed that complement deficiency leads to ineffective processing of immune complexes, followed by defective clearance of complexes by the mononuclear phagocytic system. This allows deposition of immune complexes in tissues causing inflammation. This inflammation leads in turn to release of autoantigens and the stimulation of a self-perpetuating autoantibody response. Some evidence in support of this hypothesis has come from studies of the mechanisms of immune-complex clearance in rabbits, monkeys, and man.

A second hypothesis associating complement deficiency with systemic lupus erythematosus flows from the role of the complement system in host defence against infectious disease. There is unequivocal evidence that deficiency of classical-pathway proteins is associated with increased susceptibility to infections by pyogenic bacteria, but no evidence for a role for such bacteria in the causation of systemic lupus. In contrast there is little evidence associating complement deficiency to increased susceptibility to virus infections. However, complement is effective in the lysis of certain retrovirally infected cells (though not those infected by the human immunodeficiency virus). There is enthusiasm for a possible role for retroviruses in the induction of systemic lupus, though little hard evidence. It may be postulated that complement deficiency could promote infection and the induction of systemic lupus by some, as yet unrecognized, viral agent.

What is the basis for tissue specificity in rheumatic disease?

The question of which factors determine the localization of inflammatory disease to joints remains unanswered. Possible explanations requiring consideration are that: (i) there are antigens present that are largely restricted to the complicated extracellular matrix of synovium, cartilage, and bone, which may behave as autoantigens or cross-react antigenically with foreign antigens; (ii) trapping of foreign antigens or infectious agents occurs as a result of an affinity to matrix or cellular components of joints or because of peculiarities of the vascular and lymphatic anatomy of joints; (iii) there are 'non-immunological' inflammatory responses to particular components of joints, for example the responses to calcium pyrophosphate and uric acid crystals; and (iv) the mechanical properties of joints, variable load bearing and repetitive movement, contribute to inflammation by mechanisms such as trauma and hypoxic reperfusion injury.

If the question 'why synovium' is difficult to answer, the answer to 'why particular synovial joints' is even more obscure, though there are a number of clues that might help. No very general explanation is likely to explain the distribution of inflammatory synovitis, since the pattern of joint involvement is so variable between different diseases. It is self-evident that joints are distinguished from many other structures in the body by their pattern of repetitive motion and structural loading. There are several pieces of evidence that mechanical factors play a part in joint inflammation: (i) rest is associated with amelioration of synovitis (Partridge and Duthie 1963); (ii) the presence of a hemiplegia may protect the paralysed side of the body from rheumatoid arthritis (Thompson and Bywaters 1962); (iii) mechanical damage to joints, for example cartilaginous injury in the knee, is associated with the onset later in life of osteoarthritis and crystal arthritis (Docherty *et al*. 1982).

The nature of the symmetrical involvement of joints in rheumatoid arthritis poses a further series of interesting scientific questions. One simple possibility is that symmetrically disposed joints share anatomical features that predispose to parallel mechanical injury or deposition of foreign antigen. A second intriguing possibility that is attracting much interest is that there may be reflex neuroendocrine stimuli that lead to symmetrical joint involvement.

What is special about inflammation within synovium?

The nature of the resident cells in an organ, the extracellular matrix, and its anatomical organization may determine the particular form of expression of an inflammatory response. The resident cells of joints include the cells comprising the synovium, chondrocytes, osteoblasts, and osteoclasts. The extracellular matrix is highly organized, comprising a network of collagens and proteoglycans.

The special effects of inflammation on the cells of any particular tissue depend, firstly, on the contribution of the resident cells themselves to the inflammatory process, and, secondly, on the consequences to that tissue of the loss of their specialized functions. In joints, inflammatory damage may lead to loss of lubrication, cartilage and bone, and secondary muscle wasting. The cells of the synovial membrane include monocytic phagocytes (A lining cells) and fibroblasts (B lining cells) (Henderson and Edwards 1987; Edwards 1994). Both types of cell are capable of elaborating cytokines such as IL-1 and TNF.

These factors amplify the inflammatory reaction by several mechanisms. They stimulate the local production of chemotactic peptides such as IL-8, and activate the endothelial lining cells of blood vessels—actions that increase the recruitment of blood leucocytes to the site of inflammation. Both IL-1 and TNF increase the activation of leucocytes (including an autocrine action on the macrophages responsible for their own production).

Such actions lead to production of proinflammatory prostanoids (prostaglandins and leukotrienes) through the mobilization of membrane phospholipids to generate arachidonic acid. Thus, indirectly, IL-1 and TNF contribute to local hyperaemia and increased vascular permeability leading to joint swelling. Through their induction of collagenases and proteoglycanases, and their effects on proteoglycan synthesis, the IL-1s have dramatic effects on the rate of turnover of extracellular matrix components. These actions are thought to be central to the destruction of cartilage and loss of bone density that accompany inflammation in a joint (Duff 1989a; Duff 1989b) (Fig. 4).

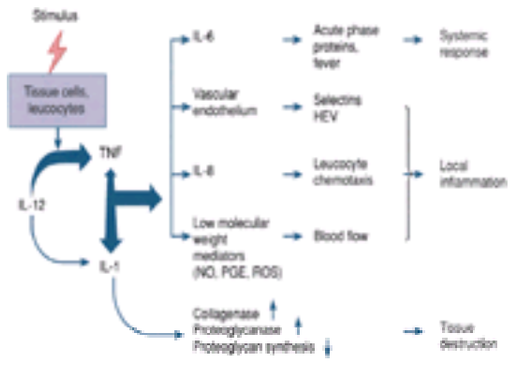


Fig. 4 Stimuli that can initiate inflammation include microbial products (e.g. bacterial lipopolysaccharides and exotoxins), irritant chemical agents (e.g. monosodium urate crystals), radiation (e.g. ultraviolet light on the skin), and endogenous molecules such as immune complexes and complement fragments. Polymorphonuclear and mononuclear phagocytes release IL-12 and this seems to be a very early event in the inflammatory response. Both IL-1b and TNF-a are inducible by IL-12 as well as being mutually inducible. In particular, TNF-a is a strong inducer of IL-1. These three cytokines trigger a proinflammatory cascade. By stimulating the production of other cytokines such as IL-6, a systemic acute-phase response ensues while local tissue inflammation progresses by the activation of vascular endothelium and the generation of high endothelial vessels (HEV). Selectin molecules expressed on endothelial cells mediate the early adherence of blood leucocytes, which then migrate into the surrounding tissue along a concentration gradient of chemokines (chemotactic cytokines, such as IL-8) and other chemotaxins towards the source of the inflammatory stimulus. Low molecular-weight mediators such as nitric oxide, prostaglandins, and reactive oxygen molecules (ROS) are also stimulated by the major

proinflammatory cytokines and cause vasodilation and local hyperaemia. Prolonged inflammation causes destruction of connective tissues, which can lead to resorption of nearby bone and cartilage. In synovial inflammation (for example in rheumatoid arthritis) IL-1b seems to be a particularly potent catabolic agent for connective tissues, inhibiting the synthesis of new extracellular matrix while accelerating the enzymatic destruction of existing components.

Many of the inflammatory diseases of synovium lead to synovial hyperplasia with pannus formation, and much of this hyperplasia appears to be due to immigration of cells of bone marrow origin with only a small contribution from division of resident cells (Henderson *et al*. 1988). The resident cells of cartilage and bone—chondrocytes, osteoblasts and osteoclasts—have the capacity both to synthesize and degrade their extracellular matrix. The balance between these activities may be tipped towards degradation by inflammatory stimuli.

One important characteristic of inflammation in synovial joints is the potential for the accumulation of inflammatory fluid, with consequences both for joint function and also as a modifier of the inflammatory response. Synovial cavities in their resting state contain only very small amounts of synovial fluid and are at atmospheric or subatmospheric pressure. The accumulation of synovial fluid following inflammation or injury is associated with a dramatic rise in intra-articular pressure, often seen clinically as the rupture of a Baker's cyst into the calf following flexion of the knee, a manoeuvre that causes a further rise in intra-articular pressure (Jayson and Dixon 1970). Such pressure changes are sufficient to overcome the capillary perfusion pressure and may render joints temporarily hypoxaemic, followed by a vascular hyperaemia on cessation of exercise (Blake *et al*. 1989). This may be associated with the generation of oxygen-derived free radicals, hypoxic reperfusion injury, and is discussed below in the context of injury caused by such radicals.

Angiogenesis (reviewed inRemmers *et al*. 1991; Folkman and Shing 1992)

An important feature of inflammatory arthritis is the formation of new blood vessels in the synovial membrane, which invade and destroy cartilage. Many peptides that can stimulate angiogenesis have been described. They include: (i) growth factors with activities on many cell types—including acidic and basic fibroblast growth factors, transforming growth factors-a and -b, and platelet-derived growth factors; (ii) growth factors with apparent specificity for endothelial cells—vascular endothelial growth factor, a family of related molecules; and (iii) cytokines and chemokines—including TNF-a and IL-8. There is evidence that several of these molecules collaborate to produce the angiogenesis that contributes to pannus formation. Both the resident synovial lining cells of the joint and invading inflammatory cells act as a source of angiogenic peptides (Fava *et al*. 1994; Koch *et al*. 1994). Angiogenesis is a potential therapeutic target and there is a report of an experimental inhibitor of angiogenesis suppressing collagen-induced arthritis in rats (Peacock *et al*. 1992).

Neuroendocrine influences

The endocrine, nervous, and immune systems are closely interwoven. Cells of each of these systems intercommunicate by the expression of molecules whose effects may be autocrine (i.e., on the same cell that secretes the molecule), paracrine (on other local cells) or endocrine (on other cells at a distance in the body). These molecules have been labelled as neurotransmitters in the nervous system, cytokines in the immune system, and hormones in the endocrine system. However, this terminology conceals the fact that there is enormous overlap in the activities of these cellular messengers. Lymphocytes bear receptors for 'neurotransmitters' and 'hormones'. Cytokines have wide-ranging effects within the nervous system. The chemical basis of these molecules is diverse. The recent discovery that nitric oxide, an inorganic gas and free radical, is a neurotransmitter and mediator of inflammation (reviewed in Moncada and Higgs 1993) has opened a whole new field of investigation. Dissection of the links between the nervous, endocrine, and immune systems in a meaningful fashion poses an enormous scientific challenge.

Endocrine influences (reviewed in Sternberg et al. 1992)

Three types of endocrine links with rheumatic disease may be distinguished, the first two physiological, the third pathological: (i) a role for sex hormones in explaining the predominantly female associations of most inflammatory rheumatic diseases (and also in the maintenance of bone density); (ii) the role of adrenal hormones as anti-inflammatory mediators; and (iii) specific associations between pathological endocrine over- or underactivity and rheumatic disease, for example hypothyroidism and Hoffman syndrome or acromegalic arthropathy. This third group of associations is not considered further in this chapter.

Sex hormones and arthritis

Sex steroids

There is a striking predominance of female sufferers from rheumatoid arthritis and systemic lupus erythematosus. The great majority of lupus patients are female, and symptoms tend to develop around puberty and improve after the menopause, which suggests an important role for sex hormones in the development of human systemic lupus. Furthermore, systemic lupus is sometimes exacerbated by oestrogen-containing oral contraceptives. Lahita and his colleagues (reviewed in Lahita 1993) have made a number of studies showing abnormalities of sex hormone metabolism in patients with systemic lupus. In particular they have observed that such patients are hyperoestrogenic and have decreased serum concentrations of testosterone. These findings are consistent with the observed effects of sex hormones on the lupus-like illness of (NZB × NZW) F_1 mice. Androgens protect these mice against disease whereas oestrogens accelerate its onset and severity. A small number of human kindreds have also been observed in which systemic lupus erythematosus is found in multiple male family members (Lahita *et al*. 1983). This may be the human counterpart of the BXSB mouse in which a Y-chromosome gene, whose effects are not mediated by sex hormone levels, confers disease susceptibility.

In rheumatoid arthritis, evidence for the involvement of sex hormones in explaining the enhanced predisposition of females to disease is: (i) an increased risk in nulliparous females; (ii) a protective effect of the oestrogen-containing oral contraceptive pill; (iii) improvement of established disease during pregnancy; and (iv) an increased susceptibility to the onset of disease during the first 3 months post partum (reviewed in Silman 1994).

Prolactin

A role has been postulated for prolactin in rheumatoid arthritis and other autoimmune diseases. Lymphocytes and monocytes bear high-affinity prolactin receptors (reviewed in Reber 1993). Hyperprolactinaemia in experimental animals is associated with worsening of systemic lupus (McMurray *et al*. 1994) and collagen-induced arthritis (Mattsson *et al*. 1992). There is evidence of excessive secretion of prolactin in patients with rheumatoid arthritis (Chikanza *et al*. 1993). Whether antagonism of prolactin will have any beneficial effects in rheumatoid arthritis in humans is not known.

Hypothalamic–pituitary–adrenal axis

The hypothalamic–pituitary axis plays a major part in the regulation of inflammation. A number of proinflammatory cytokines, produced at and released from sites of inflammation, such as TNF-a, IL-1b and IL-6, may stimulate, directly or indirectly, production of corticotrophin-releasing hormone by the hypothalamus. The corticotrophin-releasing hormone stimulates production of ACTH by the pituitary, which in turn stimulates production of glucocorticoids by the adrenal glands. Increased amounts of glucocorticoids down-regulate the inflammatory response. Defects in this homeostatic mechanism would be good candidates for increasing the severity of inflammatory responses.

There is substantial variability between different animal strains in their susceptibility to the development of experimental arthritis. The histocompatible Lewis and Fischer strains of rats show marked differences in susceptibility to the induction of several inflammatory diseases including adjuvant- and streptococcal cell wall-induced arthritis. Lewis rats are susceptible to disease induction; Fischer rats are resistant. The corticotrophin-releasing hormone response to stress is blunted in Lewis rats. A consequence of this is reduced glucocorticoid production following stress; that is, reduced expression of an endogenous anti-inflammatory molecule. Correction of this defect by infusion of corticosteroids can protect against the induction of inflammatory disease. In contrast, Fischer rats show a vigorous corticotrophin-releasing hormone response to stress. Antagonism of the corticotrophin-releasing hormone–ACTH–cortisol axis in Fischer rats by hypophysectomy or adrenalectomy increases their susceptibility to the induction of inflammatory disease.

There is preliminary evidence that the hypothalamic–pituitary axis may be defective in humans with rheumatoid arthritis. Although diurnal variation of cortisol in rheumatoid arthritis was within normal limits, it was inappropriately low for the amount of inflammation (Neeck *et al*. 1990). In keeping with this finding, the cortisol response to surgical stress was blunted in a group of patients with rheumatoid arthritis undergoing hip replacement compared with a group of osteoarthritic patients

undergoing the same operation (Chikanza *et al*. 1992).

The dominant effect on inflammation of corticotrophin-releasing hormone expressed in the hypothalamus appears to be inhibitory by downstream stimulation of glucocorticoid production by the adrenal glands. In contrast to its central action, peripheral expression of corticotrophin-releasing hormone is associated with proinflammatory effects (Karalis *et al*. 1991). Corticotrophin-releasing hormone can be produced by peripheral blood mononuclear cells and its expression is increased at sites of inflammation (Crofford *et al*. 1992; Crofford *et al*. 1993). Corticotrophin-releasing hormone receptors are present on a variety of cells of the immune system including lymphocytes and monocytes. In Lewis rats, which are susceptible to inflammatory disease and show reduced hypothalamic production of corticotrophin-releasing hormone, peripheral production was increased following stress and the hormone was expressed at sites of inflammation (Crofford *et al*. 1992).

Neural influences (reviewed in Basbaum and Levine 1991)

Synovium is richly innervated by neurones containing a large variety of neurotransmitters (reviewed in Kidd *et al*. 1990). There is a network of small, unmyelinated, afferent C-fibres containing many neuropeptides including substance P, neurokinin A, and calcitonin gene-related peptide. Efferent noradrenergic sympathetic fibres surrounding blood vessels contain neurotransmitters including neuropeptide Y. It was Bayliss in 1901 and Lewis in 1927 who observed that antidromic conduction in sensory neurones was responsible for the 'triple response' (redness, flare, and weal) seen after mechanical, thermal, or chemical injury to the skin.

This response has been shown to be largely mediated by the release of substance P, which causes vasodilatation and increased vascular permeability (reviewed in Holzer 1988; Garrett *et al*. 1992). An analogous response was described in synovium after antidromic stimulation of C-fibres, causing increased vascular permeability and exudation of protein across the synovial membrane.

Neuropeptides are analogous to cytokines in producing pleiotropic effects, and these are shown for substance P in Table 1. A teleological analysis of the inflammatory response provoked by nociceptive stimulation of C-fibres would argue that this would promote resolution of the injury responsible for the nociceptive stimulus. Inability, or failure, to remove the injurious stimulus could render the neural response maladaptive and augment chronic inflammation. A further neural consequence of injury to a joint is reflex inhibition of the surrounding musculature. This phenomenon promotes the periarticular muscular wasting that rapidly develops around an inflamed joint, an adaptive response promoting healing after acute injury, but maladaptive in the context of chronic inflammatory arthritis.

Target cell	Effect of substance P
Synoviocyte	Prostaglandin E_2 synthesis
	Collagenase synthesis
	Cell division
Neutrophil	Promotion of chemotaxis
	Degranulation
	Promotion of phagocytosis
Mast cell	Degranulation
Macrophage	Thromboxane-B_2 synthesis
T lymphocyte	Proliferation

Table 1 The pleiotropic effects of substance P

In addition to local neurogenic effects, the systemic level of activation of the sympathetic nervous system may influence the severity of arthritis (reviewed in Basbaum and Levine 1991 and Kidd *et al*. 1992. The spontaneously hypertensive rat has a high basal sympathetic drive. Experimental adjuvant arthritis in this strain was worse than in animals with normal sympathetic tone. As a corollary to these observations, adrenal medullectomy reduced the severity of adjuvant arthritis, with infusion of adrenaline restoring the normal severity of induced disease. These findings could help to explain a role for psychological and other stresses that cause an increase in sympathetic tone, inducing flares of inflammatory disease.

The field of neurogenic inflammation is still in its infancy despite the antiquity, in scientific terms, of the first observations. A reflex neurogenic hypothesis is attractive to explain the symmetry of joint involvement in rheumatoid arthritis. However, any hypothesis to explain symmetry in rheumatoid arthritis would have to cope with the asymmetry seen in other inflammatory arthritides such as reactive or psoriatic arthritis.

Nitric oxide (reviewed in Moncada and Higgs 1993; Stefanovic-Racic et al. 1993)

Nitric oxide was first identified as the molecular basis of endothelial-derived relaxing factor, a vasodilator synthesized by endothelial cells. It is a gas and a short-lived free radical, and has an enormous range of activities that may promote or inhibit inflammation under different circumstances. There are two classes of enzyme responsible for the synthesis of nitric oxide from the amino acid arginine, constitutive and inducible nitric oxide synthases. The latter enzyme is induced by several proinflammatory cytokines and is responsible for the local production of nitric oxide at sites of inflammation.

It has recently been discovered that cyclo-oxygenase also exists in constitutive and cytokine-inducible isoforms (**COX-1** and **COX-2**), see below. The discovery of cytokine-inducible enzymes that catalyse the formation of powerful mediators of inflammation adds to the pleiotropic mechanisms by which cytokines exert their effects on inflammation. To complicate matters further, it turns out that nitric oxide is an inducer of COX-2 (Salvemini *et al*. 1993), and the regulation of inducible nitric oxide synthetase and COX-2 at sites of inflammation is extremely complex (Vane *et al*. 1994).

There is good evidence for increased expression of nitric oxide in humans with inflammatory arthritis (Farrell *et al*. 1992; Kaure and Halliwell, 1994). Adjuvant arthritis in rats (Ialenti *et al*. 1993; Stefanovic-Racic *et al*. 1994) and arthritis and glomerulonephritis in MRL-lpr/lpr mice (Weinberg *et al*. 1994) could be suppressed by inhibition of nitric oxide synthesis using analogues of arginine. Nitric oxide production by osteoblasts was stimulated by combinations of cytokines and could be involved in the osteoporosis associated with inflammatory rheumatic disease (Ralston *et al*. 1994).

Humoral inflammation and rheumatic disease

There are several triggered enzyme cascades that play a major role in inflammation—the complement, coagulation and fibrinolysis, and kinin (contact activation) systems. These share many features in common and have interdependent activation pathways. Analysis of the activities of the triggered enzyme cascades can be subdivided into several major categories: (i) recognition of foreign molecules or of altered self; (ii) amplification of the recognition signal; (iii) localization of the amplified response; (iv) effector phase (Table 2); and (v) regulation. A detailed consideration of the physiology of each of these pathways is beyond the scope of this chapter and the reader is referred to the following reviews: complement (Walport and Lachmann 1993); coagulation and fibrinolysis (Hart and Fritzler 1989); contact activation (Bhoola *et al*. 1992).

System	Effect function
Complement	Opsonization
	Chemotaxis
	Lysis
	Priming and activation of leucocytes
	Anaphylatoxic activity
	Stimulation of lymphocytes
	Immune-complex lattice modification
Coagulation	Thrombosis
	Wound healing
	Cell proliferation
	Platelet activation
	Activation of fibrinolysis
Fibrinolysis	Fibrinolysis
	Complement activation
	Activation of extracellular proteinases
Kinin	Increased vascular permeability
	Pain production
	Smooth muscle contraction
	Complement activation
	Activation of fibrinolysis
	Stimulation of endothelium
	Activation of extracellular proteinases

Table 2 Humoral pathways of inflammation

The principles of inflammatory damage mediated by a triggered enzyme cascade can be illustrated briefly by consideration of the complement system. The effects of activation of this pathway result from ligation of receptors for split products of complement proteins on leucocytes and from insertion of the membrane-attack complex into cell membranes. Activation and cleavage of C3, C4, and C5 result in the production of anaphylatoxins, C3a and C5a, and of opsonins, C4b, C3b, iC3b, and C3dg (Fig. 3). Receptors for these products are present on most leucocytes. Ligation of anaphylatoxin receptors results in chemotaxis and cellular activation, including respiratory burst and degranulation; ligation of opsonic receptors leads to phagocytosis and cellular activation (also including respiratory burst and degranulation). In recent years it has been appreciated that insertion of the membrane-attack complex into nucleated cell membranes is frequently not accompanied by lysis but instead causes cellular activation, which includes mobilization of arachidonic acid to produce prostaglandins and leukotrienes, and stimulation of the release of cytokines.

There is good evidence implicating each of the triggered enzyme cascades as a contributor to synovial inflammatory injury. Fibrin is a prominent constituent of inflamed joints and the influx of monocytes into rheumatoid pannus, with their capacity to produce procoagulant, may be one important mechanism for activation of the coagulation system. Similarly, complement activation products, active tissue and plasma kallikrein, and evidence of fibrinolytic activity can all be assayed in effusions from inflamed joints. There is no definite evidence to allow assertion of the primacy of one triggered enzyme cascade over another in inducing inflammatory damage in joints. Although there is evidence for the presence of kinins in inflamed synovium, a patient with rheumatoid arthritis and Hageman factor (factor XII) deficiency has been described, a deficiency associated with a severely reduced capacity to produce kinins (Donaldson *et al*. 1972). Similarly, inherited complement deficiency is a strong predisposing factor for the development of systemic lupus erythematosus (Morgan and Walport 1991), although there is abundant evidence of a role for complement in producing inflammatory damage in systemic lupus.

Cellular inflammation and rheumatic disease

Endothelium (reviewed in Pober and Cotran 1990)

In the past it would have been considered eccentric to consider the endothelium first in a section on cellular contributors to inflammation. However, endothelium represents the cellular barrier between the blood circulation and the synovium, and has an extremely important role in controlling the delivery of blood to sites of inflammation and in regulating the traffic of cells and plasma proteins from the circulation (see Chapter 3.2 for a fuller discussion). The emigration of lymphocytes and leucocytes from blood vessels involves adhesion to endothelium followed by transmigration. Activation of both leucocytes and endothelium leads to the expression of adhesion molecules, which allow this process to occur. The importance of the adhesive process is illustrated by the consequences of hereditary deficiency of one of the families of adhesive molecules, the b_2-integrins, which leads to a severe immunodeficiency associated *inter alia* with failure of neutrophil adhesion and emigration from the circulation to sites of bacterial infection.

A second role for endothelium is in the regulation of blood flow, which can be both increased and decreased by molecules synthesized by endothelium. Tissue oxygenation, the immigration of inflammatory cells, and the exudation of plasma proteins are promoted by mediators, such as prostacyclin and endothelial-derived relaxing factor (which is nitric oxide), that increase local blood flow, and thrombin and histamine that cause shape changes in endothelial cells. Thrombosis, leading to the isolation of an inflammatory site, is mediated by factors, such as endothelin and platelet-derived growth factor, that cause contraction of vascular smooth muscle and by tissue factor expression, activating the triggered enzyme cascade of thrombosis, and leading to thrombin generation and fibrinogen cleavage.

Granulocytes (reviewed in Weiss 1989)

Neutrophils and their specialized relatives, eosinophils (reviewed in Gleich and Adolphson 1986; Spry 1988), basophils and mast cells (Crisp *et al*. 1984), collectively play a vital part in host defence against bacteria and parasites, and provide an important mechanism for the removal of autologous and exogenous debris from sites of inflammation. Their vital role in host defence is illustrated by the consequences of deficiencies of neutrophil activity. The resulting immunodeficiency state is characterized by recurrent infections with bacteria that are usually associated with the formation of pus, i.e., pyogenic bacteria, and these include Gram-positive cocci such as staphylococci and streptococci, and Gram-negative rods such as *Escherichia coli, Klebsiella*, and *Serratia*. Patients with neutropenia are also at risk of opportunistic infections with the fungal *Candida* and *Aspergillus* spp.

The interaction of neutrophils with the external milieu is mediated through the binding of extracellular ligands to specific cell-surface receptors. Neutrophil receptors may be classified into five major classes, according to their ligands and the nature of the evoked response (Table 3). Ligation of the majority of these receptors leads to priming and expression of neutrophil effector responses. Multivalent ligation and ligation of two different types of neutrophil surface receptors are often needed to evoke the full expression of neutrophil responses—respiratory burst and granule release. This can be viewed as a protective response to prevent inappropriate cell triggering.

Receptor	Ligand
Growth and differentiation	
Colony-stimulating factors (CSF)	Granulocyte–macrophage-CSF, granulocyte-CSF
Chemotaxis, priming, and activation	
C5a receptor	C5a
f-Met-Leu-Phe	Bacterial formylated
Leukotriene B_4	Leukotriene B_4
Cytokine receptors	Interleukin-1, interferon-γ, tumour necrosis factor-α and -β
Opsonic	
Fc receptors	Complexed IgG
CR1 and CR3	C3b, iC3b
Adhesion	
CR3	?
Laminin	Laminin

Table 3 Receptors on neutrophils

The cytoplasmic granules consist of an outer membrane surrounding the contents of densely packed proteins assembled in a mucopolysaccharide matrix. The granules are heterogeneous, with at least two, and probably several more, different types in neutrophils. They are packed with enzymes including elastase, cathepsin G, myeloperoxidase, collagenase, gelatinase and lysozyme, whose activities on the one hand directly promote bacterial killing by digestion of their cell walls, and on the other allow free migration of neutrophils outside blood vessels by digesting extracellular matrix proteins, leading to the formation of abscess cavities. Granules also contain peptides directly toxic to bacteria and parasites, such as eosinophil cationic protein (only eosinophils), defensins, and bactericidal/permeability-increasing protein; and molecules such as lactoferrin and vitamin B_{12}-binding protein that may act to starve bacteria of essential nutrients. The granules also act as a store of membrane receptors, such as the complement receptors, CR1 and CR3, whose expression can be rapidly up-regulated following neutrophil stimulation.

The 'professional' phagocytic cells, neutrophils, monocytes, macrophages and eosinophils, demonstrate an unusual process when they engulf microbes. They rapidly consume a large amount of oxygen. This 'extra respiration of phagocytosis' is not used for the normal purpose of generating energy by mitochondrial oxidative phosphorylation, for which the bulk of our inhaled oxygen is used. The oxygen consumption of the respiratory burst is required to produce the optimal conditions for the killing of most common bacterial and fungal pathogens as well as a variety of commensals. Cells deprived of oxygen engulf but fail to kill some microbes efficiently. The same is seen in the syndrome of chronic granulomatous disease, the hallmark of which is complete absence of this respiratory burst by cells that appear to be functionally normal in all other respects. There are three ways in which the respiratory burst may produce bacterial killing—free-radical production, myeloperoxidase-catalysed halogenation, and alkalinization of the phagocytic vacuole.

Apparently, whilst most of the receptors, the granule contents, and the respiratory burst of granulocytes are adapted to the recognition and destruction of microbes,

they also carry the potential to damage the host if triggered inappropriately or chronically. Most of the stimuli to chronic inflammation, outlined above, cause neutrophil migration and activation. Neutrophils are usually present at sites of rheumatic inflammation and their numbers often correlate with the severity of disease.

The formidable array of toxic contents of neutrophils, coupled with their short half-life of approx. 7 h and correspondingly high rate of production of about 1.63×10^{9} cells/kg per day, poses the question as to what happens to effete cells, especially at sites of inflammation? The traditional view was that neutrophils at sites of inflammation disintegrated, contributing all of their toxic contents to exacerbate the inflammatory process. Recent studies have shown that this view is largely erroneous and that neutrophils undergo a form of programmed cell death, known as apoptosis, in which the nucleus condenses and chromatin is cleaved in an ordered fashion. Apoptotic neutrophils display a change in surface phenotype that renders them susceptible to recognition and uptake by mononuclear phagocytic cells, where their contents may be harmlessly degraded. This mechanism of neutrophil death appears to operate within inflamed synovial cavities where monocytes that have phagocytosed neutrophils have been described for many years as Reiter's cells (Savill *et al*. 1989). Discovery of the physiological mechanism of disposal of granulocytes raises further hypotheses about mechanisms of inflammatory tissue damage. It is possible that defects in the process of apoptosis and the engulfing of apoptotic cells may contribute to inflammatory processes by increasing the available time for neutrophil-mediated injury or even, in some circumstances, allowing neutrophil disintegration.

Monocytes and macrophages

Many of the activities of monocytes and macrophages overlap with those of granulocytes. However, monocytes and macrophages are much longer-lived cells and they display a complex range of phenotypes depending on many factors, including their location, matrix attachment, and the soluble mediators to which they are exposed.

Recruitment of macrophages to sites of inflammation is a slower process than the accumulation of granulocytes, reflecting a smaller circulating precursor pool. However, as previously noted, the synovium also contains resident macrophages as type-A lining cells. The activation of monocytes/macrophages is complex and involves the induction of many different genes. The T-cell lymphokine, interferon-g, is a potent macrophage activator and differentiator. It 'primes' the macrophage to respond to other activating agents such as bacterial lipopolysaccharide and particles opsonized with IgG or C3, probably by increasing the number of surface receptors for these agents. Macrophage activation is associated with increased metabolic activity generating reactive oxygen radicals and an increase in phagocytic and cytolytic capacity. A range of proinflammatory cytokine genes is expressed, including those for IL-1a, IL-1b, TNF-a, IL-6, IL-8, and interferon-a. Activated macrophages release components of the complement pathway such as C2 and factor B, and, under the influence of interferons, they display on their surface increased numbers of class I and class II (interferon-g) MHC molecules.

Fibroblasts

The normal synovium contains mesodermal fibroblasts (type-B lining cells) and fibroblast-like cells are often prominent in the proliferating synovium during chronic synovitis. The extent to which these are derived from resident synovial cells or from adjacent connective tissue is unknown.

It seems likely that local signals for fibroblast proliferation would include messengers such as the fibroblast growth factors, platelet-derived growth factor, and transforming growth factor-b derived from platelets, macrophages, and other infiltrating inflammatory cells. Fibroblasts, themselves, can release proinflammatory prostanoids and cytokines such as IL-1, IL-6 and interferon-b (Meager 1990).

How do cells cause injury in joints? (see also Chapter 2.5)

Free radicals (reviewed in Merry *et al*. 1989; Greenwald 1991)

Oxygen-derived free radicals [including superoxide (O^{2-}), hydroxyl (OH), and hydrogen peroxide (H_2O_2)] may play direct and indirect roles in promoting inflammatory tissue injury. Their existence in biological fluids is transient and their activity in inflammation is largely inferred indirectly by the finding of evidence of molecular injury compatible with attack by free radicals. A low viscosity is an important characteristic that distinguishes an 'inflammatory' from a 'non-inflammatory' synovial effusion. The high viscosity of normal synovial fluid is due to hyaluronic acid, and degradation of the polymeric structure of hyaluronic acid is a characteristic of 'inflammatory' effusions. Both free radicals and the enzyme hyaluronidase can break down hyaluronic acid; however, there is no evidence for the existence of hyaluronidase in joint effusions *in vivo* and free radicals are the strongest candidates as mediators of the extracellular degradation of hyaluronic acid (McCord 1974; Greenwald and Moy 1980). There is similar evidence for a role of free radicals in causing lipid peroxidation, immunoglobulin aggregation, and collagen degradation.

Free radicals may contribute to inflammatory change in two important indirect ways. Neutrophil elastase is a broad-spectrum serine esterase enzyme capable of cleaving a range of molecules including elastin. However, it is effectively inactivated by a $_1$-proteinase inhibitor (previously known as a $_1$-antitrypsin). It has recently been appreciated that a methionine residue at the active centre of a $_1$-proteinase inhibitor is susceptible to oxidation by free radicals, leading to inactivation of enzyme inhibitory activity. Neutrophils may create a local 'zone of activity' for elastin by inactivation of a $_1$-proteinase inhibitor by oxygen-derived free radicals. A second way in which oxygen-derived free radicals may promote phagocyte-mediated injury is via the activation of latent metalloproteinase enzymes such as collagenase.

What is the source of free radicals in joints? The main cellular source in synovial effusions is phagocytic cells, both resident and immigrant. Stimulation of phagocytes, especially those partially activated by adherence to matrix proteins, by ligation of receptors, including those for Fc, C3b and iC3b, f-Met-Leu-Phe, and anaphylatoxins, may individually, or in combination, trigger oxidative metabolism and the production of a transient burst of free radicals. The chemistry of free radicals is complex and the chemical route of generation of biologically relevant oxygen radicals is not established beyond doubt. The chemical microenvironment within the synovium may be important. In particular, ferric iron (Fe^{2+}) may have a physiological role in catalysing the production of hydroxyl radicals from hydrogen peroxide, and supportive evidence for a role of iron in promoting inflammatory synovitis comes from observations of: (i) a protective role for iron chelators in animal models of inflammation (Andrews *et al*. 1987), though there are no similar data as yet from studies in humans; (ii) a correlation between the presence of free iron and rheumatoid synovial inflammation (Blake *et al*. 1984); and (iii) a flare of synovitis in a patient given an infusion of iron (Winyard *et al*. 1987).

A second potential source for oxygen-derived free radicals is from hypoxic reperfusion injury (McCord 1985). It was observed by Blake and his collaborators (Blake *et al*. 1989) that the increase in intrasynovial pressure related to exercise of knee joints containing effusions was associated with a reduction of synovial capillary blood flow, followed by reactive hyperaemia on cessation of exercise. A small increase was measured, after exercise, in oxidatively damaged lipids and IgG. The intracellular enzyme xanthine oxidase/dehydrogenase plays a major part in the generation of free radicals and this raises the speculation that allopurinol may be useful for more than the inhibition of uric acid production, a hypothesis that awaits full exploration (Puig *et al*. 1989). It is unclear to what extent this mechanism of generating oxygen radicals contributes to inflammation in joints, or whether it helps to explain the observation that inflammation tends to get better in a rested joint.

Enzymes

Tissue morphogenesis requires organized synthesis of complex matrices that are found in bone and cartilage. Growth of tissues and their remodelling after injury involves breakdown and dissociation of matrix proteins as well as synthesis of new protein. The tissues of joints and bones, therefore, are endowed with enzymes with the capacity to break down collagen as well as synthesize new collagen. Bone integrity is regulated by a delicate balance between osteoblasts and osteoclasts. Apparently any dysregulation of the balance between synthesis and breakdown of extracellular matrix may result in an increase or decrease of tissue mass. Osteoporosis following the menopause or glucocorticoid therapy is an example of such an imbalance leading to diffuse loss of bone.

Metalloproteinases and their inhibitors (reviewed in Weiss and Peppin 1986; Docherty and Murphy 1990; Sorsa et al. 1992)

Collagen, a major component of many tissues including joints, is resistant to the activity of most proteolytic enzymes because of its unusual amino acid sequence and triple helical structure, but sensitive to digestion by members of a family of matrix metalloproteinase enzymes, the collagenases and gelatinases, whose substrates are native and denatured collagen, respectively. A third member of the family, the stromelysins, cleaves proteoglycan core protein, laminin, fibronectin, and some collagen types. These enzymes are secreted in latent form by many cell types, including fibroblasts, bone cells, chondrocytes, macrophages, endothelium, and neutrophils. Their physiological activity is regulated in two ways, by activation from a latent to an active form and by inhibitors in the fluid phase, such as the tissue inhibitors of metalloproteinases, which prevent the geographical spread of enzyme activity and are synthesized and secreted by a similar range of cells.

By analogy with the other mediators of inflammation, the metalloproteinases may have deleterious effects if they are inappropriately secreted or activated, or

ineffectively inhibited. The chronic recruitment of inflammatory cells to joints, including fibroblasts, monocytes, and neutrophils, may lead to an increased tissue mass capable of secreting collagenases. Cytokine stimulation of each of these cell types promotes release of metalloproteinases. Oxidative metabolism and the activity of other protease enzymes may activate metalloproteinases, which in turn leads to collagen degradation and tissue injury.

A potential therapeutic role in inflammatory arthritis for inhibitors of matrix metalloproteinases has been proposed (reviewed in Vincenti *et al*. 1994). Experimental inhibitors of matrix metalloproteinases have proved effective in animal models of arthritis. However, it was found that a number of them inhibited the proteolytic processing of pro-TNF-a to mature, secreted TNF-a, a step presumably mediated by a zinc-dependent neutral protease (Gearing *et al*. 1994; McGeehan *et al*. 1994; Mohler *et al*. 1994). It has recently been found that one of the soluble TNF receptors, a fluid-phase inhibitor of TNF-a, is also cleaved from cell surfaces by a similar or identical metalloproteinase enzyme (Crow *et al*. 1995).

Other neutral proteinases

Many other proteinase enzymes are potential mediators of injury to joints, for example the neutral proteinases of neutrophils, cathepsin G and elastase. Elastase complexed with a_1-proteinase inhibitor has been identified in synovial effusions, providing evidence for the release of this enzyme from neutrophils. There is also evidence of oxidative injury to a_1-proteinase, which might prevent its efficient inhibition of neutrophil elastase. However, interesting experiments in beige mice genetically deficient in neutrophil elastase and cathepsin G did not show any reduction in the severity of the synovitis associated with antigen (bovine serum albumin)-induced arthritis (Pettipher *et al*. 1990). Whilst this experiment does not exclude a role for these enzymes in mediating tissue injury in synovial inflammation, it once again emphasizes the enormous redundancy of the mechanisms that potentially cause damage to joints.

Prostaglandins and leukotrienes (reviewed in Goetzl *et al*. 1995)

An important role for eicosanoids in the mediation of inflammatory damage in rheumatic disease is suggested by the clinical effectiveness of inhibitors of cyclo-oxygenase. Many of the cells present in inflamed joints mobilize membrane arachidonic acid and release the potent mediators, prostaglandin E_2, thromboxane A_2, prostacyclin, and leukotriene B_4 (Fig. 5). These mediators are particularly important in the generation of acute inflammation but some, such as prostaglandin I_2 (prostacyclin) and prostaglandin E_2, can also have suppressive effects on inflammation.

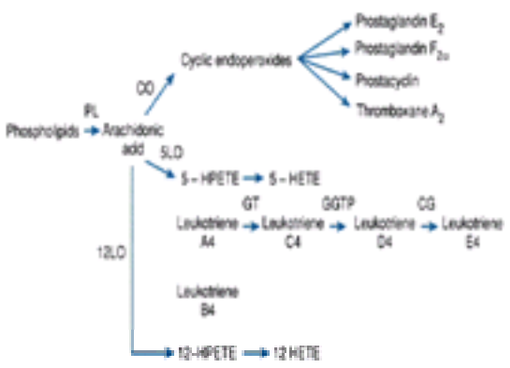


Fig. 5 Arachidonic acid metabolism: PL, phospholipases; CO, cyclo-oxygenase; 5LO, 5-lipoxygenase; 12LO, 12-lipoxygenase; H, hydrolase; GT, glutathione transferase; GGTP, g-glutamyl transpeptidase, CG, cysteinyl-glycinase.

Activation of phospholipase A_2 is an early event in the control of prostanoid production. This enzyme (stimulated by IL-1, TNF, and other factors) releases arachidonic acid from membrane phospholipids. Certain glucocorticoid-induced proteins (lipocortins) are thought to inhibit phospholipase A_2 as part of the anti-inflammatory action of steroids. Arachidonic acid is then metabolized further by one of the two major enzyme pathways, cyclo-oxygenase or lipoxygenase. The activity of the former results in prostaglandins while the latter produces the leukotrienes. Cyclo-oxygenase is a ubiquitous enzyme pathway in mammals but lipoxygenase activity is more or less restricted to monocytes, macrophages, neutrophils, eosinophils, and mast cells. Arachidonic acid is not the only substrate for these enzymes, and, for example, eicosapentaenoic acid can enter the lipoxygenase pathway, resulting in the production of pentaenoic leukotrienes, which are biologically less potent. This is probably the biochemical basis of the reported anti-inflammatory effects of diets rich in eicosapentaenoic acid (e.g. fish oil diet) (see also Chapter 1.3.6).

The proinflammatory effects of prostaglandins include vasodilation of the microcirculation (especially prostaglandins E_2 and I_2), amplification of oedema caused by other mediators (e.g. bradykinin, C5a), and increased nociception by sensitization of pain receptors. Leukotrienes can induce vascular permeability directly and leukotriene B_4 acts as a chemotaxin for neutrophils and increases their adhesion to endothelial cells. It has recently been found that a selective antagonist of the leukotriene-B_4 receptor inhibited the inflammation of collagen arthritis in mice (Griffiths *et al*. 1995). As noted above, prostaglandins have certain actions that may suppress chronic inflammation. For example, some leucocyte (macrophage and T lymphocyte) functions can be inhibited by prostaglandin E_2 and there are reports that prostaglandin E can inhibit cartilage degradation. Many of these effects may result from the inhibition, by prostaglandin (E_2), of the synthesis of IL-1 and TNF (Pettifer *et al*. 1992).

It has recently been appreciated that, as outlined above, there are two forms of cyclo-oxygenase: (i) COX-1, constitutively expressed in many tissues; and (ii) COX-2, inducible by proinflammatory cytokines and expressed relatively selectively at sites of inflammation, including rheumatoid synovium (Crofford *et al*. 1994). This discovery has substantial therapeutic implications, as selective inhibitors of COX-2 may have anti-inflammatory effects without producing gastrointestinal side-effects (Mitchell *et al*. 1993; Masferrer *et al*. 1994).

Cytokines (reviewed in Arai et al. 1990)

Cytokines can be defined as extracellular peptide mediators that regulate cellular growth differentiation and activation through specific receptor molecules on target cells. Within the immune system, cytokines that are produced by leucocytes and are active in inflammation or immune responses are designated as interleukins. There are now 17 interleukins, some of the features of which are shown in the Table 4, Table 5, Table 6, Table 7, Table 8, and Table 9. A related cytokine, TNF, is not classified as an interleukin for historical reasons but displays a very similar range of biological activities (Table 10) and has also been implicated in inflammatory diseases.

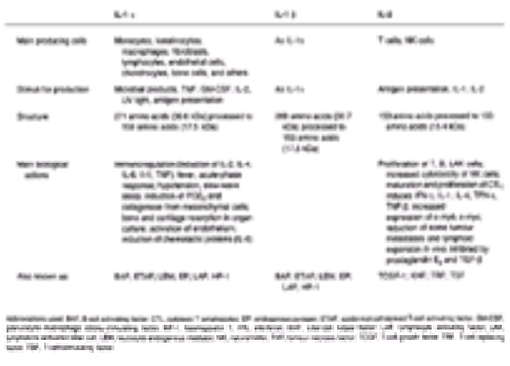


Table 4 Human interleukins (IL)

	IL-1	IL-4	IL-6
Main producing cells	Activated T cells	Activated T cells (T_H2)	Activated T cells (T_H2)
Stimulus for production	Antigen presentation	Antigen presentation	Antigen presentation
Structure	153 amino acids (14.6 kDa)	129 amino acids (15 kDa)	185 amino acids (153.2 kDa)
Main biological actions	Growth stimuli of all leucocyte lineages	T cells: TCGF activity (+++); thymocyte proliferation (antigen); $\downarrow$Ant activity B cells: BCGF (antigen) (+++); MHC class II (++); IgG, enhancer (+++); IgE enhancer (+++); IFN-γ inhibitor; (+++); IL-4R enhancer (+++) (all blocked by IFN-γ) Synergy with G-CSF, M-CSF, EPO, mast cell growth factor	Eosinophils, maturation, growth Synergy with IL-4 for IgG, IL-2 for CTL, genes and IL-2R
Also known as	Mob-CSF, MCSF-1 Pgl-CSF	TCGF-II, BCGF-I, BSF-1, BCDF-1, BCDF-γ, MCGF-II	BNF, BCDF-II, BCDF-α, EDF, HPF, BCDF

Abbreviations used: BCDF, B-cell differentiation factor; BCGF, B-cell growth factor; BCL, B-cell lymphoma; BSF, B-cell stimulating factor; CSF, colony-stimulating factor; CTL, cytotoxic T lymphocyte; EDF, eosinophil differentiation factor; EPO, erythropoietin; G-CSF, granulocyte colony-stimulating factor; IFN, interferon; IGIF, interleukin factor; MCGF, mast-cell growth factor; M-CSF, macrophage colony-stimulating factor; TCGF, T-cell growth factor; TNF, T-cell replacing factor

Table 5 Human interleukins (IL)

	IL-4	IL-7
Main producing cells	Monocytes, macrophages, activated T cells, glioblastoma cells, fibroblasts, cardiac myxoma cells, cervical adenocarcinoma lines	Stromal adherent cell from bone marrow Physiological stimuli unknown
Stimuli for production	IL-1, TNF, PDGF, poly I:C autocrine in some lines	LPS is active
Structure	212 amino acids processed to 184 amino acids (20.7 kDa)	177 amino acids processed to 152 amino acids (17.4 kDa)
Main biological actions	Proliferation: plasmacytoma lines, hybridoma lines, haemopoietic cells (IL-3 and GM-CSF-like action) EBV-transformed cells, thymocytes (antigen) T cells: reduce IL-2R; with IL-2: CTL induction B cells: induction of immunoglobulin production (differentiation without growth) Antiviral: CTL induction of acute-phase proteins in liver cells, former in vivo	Long-term growth of pre-B cells and pro-B cells Mature T-cell activation in presence of lectin Early T-cell activation Induction of IL-2/IL-2R on some T-cell subsets
Also known as	BCDF, BSF-2, IFN-β_2, HPGF, HSF, CDF, LP-1	

Abbreviations used: BCDF, B-cell differentiation factor; BSF-2, B-cell stimulatory factor; CDF, cytotoxic cell-differentiation factor; CTL, cytotoxic T lymphocyte; EBV, Epstein-Barr virus; GM-CSF, granulocyte–macrophage colony-stimulating factor; HPGF, hybridoma/plasmacytoma growth factor; HSF, hepatocyte-stimulating factor; IFN, interferon; LP, lymphopenin; LPS, lipopolysaccharide; PDGF, platelet-derived growth factor; TNF, tumour necrosis factor.

Table 6 Human interleukins (IL)

	IL-4	IL-6	IL-10
Main producing cells	Monocytes, T cells (?), endothelial cells, dermal, fibroblast, epithelial lines, Kupffer cells	T cells	T_H2 cell lines, mast cell lines; B lymphomas
Stimuli for production	LPS, IL-1β, IL-1α TNF-α, IL-3 GM-CSF, ?lymphocyte activation	Antigen presentation	Antigen presentation
Structure	72 amino acids (8 kDa)	184 amino acid propeptide cleaved to 139 amino acids (19.1 kDa)	178 amino acids (18.7 kDa)
Main biological actions	Neutrophil activation: respiratory burst; release lysosomal enzymes Neutrophil chemotaxis: in vitro and in vivo T-cell and macrophage chemotaxis	Synergy with IL-3 on mast cell proliferation and activity Proliferation of megakaryoblastic leukaemia cells Enhancement of survival of T helper cell lines Proliferation of fetal thymocytes	Inhibits production of IFN-γ and IL-2; suppressor factor for T_H1 response
Also known as	MONCF, NAF, MDNAP, LYAAP, GCP, MIP-1	PAS, MFA, TCGF-II	CSIF

Abbreviations used: CSIF, cytokine synthesis inhibitory factor; ED, N-acetyladibin protein; EPO, erythropoietin; GCP, granulocyte chemotactic peptide; GM-CSF, granulocyte–macrophage colony-stimulating factor; MFA, mast-cell enhancing activity; MIP, macrophage-derived inflammatory protein; MONAP, monocyte-derived neutrophil activating peptide; NAF, neutrophil-activating factor; PBP, platelet-basic protein; TCGF, T-cell growth factor

Table 7 Human interleukins (IL)

	IL-11	IL-12	IL-13
Main producing cells	Bone marrow, stromal fibroblasts, trophoblasts	Monocytes/macrophages, B cells	Activated CD4 T cells
Structure	178 amino acids (23 kDa) monomer	219 amino acids (35 kDa) 328 amino acids (40 kDa) heterodimer	132 amino acids (10 kDa) monomer
Main biological actions	Synergism with IL-3 to expand megakaryocyte colonies Hepatic APP synthesis	Induction of T_H1 responses in TH0 cells IFN-γ induction in NK cells	B-cell proliferation and antibody class-switching to IgE Inhibition of proinflammatory cytokine production in macrophages/monocytes
Also known as	Adipogenesis inhibitory factor	Natural killer-cell stimulatory factor Cytotoxic lymphocyte maturation factor	

Abbreviations: APP, acute-phase protein; IFN, interferon; NK, natural killer.

Table 8 Human interleukins (IL)

	IL-15	IL-16	IL-18
Main producing cells	T cells, some malignant B cells	Fibroblasts, activated eosinophils	CD8 T cells, endothelial cells
Structure	4SB amino acids (52–64 kDa) monomer	70 kDa	192 kDa
Main biological actions	Inhibits Ig secretion and induces proliferation in B cells Induces Fas on B cells	Stimulates T- and B-cell proliferation Skeletal muscle hypertrophy Shares β- and γ-receptors with IL-2	Chemotaxis and activation of CD4 T cells, eosinophils and neutrophils IL-2 receptor induction in CD4 T cells
Also known as	High molecular-weight B-cell growth factor	IL-16	Lymphocyte chemoattractant factor

	IL-17
Main producing cells	CD4+ T cells
Structure	155 amino acids (17.3 kDa) predicted
Main biological actions	Stimulates T-cell proliferation IL-6 secretion and surface ICAM-1 in fibroblasts Activates NFkB

Abbreviations used: ICAM, intercellular adhesion molecule; NFKB, nuclear factor kappa B

Table 9 Human interleukins (IL)

	TNF-α	TNF-β
Also known as	Cachectin	Lymphotoxin
Molecular mass	Trimer of 3 × 17 kDa subunits	Subunit 18.5 kDa
Chromosomal location	Human chromosome 6 within MHC	Human chromosome 6 within MHC
Main cellular sources	Macrophages, T cells, NK cells, mast cells, fibroblasts	T cells, B cells
Main stimuli for production	Microbial products, LPS, viruses, IL-1, IL-2, antigen presentation	Mitogenic or antigenic activation
Receptors	55 kDa, 75 kDa	As TBF-1
Major effects	Induction of c-fos, c-myc, and c-ras in fibroblasts Induction of cytokines, IL-1, IL-6, IL-8, PDGF, IL-2 receptor, M-CSF, GM-CSF Increased expression of MHC molecules (especially class I) Inhibition of adipocyte lipoprotein lipase Autocrine macrophage activation Stimulation of neutrophil functions Activation of endothelial cells *In vivo* hypotension and fever Selective tumoricidal activity Catabolic effects on bone and cartilage Modification of liver acute-phase protein synthesis	

Abbreviations used: GM-CSF, granulocyte–macrophage colony-stimulating factor; IL, interleukin; LPS, lipopolysaccharide; M-CSF, macrophage colony-stimulating factor; NK, natural killer; PDGF, platelet-derived growth factor.

Table 10 Human tumour necrosis factors (TNF)

The findings that IL-1 and TNF have biological properties which include stimulation of inflammatory prostanoids, activation of vascular endothelial cells, induction of bone and cartilage metabolism, and activation of lymphocytes elucidate many aspects of inflammation in joint diseases, and also the associated destruction of osteoarticular tissues. Indeed, injection of IL-1 into animal joints produces an acute arthritis with leucocytes in synovial fluid and the breakdown of connective tissue proteoglycans. Since these cytokines are immunopotentiating and are themselves produced during immune responses, clearly they may contribute to the self-perpetuating mechanisms of disease chronicity (Symons *et al*. 1992).

The advent of specific immunoassays for IL-1a, IL-1b, TNF-a, IL-6, and other cytokines was followed by rapid progress in detecting these factors *in vivo* and relating them to different indices of disease activity. The best correlations with systemic inflammatory disease (erythrocyte sedimentation rate, peripheral blood platelet count, C-reactive protein) have been obtained with IL-1b, IL-6, and TNF-a. *In situ* hybridization studies have shown that the main source of IL-1b in the synovium of a rheumatoid joint appears to be the CD14+ tissue macrophage. What the stimulus or stimuli for macrophage activation might be in rheumatoid arthritis is unknown, but it is clear that the production of IL-1 by macrophages can be activated in an autocrine fashion by IL-1 itself, by TNF, other cytokines, immune complexes, and probably certain breakdown products of connective tissue. The initial triggering signal in rheumatoid arthritis remains unknown. The stimulation of IL-1 and TNF production by urate crystals is thought to be central in the development of inflammation in gout (Di Giovine *et al*. 1987; Di Giovine *et al*. 1991).

The biological activity of a cytokine *in vivo* is a net effect related both to its concentration and to the presence of specific inhibitory molecules. For example, in the cytokine system the shedding of cell-surface receptors has been found to be a fairly general phenomenon. Shed receptors retain affinity for their ligand, even in soluble form. In several cases, they have been shown to act as binding proteins that compete with the cell-surface receptor for any free ligand and act, therefore, as biological inhibitors (Symons *et al*. 1991). Some soluble receptors may be stimulatory rather than inhibitory. For example, the complex formed between IL-6 and its soluble receptor can stimulate the signalling component of the cell-surface IL-6 receptor. The soluble form of the IL-2 receptor (a-chain, tac protein) has been especially well studied in immune diseases. The level of soluble IL-2 receptor in body fluids has been found to reflect immune cell activation, and in many cases it correlates with clinical disease activity in autoimmune diseases and in other diseases with an immunopathological component (Symons *et al*. 1988). Apart from soluble receptors, there are examples of other cytokine inhibitory molecules. In particular, there is a third member of the *IL1* gene family that has the same affinity for the two IL-1 surface receptors but receptor occupancy does not transduce an activating signal. This molecule therefore acts as a receptor antagonist and is a potent biological inhibitor. It would seem that such molecules may have potential therapeutic applications (Meager 1990; Symons *et al*. 1992; Duff 1993).

Many of the local and systemic manifestations of the inflammatory response can be explained by the actions of the major monocyte-derived cytokines IL-1a, IL-1b, IL-6, IL-8, and TNF-a. The other interleukins (IL-2, IL-3, IL-4, IL-5, IL-9, IL-10) are mainly lymphocyte-derived and their predominant actions are in the pathways of clonal expansion and differentiation of different classes of lymphocyte (Callard 1990).

Cytokines and T-cell subsets

Helper T cells (Th–CD4+) can be divided into two major types with respect to function and the range of lymphokines they produce. Those that secrete IL-2, interferon-g, IL-3, granulocyte–macrophage-colony-stimulating factor (**GM-CSF**), and TNF-a and -b are known as Th1 cells. They are involved predominantly in the generation of cell-mediated immunity aimed at intracellular pathogens, and they generate the delayed-type hypersensitivity response of the kind seen in chronic inflammatory diseases such as rheumatoid arthritis.

The Th2 cell secretes B-cell growth and differentiation factors (IL-4, IL-5, IL-6, IL-10, IL-3, GM-CSF, TNF), leading to a predominantly humoral immune response that generates antibodies (e.g. allergic response with IgE production). These two major pathways of T-cell differentiation (and therefore types of immune responses) are mutually antagonistic (e.g. IL-4 inhibits Th1 development) so that one type predominates in response to an antigenic challenge (Fig. 6).

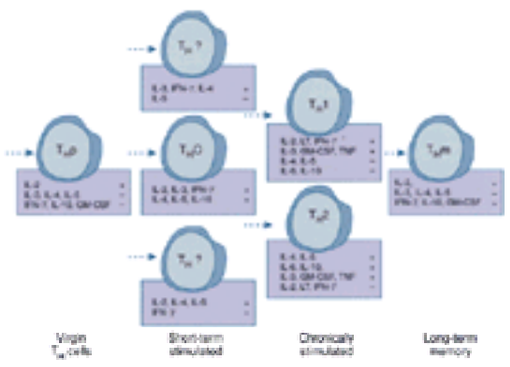


Fig. 6 Relation between the various cytokine secretion phenotypes. A scheme of the T helper (Th) cell functional subdivision is shown. Cytokines produced following activation are indicated in the box below each cell type. Resting cells ('virgin' T cells and 'memory' T cells) only express IL-2 protein; all intermediate differentiation/activation stages have definite patterns of cytokine production.

Since chronic inflammation in diseases such as rheumatoid arthritis has a predominantly cellular immune component, the prospect of suppressing this with Th2 lymphokines (such as IL-4, IL-10, and the related IL-13) has been suggested and clinical trials are in progress. It has been well established that IL-4, IL-10, and IL-13 suppress the production of IL-1 and TNF by stimulated blood monocytes *in vitro*. Whether this is a general phenomenon applicable to any cell type or is unique to blood mononuclear cells is not known.

Cytokine and cytokine receptor families

There are well over 100 characterized cytokines, but it is likely that more will be discovered. It is possible to define large families of similar molecules based on DNA and/or protein sequence homology. In many cases the biological roles of these homologies are not yet understood. A good example is the TNF group of ligands, which, in addition to TNF-a, TNF-b (lymphotoxin-a) and lymphotoxin-b, includes FAS ligand, CD27 ligand, CD30 ligand, CD40 ligand, and OX40 ligand. There are also several sequences with similarity to the known *IL1* genes.

Receptor molecules for cytokines are structurally dissimilar (Fig. 7(a)), but a major subgroup, the 'haematopoietin receptor family', can be distinguished on the basis of shared structural features (Fig. 7(b)). For several individual cytokine receptors, it is also possible to recognize homologous sequences that may represent an extended family of similar receptors. Good examples of this phenomenon are the IL-1 (Fig. 7(c)) and TNF (Fig. 7(d)) receptor families. These sequences have been discovered in vertebrates and in invertebrates (e.g. T cell in clinocephila) suggesting important and highly conserved functions. These observations highlight the limitations of our present understanding of the cytokine system.

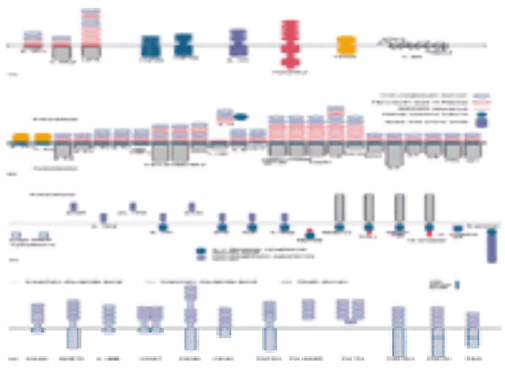


Fig. 7 (a) Cytokine receptor families. (b) The haematopoietin receptor family. (c) Type I IL-1 receptor gene family. (d) Tumour necrosis factor receptor family.

Cytokines in rheumatic diseases

There is little doubt that the pathology of all rheumatic diseases involves cytokine mediators, but the roles of the major proinflammatory cytokines IL-1 and TNF in rheumatoid arthritis has been the most intensively studied area to date.

Activity of proinflammatory cytokines in rheumatoid arthritis was first reported almost 30 years ago (Bodel and Hollingsworth 1969). Recognition of specific IL-1 activity dates back to the early 1980s, when interest was stimulated by reports of IL-1 in exudates in rheumatoid arthritis (Fontana *et al*. 1982; Wood *et al*. 1983), coinciding with the demonstration that IL-1 was produced by mononuclear cells *in vitro* when challenged with arthritogenic stimuli (Duff *et al*. 1983). *In vitro*, synovial cells from rheumatoid arthritic joints were found to release IL-1 (Duff *et al*. 1985) and the concentration of immunoreactive IL-1b in the blood of patients with rheumatoid arthritis was correlated with markers of disease activity (Eastgate *et al*. 1988). The main source of IL-1 in rheumatoid arthritis is probably the CD14+ synovial macrophage (Ogilvie *et al*. 1990; Wood *et al*. 1992). Several trials of anti-IL-1 therapy in rheumatoid arthritis (including small molecular-weight agents and recombinant IL-1 receptor antagonist) are in progress.

The other major cytokine implicated in rheumatoid arthritis is TNF-a, which, in addition to its own proinflammatory activities, is also a potent inducer of IL-1. The first reports implicating biologically active and immunoreactive TNF-a in rheumatoid arthritis are more recent than those for IL-1 (Di Giovine *et al*. 1986; Di Giovine *et al*. 1988a; Di Giovine *et al*. 1988b; Hopkins *et al*. 1988; Saxne *et al*. 1988).

The fact that the spontaneous production of IL-1 by explanted synovial cells from rheumatoid joints could be inhibited by an antibody to TNF-a (Brennan *et al*. 1989) was an important observation. It demonstrated that the suppression of a wider cytokine cascade could be achieved by the specific inhibition of a single cytokine. This work led to clinical trials of 'humanized' antibody to TNF-a in patients with rheumatoid arthritis. Excellent and sustained improvements in both laboratory markers of inflammation and the clinical condition of the patients were reported (Elliot *et al*. 1993) and have been confirmed independently with a different antibody by others (Rankin *et al*. 1995).

This, and similar work by many laboratories, represents a proof of the important principle that specific anticytokine inhibition is potentially a major new therapeutic strategy in rheumatoid arthritis. In the longer term it remains to be shown that efficacy extends to the prevention of erosive disease of cartilage and bone (the most important cause of disability in rheumatoid arthritis), but this seems a reasonable prediction given the activities of TNF, and particularly IL-1, in connective tissue catabolism. It will also be important to demonstrate long-term safety—especially in relation to infections and/or tumours—given the innate immunological activities of IL-1, TNF, and similar cytokines. However, it seems certain that a wide range of anticytokine therapies will enter clinical trials in rheumatic diseases, including biological agents (monoclonal antibodies, soluble receptor proteins, receptor antagonists, and 'anti-inflammatory cytokines' such as IL-10, IL-13), and also conventional, low molecular-weight pharmaceuticals aimed at suppressing the release of cytokines or their activities on target cells.

There are many other cytokines and growth factors that might potentially contribute to the pathology of rheumatic diseases but for which we have, at present, insufficient information. Of great interest is the relatively recent insight that many pathogenic viruses direct the production of factors that transactivate cytokine genes of host cells, while others seem to have captured mammalian cytokine genes (e.g. the Epstein–Barr virus and IL-10) or cytokine receptors (e.g. the poxvirus and IL-1, IL-6, TNF receptors). Such observations may be of relevance to explanations of polyarthritis during known viral infections and may also facilitate our understanding of some idiopathic inflammatory diseases.

Conclusion

The complexity of the mechanisms of inflammation is daunting and the cascade of interactions between the different effector pathways is almost impenetrable. Light is cast by analysis of the rare patient with an inherited deficiency of a single component of inflammation in whom the pathophysiological effects, or otherwise, provide genuine insight into the role of an inflammatory pathway *in vivo*. Our new capacity to delete genes from experimental animals by the method of homologous recombination will in the future allow the systematic analysis of the mechanisms of inflammation in animals deficient in specific pathways. Our present knowledge shows that the mechanisms of inflammation are highly redundant, and the notion that the pharmacological manipulation of a single pathway will have a major therapeutic effect is likely to be regularly frustrated.

Chapter References

Andrews, F. J., Morris, C. J., Kondratowicz, G., and Blake, D. R. (1987). Effect of iron chelation on inflammatory joint disease. *Annals of the Rheumatic Diseases*, **46**, 327–33.

Arai, K-I., Lee, F., Miyajima, A., Miyatake, S., Arai, N., and Yokota, T. (1990). Cytokines: coordinators of immune and inflammatory responses. *Annual Reviews of Biochemistry*, **59**, 783–836.

Basbaum, A. I. and Levine, J. D. (1991). The contribution of the nervous system to inflammation and inflammatory disease. *Canadian Journal of Physiology and Pharmacology*, **69**, 647–51.

Bhoola, K. D., Elson, C. J., and Dieppe, P. A. (1992). Kinins—key mediators in inflammatory arthritis? *British Journal of Rheumatology*, **31**, 509–18.

Blake, D. R., Gallagher, P. J., Potter, A. R., Bell, M. J., and Bacon, P. A. (1984). The effect of synovial iron on the progression of rheumatoid disease. *Arthritis and Rheumatism*, **27**, 495–501.

Blake, D. R. *et al*. (1989). Hypoxic-reperfusion injury in the inflamed human joint. *Lancet*, **i**, 289–92.

Blakemore, A. I. *et al*. (1994). Interleukin-1 receptor antagonist gene polymorphism as a disease severity factor in systemic lupus erythematosus. *Arthritis and Rheumatism*, **37**, 1380–5.

Bodel, P. and Hollingsworth, J. W. (1969). Pyrogen release from human synovial exudate cells. *British Journal of Experimental Pathology*, **49**, 11–19.

Botto, M. *et al*. (1993). Hereditary deficiency of C3 in animals and humans. *International Review of Immunology*, **10**, 37–50.

Bowness, P. *et al*. (1994). Hereditary C1q deficiency and systemic lupus erythematosus. *Quarterly Journal of Medicine*, **87**, 455–64.

Brennan, I., Chantry, D., Jackson, A., Maini, R., and Feldmann, M. (1989). Inhibitory effect of TNF alpha antibodies on synovial cell IL1 production in rheumatoid arthritis. *Lancet*, **ii**, 244–7.

Brouet, J-C. *et al*. (1974). Biological and clinical significance of cryoglobulins: a report of 86 cases. *American Journal of Medicine*, **57**, 775–88.

Butler, R. C., Thompson, J. M., and Keat, A. C. S. (1987). Paraneoplastic rheumatic disorders: a review. *Journal of the Royal Society of Medicine*, **80**, 168–72.

Callard, R. E (ed.). (1990). *Cytokines and B lymphocytes.* Academic Press, London.

Catovsky, D. *et al*. (1980). The association of eosinophilia with lymphoblastic leukaemia or lymphoma: a study of seven patients. *British Journal of Haematology*, **45**, 523–34.

Chikanza, I. C., Petrou. P., Kingsley, G., Chrousos, G., and Panayi, G. S. (1992). Defective hypothalamic response to immune and inflammatory stimuli in patients with rheumatoid arthritis. *Arthritis*

and Rheumatism, **35**, 1281–8.

Chikanza, I. C., Petrou. P., Chrousos, G., Kingsley, G., and Panayi, G. S. (1993). Excessive and dysregulated secretion of prolactin in rheumatoid arthritis: immunopathogenetic and therapeutic implications. *British Journal of Rheumatology*, **32**, 445–8.

Clay, F. E. *et al*. (1994). Interleukin 1 receptor antagonist gene polymorphism association with lichen sclerosis. *Human Genetics*, **94**, 407–10.

Crisp, A. J., Chapman, C. M., Kirkham, S. E., Schiller, A. L., and Krane, S. M. (1984). Articular mastocytosis in rheumatoid arthritis. *Arthritis and Rheumatism*, **27**, 845–51.

Crofford, L. J. *et al*. (1992). Local secretion of corticotrophin-releasing hormone in the joints of Lewis rats with inflammatory arthritis. *Journal of Clinical Investigation*, **90**, 2555–64.

Crofford, L. J. *et al*. (1993). Corticotrophin-releasing hormone in synovial fluids and tissues of patients with rheumatoid arthritis and osteoarthritis. *Journal of Immunology*, **151**,1587–96.

Crofford, L. J. *et al*. (1994). Cyclooxygenase-1 and -2 expression in rheumatoid synovial tissues. Effects of interleukin-1 beta, phorbol ester, and corticosteroids. *Journal of Clinical Investigation*, **93**, 1095–101.

Crowe, P. D., Walter, B. N., Mohler, K. M., Otten Evans, C., Black, R. A., and Ware, C. F. (1995). A metalloproteinase inhibitor blocks shedding of the 80-kD TNF receptor and TNF processing in T lymphocytes. *Journal of Experimental Medicine*, **181**, 1205–10.

Crusius, J. B. *et al*. (1995). Interleukin 1 receptor antagonist gene polymorphism and multiple sclerosis. *Lancet*, **346**, 979–80.

Dhodapkar, M. V., Li, C. Y., Lust, J. A., Tefferi, A., and Phyliky, R. L. (1994). Clinical spectrum of clonal proliferations of T-large granular lymphocytes: a T-cell clonopathy of undetermined significance? *Blood*, **84**, 1620–7.

Di Giovine, F. S., Mansaw, S., Nuki, G., and Duff, G. W. (1986). Tumour necrosis factor activity in synovial exudate fluids from patients with osteoarthritis or rheumatoid arthritis. *British Journal of Rheumatology*, **25**, A108.

Di Giovine, F. S., Malawista, S. E., Nuki, G., and Duff, G. W. (1987). Interleukin 1 as a mediator of crystal arthritis. *Journal of Immunology*, **138**, 3213–18.

Di Giovine, F. S., Nuki, G., and Duff, G. W. (1988 *a*). Tumour necrosis factor in synovial exudates. *Annals of the Rheumatic Diseases*, **47**, 768–72.

Di Giovine, F. S., Meager, A., Leung, H., and Duff, G. W. (1988 *b*). Immunoreactive TNF alpha and biological inhibitor(s) in synovial fluids from rheumatoid arthritis patients. *International Journal of Immunopathology and Pharmacology*, **1**, 17–26.

Di Giovine, F. S., Malawista, S. E., Thornton, E., and Duff, G. W. (1991). Urate crystals stimulate production of TNFa from human blood monocytes and synovial cells: cytokine mRNA and protein kinetics and cellular distribution. *Journal of Clinical Investigation*, **87**, 1375–81.

Di Giovine, F. S., Cork, M.J., Crane, A., Mee, J.B., and Duff, G.W. (1996). New genetic association of an IL1b gene variation at +3953 with IL1b production and psoriasis. *Cytokine*, **7**, 606.

Docherty, A. J. P. and Murphy, G. (1990). The tissue metalloproteinase family and the inhibitor TIMP: a study using cDNAs and recombinant proteins. *Annals of the Rheumatic Diseases*, **49**, 469–79.

Docherty, M., Watt, I., and Dieppe, P. (1982). A. Localised chondrocalcinosis in post-meniscetomy knees. *Lancet*, 1207–10.

Donaldson, V. H., Glueck, H. I., and Fleming, T. (1972). Rheumatoid arthritis in a patient with Hageman trait. *New England Journal of Medicine*, **286**, 528–30.

Duff, G. W. (1989 *a*). Interleukin 1 in inflammatory joint disease. In *Interleukin 1, inflammation and disease* (ed. R. Bomford and B. Henderson), pp. 243–56. Elsevier, Amsterdam.

Duff, G. W. (1989 *b*). Protein regulatory factors in non-malignant diseases. *Lancet*, **ii**, 1432–5.

Duff, G. W. (1993). Cytokines and anti-cytokines. *British Journal of Rheumatology*, **32**, 15–20.

Duff, G. W. (1994). Molecular genetics of cytokines: cytokines in chronic inflammatory disease. In *The cytokine handbook*, (2nd edn) (ed. A. Thompson), pp. 21–30. Academic Press, London.

Duff, G. W., Atkins, E., and Malawista, S. E. (1983). The fever of gout: urate crystals activate endogenous pyrogen production from human and rabbit mononuclear cells. *Transactions of the Association of American Physicians*, **96**, 234–45.

Duff, G. W., Forre, O., Waalen, K., Dickens, E., and Nuki, G. (1985). Rheumatoid arthritis synovial dendritic cells produce IL1. *British Journal of Rheumatology*, **24**, 94–8.

Eastgate, J. A., Symons, J. A., Wood, N. C., Grinlington, F. M., Di Giovine, F. S., and Duff, G. W. (1988). Correlation of plasma IL1 levels with disease activity in rheumatoid arthritis. *Lancet*, **ii**, 706–9.

Edwards, J. C. (1994). The nature and origins of synovium: experimental approaches to the study of synoviocyte differentiation. *Journal of Anatomy*, **184**, 493–501.

Elliot, M. J. *et al*. (1993). Treatment of rheumatoid arthritis with chimeric monoclonal antibodies to tumour necrosis factor alpha. *Arthritis and Rheumatism*, **36**, 1681–90.

Farrell, A. J., Blake, D. R., Palmer, R. M., and Moncada, S. (1992). Increased concentrations of nitrite in synovial fluid and serum samples suggest increased nitric oxide synthesis in rheumatic diseases. *Annals of the Rheumatic Diseases*, **51**, 1219–22.

Fauci, A. S. (1982). The idiopathic hypereosinophilic syndrome. *Annals of Internal Medicine*, **97**, 78–92.

Fava, R. A. *et al*. (1994). Vascular permeability factor/endothelial growth factor (VPF/VEGF): accumulation and expression in human synovial fluids and rheumatoid synovial tissue. *Journal of Experimental Medicine*, **180**, 341–6.

Folkman, J. and Shing, Y. (1992). Angiogenesis. *Journal of Biological Chemistry*, **267**, 10931–4.

Fontana, A., Hengartner, H., Weber, E., Fehr, K., Frob, P. J., and Cohen, G. (1982). Interleukin-1 activity in the synovial fluid of patients with rheumatoid arthritis. *Rheumatology International*, **2**, 49–53.

Garrett, N. E. Mapp, P. I. Cruwys, S. C. Kidd, B. L., and Blake, D. R. (1992). Role of substance P in inflammatory arthritis. *Annals of the Rheumatic Diseases*, **51**, 1014–18.

Gearing, A. J. *et al*. (1994). Processing of tumour necrosis factor-alpha precursor by metalloproteinases. *Nature*, **370**, 555–7.

Gleich, G. J. and Adolphson, C. R. (1986). The eosinophilic leukocyte: structure and function. *Advances in Immunology*, **39**, 177–253.

Goetzl, E. J., An, S., and Smith, W. L. (1995). Specificity of expression and effects of eicosanoid mediators in normal physiology and human diseases. *FASEB Journal*, **9**, 1051–8

Gorevic, P. D. *et al*. (1980). Mixed cryoglobulinemia: clinical aspects of long-term follow-up of 40 patients. *American Journal of Medicine*, **69**, 287–308.

Greenwald, R. A. (1991). Oxygen radicals, inflammation, and arthritis: pathophysiological considerations and implications for treatment. *Seminars in Arthritis and Rheumatism*, **20**, 219–40.

Greenwald, R. A. and Moy, W. W. (1980). Effect of oxygen-derived free radicals on hyaluronic acid. *Arthritis and Rheumatism*, **23**, 455–63.

Griffiths, R. J. *et al*. (1995). Leukotriene B4 plays a critical role in the progression of collagen-induced arthritis. *Proceedings of the National Academy of Sciences (USA)*, **92**, 517–21.

Groves, R. W. and Kupper, T. S. (1994). Spontaneous inflammatory skin disease in mice which overexpress IL1 alpha in basal keratinocytes. *Journal of Investigative Dermatology*, **103**, 442.

Hart, D. A. and Fritzler, M. J. (1989). Regulation of plasminogen activators and their inhibitors in rheumatic diseases: new understanding and the potential for new directions. *Journal of Rheumatology*, **169**, 1184–91.

Hauptmann. G. *et al*. (1986). Component deficiencies. 2. The fourth component. *Progress in Allergy*, **39**, 232–49.

Henderson, B. and Edwards, J. C. W. (1987). *The synovial lining in health and disease*. Chapman and Hall, London.

Henderson, B., Revell, P. A., and Edwards, J. C. W. (1988). Synovial lining cell hyperplasia in rheumatoid arthritis, dogma and fact. *Annals of the Rheumatic Diseases*, **47**, 348–9.

Holzer, P. (1988). Local effector functions of capsaicin-sensitive sensory nerve endings: involvement of tachykinins, calcitonin gene-related peptide and other neuropeptides. *Neuroscience*, **24**,

739–68.

Hopkins, S.J. and Meager, A. (1988). Cytokines in synovial fluid: the presence of tumour necrosis factor and interferon. *Clinical and Experimental Immunology*, **73**, 88–92.

Ialenti, A., Moncada, S., and Di Rosa, M. (1993). Modulation of adjuvant arthritis by endogenous nitric oxide. *British Journal of Pharmacology*, **110**, 701–6.

Jacob, C. O. and McDevitt, H. O. (1988). Tumour necrosis factor-a in murine autoimmune 'lupus' nephritis. *Nature*, **331**, 356–8.

Jacob, C. O., Fronek, Z., Lewis, G. D., Koo, M., Hansen, J. A., and McDevitt, H. O. (1990). Heritable major histocompatibility complex class II-associated differences in production of tumor necrosis factor a: relevance to genetic predisposition to systemic lupus erythematosus. *Proceedings of the National Academy of Sciences (USA)*, **87**, 1233–7.

Jayson, M. I. V. and Dixon, A. St J. (1970). Intra-articular pressure in rheumatoid arthritis of the knee. *Annals of the Rheumatic Diseases*, **29**, 401–8.

Karalis, K., Sano, H., Redwine, J., Listwak, S., Wilder, R. L., and Chrousos, G. P. (1991). Autocrine or paracrine inflammatory actions of corticotrophin-releasing hormone *in vivo*. *Science*, **254**, 421–3.

Kaur, H. and Halliwell, B. (1994). Evidence for nitric oxide-mediated oxidative damage in chronic inflammation. Nitrotyrosine in serum and synovial fluid from rheumatoid patients. *FEBS Letters*, **350**, 9–12.

Kidd, B. L., Mapp, P. I., Blake, D. R., Gibson, S. J., and Polak, J. M. (1990). Neurogenic influences in arthritis. *Annals of the Rheumatic Diseases*, **49**, 649–52.

Kidd, B. L., Cruwys, S., Mapp, P. I., and Blake, D. R. (1992). Role of the sympathetic nervous system in chronic joint pain and inflammation. *Annals of the Rheumatic Diseases*, **51**, 1188–91.

Koch, A. E. *et al*. (1994). Vascular endothelial growth factor. A cytokine modulating endothelial function in rheumatoid arthritis. *Journal of Immunology*, **152**, 4149–56.

Lachmann, P. J. and Walport, M. J. (1987). Deficiency of the effector mechanisms of the immune response and autoimmunity. In *Autoimmunity and autoimmune disease, Ciba Foundation symposium 129* (ed. J. Whelan), pp. 149–71. Wiley, Chichester.

Lahita, R. G. (1993). Sex hormones as immunomodulators of disease. *Annals of the New York Academy of Science*, **685**, 278–87.

Lahita, R. G., Chiorazzi, N., Gibovsky, A., Winchester, R. J., and Kunkel, H. G. (1983). Familial systemic lupus erythematosus in males. *Arthritis and Rheumatism*, **26**, 39–44.

McCord, J. M. (1974). Free radicals and inflammation: protection of synovial fluid by superoxide dismutase. *Science*, **185**, 529–31.

McCord, J. M. (1985). Oxygen-derived free radicals in postischemic tissue injury. *New England Journal of Medicine*, **312**, 159–63.

McDowell, T. L., Symona, J. A., Ploski, R., Forre, O., and Duff, G. W. (1995). A genetic association between juvenile rheumatoid arthritis and a novel interleukin-1 alpha polymorphism. *Arthritis and Rheumatism*, **38**, 221–8.

McGeehan, G. M. *et al*. (1994). Regulation of tumour necrosis factor-alpha processing by a metalloproteinase inhibitor. *Nature*, **370**, 558–61.

McGuire, W., Hill, A. V. S., Allsopp, C. E. M., Greenwood, B. M., and Kwiatkowsky, D. (1994). Variation in the TNF alpha promoter region associated with susceptibility to cerebral malaria. *Nature*, **371**, 508–11.

McMurray, R., Keisler, D., Izui, S., and Walker, S. E. (1994). Hyperprolactinemia in male NZB/NZW (B/W) F1 mice: accelerated autoimmune disease with normal circulating testosterone. *Clinical Immunology and Immunopathology*, **71**, 338–43.

Mansfield, J. C. *et al*. (1994). Novel genetic association between ulcerative colitis and the anti-inflammatory cytokine interleukin 1 receptor antagonist. *Gastroenterology*, **106**, 637–42.

Marsh, D. G. *et al*. (1994). Linkage analysis of IL4 and other chromosome 5Q31.1 markers and total serum immunoglobulin-E concentrations. *Science*, **264**, 1152–6.

Masferrer, J. L. *et al*. (1994). Selective inhibition of inducible cyclooxygenase 2 *in vivo* is antiinflammatory and nonulcerogenic. *Proceedings of the National Academy of Sciences (USA)*, **91**, 3228–32.

Mattsson, R., Mattsson, A., Hansson, I., Holmdahl, R., Rook, G. A., and Whyte, A. (1992). Increased levels of prolactin during, but not after, the immunisation with rat collagen II enhances the course of arthritis in DBA/1 mice. *Autoimmunity*, **11**, 163–70.

Meager, A. (1990). *Cytokines*. Open University Press, Milton Keynes.

Merry, P., Winyard, P. G., Morris, C. J., Grootveld, M., and Blake, D. R. (1989). Oxygen free radicals, inflammation and synovitis: the current status. *Annals of the Rheumatic Diseases*, **48**, 864–70.

Mitchell, J. A., Akarasereenont, P., Thiemermann, C., Flower, R. J., and Vane, J. R. (1993). Selectivity of nonsteroidal antiinflammatory drugs as inhibitors of constitutive and inducible cyclooxygenase. *Proceedings of the National Academy of Sciences (USA)*, **90**, 11693–7.

Mohler, K. M. *et al*. (1994). Protection against a lethal dose of endotoxin by an inhibitor of tumour necrosis factor processing. *Nature*, **370**, 218–20.

Moncada, S. and Higgs, A. (1993). The l-arginine–nitric oxide pathway. *New England Journal of Medicine*, **329**, 2002–12.

Morgan, B. P. and Walport, M. J. (1991). Complement deficiency and disease. *Immunology Today*, **12**, 301–6.

Neeck, G., Federlin, K., Graef, V., Rusch, D., and Schmidt, K. L. (1990). Adrenal secretion of cortisol in patients with rheumatoid arthritis. *Journal of Rheumatology*, **17**, 24–9.

Nicklin, M. J. H., Weith, A., and Duff, G. W. (1994). A physical map of the region encompassing the human interleukin-1a, b and the interleukin-1 receptor antagonist genes. *Genomics*, **19**, 382–4.

Ogilvie, A.D., Wood, N.C., Dickens, E., Wajtacha, D., and Duff, G.W. (1990). *In situ* hybridisation. *Annals of the Rheumatic Diseases*, **49**, 434–9.

O'Shea, J. J., Jaffe, E. S., Lane, H. C., MacDermott, R. P., and Fauci, A. S. (1987). Peripheral T cell lymphoma presenting as hypereosinophilia with vasculitis. *American Journal of Medicine*, **82**, 539–45.

Partridge, R. E. H. and Duthie, J. J. R. (1963). Controlled trial of immobilisation of the joints in rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **22**, 91–9.

Peacock, D. J., Banquerigo, M. L, and Brahn, E. (1992). Angiogenesis inhibition suppresses collagen arthritis. *Journal of Experimental Medicine*, **175**, 1135–8.

Pettifer, E. R., Higgs, G. A., and Salmon, J. A. (1992). Eicosanoids (prostaglandins and leukotrienes). In *Biochemistry of inflammation* (ed. J. Whicher and S. Evans), pp. 91–108. Kluwer, Dordrecht,.

Pettipher, R., Edwards, J., Cruwys, S., Jessup, E., Beesley, J., and Henderson, B. (1990). Pathogenesis of antigen-induced arthritis in mice deficient in neutrophil elastase and cathepsin G. *American Journal of Pathology*, **137**, 1077–82.

Pober, J. S. and Cotran, R. S. (1990). The role of endothelial cells in inflammation. *Transplantation*, **50**, 537–44.

Puig, J. G., Mateos, F. A., and Diaz, V. D. (1989). Inhibition of xanthine oxidase by allopurinol: a therapeutic option for ischaemia induced pathological processes? *Annals of the Rheumatic Diseases*, **48**, 883–8.

Ralston, S. H., Todd, D., Helfrich, M., Benjamin, N., and Grabowski, P. S. (1994). Human osteoblast-like cells produce nitric oxide and express inducible nitric oxide synthase. *Endocrinology*, **135**, 330–6.

Rankin, E. C. C. *et al*. (1995). The therapeutic effects of an engineered human anti-tumour necrosis factor alpha antibody (CDP571) in rheumatoid arthritis. *British Journal of Rheumatology*, **34**, 334–42.

Reber, P. M. (1993). Prolactin and immunomodulation. *American Journal of Medicine*, **95**, 637–44.

Reid, K. B. M. (1989). Deficiency of the first component of human complement. *Immunodeficiency Reviews*, **1**, 247–60.

Remmers, E. F., Sano, H., and Wilder, R. L. (1991). Platelet-derived growth factors and heparin-binding (fibroblast) growth factors in the synovial tissue pathology of rheumatoid arthritis. *Seminars in Arthritis and Rheumatism*, **21**, 191–9.

Ruddy, S. (1986). Component deficiencies 3. The second component. *Progress in Allergy*, **39**, 250–66.

Salvemini, D., Misko, T. P., Masferrer, J. L., Seibert, K., Currie, M. G., and Needleman, P. (1993). Nitric oxide activates cyclooxygenase enzymes. *Proceedings of the National Academy of Sciences (USA)*, **90**, 7240–4.

Sanchez-Guerrero, J., Gutierrez-Urena, S., Vidaller, A., Reyes, E., Iglesias, A., and Alarcon-Segovia, D. (1990). Vasculitis as a paraneoplastic syndrome. Report of 11 cases and review of the literature. *Journal of Rheumatology*, **17**, 1458–62.

Savill, J. S., Wyllie, A. H., Henson, J. E., Walport, M. J., Henson, P. M., and Haslett, C. (1989). Macrophage phagocytosis of aging neutrophils in inflammation. Programmed cell death in the neutrophil leads to its recognition by macrophages. *Journal of Clinical Investigation*, **83**, 865–75.

Saxne, T. M., Palladino, M. A., Heinegard, D., Talal, N., and Wollheim, F. A. (1988). Detection of TNF alpha but not TNF beta in rheumatoid arthritis. Synovial fluid and serum. *Arthritis and Rheumatism*, **31**, 1041–4.

Silman, A. J. (1994). Epidemiology of rheumatoid arthritis. *APMIS*, **102**, 721–8.

Sorsa, T. *et al*. (1992). Collagenase in synovitis of rheumatoid arthritis. *Seminars in Arthritis and Rheumatism*, **22**, 44–53.

Spry, C. J. F. (1988). *Eosinophils. A comprehensive review and guide to the scientific and medical literature*. Oxford University Press.

Stefanovic-Racic, M., Stadler, J., and Evans, C. H. (1993). Nitric oxide and arthritis. *Arthritis and Rheumatism*, **36**, 1036–44.

Stefanovic-Racic, M., Meyers, K., Meschter, C., Coffey, J. W., Hoffman, R. A., and Evans, C. H. (1994). *N*-monomethyl arginine, an inhibitor of nitric oxide synthase, suppresses the development of adjuvant arthritis in rats. *Arthritis and Rheumatism*, **37**, 1062–9.

Sternberg, E. M., Chrousos, G. P., Wilder, R. L., and Gold, P. W. (1992). The stress response and the regulation of inflammatory disease. *Annals of Internal Medicine*, **117**, 854–66.

Symons, J. A., Wood, N. C., Di Giovine, F. S., and Duff, G. W. (1988). Soluble IL2 receptor in rheumatoid arthritis: correlation with disease activity, IL1 and IL2—inhibition. *Journal of Immunology*, **1**, 2612–18.

Symons, J. A., Eastgate, J. A., and Duff, G. W. (1991). Purification and characterisation of a novel soluble receptor for interleukin-1. *Journal of Experimental Medicine*, **174**, 1251–4.

Symons, J. A., Eastgate, J. A., and Duff, G. W. (1992). Interleukin 1 as an inflammatory mediator. In *Biochemistry of inflammation* (ed. J. Whicher and S. Evans), pp. 183–210. Kluwer, Dordrecht.

Tarlow, J. K. *et al*. (1993). Polymorphism in human IL-1 receptor antagonist gene intron 2 is caused by variable numbers of an 86-bp tandem repeat. *Human Genetics*, **91**, 403–4.

Tarlow, J. K *et al*. (1994). Severity of alopecia areata is associated with a polymorphism in the interleukin-1 receptor antagonist gene. *Journal of Investigative Dermatology*, **103**, 387–9.

Thompson, M. and Bywaters, E. G. L. (1962). Unilateral rheumatoid arthritis following hemiplegia. *Annals of the Rheumatic Diseases*, **21**, 370–7.

Vane, J. R. *et al*. (1994). Inducible isoforms of cyclooxygenase and nitric-oxide synthase in inflammation. *Proceedings of the National Academy of Sciences (USA)*, **91**, 2046–50.

Vincenti, M. P., Clark, I. M., and Brinckerhoff, C. E. (1994). Using inhibitors of metalloproteinases to treat arthritis. Easier said than done? *Arthritis and Rheumatism*, **37**,1115–26.

Walport, M. J. and Lachmann, P. J. (1993). Complement. In *Clinical aspects of immunology*, (5th edn) (ed. P. J. Lachmann, D. K. Peters, F. S. Rosen, and M. J. Walport), pp. 347–75. Blackwell Scientific, Boston.

Weinberg, J. B. *et al*. (1994). The role of nitric oxide in the pathogenesis of spontaneous murine autoimmune disease: increased nitric oxide production and nitric oxide synthase expression in MRL-lpr/lpr mice, and reduction of spontaneous glomerulonephritis and arthritis by orally administered NG-monomethyl-l-arginine. *Journal of Experimental Medicine*, **179**, 651–60.

Weiss, S. J. (1989). Mechanisms of disease. Tissue destruction by neutrophils. *New England Journal of Medicine*, **320**, 365–76.

Weiss, S. J. and Peppin, G. J. (1986). Collagenolytic metalloenzymes of the human neutrophil. *Biochemistry and Pharmacology*, **35**, 3189–97.

Wilson, A. G., Di Giovine, F. S., Blakemore, A. I. F., and Duff, G. W. (1992). Single base polymorphism in the human tumour necrosis factor alpha (TNFa) gene detectable by NcoI restriction of PCR product. *Human Molecular Genetics*, **1**, 353.

Wilson, A. G., de Vries, N., Pociot, F., Di Giovine, F. S., van der Putte, L. B. A., and Duff, G. W. (1993). An allelic polymorphism within the human tumor necrosis factor a promoter region is strongly associated with HLA A1, B8, and DR3 alleles. *Journal of Experimental Medicine*, **2**, 557–60.

Winyard, P. G., Blake, D. R., Chirico, S., Gutteridge, J. M. C., and Lunec, J. (1987). Mechanism of exacerbation of rheumatoid synovitis by total-dose iron-dextran infusion: *in-vivo* demonstration of iron-promoted oxidant stress. *Lancet*, **i**, 69–72.

Wood, D. D., Ihrie, E. J., Dinarello, C. A., and Cohen, P. L. (1983). Isolation of an IL-1-like factor from human joint effusions. *Arthritis and Rheumatism*, **26**, 975–83.

Wood, J. C., Dickens, E., Symons, J. A., and Duff, G. W. (1992). *In situ* hybridisation of IL-1 in CD14-positive cells in rheumatoid arthritis. *Clinical Immunology and Immunopathology*, **62**, 295–300.

3.2 Lymphocyte traffic in inflammation

Dorian O. Haskard

Introduction

Inflammation is traditionally divided into two patterns (Robbins *et al*. 1984). Whilst acute inflammation is dominated by a predominantly polymorphonuclear leucocyte infiltration of the tissues and by the exudation of fluid and plasma proteins, chronic inflammation is characterized by the presence of lymphocytes and monocyte/macrophages, and by the activation and proliferation of connective tissue. In some complex forms of inflammation such as rheumatoid synovitis, acute inflammation may be periodically superimposed upon a background of chronic inflammation. This chapter focuses on lymphocyte traffic to explain how an understanding of these inflammatory processes may be improved by analysing the mechanisms of leucocyte migration.

Until the second half of this century the origin and fate of lymphocytes was as much an enigma as was their function (Nature 1957). Lymphocytes could be found in large numbers in thoracic duct lymph and these were thought to be the products of recent cell division within lymph nodes. Moreover, lymphocytes that had been collected by thoracic duct drainage and then injected intravenously were found to disappear from the circulation within an hour, giving rise to the assumption that lymphocytes were short-lived non-dividing cells. The classical experiments of Sir James Gowans in the rat provided a completely altered view of lymphocyte physiology (Gowans and Knight 1964). Gowans showed that the reason why thoracic duct lymphocytes rapidly leave the circulation when injected intravenously is that they enter lymphatic tissues, through which they pass on a recirculation pathway back to blood (Fig. 1).

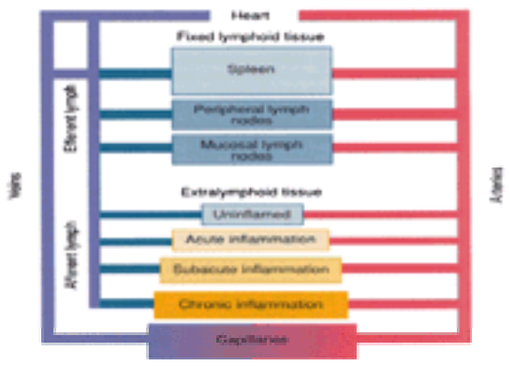


Fig. 1 The pathways of ongoing lymphocyte recirculation through spleen and lymph nodes, and the acquired pathway of recirculation through inflammatory tissues. Whilst passage from the blood into spleen and uninflamed extralymphoid tissue takes place through blood vessels with flat endothelium, passage into lymph nodes and some chronic inflammatory tissues takes place across blood vessels with 'high endothelial' morphology.

Besides recirculating through lymph nodes, lymphocytes also pass in large numbers through the spleen and bone marrow. Indeed it has been demonstrated in the rat that more lymphocytes circulate through the spleen than through all the lymph nodes put together (Pabst 1988).

The demonstration of lymphocyte recirculation implied that small lymphocytes have much longer lifespans than had previously been realized. Studies on the longevity of chromosomal alterations in patients with ankylosing spondylitis who had undergone spinal irradiation showed that a population of small lymphocytes recirculates for years (Buckton *et al*. 1967), although the median survival time for recirculating lymphocytes in man is probably more like 2 to 3 weeks.

The mean size of the recirculating lymphocyte pool in man has been estimated as 2.3×10^{10}, with approximately 50 per cent of recirculating cells being present in the blood (Scott *et al*. 1972). The mean blood transit of recirculating lymphocytes may be as fast as 30 min, resulting in an exchange rate between blood and lymphatic tissues of up to 48 times daily (Schick *et al*. 1975).

Lymphocyte recirculation is a vital element in the function of the immune system since it enables immunological memory to be shared effectively by many anatomically discrete lymphoid organs. Immobilized antigen is exposed to a wide range of lymphocytes of varying clonal specificity because of lymphocyte recirculation. Moreover, lymphocyte recirculation promotes interactions between different functional subsets of lymphocyte and accessory cell, and thus aids the modulation of ongoing immune reactions.

Lymphocyte migration in inflammation

In addition to the large-scale constitutive traffic of lymphocytes through lymphoid tissues, a small but significant steady-state migration of lymphocytes also occurs through most non-lymphoid tissues, as demonstrated histologically or by the migration of radiolabelled lymphocytes in animal models. Lymphocytes are found in afferent lymph draining uninflamed non-lymphoid tissues, although, with the exception of the liver and intestinal mucosa, the numbers of migrating lymphocytes are few compared with those found in efferent lymphatics from lymph nodes. Although modest in extent, the ongoing migration of lymphocytes through non-lymphoid tissues is probably important for routine immunosurveillance.

The presence of perivascular collections and diffuse infiltrates of lymphocytes is a hallmark of immune-mediated chronic inflammation in non-lymphoid tissues, as typified by the cutaneous delayed-type hypersensitivity response to tuberculin or by the rejection response to foreign grafts. Early studies using autoradiography demonstrated that mononuclear cells in chronic inflammatory tissues are mainly derived from the blood rather than from cells proliferating locally (Kosunen *et al.* 1963), and this view has subsequently been supported by immunocytochemical analysis of antigens related to activation and proliferation. For example, although

most rheumatoid synovial lymphocytes show evidence of partial activation (Burmester *et al*. 1984), only a small percentage are actually undergoing cell division (Bonvoisin *et al*. 1984).

Lymphocytes not only migrate in increased numbers into inflamed tissues, but an increase also occurs in the number of lymphocytes present in the afferent lymph draining sites of inflammation, indicating an induced peripheral pathway of recirculation (Hall and Morris 1963) (Fig. 1). Studies in which the peripheral lymph draining sites of chronic inflammation have been sampled by in-dwelling cannulas placed in afferent lymphatics have shown that the scale of lymphocyte migration through inflammatory tissue increases over days to weeks, eventually resembling that through lymph nodes (Smith *et al*. 1970). Peripheral recirculation of this magnitude may occur in some chronically inflamed synovial tissues, although this has never been demonstrated directly. The ability of chronic inflammatory tissues to support the large-scale traffic of lymphocytes has led to the speculation that fixed lymph nodes are an evolutionary development from chronic inflammatory tissues (Rannie *et al*. 1977). Conversely, chronically inflamed non-lymphoid tissues may acquire immunological functions that are normally confined to fixed lymphoid tissues, thereby devolving the immunological response to injury to the periphery (Janossy *et al*. 1981).

Factors promoting lymphocyte emigration and accumulation

From these considerations it will be appreciated that the lymphocytes in an inflammatory focus are in a dynamic state, governed by entry from the blood, retention in the tissues, programmed cell death, and migration into afferent lymph, with only a small proportion actually proliferating at the site of the lesion (Fig. 2).

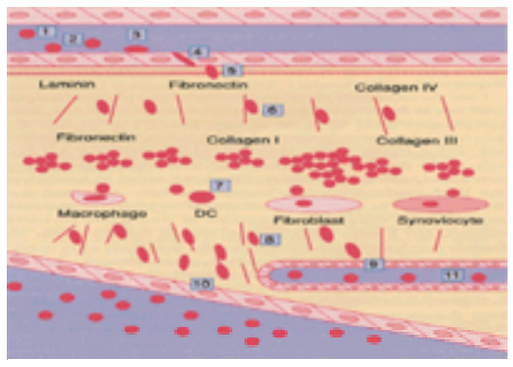


Fig. 2 Events in the migration of leucocytes from blood into and through inflammatory tissues: (1) circulating leucocytes; (2) rolling of resting unactivated leucocytes on endothelium; (3) contact with activating factors on the endothelial cell surface, resulting in up-regulation of leucocyte integrin function and firm adhesion of the leucocyte to the endothelial cell; (4) leucocyte transmigration between endothelial cells; (5) penetration of basement membrane; (6) adhesion and migration upon components of extracellular matrix (e.g. fibronectin, collagen); (7) contact interactions with resident cells; (8) programmed cell death or migration away from the inflammatory focus; (9) entry into afferent lymphatics; (10) entry into synovial fluid or other extracellular fluids; and (11) recirculation to lymph nodes and blood.

The role of vascular endothelium

Adhesion of leucocytes to the luminal surface of vascular endothelial cells is the first step in their migration from the blood into the tissues and is of critical importance in determining the number and type of leucocytes entering sites of inflammation. Interactions between leucocytes and endothelial cells therefore regulate the nature and progression of the inflammatory response.

Lymph nodes

Marchesi and Gowans (1964) observed that recirculating lymphocytes enter lymph nodes by passing through paracortical postcapillary venules, which are characterized by their plump, columnar endothelial cells. These vessels have subsequently been referred to as 'high endothelial venules' and have been found to have properties that are adapted to support lymphocyte adhesion and transmigration from the blood into the lymph node parenchyma (Yednock and Rosen 1989). The interaction between lymphocytes and high endothelial venules is specific, as neutrophils and monocytes do not migrate into the lymph node parenchyma under resting non-inflamed conditions.

Besides having an adhesive luminal surface that is suited to lymphocyte adhesion, the structure of high endothelial venules may provide a valve-like regulation of permeability to macromolecules, thereby minimizing the extravasation of plasma proteins during large-scale lymphocyte transmigration of endothelium. These venules are further characterized by the ability to synthesize and secrete sulphated glycoconjugates that are not normally synthesized by endothelial cells in uninflamed non-lymphoid tissues (Andrews *et al*. 1983). Recent work has identified some of these sulphated molecules as glycoforms of CD34 and GLYCAM-1 (see below) and shown that they act as ligands for the leucocyte adhesion molecule L-selectin (see below).

Spleen and bone marrow

There are normally no high endothelial venules in spleen and bone marrow, and lymphocytes migrate into these tissues across flat fenestrated endothelium (Goldschneider and McGregor 1968). In contrast to the effect on migration of lymphocytes into lymph nodes, lymphocyte entry into spleen and bone marrow is resistant to the digestion of the lymphocyte surface by trypsin, suggesting that different types of adhesion interactions are involved. In the case of the spleen, lymphocytes probably pass readily through fenestrated endothelium in the splenic marginal zones (Peters 1983) and retention in this organs is likely to be dependent largely upon adhesion to parenchymal cells.

Inflammation

Morphological changes in endothelial cells can be seen relatively early during cell-mediated immune reactions. Gell (1959) showed that endothelium took up intravenously injected carbon particles at an early stage during a cutaneous tuberculin reaction in guinea pigs and considered this to be evidence for endothelial cell activation. Likewise, endothelial cells appear activated in the early phase of graft rejection (Dvorak *et al*. 1979). Endothelial cells examined at the start of an immune response typically show an increase in size, abundant eosinophilic cytoplasm, enlarged nuclei, and prominent ribosomes and endoplasmic reticulum, suggestive of increased protein synthesis (Wiener *et al*. 1967). These morphological changes illustrate the fact that endothelial cells are far from inert during inflammation.

Although postcapillary venules in non-lymphoid tissues normally have a flattened endothelium, vessels can be found in established chronic inflammatory lesions that are similar or even identical histologically to the high endothelial venules of secondary lymphoid organs. Such vessels tend to develop in non-lymphoid tissues particularly in the vicinity of lymphocyte-rich inflammatory infiltrates and may occur as a consequence of lymphocyte traffic into the tissues (Freemont *et al*. 1983; Iguchi and Ziff 1986). Furthermore, high endothelial venule-like vessels in rheumatoid synovium can be shown *in vitro* to bind lymphocytes, and to have properties normally confined to high endothelial venules of peripheral lymph nodes, such as the synthesis of sulphated macromolecules, and expression of L-selectin ligands (see below) (Freemont 1987; Michie *et al*. 1993). The acquisition by chronic inflammatory tissue of high endothelial venule-associated mechanisms for the recruitment of lymphocytes may not only augment the scale of lymphocyte traffic, but may also allow the migration of lymphocyte subpopulations that do not normally enter non-lymphoid tissues in the absence of chronic inflammation.

The role of antigen

The large majority of lymphocytes entering inflammatory tissues are not clonally restricted to react with antigens responsible for the initiation or maintenance of the inflammatory response (Prendergast 1964). At the onset of immune-mediated inflammation the local activation of T lymphocytes sets up a cascade of events that results in the non-specific secondary recruitment of more lymphocytes and also other leucocyte types (Hanto *et al*. 1982). The degree to which inflammation can be induced and amplified by this means is shown by the ability to mount a delayed-type hypersensitivity response after passive transfer of a single T cell (Marchal *et al*. 1982). Any oligoclonality of antigen specificity detected in populations of lymphocytes isolated from inflamed tissues is likely, therefore, to be due predominantly to the selective retention, survival, and/or proliferation of cells activated within the tissues, rather than to the existence of a mechanism for the selective recruitment of

antigen-specific lymphocytes.

The influence of lymphocyte lineage

In the absence of inflammation, lymphocytes in human extracellular fluids are predominantly T cells rather than B cells (Manconi *et al*. 1978). Furthermore, the large majority of lymphocytes found in the tissues during a delayed-type hypersensitivity response are T cells with very few B cells (Poulter *et al*. 1982). Clearly, therefore, B cells migrate into non-lymphoid tissues less readily than T cells, at least during day-to-day constitutive traffic and in the first few days of chronic inflammation. Although both CD4+ and CD8+ T cells migrate into the skin during a delayed-type hypersensitivity response, there is an increase in the CD4/CD8 ratio in the tissue compared with peripheral blood, indicating the existence of T-cell subset selectivity in migration (Pitzalis *et al*. 1991).

A more complex picture is seen in chronic synovitis due to rheumatoid arthritis or to other causes. Here the tissue underlying the synovial lining contains perivascular aggregates of predominantly CD4+ T cells, B cells, and plasma cells (Kurosaka and Ziff 1983). Although it is difficult to distinguish the effects of preferential recruitment from preferential retention, mechanisms of lymphocyte migration must evolve in established chronic inflammation which facilitate the migration of B cells and other phenotypes of lymphocyte that do not migrate into the tissues within the limited time-frame of a cutaneous delayed-type hypersensitivity response. As discussed below, this might be because of changes in endothelial cell adhesion molecule expression or in the profile of tissue chemoattractants.

Natural killer cells and large granular lymphocytes constitute up to 30 per cent of CD2-positive lymphocytes circulating in peripheral blood and form a heterogeneous population of cells. There are few natural killer cells in thoracic duct lymph, suggesting that they are not recirculating cells (Fox *et al*. 1984). Natural killer cells tend to be found in rheumatoid synovial fluid but not synovial tissue, suggesting that they migrate rapidly into the joint cavity (Fox *et al*. 1984).

The effect of lymphocyte sensitization

The activation state of lymphocytes is one of the key factors determining their capacity to pass into inflamed non-lymphoid tissues. Lymphocytes that have recently undergone cell division in response to antigen within secondary lymphoid tissue enter efferent lymph and pass into peripheral blood but may not recirculate back through lymphoid tissue. True lymphoblasts are preferentially retained by extralymphoid tissues such as liver, lung, lamina propria, and skin (Rose *et al*. 1978; van Dinther-Jansson *et al*. 1983). Using isoforms of CD45 to distinguish unsensitized (CD45RA positive, 'naïve') from sensitized (CD45RO positive, 'memory') small lymphocytes, it is possible to show that the large majority of T cells found in inflammatory tissues in humans are memory cells (Pitzalis *et al*. 1987). In sheep, in which the traffic of lymphocytes can be studied using in-dwelling cannulas, most lymphocytes in afferent lymph are memory cells. On the other hand, lymphocytes that show little migration to inflammatory tissues are the predominant cells migrating through lymph nodes (Mackay *et al*. 1990).

The altered migration capacity of lymphocytes following sensitization is thought to be due largely to changes in the expression of adhesion molecules and responsiveness to chemoattractants that occur upon cell division. Compared with naïve lymphocytes, memory cells have an increased surface expression of a $^{L}b_{2}$, $a^{3}b_{1}$, $a^{4}b_{1}$, and $a^{5}b_{1}$ integrins, ICAM-1, and CD44 as well as the accessory activation molecules CD2 and CD58 (LFA-3) (Dougherty *et al*. 1988; Pitzalis *et al*. 1988; Sanders *et al*. 1988; Shimizu *et al*. 1990a). This increased adhesion molecule expression of recently activated lymphocytes contributes to their enhanced capacity to adhere to cultured endothelial cells and to transmigrate across endothelial cell monolayers (see below) (Pitzalis *et al*. 1988; Damle and Doyle 1990). Furthermore, memory lymphocytes have a lower threshold for further activation (Sanders *et al*. 1989; Merkenschlager *et al*. 1991) and increased effector function (Rodrigues *et al*. 1992).

Organ-selective migration

Lymphocyte traffic in animal models is not random, since lymphocytes show preferential routes of migration. For example, sheep lymphocytes collected from intestinal lymph or efferent lymph draining peripheral lymph nodes tend to recirculate selectively through intestinal and peripheral lymph nodes, respectively (Chin and Hay 1980; Mackay *et al*. 1992). Differences therefore exist amongst secondary lymphoid organs in the mechanisms by which recirculating lymphocytes are recruited across high endothelial venules into the lymphoid parenchyma. The molecular basis for the selective lymphocyte–endothelial cell interactions involved in homing to different lymphoid organs has been most extensively studied in mice (Berg *et al*. 1989).

There is an emerging picture that the cellular differentiation that occurs upon sensitization in lymph nodes not only enables lymphocytes to emigrate into inflamed tissues (see above) but also confers selective properties to migrate to particular tissues (Picker *et al*. 1990). Lymphocytes that home to skin have been found to express a carbohydrate antigen designated cutaneous lymphocyte antigen, which acts as a ligand for E-selectin (see below) (Picker *et al*. 1993a; Santamaria Babi *et al*. 1995). In contrast, the expression of $a^{4}b_{7}$ integrin (see below) is associated with the migration of lymphocytes to the gut (Hamann *et al*. 1994). Although it is likely that the exact lymphocyte–endothelial interactions that result in lymphocyte recruitment into synovium are distinct from those involved in adhesion of lymphocytes to endothelial cells in peripheral lymph nodes and mucosal tissues (Salmi *et al*. 1995), whether or not there is a particular subset of lymphocytes that selectively homes to inflamed synovium is still unclear.

Adhesion molecules

In the late nineteenth century Cohnheim (1889) postulated that a molecular change in the lining of the blood vessels in inflammation made the vessel wall more adhesive and allowed leucocytes to bind. In recent years monoclonal antibody and DNA technology has led to the identification of a number of leucocyte and endothelial cell adhesion molecules which participate in leucocyte diapedesis. Many of the same molecules are also concerned with the adhesion of leucocytes to cells other than endothelial cells and to components of extracellular matrix within the tissues (Springer 1990).

Selectins

The selectin family is a relatively recently described group of single-chain integral membrane glycoproteins that provide an adhesion mechanism for the tethering of circulating leucocytes to endothelial cells, platelets, or other leucocytes under conditions of flow (Bevilacqua and Nelson 1993). The primary structure of selectins is characterized by a terminal lectin-like domain, a proximal epidermal growth factor-like domain, and a variable number of short consensus repeats with homology to complement regulatory molecules such as decay accelerating factor, complement receptor 1 (CR1), and complement receptor 2 (CR2). The lectin domains of selectins are similar to mammalian asialoglycoprotein receptors and mannose-binding proteins, and have in common a functional dependence upon extracellular calcium ions (Drickamer 1988). There are three recognized molecules within the selectin family, each designated on the basis of the cell type on which the molecule was first characterized (Table 1). Thus L-selectin was first detected on leucocytes, P-selectin on platelets, and E-selectin on endothelial cells. Each of the three molecules is encoded by genes that lie within 300 kilobases of each other on chromosome 1, suggesting that they are derived from duplications and mutations of a single gene.

	L-selectin	P-selectin	E-selectin
Synonyms	CD62L, gp90[MEL]; Leu-8, TQ1, LECAM-1	CD62P, PADGEM, GMP-140	CD62E, ELAM-1
Molecular mass	75–100 kDa	140 kDa	110–115 kDa
Short consensus repeats	2	8–9	6
Distribution	Leucocytes	Endothelial cells, platelets	Endothelial cells
Regulation	Constitutive and shed upon activation	Rapid translocation to cell surface and transcriptional	Transcriptional
Cells bound	Endothelial cells, neutrophils	Leucocytes	Leucocytes
Carbohydrates	Sulphated sialylated and fucosylated carbohydrates	Sialyl Lewis x	Sialyl Lewis x, CLA
Counter-receptors	GLYCAM-1, CD34, ß-gp200, ?other	PSGL-1, ?other	PSGL-1, ESL-1, ?other

Table 1 Selectins

L-selectin

L-selectin is the smallest of the three selectins, consisting of an N-terminal lectin domain, an epidermal growth factor-like domain, two short consensus repeat sequences, a transmembrane domain, and a cytoplasmic domain. The cytoplasmic domain interacts with the cytoskeleton via a-actinin, and is necessary for L-selectin function (Pavalko *et al*. 1995). The molecular mass of L-selectin ranges from between 90 and 100 kDa (neutrophils) and between 74 and 90 kDa (lymphocytes) dependent upon differential post-translational modification. L-selectin was first identified by a monoclonal antibody (MEL-14), which inhibited adhesion of murine lymphocytes to the high endothelial venules of peripheral lymph nodes in an *in vitro* adhesion assay on frozen tissue sections (Gallatin *et al*. 1983). The MEL-14 antigen (mouse L-selectin) was then extensively characterized as a lymphocyte 'homing receptor', which mediates the primary adhesion of lymphocytes to the high endothelial venules of peripheral lymph nodes during recirculation. However, L-selectin was subsequently found to be expressed by most circulating leucocytes apart from a subset of memory T lymphocytes (Picker *et al*. 1993b), and this molecule is now considered to have an important role in the emigration of leucocytes into inflamed tissues as well as in lymphocyte recirculation. Apart from its role in supporting lymphocyte–endothelial cell adhesion, L-selectin also mediates the adhesion of neutrophils to other neutrophils, and may therefore play an important part in neutrophil intravascular aggregation during systemic vascular activation (Bargatze *et al*. 1994).

P-selectin

P-selectin is the largest of the three selectins, having eight or nine short consensus repeat sequences, depending upon alternative splicing of the gene. It was first characterized as a 140 kDa glycoprotein found in platelet a-granules and which appeared on the platelet surface after activation by thrombin. Subsequently, P-selectin was found to be synthesized also by endothelial cells and stored within intracellular granules known as Weibel–Palade bodies. As with platelets, P-selectin is rapidly translocated and expressed on the endothelial cell surface upon activation (see below). Additionally, recent evidence indicates that P-selectin expression by endothelial cells is cytokine-inducible (see below). P-selectin mediates leucocyte adhesion to platelets in addition to their adhesion to endothelial cells (E. Larsen *et al*. 1989).

E-selectin

The observation that neutrophils bind avidly to endothelial cells activated *in vitro* by interleukin-1 or tumour necrosis factor led to the generation of a monoclonal antibody which inhibited neutrophil adhesion to cytokine-stimulated endothelial cells (Bevilacqua *et al*. 1987). The molecule recognized by this monoclonal antibody was first designated endothelial leucocyte adhesion molecule-1 (ELAM-1), but is now known as E-selectin (CD62E). The crystal structure of E-selectin shows a rather linear molecule with little interaction between the epidermal growth factor-like domain and the lectin domain (Graves *et al*. 1994). Mutations that lead to loss of adhesive function are mainly clustered on the top face of the lectin domain close to the site of calcium binding, consistent with a primary role of calcium ions in selectin–ligand interactions. As discussed below, expression of the E-selectin gene is dependent upon endothelial activation by cytokines and other factors.

Selectin ligands

With the discovery of the primary sequences of the three selectins, it became clear that each of the extracellular N-terminal domains is similar to previously characterized C-type animal lectins. Consistent with this lectin structure, each of the three selectins has been found to bind carbohydrate determinants, expressed either on glycolipids or glycoproteins (Lasky 1992). Whilst it is now well established that selectins bind oligosaccharide motifs, the protein and/or lipid counter-receptors which present the oligosaccharides are less well defined. The regulation of expression of selectin ligands is complex since differential expression of glycosyl-transferases may result in potential protein- or lipid-counter-structures being appropriately glycosylated for selectin binding only in particular tissues or in response to specific cellular activation events (Kuijpers 1993).

Oligosaccharide ligands

Although there are differences between the selectins in their binding affinities for different carbohydrate structures, there is a considerable overlap in carbohydrate specificity. Thus each of the three selectins binds sialylated, fucosylated tetrasaccharide—sialyl-3-fucosyl- *N*-acetyllactosamine (**sialyl-Lewis x**). In the case of L-selectin, the oligosaccharide ligand also needs to be sulphated (Green *et al*. 1992; Hemmerich *et al*. 1995). The importance of sialyl-Lewis x for selectin binding is demonstrated by the recently described leucocyte adhesion deficiency type II in which an abnormality in fucosylation of glycoproteins leads to absence of sialyl-Lewis x expression by leucocytes and reduced selectin-mediated adhesion *in vivo* (Von Andrian *et al*. 1993). Affected individuals have a spectrum of clinical abnormalities which overlap with those of leucocyte adhesion deficiency type I (see below) (Etzioni *et al*. 1992).

Carbohydrate ligands for P- and E-selectin are found on neutrophils, eosinophils, natural killer cells, subsets of T lymphocytes, and activated B lymphocytes (Picker *et al*. 1991a; Shimizu *et al*. 1991; Damle *et al*. 1992a; Moore and Thompson 1992; Pinola *et al*. 1994; Postigo *et al*. 1994). In the case of CD45RO-positive memory T lymphocytes which enter skin, the carbohydrate which binds E-selectin is now thought to be the cutaneous lymphocyte antigen (see above), a complex carbohydrate related to sialyl-Lewis x (Berg *et al*. 1991). T lymphocytes isolated from either synovial tissue or synovial fluid have also been found to be capable of binding E-selectin *in vitro* (Postigo *et al*. 1992), although the carbohydrate ligand involved in this adhesion is not yet known. The relative contributions of P- and E-selectin to lymphocyte traffic are not yet established, since the T lymphocytes that bind P-selectin do not appear to be the same population as binds E-selectin (Rossiter *et al*. 1994).

Carbohydrate-bearing selectin counter-receptors

Most of the glyoproteins that have been so far characterized as presenting specific carbohydrate ligands to selectins are sialo-mucins, which have a serine-, threonine-, and proline-rich extracellular structure decorated with *O*-linked oligosaccharides which act as the selectin ligands. Glycoproteins in mouse lymph nodes which present sulphated carbohydrates to L-selectin include GLYCAM-1, CD34, and Sgp200 (Lasky 1992; Baumhueter *et al*. 1993; Hemmerich *et al*. 1995). Furthermore, L-selectin can bind *O*-linked oligosaccharides presented by MadCam-1 (see below) isolated from mucosal lymphoid tissue, suggesting that in mucosal lymphoid vessels L-selectin-mediated rolling on MadCam-1 may precede lymphocyte–Peyer's patch adhesion molecule-1(**LPAM-1**)-mediated firm adhesion (see below) (Berg *et al*. 1993). Although there is good evidence for inducible L-selectin-mediated adhesion on cultured endothelial cells *in vitro* (Spertini *et al*. 1991; Norgard-Sumnicht *et al*. 1993) and in non-lymphoid tissues in inflammation (Michie *et al*. 1993; Faveeuw *et al*. 1994; Lee and Sarvetnick 1994; Whyte *et al*. 1994), the ligands involved have not yet been characterized.

The principal counter-receptor for P-selectin is probably the recently described P-selectin binding glycoprotein-1 (**PSGL-1**). This is a 220 kDa sialylated homodimer with a disulphide bond situated extracellularly near the membrane (Sako *et al*. 1993). In addition to binding carbohydrates expressed by proteins, P-selectin may also bind myeloid cell sulphatides (Aruffo *et al*. 1991). Although PSGL-1 binds E-selectin, the affinity is weak and 50-fold lower than for P-selectin. A number of other candidate molecules have been proposed as counter-receptors for E-selectin, including L-selectin (Picker *et al*. 1991b), CD66 (Kuijpers *et al*. 1992), b_2-integrins (Asada *et al*. 1991), and a 150 kDa (reduced) glycoprotein on murine myeloid cells designated ESGL-1 (Steegmaler *et al*. 1995).

Whilst the lectin domain of selectins mediates adhesion by binding carbohydrate ligands, the epidermal growth factor-like domains and short consensus repeat sequences may also be important in determining the binding specificity of selectins. For example, the epidermal growth factor-like domain of P-selectin may be involved in protein–protein interactions between the selectin and the oligosaccharide-bearing protein counter-receptors (see below) (Kansas *et al*. 1994).

Integrins

The integrins constitute a large family of non-covalently linked ab-heterodimeric glycoproteins (subunits of 95 to 200 kDa) that are so named because of their integral function in connecting the intracellular cytoskeleton with the extracellular environment (Hynes 1992). One or more integrins is found on most animal cells, with the characteristics of the integrin(s) affecting the other types of cell and matrix components with which the cell can interact. The importance of this family of molecules is supported by their degree of conservation between species, with homologous heterodimers being found in the *Drosophila* fruit fly and in the *Xenopus* toad. A characteristic feature of integrins is their functional dependence upon divalent cations (Ca^{2+}, Mg^{2+}). Integrin a-subunits contain repeated segments which putatively act as cation binding sites and which are important not only for the association of a- and b-chains but also in influencing the affinity with which ligand can be bound (see below).

Classification of integrins

At least 15 human a-chains and 8 b-chains have been defined, with ab-associations grouping into subfamilies sharing a particular a- or b-chain (Fig. 3). The ligand specificity of individual integrins depends not only on the particular ab-association but also on the cell type in which the integrin is expressed and upon the state of cellular activation. In general there is considerable primary sequence similarity amongst the different a-chains and also amongst the different b-chains. It should be noted however that despite the similarity between extracellular regions of different integrins, there are substantial differences between the C-terminal cytoplasmic domains of different subunits. Integrins may therefore differ markedly in their signalling roles and relationships with the cytoskeletal proteins, and therefore perform distinct functions (Chan *et al*. 1992; Pasqualini and Hemler 1994).

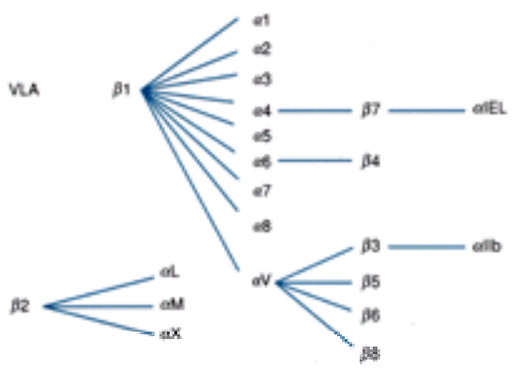


Fig. 3 The associations between integrin a- and b-subunits, resulting in integrin subfamilies.

It is convenient to classify the integrins involved in leucocyte traffic into the Leu-CAM (b_2), a^4, and VLA (b_1) sub-families (Table 2).

Integrin	Synonyms	Distribution		Ligands
		Leucocytes	Other	
$\alpha^L\beta_2$	LFA-1, CD11a/CD18	All		ICAM-1, -2, -3
$\alpha^M\beta_2$	Mac-1, CR3, CD11b/CD18	Prm, Eo, Bp, Mo, LGL		ICAM-1, iC3b, fibrinogen, factor X, β-glucan
$\alpha^X\beta_2$	p150,95, CR4, CD11c/CD18	Prm, Eo, Bp, Mo, LGL		Fibrinogen, iC3b, other
$\alpha_1\beta_1$	VLA-1, CD49a/CD29	Mo, late-activated T cells	EC, Fb	Collagens I and IV, laminin
$\alpha_2\beta_1$	VLA-2, CD49b/CD29	Mo, B, late-activated T cells	Pl, EC, Ep, Fb	Collagens I and IV, laminin
$\alpha_3\beta_1$	VLA-3, CD49c/CD29	Mo, B, late-activated T cells	EC, Fb, Ep	Laminin, fibronectin, collagen I
$\alpha_4\beta_1$	VLA-4, CD49d/CD29	Memory T cells, Bp, Eo, Mo	Fb	VCAM-1, fibronectin
$\alpha_5\beta_1$	VLA-5, CD49e/CD29	Memory T cells, B, Mo	EC, Fb, Ep	Fibronectin
$\alpha_6\beta_1$	VLA-6, CD49f/CD29	Memory T cells, B, Mo	EC	Laminin
$\alpha_7\beta_1$	LPAM-1	Memory T cells, Bp, Eo, Mo		MAdCAM-1, VCAM-1, fibronectin
$\alpha^E\beta_7$	HML-1	Mucosa-associated T cells		E-cadherin
$\alpha_V\beta_3$	Vitronectin receptor	Activated T cells, Mo, NK	EC	Vitronectin, fibronectin, laminin, Thrombospondin, fibrinogen, von Willebrand factor, CD23

(LFA-1, lymphocyte function associated antigen-1; LPAM-1, lymphocyte Peyer's patch adhesion molecule-1; VLA, very late activation antigen; Bp, basophil; Eo, eosinophil; EC, endothelial cell; Ep, epithelial cell; Fb, fibroblasts; LGL, large granular lymphocytes; Mo, monocytes; NK, natural killer cell; Prm, polymorphonuclear leucocytes; T, T cells; Pl, platelets; HML-1, human mucosal lymphocyte-1.

Table 2 Integrins involved in lymphocyte traffic

The Leu-CAM proteins (b_2-integrins)

The b_2(Leu-CAM)-integrins consist of a^Lb_2 (lymphocyte function associated antigen (**LFA**)-1, CD11a/CD18), a^Mb_2 (Mac-1, CD11b/CD18, CR3), and a^Xb_2 (p150,95, CD11c/CD18, CR4) (Larson and Springer 1990). The a-chains of b_2-integrins, as well as the a^1- and a^2-subunits (see below), are unusual in each having an inserted 'I domain'of 180 to 200 amino acids. The I domain is critically involved in the binding of metal ions (e.g. Mg^{2+}) and the mediation of the protein–protein interaction of the integrin with its ligand (Landis *et al*. 1994; Lee *et al*. 1995).

The expression of b_2-integrins is restricted to myeloid and lymphoid cells, with a $^Lb_2$ being found on all leucocytes and a $^Mb_2$ and a^Xb_2 on myeloid cells and relatively small lymphocyte subpopulations. a^Lb_2 is the b_2-integrin that is best established as being important for lymphocyte function. However, although a $^Mb_2$ is particularly involved in the extravasation of neutrophils and monocytes, a recent report suggests that this integrin also contributes to the emigration into inflamed tissues of some CD8+ T lymphocytes (Nielson *et al*. 1994).

Integrins of the b_2-subfamily are important for numerous aspects of leucocyte migration and function. Striking evidence for this can be found in the rare, hereditary, leucocyte adhesion deficiency type 1 in which leucocytes lack surface expression of all three b_2-integrins due to failure of synthesis of the b_2-subunit (Arnaout 1990). Infants with b_2-integrin deficiency are often noticed to have delayed separation of the umbilical cord, reflecting the role of phagocytes in this process. The clinical course of b_2-integrin deficiency is characterized by recurrent pyogenic infections, often leading to death from septicaemia in the first or second decade. Whilst these patients have strikingly high peripheral blood leucocyte counts (up to 150×10^9/l), infected tissue is characterized by a poverty of neutrophils, reflecting the failure of leucocytes to emigrate from the blood (Fig. 4).

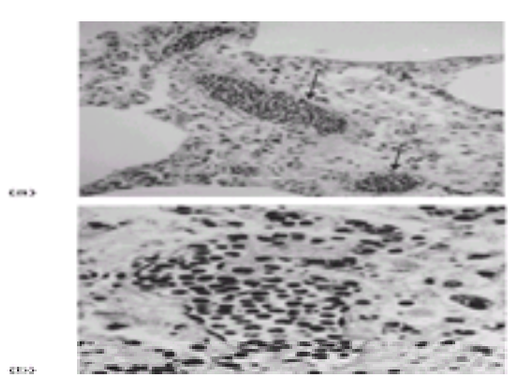


Fig. 4 Autopsy of infected lung tissue from a 19-year-old man with b_2-integrin deficiency, demonstrating the inability of neutrophils to extravasate into the tissues: (a) septal capillaries of the lung are filled with neutrophils (arrows) in contrast to the alveoli which contain macrophages and desquamated alveolar lining cells. Haematoxylin and eosin, magnification × 72; (b) the intravascular cells are mature and immature neutrophils. Haematoxylin and eosin, magnification × 220. (Davies *et al*. 1991, by courtesy of Blackwell Scientific Publications.)

a^4-integrins

In spite of the abnormalities of leucocyte migration in b_2-integrin deficiency, problematic viral infections are rare and patients are able to mount satisfactory humoral

and cell-mediated immune responses. Moreover, the chronically inflamed tissues contain some lymphocytes, monocytes, and eosinophils despite the marked reduction in neutrophils (Fig. 4). These observations, together with the finding that anti-b_2-integrin monoclonal antibodies failed to inhibit fully the adhesion of T lymphocytes to cytokine-stimulated endothelial cells *in vitro*, suggested the existence of an alternative mechanisms of integrin-mediated lymphocyte-endothelial cell adhesion (Haskard *et al*. 1986). b_2-integrin-independent leucocyte adhesion is now known to be mediated at least in part through integrins containing the a^4-subunit, which is expressed by lymphocytes, monocytes, eosinophils, basophils, and natural killer cells but only at very low levels on neutrophils (Lobb and Hemler 1994).

The a^4-subunit is unusual in having a cleavage site near the middle of the extracellular domain, leading to the generation of fragments of 70 and 80 kDa (Hemler *et al*. 1990). It can associate with either the b_1-subunit to form very late antigen-4 (**VLA-4**) (Hemler *et al*. 1990) or with the b_7-subunit to form lymphocyte–Peyer's patch adhesion molecule-1 (**LPAM-1**) (Holzmann and Weissman 1989). Interestingly, CD4+ memory T lymphocytes express either the b_1- or the b_7-subunit in conjunction with the a^4-subunit but not usually both, and which b-chain is used appears to depend on whether the lymphocyte was sensitized in peripheral (b_1) or mucosal (b_7) tissues (see above) (Schweighoffer *et al*. 1993; Erle *et al*. 1994). Both a^4-integrins bind VCAM-1 (see below), although VLA-4 on resting lymphocytes has a higher affinity than LPAM-1 for this molecule (Elices 1990; Ruegg *et al*. 1992). In contrast LPAM-1 but not VLA-4 binds MadCam-1 (see below) and is involved in selective lymphocyte traffic to mucosal tissues (Berlin *et al*. 1993). Both VLA-4 and LPAM-1 also bind the alternatively spliced form of fibronectin containing the CS-1 peptide sequence, which is therefore a possible alternative ligand to VCAM-1 on rheumatoid synovial endothelial cells (Elices 1994).

a^4-integrins are minimally expressed by neutrophils and, therefore, there is considerable interest in their role in supporting the selective traffic of mononuclear cells into chronic inflammatory tissues. Some evidence for this comes from the ability of antibodies against a^4-integrins to block the adhesion of lymphocytes to rheumatoid synovial endothelial cells in frozen-section adhesion assays (van Dinther-Janssen *et al*. 1991; Elices 1994).

VLA (very late antigen) proteins (b_1-integrins)

The b_1-integrin subunit has at least eight alternative partners in addition to the a^4-subunit, and collectively the integrins involving the b_1-integrin subunit are known as the very late antigen (**VLA**) subfamily, named after the detection of VLA-1 (a^1b_1) and VLA-2 (a^2b_1) on lymphocytes 2 to 4 weeks after the onset of T-cell proliferation *in vitro*, at a time that is late relative to the expression of other activation antigens (Hemler 1990). The VLA integrins are widely expressed on other cells besides leucocytes and, with the exception of VLA-4, are mainly concerned with adhesion to extracellular matrix components such as collagen (VLA-1, -2, -3), laminin (VLA-1, -2, -3, -6), and fibronectin (VLA-3,-4,-5). In the case of fibronectin, there are separate binding sites for VLA-4 and VLA-5, and the site utiliszd for lymphocyte adhesion may depend not only on the state of lymphocyte activation but also on differential alternative splicing of the fibronectin molecule (Guan and Hynes 1990).

Following stimulation of T-cell proliferation, there is a gradual increase over 3 to 4 weeks in expression of VLA-1 and VLA-2 and, to a lesser extent, VLA-3, -4, and -5. In contrast, the expression of VLA-6 on late-activated T cells is less than on resting cells. These alterations in VLA-integrin expression, together with variation in expression between lymphocyte subsets, leads to considerable heterogeneity between lymphocytes in their expression of receptors for the various components of extracellular matrix.

Control of leucocyte integrin function

A hallmark of integrins is their capacity to undergo rapid and reversible changes in affinity and avidity for their ligands. Changes in integrin affinity, which are thought to be related to conformational alterations in the ab-complex, are controlled in part from signals from within the cell ('inside-out signalling') and are dependent upon the cytoplasmic domains. Inside-out signalling to integrins occurs upon cross-linking a number of different surface proteins, including the T-cell receptor, CD7, CD28, CD31, and CD44 (Dustin and Springer 1989; Belitsos *et al*. 1990; Koopman *et al*. 1990; Shimizu *et al*. 1992; Tanaka *et al*. 1992), and probably also occurs in response to the action of chemoattractants on their G-protein-coupled receptors (see below). However, inside-out signalling is not the only mechanism whereby integrin affinity is increased, since affinity changes also occur upon contact of the integrin with its ligand ('ligand-induced adhesion') (Cabañas and Hogg 1993). This may be an important mechanism whereby adhesion can be strengthened for the ligand once initial contact has been made. Furthermore, changes in integrin conformation upon contact with ligand may signal changes in the affinity of the molecule for other ligands (Sanchez-Aparicio *et al*. 1994; Li *et al*. 1995), thus promoting a series of adhesion events that could guide a cell through a particular activity.

The ability of leucocytes to adhere through integrin-mediated mechanisms is also influenced by avidity changes due to clustering of the heterodimers at sites of cell contact. These are related to interactions between integrin a- and b-chain cytoplasmic domains and cytoskeletal proteins (Kupfer and Singer 1989; Peter and O'Toole 1995). Interactions between integrins and the cytoskeleton probably occur as secondary 'postreceptor' events, consequent to the initial adhesion interaction, as suggested by studies in which cross-linking of b_2-integrins leads to the polymerization of actin (Lofgren *et al*. 1993).

Integrins as costimulatory molecules

The ability of integrins to transduce signals into the cell is referred to as 'outside-in signalling', and is fundamental to their role in linking adhesion to the biology of the cell (Clark and Brugge 1995). Besides directly regulating cytoskeletal organization, outside-in signalling through integrins can provide important costimulatory signals which greatly influence lymphocyte responsiveness to other stimuli occurring in parallel. Thus surface integrins (e.g. a^Lb_2 and a^4b_1) can act as costimulatory molecules for lymphocyte proliferation upon binding to appropriate ligands (Shimizu *et al*. 1990b; Damle *et al*. 1992b). This costimulatory role of integrins may account for the prolonged immunosuppressive effects of administering antileucocyte integrin antibodies *in vivo* (see below).

Immunoglobulin-like molecules

The immunoglobulin (**Ig**) supergene family consists of a large number of molecules that are structurally related to antibodies (Williams and Barclay 1988). Within this family falls a subgroup of molecules (ICAM-1,-2,-3, VCAM-1, MadCAM-1) which act as adhesion ligands for integrins and which contain in their N-terminals variations of a conserved integrin-binding motif (Clements *et al*. 1994). Other important lymphocyte surface molecules that are contained within the immunoglobulin superfamily are HLA class I and II, CD4, CD8, CD2, CD31 (PECAM-1), and CD58 (LFA-3).

Intercellular adhesion molecule-1, -2, -3 (ICAM-1, -2, -3)

The finding that normal lymphocytes were able to adhere to lymphoid cells from a patient with b_2-integrin deficiency (see above) indicated that a ligand for LFA-1 (a^Lb_2) was present on the b_2- integrin deficient cells. By immunizing mice with b_2-integrin-deficient lymphocytes it was possible to raise a monoclonal antibody (RR1/1) that inhibited the LFA-1-dependent adhesion of normal cells (Rothlein *et al*. 1986). The antigen recognized by RR1/1 was designated intercellular adhesion molecule-1 (**ICAM-1**, CD54) and found to be a single-chain glycoprotein with a molecular mass from 88 to 110 kDa, depending upon tissue-specific glycosylation. ICAM-1 is a transmembrane protein with five C2-type immunoglobulin domains. The LFA-1 binding sites are localized on the two Ig domains nearest the N-terminal (Staunton *et al*. 1990). In contrast to their arrangement in antibodies, the Ig domains of ICAM-1 are unpaired and linear, with a potential bend between domains 2 and 3. In addition to acting as a ligand for LFA-1, ICAM-1 also binds Mac-1 (a^Mb_2) by a site on domain 3 (Diamond *et al*. 1991). The use of different binding sites on ICAM-1 by LFA-1 and Mac-1 suggests a mechanism for the strengthening of cellular adhesion to ICAM-1, which could be utilized by monocytes and other leucocytes which express both of these b_2-integrins (see above).

Most cells can express ICAM-1 if appropriately activated. Endothelial cells are unusual compared with other cell types in showing a relatively high level of constitutive ICAM-1 expression. The presence of ICAM-1 on extravascular cells in inflammation indicates that this molecule may be important not only for leucocyte–endothelial interactions but also for leucocyte migration and interaction with other cells within the tissues (see below).

Although ICAM-1 is constitutively expressed by endothelial cells, anti-ICAM-1 monoclonal antibodies were found not to be as effective as anti-LFA-1 monoclonal antibodies in inhibiting LFA-1-dependent lymphocyte adhesion, suggesting the existence of other LFA-1 ligands. Using a direct functional cDNA cloning strategy a novel LFA-1 ligand was isolated and designated ICAM-2 (Staunton *et al*. 1989). ICAM-2 (CD102) has a molecular mass of 55 kDa and can be considered as a truncated form of ICAM-1, consisting of two Ig-like domains which are 34 per cent similar to the two most N-terminal domains of ICAM-1. As might be expected from the lack of a third Ig domain, ICAM-2 is not thought to bind Mac-1. ICAM-2 has a high constitutive expression on endothelial cells, with a density 10-fold greater than that of the constitutive expression of ICAM-1 (De Fougerolles *et al*. 1991), and this may therefore be the more important LFA-1 ligand on endothelium in uninflamed tissues. However, ICAM-2 expression does not appear to be up-regulated by proinflammatory stimuli, and consequently ICAM-1 may be the predominant LFA-1 ligand

on endothelial cells in inflamed tissues.

Although it is likely that ICAM-1 and ICAM-2 are the only ligands on endothelial cells which can bind LFA-1, there is a third LFA-1 ligand in the form of ICAM-3 (De Fougerolles and Springer 1992). ICAM-3 (CD50, molecular mass 124 kDa) is a highly glycosylated protein with an extracellular portion of five Ig domains (Fawcett *et al*. 1992; Vazeux *et al*. 1992). It is highly expressed on most resting leucocytes and may be particularly important in initiating the interactions of lymphocytes with other leucocytes. Consistent with this, ICAM-3 has been found to be an active signalling molecule (Arroyo *et al*. 1994; Campanero *et al*. 1994).

Vascular cell adhesion molecule-1 (VCAM-1)

Vascular cell adhesion molecule-1 (**VCAM-1**, CD106) is a single chain glycoprotein (molecular mass 96 kDa) which is approximately 26 per cent similar to ICAM-1 (Osborn *et al*. 1989). It is encoded by a single gene on chromosome 1. Although the original form of VCAM-1 to be sequenced had six C2-type Ig domains, it is now clear that the predominant form on endothelial cells has an additional domain, which is very similar in primary structure to the N-terminal domain 1 and which is situated between domains three and four of the six domain form. Whereas VLA-4 binds domain 1 on both the six and seven domain forms, VLA-4 binding can additionally occur to the homologous fourth domain of the seven domain form, providing a mechanism for the strengthening of VLA-4-dependent adhesion. The crystal structure of the N-terminal two domains of VCAM-1 has recently been described (Jones *et al*. 1995). VCAM-1 is expressed by several cell types apart from endothelial cells, including germinal centre dendritic cells, interdigitating dendritic cells, Kupffer cells, renal proximal tubule cells, and type B synovial lining cells (Rice *et al*. 1991; Wilkinson *et al*. 1993).

Mucosal addressin cell adhesion molecule-1(MadCAM-1)

Mucosal addressin cell adhesion molecule-1 (**MadCAM-1**) is a 58 to 66 kDa glycoprotein which is expressed by endothelial cells in high endothelial venules of Peyer's patches and mesenteric lymph nodes and also on blood vessels in lamina propria of the small intestine and in some forms of extramucosal inflammation (Streeter *et al*. 1988; O'Neill *et al*. 1991; Hanninen *et al*. 1993). It is has three extracellular immunoglobulin domains which have structural similarities to ICAM-1, VCAM-1, and IgA1 (Briskin *et al*. 1993). The portion between the membrane proximal (IgA1-like) and intermediate (VCAM-1-like) Ig domains is occupied by a mucin region rich in *O*-linked sugars, which in mucosal lymphoid tissues are capable of binding L-selectin (Berg *et al*. 1993). MadCAM-1 acts as an endothelial ligand for the a^4b_7-integrin (Berlin *et al*. 1993).

Platelet–endothelial cell adhesion molecule-1(PECAM-1)

Platelet–endothelial cell adhesion molecule-1 (**PECAM-1**, CD31) is a six Ig domain glycoprotein which can self-associate in homotypic interactions between cells (Fawcett *et al*. 1995), or can bind heterophilically to the a^vb_3 -integrin on leucocytes (Piali *et al*. 1995). PECAM-1 is expressed by endothelial cells, platelets, neutrophils, monocytes, and a subset of T cells. On endothelial cells it is found at points of intercellular contact between endothelial cells, and may play a role in contact inhibition of endothelial cell migration and proliferation (Fawcett *et al*. 1995). PECAM-1 is also capable of transducing signals into cells and may be involved in modulating the function of surface integrins (Tanaka *et al*. 1992). Recent evidence points to a role of CD31 in leucocyte transmigration though endothelium into the tissues (see below).

CD44

CD44 (Pgp-1, HCAM, Hermes) is a highly polymorphic cell surface glycoprotein with similarity to cartilage link proteins. It has a wide cellular distribution which includes lymphoid and myeloid cells, endothelial cells, epithelial cells, and fibroblasts (Pals *et al*. 1989). Heterogeneity of CD44 is partly attributable to alternative splicing of 10 exons lying in tandem within the extracellular portion of the CD44 gene (Screaton *et al*. 1992), and partly to differential glycosylation (Jalkanen *et al*. 1988). Thus CD44 can be modified by *N*- and/or *O*-linked glysosylation and may also express chondroitin sulphate or heparan sulphate side-chains (Jackson *et al*. 1995). CD44 acts as a receptor for several extracellular matrix components including hyaluronan (Aruffo *et al*. 1990) and fibronectin (Carter and Wayner 1988). Furthermore, CD44 may support the adhesion of lymphocytes to synovial endothelial cells through binding an as yet unidentified ligand. Apart from acting as an adhesion molecule, CD44 is also a signalling molecule capable of mediating an up-regulation of LFA-1-dependent adhesion (Belitsos *et al*. 1990; Koopman *et al*. 1990), and also the costimulation of T-lymphocyte proliferation (Huet *et al*. 1989). Lastly, variants of CD44 that express proteoglycan side-chains may bind cytokines and growth factors to the cell surface (Tanaka *et al*. 1993; Bennett *et al*. 1995).

Vascular adhesion protein-1 (VAP-1)

Vascular adhesion protein-1 (**VAP-1**) is a newly described 90 kDa adhesion molecule for lymphocytes, recognized by a monoclonal antibody generated against stromal material from rheumatoid synovium (Salmi and Jalkanen 1992). In addition to being expressed by rheumatoid synovial endothelial cells, VAP-1 is also found on high endothelial venules of peripheral lymph nodes and on endothelium in mucosal and cutaneous inflammatory tissues (Salmi *et al*. 1993). The molecular nature of VAP-1 and the ligands it binds are not yet described.

Lymphocyte chemoattractants

The directional movement of leucocytes in response to a soluble stimulus is known as 'chemotaxis'. A large number of non-specific leucocyte chemotactic factors have been described, including activated complement components (e.g. C5a), lipid metabolites (e.g. leukotrienes, platelet activating factor), and bacterial-derived peptides (e.g. F–Met–-Leu–Phe). More recently a novel class of cytokines, known as chemokines, has been described, which differ from classical chemoattractants in having marked specificity of actions on leucocytes of particular lineage (Baggiolini *et al*. 1994). In the case of lymphocytes, additional chemotactic factors include interleukin-2 and interleukin-15 (Kornfield *et al*. 1985; Wilkinson and Newman 1994; Wilkinson and Liew 1995). Most chemoattractants act on leucocytes through a newly described family of surface receptors which span the membrane seven times ('serentine receptors') and which are coupled to G proteins (Murphy 1994).

Chemokines are 8 to10 kDa proteins that are released by leucocytes and tissue cells in response to proinflammatory stimuli, of which the best described are interleukin-1 and tumour necrosis factor. Members of the family have in common four conserved cysteine residues, with two subfamilies being distinguished by whether or not an amino acid separates the first and second cysteine (i.e. CXC or CC) (Table 3). CXC chemokines are mostly encoded by genes on human chromosome 4 (q12–21) and probably derive from a single ancestral gene. In contrast, CC chemokines are encoded by genes on human chromosome 17 (q11–32). Although IL-8 (Larsen *et al*. 1989) and particularly IP-10 (Taub *et al*. 1993) have chemotactic activity for lymphocytes, CXC chemokines are predominantly active specifically on neutrophils. Conversely, CC chemokines are chemotactic for lymphocytes, monocytes, basophils, and eosinophils but not neutrophils. CC chemokines that are chemoattractant for lymphocytes include RANTES (Schall *et al*. 1990), MIP-1a and -1b (Schall *et al*. 1993), and MCP-1, -2, and -3 (Carr *et al*. 1994; Taub *et al*. 1995). Variation between lymphocyte subpopulations in responsiveness to these chemokines is probably very important in determining the selective recruitment and distribution of lymphocyte subsets in the tissues. A critical feature of many chemokines is that they are heparin-binding proteins and have the ability to bind negatively-charged glycosaminoglycans on the surface of cells or incorporated in extracellular matrix. As discussed below, this property is thought to be fundamental to how chemokines stimulate leucocyte movement, since *in vivo* leucocytes may migrate along gradients of bound chemoattractants and adhesion molecules ('haptotaxis') rather than gradients of chemoattractants in solution.

Chemokine	Synonyms	Targets
C-X-C		
Interleukin-8	IL-8	Pmn, T
ENA-78		Pmn
Platelet factor 4	PF4	Pmn
Platelet basic protein	PBP	Pmn
β-Thromboglobulin	βTG	Pmn
Connective tissue activating peptide-III	CTAP-III	Pmn
Neutrophil activating protein-2	NAP-2	Pmn
GRO α, β and γ		Pmn, Mo
Interferon-inducible protein-10	IP-10	T
C-C		
Macrophage inflammatory protein-1α	MIP-1α	B, T, Bp, Eo, Mo
Macrophage inflammatory protein-1β	MIP-1β	T, ?other
Regulated on activation, normal T expressed and secreted	RANTES	T, Mo, Eo, Bp
Monocyte chemotactic protein-1	MCP-1	T, Mo, Bp
Monocyte chemotactic protein-2	MCP-2	T, Mo, Bp, Eo
Monocyte chemotactic protein-3	MCP-3	T, ?other

B, B lymphocyte; Bp, basophil; Eo, eosinophil; Mo, monocyte; Pmn, polymorphonuclear neutrophil; T, T lymphocyte.

Table 3 Chemokines

Control of leucocyte adhesion and migration

The balance in postcapillary venules between adhesion due to intercellular molecular bonds on the one hand and factors preventing adhesion (e.g. shear stress and surface charge) on the other, is normally set at an equilibrium which allows for the routine immunosurveillance of the tissues by a small number of migrating lymphocytes. During inflammation, lymphocyte adhesion and migration are enhanced by the locally regulated alteration in expression and/or function of a number of different adhesion molecules both on endothelial cells and on lymphocytes.

The adhesion cascade

The ultrastructural appearance of leucocyte emigration was well described by the early 1960s (Marchesi and Florey 1960). Leucocytes could be seen adherent to the luminal surface of the postcapillary venules and often appeared flattened with an increased area of membrane in contact with the endothelial cell surface. Migration through endothelium was associated with the development of pseudopodia which formed the leading edge of the leucocyte and which were seen to penetrate between adjacent endothelial cells.

Over the last few years there has been a considerable increase in understanding of the molecular and cellular mechanisms involved in leucocyte–endothelial cell interactions. The series of co-operative adhesive events which are now known to underlie the emigration of leucocytes into the tissues has become known as the 'adhesion cascade'. Although the events were first described for neutrophil–endothelial cell interactions, there is now evidence that similar rules apply for monocytes (Luskinskas *et al*. 1994) and for lymphocytes (Jones *et al*. 1994). The 'adhesion cascade' model now provides a framework with which to understand the specificity of leucocyte migration into different inflammatory tissues. Thus whether or not a leucocyte emigrates from the blood into the tissues is no longer seen as being due to a single receptor–ligand interaction but dependent on the leucocyte being able to participate in a series of potentially selective adhesion and activation events (Butcher 1991; Springer 1994). Although this model provides a reasonable paradigm with which to account for the complexity of cellular traffic in inflammation, the exact conditions of emigration may well vary with the leucocyte type, the tissue, and the precise type of inflammatory insult (Bullard *et al*. 1995; Ward 1995).

Leucocyte rolling

Atherton and Born (1972) showed by direct observation of the microvasculature *in vivo* that circulating leucocytes roll along the surface of endothelium before becoming firmly adherent to the vessel wall. As a result of rolling, the velocity with which leucocytes pass through venules drops from about 2000 to 3000 μm/s to approximately 10 to 25 μm/s.

Leucocyte rolling is now thought to be mediated largely by the interactions of selectins with their carbohydrate ligands, which have a high tensile strength but rapid on and off rates (Alon *et al*. 1995a). The involvement of selectins in leucocyte rolling has now been experimentally validated with models of neutrophil and lymphocyte interactions with endothelial cells or purified adhesion ligands under conditions of flow *in vitro* (Lawrence and Springer 1991; Abbassi *et al*. 1993; Lawrence and Springer 1993), and *in vivo* (Ley *et al*. 1991; Bargatze and Butcher 1993; Doré *et al*. 1993; Mayadas *et al*. 1993). In mice P-selectin and L-selectin dominate rolling at early and later time-points after endothelial cell activation (Ley *et al*. 1995), indicating some functional specialization amongst the three selectins. The relative contribution that each of the three selectins makes in humans to leucocyte rolling in different forms of inflammation and at different stages of the inflammatory response is still uncertain. Whilst the primary function of selectin–carbohydrate interactions is to support intercellular adhesion (see below), there is some evidence that these interactions may have cell-signalling effects, both in leucocytes and in endothelial cells (Lo *et al*. 1991; Kaplanski *et al*. 1994; Laudanna *et al*. 1994; Weyrich *et al*. 1995).

Firm adhesion of leucocytes to endothelial cells

Although early experiments studying neutrophil adhesion to endothelial cells under physiological shear stresses indicated a requirement for selectin-mediated adhesion prior to the involvement of integrins, reductions in shear rates in pathological situations may allow direct b_2-integrin-mediated attachment (Gaboury and Kubes 1994). Furthermore, it is now clear that monocytes and lymphocytes may use a^4-integrins with VCAM-1 or MAdCAM-1 for the initial attachment step (Jones *et al*. 1994; Alon *et al*. 1995b; Berlin *et al*. 1995; Luskinskas *et al*. 1995). Therefore mononuclear cells may be able adhere and emigrate in vessels in which selectin-mediated adhesion is not prominent.

During rolling the leucocyte is thought to be stimulated by local mediators that activate the adhesivity of integrins, resulting in firm integrin-mediated immobilization of the rolling cell and transmigration into the tissues. However, the precise stimuli which activate a rolling lymphocyte are not yet known but are probably associated with the endothelial surface, either in the form of integral membrane proteins, such as selectins (see above), or in the form of chemokines bound to glycosaminoglycans in the endothelial glycocalyx (Tanaka *et al*. 1993; Webb *et al*. 1993). Thus rolling leucocytes are thought to sample the endothelial surface for immobilized chemoattractants, and the subsequent activation of integrin function then leads to firm integrin-mediated adhesion to Ig-superfamily molecules such as ICAM-1 and -2 and VCAM-1.

Transmigration of leucocytes through endothelium

After adhering to the luminal surface of endothelium, leucocytes transmigrate between endothelial cells and through the endothelial basement membrane before gaining access to the tissues. Judging from work performed on rat lymph nodes, this process is very rapid, with emigrating lymphocytes passing through basement membrane within 10 min (Smith and Ford 1983). The adhesion interactions which result in leucocyte transmigration are not yet well defined but include mechanisms which facilitate de-adhesion from the luminal surface of endothelium, as well as mechanisms supporting the adhesion necessary for migration into the tissues.

In the case of granulocytes and monocytes there is good evidence that L-selectin is enzymatically cleaved from the cell surface between Lys-321 and Ser-322 as part of the sequence of events taking place on the endothelial cell surface (Kahn *et al*. 1994). This mechanism may facilitate de-adhesion from the endothelial lumenal surface and may also serve to avoid inappropriate adhesion in tissues containing carbohydrate ligands to which L-selectin could bind, as in the central nervous system (Huang *et al*. 1991). The precise events that lead to the rapid shedding of L-selectin may involve cross-linking L-selectin on the lympocyte surface (Palecanda *et al*. 1992). Although L-selectin is also cleaved from the surface of lymphocytes upon activation, *in vitro* studies indicate that this process is relatively slow, and may not occur during lymphocyte emigration through the vessel wall (Bührer *et al*. 1992).

The mechanisms which guide leucocytes through the vessel wall are still being established, but appear to involve two-way signalling interactions between leucocytes and endothelial cells (Huang *et al*. 1993; Pfau *et al*. 1995). The adhesion molecules involved in transmigration include similar integrin interactions with immunoglobulin-superfamily ligands to those supporting firm adhesion (Van Epps *et al*. 1989; Kavanaugh *et al*. 1991). A role for CD31 in leucocyte transmigration through endothelium has been suggested by studies both *in vitro* (Muller *et al*. 1993) and *in vivo* (Vaporciyan *et al*. 1993; Bogen *et al*. 1994). As discussed above, this could either involve CD31–CD31 homotypic interactions or binding of CD31 to other ligands such as a $_v b_3$ -integrin.

Endothelial cells may actively promote leucocyte transmigration, particularly after cytokine activation (Oppenheimer-Marks and Ziff 1988; Furie and McHugh 1989; Moser *et al*. 1989). This increased capacity of cytokine-activated endothelial cells to support leucocyte transmigration is probably related to the increased expression of adhesion ligands and chemokines (see below). A chemoattactant gradient appears essential for emigration, since inappropriate exposure of leucocytes to chemoattractants in the vascular compartment may prevent migration (Hechtman *et al*. 1991; Huber *et al*. 1991).

The mechanisms by which leucocytes penetrate the basal lamina, which is composed of types III, IV, and V collagen, heparan suphate, laminin, and fibronectin, are not well understood but probably involve the induced secretion or surface expression of degradative enzymes (Naparstek *et al*. 1984; Romanic and Madri 1994).

Endothelial activation

Functional changes in endothelium are of paramount importance for localizing inflammatory responses. Activation of endothelial cells not only leads to the controlled expression of adhesion ligands on the luminal surface, but also results in the generation of surface-bound and secreted factors capable of affecting leucocyte function and responsiveness to subsequent stimulation within the tissues. Endothelial activation may be classified according to the different phases of the inflammatory response, with each phase being associated with a particular profile of surface adhesion molecules, thereby contributing to the migration of the particular leucocyte subsets that characteristically infiltrate different types of inflammatory lesion.

Rapid activation

Stimulation of endothelial cells with agonists such as histamine, thrombin, C5a, or C5b to C9 terminal complement complexes elicits a rapid response, which includes the secretion of factors that regulate vascular tone (e.g. prostacyclin and nitric oxide), the surface expression of platelet-activating factor, and the translocation of Weibel–Palade bodies to the plasma membrane with the release of von Willebrand factor and the surface expression of P-selectin (Hattori *et al*. 1989a; Hattori *et al*. 1989b; Geng *et al*. 1990; Foreman *et al*. 1994). The rapid and transient expression of platelet-activating factor and P-selectin is thought to lead to the early phase of neutrophil adhesion to endothelial cells that occurs within minutes of the onset of an inflammatory reaction (Zimmerman *et al*. 1990).

Subacute activation

The effects of interleukin-1, tumour necrosis factor, and lipopolysaccharide

Injection of interleukin-1, tumour necrosis factor, or lipopolysaccharide *in vivo* stimulates a subacute inflammatory response associated with marked leucocyte emigration into the tissues (Cybulsky *et al*. 1988; Binns *et al*. 1992). To a large extent this is attributable to an orchestrated induction of endothelial function, which differs from the rapid reaction to the acute endothelial cell agonists described above in that it involves the transcription of a number of primary and secondary response genes, with subsequent *de novo* protein synthesis and surface expression (Pober and Cotran 1990). The transcriptional control of much of this response involves the translocation to the nucleus of the transcription factor NFkB, and is autoregulated by the synthesis of the inhibitory protein IkBa (Read *et al*. 1994).

The effects on endothelial cells of interleukin-1, tumour necrosis factor, or lipopolysaccharide are first detectable *in vitro* after 1 to 2 h and last for a variable time, depending upon the stimulus applied and the exact response measured. These factors stimulate the expression of a number of adhesion molecules involved in the rolling and firm adhesion stages of leucocyte–endothelial cell interaction, including E-selectin, P-selectin, ICAM-1, VCAM-1, and MadCAM-1, as well as the synthesis and secretion of cytokines such as interleukins-1and -6, colony stimulating factors (CSF) G-CSF, GM-CSF, and M-CSF, platelet-derived growth factor, and a number of chemokines such as interleukin-8 and monocyte chemotactic protein-1.

Effect of interferon-g and interleukin-4

Whilst interleukin-1 and tumour necrosis factor appear to be the dominant cytokines involved in the subacute phase of endothelial cell activation, their effects can be modulated by other cytokines, which in themselves have little or no direct action on the expression of endothelial cell activation determinants. In the context of immune-mediated inflammation, the effects of interleukin-1 and particularly tumour necrosis factor on the expression of ICAM-1 and VCAM-1 may be selectively up-regulated by the actions of interferon-g and interleukin-4 respectively, suggesting a possible differential control of endothelial cell activation during T-helper 1- and T-helper 2-type immune responses (Thornhill and Haskard 1990). Lymphokines may also have selective effects on chemokine synthesis, as shown by interferon-g inducing endothelial cell expression of monocyte chemotactic protein-1 but not interleukin-8 (Brown *et al*. 1994).

Chronic activation

Understanding the nature of endothelial cell activation in chronic inflammation has been hampered by the limitations of endothelial cell cultures, which are difficult to prolong beyond a few days. On the one hand it is likely that some endothelial cell responses are down-regulated as the inflammatory process becomes chronic, either through the endogenous control of transcriptional mechanisms (Read *et al*. 1994) or through the down-regulating actions of some cytokines and growth factors such as interleukin-4 and transforming growth factor-b (Thornhill and Haskard 1990; Gamble *et al*. 1993). On the other hand, additional adhesion molecules may be expressed at later times during the development of chronic inflammation, perhaps in connection with the acquisition of vessels with high endothelial venule-like morphology (see above). Although there is very little information on this point, Michie *et al*. (1993) have observed the expression of the peripheral lymph node addressin, a ligand for L-selectin, on endothelial cells in some forms of chronic inflammation, including rheumatoid synovitis.

Lymphocyte migration within the tissues

Once leucocytes have passed through endothelium they migrate within tissues by contact interactions with other cells and/or with components of the extracellular matrix such as fibronectin, collagen, and hyaluronate. Less is known about the molecular basis of these events than is the case for leucocyte interactions with endothelial cells, although it is likely that similar principles apply. Cytokines up-regulate the expression of ICAM-1 and VCAM-1 on synovial tissue fibroblasts, providing a mechanism for enhanced interactions of these cells with lymphocytes (En Chin *et al*. 1990; Marlor *et al*. 1992; Cicuttini *et al*. 1994; Shimada *et al*. 1994). Adhesion molecule expression by synovial fibroblasts may also be directly regulated by physical contact with T lymphocytes or monocytes (Bombara *et al*. 1993). As described above, adhesion molecules play a central role not only in the migration and arrest of lymphocytes within the tissues but also in determining their response to activating stimuli.

Identification and quantification of adhesion molecules *in situ*

Much of the new understanding of the mechanisms of intercellular adhesion has been gained from investigations at the cellular and molecular level *in vitro*. Evidence that the molecules identified as important for leucocyte function *in vitro* are relevant to the *in vivo* situation comes first from a large number of immunocytochemical studies of synovium (Table 4) and other tissues (Mason and Haskard 1994). An extension of the histological approach is to use the technique first devised by Stamper and Woodruff (1976) for studying the adhesion of lymphocytes to endothelial cells in frozen tissue sections. This technique has been used with inhibitory monoclonal antibodies to implicate the involvement of a number of adhesion molecules in lymphocyte adhesion to synovial endothelial cells *in situ*, including a^L-, a^M-, b_2- and b_1-integrin subunits, L-selectin, ICAM-1, VCAM-1, alternatively-spliced fibronectin, CD44, and VAP-1 (Jalkanen *et al*. 1987; van Dinther-Janssen *et al*. 1991; Salmi and Jalkanen 1992; Fischer *et al*. 1993; Elices 1994).

Table 4 Expression of adhesion molecules in arthritis

Owing to the constraints imposed by the availability of human tissue, there is also interest in measuring the soluble forms of the three selectins, ICAM-1, ICAM-3, and VCAM-1, as serum markers of cellular activation in tissues (Gearing and Newman 1993; Mason *et al*. 1993; Carson *et al*. 1994; Martin *et al*. 1995). In the case of E-selectin, synovial endothelial activation and E-selectin expression can also be quantified and imaged non-invasively using intravenously injected, radiolabelled,

monoclonal antibody (Keelan *et al*. 1994; Jamar *et al*. 1995; Chapman *et al*. 1996).

Therapeutic possibilities

A reduction in leucocyte traffic into tissues is to be expected as part of a general down-regulation of the inflammatory response. However, although the majority of antirheumatic drugs in routine use were originally introduced into clinical practice on an empirical basis, some agents may have direct actions that are relevant to cell migration (Chapman and Haskard 1995).

Experience with thoracic duct drainage and lymphocytephoresis gave some encouragement that direct manipulation of lymphocyte traffic might have therapeutic effects in immune-mediated inflammatory diseases (Paulus *et al*. 1979; Karsh *et al*. 1981; Wahl *et al*. 1983). The rapid development in understanding of mechanisms of leucocyte traffic described in this chapter has led to renewed interest in the possibility of inhibiting inflammatory processes by blocking white cell migration *in vivo*.

Studies in animal models of arthritis and other forms of inflammation have shown that monoclonal antibodies against a number of adhesion molecules can have marked suppressive effects on leucocyte traffic and upon the inflammatory process (reviewed in Mason and Haskard 1994). Whilst inhibiting selectins has effects that are limited to leucocyte–endothelial cell interactions, inhibiting integrin interactions may have quite widespread and lasting suppressive effects on the immune response, as shown in animal models of transplantation (Isobe *et al*. 1992; Orosz *et al*. 1993) and arthritis (Iigo *et al*. 1991; Jasin *et al*. 1992; Barbadillo *et al*. 1995) and more recently for anti-ICAM-1 monoclonal antibodies in patients with rheumatoid arthritis (Kavanaugh *et al*. 1994; Davis *et al*. 1995). Recent developments which should facilitate further understanding of the therapeutic potential of lymphocyte traffic as a target for therapy include the use of immunodeficient mice engrafted with human synovium (Rendt *et al*. 1993), and the use of murine models of arthritis in gene-targeted mice (Sligh *et al*. 1993; Wilson *et al*. 1993; Labow *et al*. 1994; Bullard *et al*. 1995; Subramaniam *et al*. 1995; Tedder *et al*. 1995).

Chapter References

Abbassi, O., Kishimoto, T.K., McIntire, L.V., Anderson, D.C., and Smith, C.W. (1993). E-selectin supports neutrophil rolling under conditions of flow. *Journal of Clinical Investigation*, **92**, 2719–30.

Allen, C.A., Highton, J., and Palmer, D.G. (1989). Increased expression of p150,95 and CR3 leukocyte adhesion molecules by mononuclear phagocytes in rheumatoid synovial membranes. *Arthritis and Rheumatism*, **32**, 947–54.

Alon, R., Hammer, D.A., and Springer, T.A. (1995 *a*). Lifetime of the P-selectin-carbohydrate bond and its response to tensile force in hydrodymanic flow. *Nature*, **374**, 539–42.

Alon, R., Kassner, P.D., Woldemar Carr, M., Finger, E.B., Hemler, M.E., and Springer, T.A. (1995 *b*). The integrin VLA-4 supports tethering and rolling in flow on VCAM-1. *Journal of Cell Biology*, **128**, 1243–53.

Andrews, P., Milsom, D.W., and Stoddart, R.W. (1983). Glycoconjugates from high endothelial cells: I. Partial characterization of a sulphated glycoconjugate from the high endothelial cells of rat lymph nodes. *Journal of Cell Science*, **59**, 231–44.

Arnaout, M.A. (1990). Leukocyte adhesion molecules deficiency: its structural basis, pathophysiology and implications for modulating the inflammatory response. *Immunological Reviews*, **114**, 145–79.

Arroyo, A.G. *et al*. (1994). Induction of tyrosine phosphorylation during ICAM-3 and LFA-1-mediated intercellular adhesion, and its regulation by the CD45 tyrosine phosphatase. *Journal of Cell Biology*, **126**, 1277–86.

Aruffo, A., Stamenkovic, I., Melnick, M., Underhill, C.B., and Seed, B. (1990). CD44 is the principal cell surface receptor for hyaluronate. *Cell*, **61**, 1303–13.

Aruffo, A., Kolanus, W., Walz, G., Fredman, P., and Seed, B. (1991). CD62/P-Selectin recognition of myeloid and tumour cell sulfatides. *Cell*, **67**, 35–44.

Asada, M., Furukawa, K., Kantor, C., Gahmberg, C.G., and Kobata, A. (1991). Structural study of the sugar chains of human leukocyte cell adhesion molecules CD11/CD18. *Biochemistry*, **30**, 1561–71.

Atherton, A. and Born, G.V.R. (1972). Quantitative investigations of the adhesiveness of circulating polymorphonuclear leukocytes to blood vessel walls. *Journal of Physiology*, **222**, 447–74.

Baggiolini, M., Dewald, B., and Moser, B. (1994). Interleukin-8 and related chemotactic cytokines—CXC and CC chemokines. *Advances in Immunology*, **55**, 97–179.

Barbadillo, C., G-Arroyo, A., Salas, C., Mulero, J., Sanchez-Madrid, F., and Andreu, J.L. (1995). Anti-integrin immunotherapy in rheumatoid arthritis: protective effect of anti-a^4 antibody in adjuvant arthritis. *Springer Seminars in Immunopathology*, **16**, 427–36.

Bargatze, R.F. and Butcher, E.C. (1993). Rapid G protein-regulated activation event involved in lymphocyte binding to high endothelial venules. *Journal of Experimental Medicine*, **178**, 367–72.

Bargatze, R.F., Kurk, S., Butcher, E.C., and Jutila, M.A. (1994). Neutrophils roll on adherent neutrophils bound to cytokine-induced endothelial cells via L-selectin on rolling cells. *Journal of Experimental Medicine*, **180**, 1785–92.

Baumhueter, S. *et al*. (1993). Binding of L-selectin to the vascular sialomucoin CD34. *Science*, **262**, 436–8.

Belitsos, P.C., Hildreth, J.E.K., and August, J.T. (1990). Homotypic cell aggregation induced by anti-CD44 (Pgp-1) monoclonal antibodies and related to CD44 (Pgp-1) expression. *Journal of Immunology*, **144**, 1661–70.

Bennett, K.L. *et al*. (1995). CD44 isoforms containing exon V3 are responsible for the presentation of heparin-binding growth factor. *Journal of Cell Biology*, **128**, 687–98.

Berg, E.L. *et al*. (1989). Homing receptors and vascular addressins: cell adhesion molecules that direct lymphocyte traffic. *Immunological Reviews*, **108**, 5–18.

Berg, E.L. *et al*. (1991). The cutaneous lymphocyte antigen is a skin lymphocyte homing receptor for the vascular lectin endothelial cell-leukocyte adhesion molecule-1. *Journal of Experimental Medicine*, **174**, 1461–6.

Berg, E.L., McEvoy, L.M., Berlin, C., Bargatze, R.F., and Butcher, E.C. (1993). L-selectin-mediated lymphocyte rolling on MadCAM-1. *Nature*, **366**, 695–8.

Berlin, C. *et al*. (1993). a^4b_7 integrin mediates lymphocyte binding to the mucosal vascular addressin MAdCAM-1. *Cell*, **74**, 185–95.

Berlin, C. *et al*. (1995). a^2-integrins mediate lymphocyte attachment and rolling under physiologic flow. *Cell*, **80**, 413–22.

Bevilacqua, M.P. and Nelson, R.M. (1993). Selectins. *Journal of Clinical Investigation*, **91**, 379–87.

Bevilacqua, M.P., Pober, J.S., Mendrick, D.L, Cotran, R.S., and Gimbrone, M.A. (1987). Identification of an inducible endothelial-leukocyte adhesion molecule. *Proceedings of the National Academy of Sciences* (*USA*), **84**, 9238–42.

Binns, R.M., Licence, S.T., Wooding, F.B.P., and Duffus, W.P.H. (1992). Active lymphocyte traffic induced in the periphery by cytokines and phytohemagglutinin: three different mechanisms? *European Journal of Immunology*, **22**, 2195–203.

Bogen, S., Pak, J., Garifallou, M., Deng, X., and Muller, W.A. (1994). Monoclonal antibody to murine PECAM-1 (CD31) blocks acute inflammation *in vivo*. *Journal of Experimental Medicine*, **179**, 1059–64.

Bombara, M.P. *et al*. (1993). Cell contact between T cells and synovial fibroblasts causes induction of adhesion molecules and cytokines. *Journal of Leukocyte Biology*, **54**, 399–406.

Bonvoisin, B., Cordier, G., Revillard, J.P., Lejeune, E., and Bouvier, M. (1984). Increased DNA and/or RNA content of synovial fluid cells in rheumatoid arthritis: a flow-cytometry study. *Annals of the Rheumatic Diseases*, **43**, 222–7.

Briskin, M.J., McEvoy, L.M., and Butcher, E.C. (1993). MAdCAM-1 has homology to immunoglobulin and mucin-like adhesion receptors and to IgA1. *Nature*, **363**, 461–4.

Brown, Z., Gerritsen, M.E., Carley, W.W., Strieter, R.M., Kunkel, S.L., and Westwick, J. (1994). Chemokine gene expression and secretion by cytokine-activated human microvascular endothelial cells. *American Journal of Pathology*, **145**, 913–21.

Buckton, K.E., Court Brown, W.M., and Smith, P.G. (1967). Lymphocyte survival in men treated with X-rays for ankylosing spondylitis. *Nature*, **214**, 470–3.

Bührer, C., Berlin, C., Jablonski-Westrich, D., Holzmann, B., Thiele, H.-G., and Hamann, A. (1992). Lymphocyte activation and regulation of three adhesion molecules with supposed function in homing: LECAM-1 (MEL-14 antigen), LPAM-1/2 (a^4-integrin) and CD44 (Pgp-1). *Scandinavian Journal of Immunology*, **35**, 107–20.

Bullard, D.C. *et al*. (1995). P-selectin/ICAM-1 double mutant mice: acute emigration of neutrophils into peritoneum is completely absent but is normal into pulmonary alveoli. *Journal of Clinical Investigation*, **95**, 1782–8.

Burmester, G.R., Jahn, B., Gramatzki, M., Zacher, J., and Kalden, J.R. (1984). Activated T cells *in vivo* and *in vitro*: divergence in expression of Tac and Ia antigens in the nonblastoid small T cells of inflammation and normal T cells activated *in vitro*. *Journal of Immunology*, **133**, 1230–4.

Butcher, E.C. (1991). Leukocyte-endothelial cell recognition: three (or more) steps to specificity and diversity. *Cell*, **67**, 1033–6.

Cabañas, C. and Hogg, N. (1993). Ligand intercellular adhesion molecule 1 has a necessary role in activation of integrin lymphocyte function-associated molecule 1. *Proceedings of the National Academy of Sciences* (*USA*), **90**, 5838–42.

Campanero, M.R., Sanchez-Mateos, P., de Pozo, M.A., and Sanchez-Madrid, F. (1994). ICAM-3 regulates lymphocyte morphology and integrin-mediated T cell interaction with endothelial cell and extracellular matrix ligands. *Journal of Cell Biology*, **127**, 867–78.

Carr, M.W., Roth, S.J., Luther, E., Rose, S.S., and Springer, T.A. (1994). Monocyte chemoattractant protein 1 acts as a T-lymphocyte chemoattractant. *Proceedings of the National Academy of Sciences* (*USA*), **91**, 3652–6.

Carson, C.W., Beall, L.D., Hunder, G.G., Johnson, C.M., and Newman, W. (1994). Soluble E-selectin is increased in inflammatory synovial fluid. *Journal of Rheumatology*, **21**, 605–11.

Carter, W.G. and Wayner, E.A. (1988). Characterization of the class III collagen receptor, a phosphorylated, transmembrane glycoprotein expressed in nucleated human cells. *Journal of Biological Chemistry*, **263**, 4193–201.

Chan, B.M.C., Kassner, P.D., Schiro, J.A., Byers, H.R., Kupper, T.S, and Hemler, M.E. (1992). Distinct cellular functions mediated by different VLA integrin a subunit cytoplasmic domains. *Cell*, **68**, 1051–60.

Chapman, P.T. and Haskard, D.O. (1995). Leukocyte adhesion molecules. *British Medical Bulletin*, **51**, 296–311.

Chapman, P.T., Jamar, F., Keelan, E.T.M., Peters, A.M., and Haskard, D.O. (1996). Use of a radiolabeled monoclonal antibody against E-selectin for imaging endothelial activation in rheumatoid arthritis. *Arthritis and Rheumatism*, **39**, 1371–5.

Chin, W. and Hay, J.B. (1980). A comparison of lymphocyte migration through intestinal lymph nodes, subcutaneous lymph nodes, and chronic inflammatory sites of sheep. *Gastroenterology*, **79**, 1231–42.

Cicuttini, F.M., Martin, M., and Boyd, A.W. (1994). Cytokine induction of adhesion molecules on synovial Type B cells. *Journal of Rheumatology*, **21**, 406–12.

Clark, E.A. and Brugge, J.S. (1995). Integrins and signal transduction pathways: the road taken. *Science*, **268**, 233–9.

Clements, J.M. *et al*. (1994). Identification of a key integrin-binding sequence in VCAM-1 homologous to the LDV active site in fibronectin. *Journal of Cell Science*, **107**, 2127–35.

Cohnheim, J. (1889). In *Lectures in general pathology* (ed. J. Cohnheim), pp. 242–382. New Sydenham Society, London.

Corkill, M.M., Kirkham, B.W., Haskard, D.O., Barbatis, C., Gibson, T., and Panayi, G.S. (1991). Gold treatment of rheumatoid arthritis decreases synovial expression of the endothelial leukocyte adhesion receptor ELAM-1. *Journal of Rheumatology*, **18**, 1453–60.

Cybulsky, M.I., Chan, M.K.W., and Movat, H.Z. (1988). Acute inflammation and microthrombosis induced by endotoxin, interleukin-1, and tumour necrosis factor and their implication in Gram-negative infection. *Laboratory Investigation*, **58**, 365–78.

Damle, N.K. and Doyle, L.V. (1990). Ability of human T lymphocytes to adhere to vascular endothelial cells and to augment endothelial permeability to macromolecules is linked to their state of post-thymic maturation. *Journal of Immunology*, **144**, 1233–40.

Damle, N.K., Klussman, K., Dietsch, M.T., Mohagheghpour, N., and Aruffo, A. (1992 *a*). GMP-140 (P-selectin/CD62) binds to chronically stimulated but not resting CD4+ T lymphocytes and regulates their production of proinflammatory cytokines. *European Journal of Immunology*, **22**, 1789–93.

Damle, N.K., Klussman, K., and Aruffo, A. (1992 *b*). Intercellular adhesion molecule-2, a second counter-receptor for CD11a/CD18 (leukocyte function-associated antigen-1), provides a costimulatory signal for T-cell receptor-initiated activation of human T cells. *Journal of Immunology*, **148**, 665–71.

Davies, K.A. *et al*. (1991). A 19-year-old man with leukocyte adhesion deficiency. *In vitro* and *in vivo* studies of leucocyte function. *Clinical and Experimental Immunology*, **84**, 223–31.

Davis, L.S., Kavanaugh, A.F., Nichols, L.A., and Lipsky, P.E. (1995). Induction by persistent T cell hyporesponsiveness *in vivo* by monoclonal antibody to ICAM-1 in patients with rheumatoid arthritis. *Journal of Immunology*, **154**, 3525–37.

De Fougerolles, A.R. and Springer, T.A. (1992). Intercellular adhesion molecule 3, a third adhesion counter-receptor for lymphocyte function-associated molecule 1 on resting lymphocytes. *Journal of Experimental Medicine*, **175**, 185–90.

De Fougerolles, A.R., Stacker, S.A., Schwarting, R., and Springer, T.A. (1991). Characterization of ICAM-2 and evidence for a third counter-receptor for LFA-1. *Journal of Experimental Medicine*, **174**, 253–67.

Diamond, M.S., Staunton, D.E., Marlin, S.D., and Springer, T.A. (1991). Binding of the integrin Mac-1 (CD11b/CD18) to the third immunoglobulin-like domain of ICAM-1 (CD54) and its regulation by glycosylation. *Cell*, **65**, 961–71.

Doré, M., Korthuis, R.J., Granger, D.N., Entman, M.L., and Smith, C.W. (1993). P-selectin mediates spontaneous leukocyte rolling *in vivo*. *Blood*, **82**, 1308–16.

Dougherty, G.J., Murdoch, S., and Hogg, N. (1988). The function of intercellular adhesion molecule-1 (ICAM-1) in the generation of an immune response. *European Journal of Immunology*, **18**, 35–9.

Drickamer, K. (1988). Two distinct classes of carobohydrate-recognition domains in animal lectins. *Journal of Biological Chemistry*, **263**, 9557–60.

Dustin, M.L. and Springer, T.A. (1989). T-cell receptor cross-linking transiently stimulates adhesiveness through LFA-1. *Nature*, **341**, 619–24.

Dvorak, H.F., Mihm, M.C.J., Dvorak, A.M., Barnes, B.A., Manseau, E.J., and Galli, S.J. (1979). Rejection of first-set skin allografts in man. The microvasculature is the critical target of the immune response. *Journal of Experimental Medicine*, **150**, 322–37.

El-Gabalawy, H. and Wilkins, J. (1993). b_1 (CD29) integrin expression in rheumatoid synovial membranes: an immunohistologic study of distribution patterns. *Journal of Rheumatology*, **20**, 231–7.

El-Gabalawy, H., Gallatin, M., Vazeux, R., Peterman, G., and Wilkins, J. (1994). Expression of ICAM-R (ICAM-3) a novel counter-receptor for LFA-1, in rheumatoid and non-rheumatoid synovium. *Arthritis and Rheumatism*, **37**, 846–54.

Elices, M.J. (1990). VCAM-1 on activated endothelium interacts with the leukocyte integrin VLA-4 at a site distinct from the VLA-4/fibronectin binding site. *Cell*, **60**, 577–84.

Elices, M.J. (1994). Expression and functional significance of alternatively spliced CS1 fibronectin in rheumatoid arthritis microvasculature. *Journal of Clinical Investigation*, **93**, 405–16.

En Chin, J., Winterrowd, G.E., Krzesicki, R.F., and Sanders, M.E. (1990). Role of cytokines in inflammatory synovitis: the coordinate regulation of intercellular adhesion molecule 1 and HLA class I and class II antigens in rheumatoid synovial fibroblasts. *Arthritis and Rheumatism*, **33**, 1776–86.

Erle, D.J., Briskin, M.J., Butcher, E.C., Garcia-Pardo, A., Lazarovits, A.I., and Tidswell, M. (1994). Expression and function of the MAdCAM-1 receptor, integrin a^4b_7, on human leukocytes. *Journal of Immunology*, **153**, 517–28.

Etzioni, A. *et al*. (1992). Brief report: recurrent severe infections caused by a novel leukocyte adhesion deficiency. *New England Journal of Medicine*, **327**, 1789–92.

Fairburn, K., Kunaver, M., Wilkinson, L.S., Cambridge, G., Haskard, D.O., and Edwards, J.C.W. (1992). Intercellular adhesion molecules in normal synovium. *British Journal of Rheumatology*, **32**, 302–6.

Faveeuw, C., Gagnerault, M.C., and Lepault, F. (1994). Expression of homing and adhesion molecules in infiltrated islets of Langerhans and salivary glands of non-obese diabetic mice. *Journal of*

Immunology, **152**, 5969–78.

Fawcett, J. *et al*. (1992). Molecular cloning of ICAM-3, a third ligand for LFA-1, constitutively expressed on resting leukocytes. *Nature*, **360**, 481–4.

Fawcett, J. *et al*. (1995). Mapping the homotypic binding sites in CD31 and the role of CD31 adhesion in the formation of interendothelial cell contacts. *Journal of Cell Biology*, **128**, 1299–41.

Fischer, C., Thiele, H.-G., and Hamann, A. (1993). Lymphocyte-endothelial interactions in inflamed synovia: involvement of several adhesion molecules and integrin epitopes. *Scandinavian Journal of Immunology*, **38**, 158–66.

Foreman, K.E. *et al*. (1994). C5a-induced expression of P-selectin in endothelial cells. *Journal of Clinical Investigation*, **94**, 1147–55.

Fox, R.I., Fong, S., Tsoukas, C., and Vaughan, J.H. (1984). Characterization of recirculating lymphocytes in rheumatoid arthritis patients: selective deficiency of natural killer cells in thoracic duct lymph. *Journal of Immunology*, **132**, 2883–7.

Freemont, A.J. (1987). Molecules controlling lymphocyte-endothelial interactions in lymph nodes are produced in vessels of inflamed synovium. *Annals of the Rheumatic Diseases*, **46**, 924–8.

Freemont, A.J., Jones, C.J.P., Bromley, M., and Andrews, P. (1983). Changes in vascular endothelium related to lymphocyte collections in diseased synovia. *Arthritis and Rheumatism*, **26**, 1427–33.

Furie, M.B. and McHugh, D.D. (1989). Migration of neutrophils across endothelial monolayers is stimulated by treatment of the monolayers with interleukin-1 or tumor necrosis factor-a. *Journal of Immunology*, **143**, 3309–17.

Gaboury, J.P. and Kubes, P. (1994). Reductions in physiologic shear rates lead to CD11/CD18-dependent, selectin-independent leukocyte rolling *in vivo. Blood*, **83**, 345.

Gallatin, W.M., Weissman, I.L., and Butcher, E.C. (1983). A cell-surface molecule involved in organ-specific homing of lymphocytes. *Nature*, **304**, 30–34.

Gamble, J.R., Khew-Goodall, Y., and Vadas, M.A. (1993). Transforming growth factor-b inhibits E-selectin expression on human endothelial cells. *Journal of Immunology*, **150**, 4494–503.

Gearing, A.J.H. and Newman, W. (1993). Circulating adhesion molecules in disease. *Immunology Today*, **14**, 506–12.

Gell, P.G.H. (1959). Cytologic events in hypersensitivity reactions. In *Cellular and humoral aspects of the hypersensitive states* (ed. S. Lawrence), pp. 43–66. Hoeber-Harper.

Geng, J.-G. *et al*. (1990). Rapid neutrophil adhesion to activated endothelium mediated by GMP-140. *Nature*, **343**, 757–60.

Goldschneider, I. and McGregor, D.D. (1968). Migration of lymphocytes and thymocytes in the rat: I. The route of migration from blood to spleen and lymph nodes. *Journal of Experimental Medicine*, **127**, 155–68.

Gowans, J.L. and Knight, E.J. (1964). The route of recirculation of lymphocytes in the rat. *Proceedings of the Royal Society of London* (*Biology*), **159**, 257–82.

Graves, B.J. *et al*. (1994). Insight from E-selectin/ligand interaction from the crystal structure and mutagenesis of the lec/EGF domains. *Nature*, **367**, 532–8.

Green, P.J. *et al*. (1992). High affinity binding of the leucocyte adhesion molecule L-selectin to 3'-sulphated-Lea and -Lex oligosaccharides and the predominance of sulphate in this interaction demonstrated by binding studies with a series of lipid-linked oligosaccharides. *Biochemical and Biophysical Research Communications*, **188**, 244–51.

Grober, J.S. *et al*. (1993). Monocyte-endothelial adhesion in chronic rheumatoid arthritis. *In situ* detection of selectin and integrin-dependent interactions. *Journal of Clinical Investigation*, **91**, 2609–19.

Guan, J.L. and Hynes, R.O. (1990). Lymphoid cells recognize an alternatively spliced segment of fibronectin via the integrin receptor a $^{4}b_{1}$. *Cell*, **60**, 53–61.

Hale, L.P. *et al*. (1989). Immunohistologic analysis of the distribution of cell adhesion molecules within the inflammatory synovial microenvironment. *Arthritis and Rheumatism*, **32**, 22–30.

Hall, J.G. and Morris, B. (1963). The lymph-borne cells of the immune response. *Quarterly Journal of Experimental Physiology*, **48**, 235–47.

Hamann, A., Andrew, D.P., Jablonski-Westrich, D., Holzmann, B., and Butcher, E.C. (1994). Role of a ^{4}-integrins in lymphocyte homing to mucosal tissues *in vivo*. *Journal of Immunology*, **152**, 3282–93.

Hanninen, A. *et al*. (1993). Vascular addressins are induced on islet vessels during insulitis in nonobese diabetic mice and are involved in lymphoid cell binding to islet endothelium. *Journal of Clinical Investigation*, **92**, 2509–15.

Hanto, D.W., Hopt, U.T., Hoffman, R., and Simmons, R.L. (1982). Recruitment of unsensitized circulating lymphocytes to sites of allogeneic cellular interactions. *Transplantation*, **33**, 541–6.

Haskard, D.O., Cavender, D.E., Beatty, P., Springer, T.A., and Ziff, M. (1986). T cell adhesion to endothelial cells: mechanisms demonstrated by anti-LFA-1 monoclonal antibodies. *Journal of Immunology*, **137**, 2901–6.

Hattori, R., Hamilton, K.K., Fugate, R.D., McEver, R.D., and Sims, P.J. (1989 *a*). Stimulated secretion of endothelial von Willebrand factor is accompanied by rapid redistribution to the cell surface of the intracellular granule membrane protein GMP-140. *Journal of Biological Chemistry*, **264**, 7768–71.

Hattori, R., Hamilton, K.K., McEver, R.P., and Sims, P.J. (1989 *b*). Complement proteins C5b-9 induce secretion of high molecular weight multimers of endothelial von Willebrand factor and translocation of granule membrane protein GMP-140 to the cell surface. *Journal of Biological Chemistry*, **264**, 9053–60.

Haynes, B.F., Hale, L.P., Patton, K.L., Martin, M.E., and McCallum, R.M. (1991). Measurement of an adhesion molecule as an indicator of inflammatory disease activity: up-regulation of the receptor for hyaluronate (CD44) in rheumatoid arthritis. *Arthritis and Rheumatism*, **34**, 1434–43.

Hechtman, D.H., Cybulsky, M.I., Fuchs, H.J., Baker, J.B., and Gimbrone, M.A.J (1991). Intravascular IL-8: inhibitor of polymorphonuclear leukocyte accumulation at sites of acute inflammation. *Journal of Immunology*, **147**, 883–92.

Hemler, M.E. (1990). VLA proteins in the integrin family: structures, functions, and their role on leukocytes. *Annual Review of Immunology*, **8**, 365–400.

Hemler, M.E., Elices, M.J., Parker, C., and Takada, Y. (1990). Structure of the integrin VLA-4 and its cell–cell and cell–matrix adhesion functions. *Immunological Reviews*, **114**, 45–65.

Hemmerich, S., Butcher, E.C., and Rosen, S.R. (1995). Sulfation-dependent recognition of high endothelial venules (HEV)-ligands by L-selectin and MECA 79, an adhesion-blocking monoclonal antibody. *Journal of Experimental Medicine*, **180**, 2219–26.

Holzmann, B. and Weissman, I.L. (1989). Integrin molecules involved in lymphocyte homing to Peyer's Patches. *Immunological Reviews*, **108**, 45–61.

Huang, K., Geoffroy, J.S., Singer, M.S., and Rosen, S.D. (1991). A lymphocyte homing receptor (L-selectin) mediates the *in vitro* attachment of lymphocytes to myelinated tracts of the central nervous system. *Journal of Clinical Investigation*, **88**, 1778–83.

Huang, A.J., Manning, J.E., Bandak, T.M., Ratau, M.C., Hanser, K.R., and Silverstein, subcutaneous (1993). Endothelial cell cytosolic free calcium regulates neutrophil migration across monolayers of endothelial cells. *Journal of Cell Biology*, **120**, 1371–80.

Huber, A.R., Kunkel, S.L., Todd, R.F., III, and Weiss, S.J. (1991). Regulation of transendothelial neutrophil migration by endogenous interleukin-8. *Science*, **254**, 99–102.

Huet, S., Groux, H., Caillou, B., Valentin, H., Prieur, M., and Bernard, A. (1989). CD44 contibutes to T cell activation. *Journal of Immunology*, **143**, 798–801.

Hynes, R.O. (1992). Integrins: versatility, modulation, and signaling in cell adhesion. *Cell*, **69**, 11–25.

Iguchi, T. and Ziff, M. (1986). Electron microscopic study of rheumatoid synovial vasculature. Intimate relationship between tall endothelium and lymphoid aggregation. *Journal of Clinical Investigation*, **77**, 355–61.

Iigo, Y. *et al*. (1991). ICAM-1-dependent pathway is critically involved in the pathogenesis of adjuvant arthritis in rats. *Journal of Immunology*, **147**, 4167–71.

Isobe, M., Yagita, H., Okumura, K., and Ihara, A. (1992). Specific acceptance of cardiac allograft after treatment with antibodies to ICAM-1 and LFA-1. *Science*, **255**, 1125–7.

Jackson, D.G., Bell, J.I., Dickinson, R., Timans, J., Shields, J., and Whittle, N. (1995). Proteoglycan forms of the lymphocyte homing receptor CD44 are alternatively spliced variants containing the V3 exon. *Journal of Cell Biology*, **128**, 673–85.

Jalkanen, S., Bargatze, R.F., Toyos, J., de los., and Butcher, E.C. (1987). Lymphocyte recognition of high endothelium: antibodies to distinct epitopes of an 85–95 kDa glycoprotein antigen

differentially inhibit lymphocyte binding to lymph node, mucosal, or synovial endothelial cells. *Journal of Cell Biology*, **105**, 983–90.

Jalkanen, S., Jalkanen, M., Bargatze, R., Tammi, M., and Butcher, E.C. (1988). Biochemical properties of glycoproteins involved in lymphocyte recognition of high endothelial venules in man. *Journal of Immunology*, **141**, 1615–23.

Jamar, F., Chapman, P.T., Harrison, A.A., Binns, R.M., Haskard, D.O., and Peters, A.M. (1995). Inflammatory arthritis: imaging of endothelial activation with an indium-111-labeled F(ab')2 fragment of anti-E-selectin monoclonal antibody. *Radiology*, **194**, 843–50.

Janossy, G., Panayi, G., Duke, O., Bofill, M., Poulter, L.W., and Goldstein, G. (1981). Rheumatoid arthritis: a disease of T-lymphocyte/macrophage immunoregulation. *Lancet*, **ii**, 839–42.

Jasin, H.E., Lightfoot, E., Davis, L.S., Rothlein, R., Faanes, R.B., and Lipsky, P.E. (1992). Amelioration of antigen-induced arthritis in rabbits treated with monoclonal antibodies to leukocyte adhesion molecules. *Arthritis and Rheumatism*, **35**, 541–9.

Johnson, B.A., Haines, G.K., Harlow, L.A., and Koch, A.E. (1993). Adhesion molecule expression in human synovial tissue. *Arthritis and Rheumatism*, **36**, 137–46.

Jones, D.A., McIntyre, L.V., Smith, C.W., and Picker, L.J. (1994). A two-step adhesion cascade for T cell/endothelial cell interactions under flow conditions. *Journal of Clinical Investigation*, **94**, 2443–50.

Jones, E.Y. *et al*. (1995). Crystal structure of an integrin-binding fragment of vascular cell adhesion molecule-1 at 1.8 resolution. *Nature*, **373**, 539–44.

Kahn, J., Ingraham, R.H., Shirley, F., Migaki, G.I., and Kishimoto, T.K. (1994). Membrane proximal cleavage of L-selectin: identification of the cleavage site and a 6 kD transmembrane peptide fragment of L-selectin. *Journal of Cell Biology*, **125**, 461–70.

Kansas, G.S. *et al*. (1994). A role for the epidermal growth factor-like domain of P-selectin in ligand recognition and cell adhesion. *Journal of Cell Biology*, **124**, 609–18.

Kaplanski, G., Farnarier, C., Benoliel, A.-M., Foa, C., Kaplanski, S., and Bongrand, P. (1994). A novel role for E- and P-selectins: shape control of endothelial cell monolayers. *Journal of Cell Science*, **107**, 2449–57.

Karsh, J., Klippel, J.H., Plotz, P.H., Decker, J.L., Wright, D.G., and Wayne Flye, M. (1981). Lymphapheresis in rheumatoid arthritis: a randomized trial. *Arthritis and Rheumatism*, **24**, 867–73.

Kavanaugh, A.F., Lightfoot, E., Lipsky, P.E., and Oppenheimer-Marks, N. (1991). Role of CD11/CD18 in adhesion and transendothelial migration of T cells: analysis utilizing CD18-deficient T cell clones. *Journal of Immunology*, **146**, 4149–56.

Kavanaugh, A.F. *et al*. (1994). Treatment of refractory rheumatoid arthritis with a monoclonal antibody to intercellular adhesion molecule-1. *Arthritis and Rheumatism*, **37**, 992–9.

Keelan, E.T.M., Harrison, A.A., Chapman, P.T., Binns, R.M., Peters, A.M., and Haskard, D.O. (1994). Imaging vascular endothelial activation: an approach using radiolabelled monoclonal antibody against the endothelial cell adhesion molecule E-selectin. *Journal of Nuclear Medicine*, **35**, 276–81.

Koch, A.E., Burrows, J.C., Haines, G.K., Carlos, T.M., Harlan, J.M., and Leibovich, S.J. (1991). Immunolocalization of endothelial and leukocyte adhesion molecules in human rheumatoid and osteoarthritic synovial tissues. *Laboratory Investigation*, **64**, 313–20.

Koopman, G., van Kooyk, Y., De Graaff, M., Meyer, C.J.L.M., Figdor, C.G., and Pals, S.T. (1990). Triggering of the CD44 antigen on T lymphocytes promotes T cell adhesion through the LFA-1 pathway. *Journal of Immunology*, **145**, 3589–93.

Kornfield, H., Berman, J.S., Beer, D.J., and Center, D.M. (1985). Induction of human T lymphocyte motility by interleukin 2. *Journal of Immunology*, **134**, 3887.

Kosunen, T.U., Waksman, B.H., Flax, M.H., and Tihen, W.S. (1963). Radioautographic study of cellular mechanisms in delayed hypersensitivity. *Immunology*, **6**, 276–90.

Kuijpers, T.W. (1993). Terminal glycosyltransferase activity: a selective role in cell adhesion. *Blood*, **81**, 873–82.

Kuijpers, T.W. *et al*. (1992). CD66 nonspecific cross-reacting antigens are involved in neutrophil adherence to cytokine-activated endothelial cells. *Journal of Cell Biology*, **118**, 457–66.

Kupfer, A. and Singer, S.J. (1989). The specific interaction of helper T cells and antigen-presenting B cells: IV. Membrane and cytoskeletal reorganizations in the bound T cell as a function of antigen dose. *Journal of Experimental Medicine*, **170**, 1697–713.

Kurosaka, M. and Ziff, M. (1983). Immunoelectron microscopic study of the distribution of T cell subsets in rheumatoid arthritis. *Journal of Experimental Medicine*, **158**, 1191–210.

Labow, M.A. *et al*. (1994). Characterization of E-selectin-deficient mice: demonstration of overlapping function of the endothelial selectins. *Immunity*, **1**, 709–20.

Laffon, A. *et al*. (1992). Upregulated expression and function of VLA-4 fibronectin receptors on human activated T cells in rheumatoid arthritis. *Journal of Clinical Investigation*, **88**, 546–52.

Landis, R.C., McDowell, A., Holness, C.L.L., Littler, A.J., Simmons, D.L,. and Hogg, N. (1994). Involvement of the 'I' domain of LFA-1 in selective binding to ligands ICAM-1 and ICAM-3. *Journal of Cell Biology*, **126**, 529–37.

Larsen, C.G., Anderson, A.O., Appella, E., Oppenheim, J.J., and Matsushima, K. (1989). The neutrophil activating protein (NAP-1) is also chemotactic for T lymphocytes. *Science*, **243**, 1464–6.

Larsen, E. *et al*. (1989). PADGEM protein: a receptor that mediates the interaction of activated platelets with neutrophils and monocytes. *Cell*, **59**, 305–12.

Larson, R.S. and Springer, T.A. (1990). Structure and function of leukocyte integrins. *Immunological Reviews*, **114**, 181–217.

Lasky, L.A. (1992). Selectins: interpreters of cell-specific carbohydrate information during inflammation. *Science*, **258**, 964–9.

Laudanna, C. *et al*. (1994). Sulfatides trigger increase of cytosolic free calcium and enhanced expression of tumor necrosis factor-a and interleukin-8 mRNA in human neutrophils: evidence for a role of L-selectin as a signalling. *Journal of Biological Chemistry*, **269**, 4021–6.

Lawrence, M.B. and Springer, T.A. (1991). Leukocytes roll on an a selectin at physiologic flow rates: distinction from and prerequisite for adhesion through integrins. *Cell*, **65**, 859–73.

Lawrence, M.B. and Springer, T.A. (1993). Neutrophils roll on E-selectin. *Journal of Immunology*, **151**, 6338–46.

Lazarovits, A.I. and Karsh, J. (1993). Differential expression in rheumatoid synovium and synovial fluid of a $^{4}b_{7}$ integrin: a novel receptor for fibronectin and vascular cell adhesion molecule-1. *Journal of Immunology*, **151**, 6482–9.

Lee, J.-O., Rieu, P., Arnaout, M.A., and Liddington, R. (1995). Crystal structure of the A domain from the a subunit of integrin CD3 (CD11b/CD18). *Cell*, **80**, 631–8.

Lee, M.-S. and Sarvetnick, N. (1994). Induction of vascular addressins and adhesion molecules in the pancreas of IFNg transgenic mice. *Journal of Immunology*, **152**, 4597–603.

Ley, K., Gaehtgens, P., Fennie, C., Singer, M.S., Lasky, L.A., and Rosen, S.D. (1991). Lectin-like cell adhesion molecule 1 mediates leukocyte rolling in mesenteric venules *in vivo. Blood*, **77**, 2553–5.

Ley, K. *et al*. (1995). Sequential contribution of L- and P-selectin to leukocyte rolling *in vivo. Journal of Experimental Medicine*, **181**, 669–75.

Li, R. *et al*. (1995). A peptide derived from the intercellular adhesion molecule-2 regulates the avidity of the leukocyte integrins CD11b/CD18 and CD11c/CD18. *Journal of Cell Biology*, **129**, 1143–53.

Lo, S.K. *et al*. (1991). Endothelial-leukocyte adhesion molecule 1 stimulates the adhesive activity of leukocyte integrin CR3 (CD11b/CD18, Mac-1, a $^{M}b_{2}$) on human neutrophils. *Journal of Experimental Medicine*, **173**, 1493–500.

Lobb, R.R. and Hemler, M.E. (1994). The pathophysiologic role of a ^{4} integrins *in vivo*. *Journal of Clinical Investigation*, **94**, 1722–8.

Lofgren, R., Ng-Sikorski, J., Sjolander, A., and Andersson, T. (1993). b_{2} integrin engagement triggers actin polymerization and phosphatidylinositol triphosphate formation in non-adherent human neutrophils. *Journal of Cell Biology*, **123**, 1597–605.

Luskinskas, F.W. *et al*. (1994). Monocyte rolling, arrest and spreading on IL-4-activated vascular endothelium under flow is mediated via sequential action of L-selectin, b_{1}-integrins, and b_{2}-integrins. *Journal of Cell Biology*, **125**, 1417–27.

Luskinskas, F.W., Ding, H., and Lichtman, A.H. (1995). P-selectin and vascular cell adhesion molecule 1 mediate rolling and arrest, rspectively, of CD4+ T lymphocytes on tumour necrosis factor

a-activated vascular endothelium under flow. *Journal of Experimental Medicine*, **181**, 1179–86.

Mackay, C.R., Marston, W.L., and Dudler, L. (1990). Naive and memory T cells show distinct pathways of lymphocyte recirculation. *Journal of Experimental Medicine*, **171**, 801–17.

Mackay, C.R., Marston, W.L., Dudler, L., Spertini, O., Tedder, T.F., and Hein, W.R. (1992). Tissue-specific migration pathways by phenotypically distinct subpopulations of memory T cells. *European Journal of Immunology*, **22**, 887–95.

Manconi, P.E., Fadda, M.F., Cadoni, A., Cornaglia, P., Zaccheo, D., and Grifoni, V. (1978). Subpopulations of T lymphocytes in human extravascular fluids. *International Archives of Allergy and Applied Immunology*, **56**, 385–95.

Marchal, G., Seman, M., Milon, G., Truffa-Bachi, P., and Zilberfarb, V. (1982). Local adoptive transfer of skin delayed-type hypersensitivity initiated by a single T lymphocyte. *Journal of Immunology*, **129**, 954–8.

Marchesi, V.T. and Florey, H.W. (1960). Electron micrographic observations on the emigration of leukocytes. *Quarterly Journal of Experimental Physiology*, **45**, 343–8.

Marchesi, V.T. and Gowans, J.L. (1964). The migration of lymphocytes through the endothelium of venules in lymph nodes: an electron microscope study. *Proceedings of the Royal Society of London* (*Biology*), **159**, 283–90.

Marlor, C.W., Webb, D.L., Bombara, M.P., Greve, J.M., and Blue, M.-L. (1992). Expression of vascular cell adhesion molecule-1 in fibroblastlike synoviocytes after stimulation with tumor necrosis factor. *American Journal of Pathology*, **140**, 1055–60.

Martin, S. *et al*. (1995). Circulating forms of ICAM-3 (cICAM-3). Elevated levels in autoimmune diseases and lack of association with cICAM-1. *Journal of Immunology*, **154**, 1951–5.

Mason, J.C. and Haskard, D.O. (1994). The clinical importance of leukocyte and endothelial cell adhesion molecules in inflammation. *Vascular Medicine Reviews*, **5**, 249–75.

Mason, J.C., Kapahi, P., and Haskard, D.O. (1993). Detection of raised levels of circulating ICAM-1 in some patients with rheumatoid arthritis but not in systemic lupus erythematosus: lack of correlation with circulating VCAM-1. *Arthritis and Rheumatism*, **36**, 519–27.

Mayadas, T.N., Johnson, R.C., Rayburn, H., Hynes, R.O., and Wagner, D.D. (1993). Leukocyte rolling and extravasation are severely compromised in P selectin-deficient mice. *Cell*, **74**, 541–54.

Merkenschlager, M. *et al*. (1991). Allorecognition of HLA-DR and -DQ transfectants by human CD45RA and CD45RO CD4 T cells: repertoire analysis and activation requirements. *European Journal of Immunology*, **21**, 79–88.

Michie, S.A., Streeter, P.R., Bolt, P.A., Butcher, E.C., and Picker, L.J. (1993). The human peripheral lymph node vascular addressin: an inducible endothelial antigen involved in lymphocyte homing. *American Journal of Pathology*, **143**, 1688–98.

Moore, K.L. and Thompson, L.F. (1992). P-Selectin (CD62) binds to subpopulations of human memory T lymphocytes and natural killer cells. *Biochemical and Biophysical Research Communications*, **186**, 173–81.

Morales-Ducret, J., Wayner, E., Elices, M.J., Alvaro-Gracia, J.M., Zvaifler, N.J., and Firestein, G.S. (1992). a $^{4}b_{1}$ Integrin (VLA-4) ligands in arthritis: vascular cell adhesion molecule-1 expression in synovium and on fibroblast-like synoviocytes. *Journal of Immunology*, **149**, 1424–31.

Moser, R., Schleiffenbaum, B., Groscurth, P., and Fehr, J. (1989). Interleukin 1 and tumour necrosis factor stimulate human vascular endothelial cells to promote transendothelial neutrophil passage. *Journal of Clinical Investigation*, **83**, 444–55.

Muller, W.A., Weigl, S.A., Deng, X., and Phillips, D.M. (1993). PECAM-1 is required for transendothelial migration of leukocytes. *Journal of Experimental Medicine*, **178**, 449–60.

Murphy, P.M. (1994). The molecular biology of leukocyte chemoattractant receptors. *Annual Review of Immunology*, **12**, 593–633.

Naparstek, Y., Cohen, I.R., Fuks, Z., and Vlodavsky, I. (1984). Activated T lymphocytes produce a matrix-degrading heparan sulphate endoglycosidase. *Nature*, **310**, 241–44.

Nature (1957). Behaviour and fate of lymphocytes. *Nature*, **180**, 1016–17.

Nielson, H.V., Christensen, J.P., Anderson, E.C., Marker, O., and Thomsen, A.R. (1994). Expression of type 3 complement receptor on activated CD8+ T cells facilitates homing to inflammatory sites. *Journal of Immunology*, **153**, 2021–8.

Nikkari, L., Aho, H., Yli-Jama, T., Larjava, H., Jalkanen, M., and Heino, J. (1993). Expression of integrin family of cell adhesion receptors in rheumatoid synovium. *American Journal of Pathology*, **142**, 1019–27.

Norgard-Sumnicht, K.E., Varki, N.M., and Varki, A. (1993). Calcium-dependent heparin-like ligands for L-selectin in nonlymphoid endothelial cells. *Science*, **261**, 480–3.

O'Neill, J.K. *et al*. (1991). Expression of vascular addressins and ICAM-1 by endothelial cells in the spinal cord during chronic relapsing experimental allergic encephalomyelitis in the Biozzi AB/H mouse. *Immunology*, **72**, 520–5.

Oppenheimer-Marks, N. and Ziff, M. (1988). Migration of lymphocytes through endothelial cell monolayers: augmentation by interferon-g. *Cellular Immunology*, **114**, 307–23.

Orosz, C.G. *et al*. (1993). Treatment with anti-vascular cell adhesion molecule 1 monoclonal antibody induces long-term murine cardiac allograft acceptance. *Transplantation*, **56**, 453–60.

Osborn, L. *et al*. (1989). Direct expression cloning of vascular cell adhesion molecule 1 (VCAM1), a cytokine-induced endothelial protein that binds to lymphocytes. *Cell*, **59**, 1203–11.

Pabst, R. (1988). The spleen in lymphocyte migration. *Immunology Today*, **9**, 43–45.

Palecanda, A., Walcheck, B., Bishop, D.K., and Jutila, M.A. (1992). Rapid activation-independent shedding of leukocyte L-selectin induced by cross-linking of the surface antigen. *European Journal of Immunology*, **22**, 1279–86.

Pals, S.T., Hogervorst, F., Keizer, G.D., Thepen, T., Horst, E., and Figdor, C.G. (1989). Identification of a widely distributed 90-kDa glycoprotein that is homologous to the Hermes-1 human lymphocyte homing receptor. *Journal of Immunology*, **143**, 851–7.

Pasqualini, R. and Hemler, M.E. (1994). Contrasting roles for integrin b $_{1}$ and b_{5} cytoplasmic domains in subcellular localization, cell proliferation, and cell migration. *Journal of Cell Biology*, **125**, 447–60.

Paulus, H.E., Machleder, H.I., Clements, P.J., Nyman, K., and Levine, S. (1979). Prolonged thoracic duct drainage in rheumatoid arthritis and systemic lupus erythemmatosus. *Western Journal of Medicine*, **130**, 309–24.

Pavalko, F.M., Walker, D.M., Graham, L., Goheen, M., Doerschuk, C.M., and Kansas, G.S. (1995). The cytoplasmic domain of L-selectin interacts with cytoskeletal proteins via a-actinin: receptor positioning in microvilli does not require interaction with a-actinin. *Journal of Cell Biology*, **129**, 1155–64.

Peter, K. and O'Toole, T.E. (1995). Modulation of cell adhesion by changes in a $^{L}b_{2}$ (LFA-1. CD11a/CD18) cytoplasmic domain/cytoskeletal interaction. *Journal of Experimental Medicine*, **181**, 315–26.

Peters, A.M. (1983). Splenic blood flow and blood cell kinetics. *Clinics in Haematology*, **12**, 421–47.

Pfau, S., Leitenberg, D., Rinder, H., Smith, B.R., Pardi, R., and Bender, J.R. (1995). Lymphocyte adhesion-dependent calcium signaling in human endothelial cells. *Journal of Cell Biology*, **128**, 969–78.

Piali, L. *et al*. (1995). CD31/PECAM-1 is a ligand for a $^{v}b_{3}$ integrin involved in adhesion of leukocytes to endothelium. *Journal of Cell Biology*, **130**, 451–60.

Picker, L.J., Terstappen, L.W.M.M., Rott, L.S., Streeter, P.R., Stein, H., and Butcher, E.C. (1990). Differential expression of homing-associated adhesion molecules by T cell subsets in man. *Journal of Immunology*, **145**, 3247–55.

Picker, L.J., Kishimoto, T.K., Smith, C.W., Warnock, R.A., and Butcher, E.C. (1991 *a*). ELAM-1 is an adhesion molecule for skin-homing T cells. *Nature*, **349**, 796–9.

Picker, L.J., Warnock, R.A., Burns, A.R., Doerschuk, C.M., Berg, E.L., and Butcher, E.C. (1991 *b*). The neutrophil selectin LECAM-1 presents carbohydrate ligands to the vascular selectins ELAM-1 and GMP-140. *Cell*, **66**, 921–33.

Picker, L.J., Treer, J.R., Ferguson-Darnell, B., Collins, P.A., Bergstresser, P.R., and Terstappen, L.W.M.M. (1993 *a*). Control of lymphocyte recirculation in man: II. Differential regulation of the

cutaneous lymphocyte-associated antigen, a tissue-selective homing receptor for skin-homing T cells. *Journal of Immunology*, **150**, 1122–36.

Picker, L.J., Treer, J.R., Ferguson-Darnell, B., Collins, P.A., Buck, D., and Terstappen, L.W.M.M. (1993 *b*). Control of lymphocyte recirculation in man: I. Differential regulation of the peripheral lymph node homing receptor L-selectin on T cells during the virgin to memory cell transition. *Journal of Immunology*, **150**, 1105–21.

Pinola, M., Renkonen, R., Majuri, M.-L., Tiisala, S., and Saksela, E. (1994). Characterization of the E-selectin ligand on NK cells. *Journal of Immunology*, **152**, 3586–94.

Pitzalis, C., Kingsley, G., Murphy, J., and Panayi, G.S. (1987). Abnormal distribution of the helper-inducer and suppressor-inducer T-lymphocyte subsets in the rheumatoid joint. *Clinical Immunology and Immunopathology*, **45**, 252–8.

Pitzalis, C., Kingsley, G., Haskard, D.O., and Panayi, G.S. (1988). The preferential accumulation of helper-inducer T lymphocytes in inflammatory lesions: evidence for regulation by selective endothelial and homotypic adhesion. *European Journal of Immunology*, **18**, 1397–404.

Pitzalis, C., Kingsley, G.H., Covelli, M., Meliconi, R., Markey, A., and Panayi, G.S. (1991). Selective migration of the human helper-inducer memory T cell subset: confirmation by *in vivo* cellular kinetic studies. *European Journal of Immunology*, **21**, 369–76.

Pober, J.S. and Cotran, R.S. (1990). Cytokines and endothelial cell biology. *Physiological Reviews*, **70**, 427–51.

Postigo, A.A. *et al*. (1992). Increased binding of synovial T lymphocytes from rheumatoid arthritis to endothelial-leukocyte adhesion molecule-1 (ELAM-1) and vascular cell adhesion molecule-1 (VCAM-1). *Journal of Clinical Investigation*, **89**, 1445–52.

Postigo, A.A., Marazuela, M., Sanchez-Madrid, F., and De Landazuri, M.O. (1994). B lymphocyte binding to E- and P-selectins is mediated through the *de novo* expression of carbohydrates on *in vitro* and *in vivo* activated human B cells. *Journal of Clinical Investigation*, **94**, 1585–96.

Poulter, L.W., Seymour, G.J., Duke, O., Janossy, G., and Panayi, G.S. (1982). Immunohistological analysis of delayed-type hypersensitivity in man. *Cellular Immunology*, **74**, 358–69.

Prendergast, R.A. (1964). Cellular specificity in the homograft reaction. *Journal of Experimental Medicine*, **119**, 377–87.

Rannie, G.H., Smith, M.E., and Ford, W.L. (1977). Lymphocyte migration into cell-mediated immune lesions is inhibited by trypsin. *Nature*, **267**, 520–2.

Read, M.A., Whitley, M.Z., Williams, A.J., and Collins, T. (1994). NF_B and I_Bb: an inducible regulatory system in endothelial activation. *Journal of Experimental Medicine*, **179**, 503–12.

Rendt, K.E., Barry, T.S., Jones, D.M., Tichter, C.B., McCachren, S.S., and Haynes, B.F. (1993). Engraftment of human synovium into severe combined immune deficient mice. *Journal of Immunology*, **151**, 7324–6.

Rice, G.E., Munro, J.M., Corless, C., and Bevilacqua, M.P. (1991). Vascular and nonvascular expression of INCAM-110: a target for mononuclear leukocyte adhesion in normal and inflamed human tissues. *American Journal of Pathology*, **138**, 385–93.

Robbins, S.L., Cotran, R.S., and Kumar, V. (1984). *Pathologic basis of disease*. W.B. Saunders, Philadelphia.

Rodrigues, M., Nussenzweig, R.S., Romero, P., and Zavala, F. (1992). The *in vivo* cytotoxic activity of CD8+ T cell clones correlates with their levels of expression of adhesion molecules. *Journal of Experimental Medicine*, **175**, 895–905.

Romanic, A.M. and Madri, J.A. (1994). The induction of 72-kD gelatinase in T cells upon adhesion to endothelial cells is VCAM-1 dependent. *Journal of Cell Biology*, **125**, 1165–78.

Rose, M.L., Parrott, D.M., and Bruce, R.G. (1978). The accumulation of immunoblasts in extravascular tissues including mammary gland, peritoneal cavity, gut and skin. *Immunology*, **35**, 415–23.

Rossiter, H. *et al*. (1994). Skin disease-related T cells bind to endothelial selectins: expression of cutaneous lymphocyte antigen (CLA) predicts E-selectin but not P-selectin binding. *European Journal of Immunology*, **24**, 205–10.

Rothlein, R., Dustin, M.L., Marlin, S.D., and Springer, T.A. (1986). A human intercellular adhesion molecule (ICAM-1) distinct from LFA-1. *Journal of Immunology*, **137**, 1270–4.

Ruegg, C., Postigo, A.A., Sikorski, E.E., Butcher, E.C., Pytela, R., and Erle, D.J. (1992). Role of integrin a $^4b_7$/a^4b_P in lymphocyte adherence to fibronectin and VCAM-1 and in homotypic cell clustering. *Journal of Cell Biology*, **117**, 179–89.

Sako, D. *et al*. (1993). Expression cloning of a functional glycoprotein ligand for P-selectin. *Cell*, **75**, 1179–86.

Salmi, M. and Jalkanen, S. (1992). A 90-kilodalton endothelial cell molecule mediating lymphocyte binding in humans. *Science*, **257**, 1407–9.

Salmi, M., Kalimo, K., and Jalkanen, S. (1993). Induction and function of vascular adhesion protein-1 at sites of inflammation. *Journal of Experimental Medicine*, **178**, 2255–60.

Salmi, M., Andrew, D.P., Butcher, E.C., and Jalkanen, S. (1995). Dual binding capacity of mucosal immunoblasts to mucosal and synovial endothelium in humans: dissection of the molecular mechansisms. *Journal of Experimental Medicine*, **181**, 137–49.

Sanchez-Aparacio, P., Dominguez-Jimenez, C., and Garcia-Pardo, A. (1994). Activation of the a $^4b_1$ integrin through the b_1 subunit induces recognition of the RGDS sequence in fibronectin. *Journal of Cell Biology*, **126**, 271–9.

Sanders, M.E. *et al*. (1988). Human memory T lymphocytes express increased levels of three cell adhesion molecules (LFA-3, CD2, and LFA-1) and three other molecules (UCHL1, CDw29, and Pgp-1) and have enhanced IFN-gamma production. *Journal of Immunology*, **140**, 1401–7.

Sanders, M.E., Makgoba, M.W., June, C.H., Young, H.A., and Shaw, S. (1989). Enhanced responsiveness of human memory T cells to CD2 and CD3 receptor-mediated activation. *European Journal of Immunology*, **19**, 803–8.

Santamaria Babi, L.F. *et al*. (1995). Circulating allergen-reactive T cells from patients with atopic dermatitis and allergic contact dermatitis express the skin-selective homing receptor, the cutaneous lymphocyte-associated antigen. *Journal of Experimental Medicine*, **181**, 1935–40.

Schall, T.J., Bacon, K., Toy, K.J., and Goeddel, D.V. (1990). Selective attraction of monocytes and T lymphocytes of the memory phenotype by cytokine RANTES. *Nature*, **347**, 669–71.

Schall, T.J., Bacon, K., Camp, R.D., Kaspari, J.W., and Goeddel, D.V. (1993) Human macrophage inflammatory protein alpha (MIP-1 alpha) and MIP-1 beta chemokines attract distinct populations of lymphocytes. *Journal of Experimental Medicine*, **177**, 1821–6.

Schick, P. *et al*. (1975). Autotransfusion of 3H-cytidine-labelled blood lymphocytes in patients with Hodgkin's disease and non-Hodgkin patients. II. Exchangeable lymphocyte pools. *Acta Haematologica Basel*, **53**, 206–18.

Schweighoffer, T. *et al*. (1993). Selective expression of integrin a $^4b_7$ on a subset of human CD4+ memory T cells with hallmarks of gut-trophism. *Journal of Immunology*, **151**, 717–29.

Scott, J.L., Davidson, J.G., Marino, J.V., and McMillan, R. (1972). Leukocyte labeling with 51-chromium. 3. The kinetics of normal lymphocytes. *Blood*, **40**, 276–81.

Screaton, G.R., Bell, M.V., Jackson, D.G., Cornelis, F.B., Gerth, U., and Bell, J.I. (1992). Genomic structure of DNA encoding the lymphocyte homing receptor CD44 reveals at least 12 alternatively spliced exons. *Proceedings of the National Academy of Sciences* (*USA*), **89**, 12160–4.

Shimada, H. *et al*. (1994). Interleukin-4 increases human synovial expression of VCAM-1 and T cell binding. *Annals of the Rheumatic Diseases*, **53**, 601–7.

Shimizu, Y., Van Seventer, G.A., Horgan, K.J., and Shaw, S. (1990 *a*). Regulate expression and binding of three VLA (beta 1) integrin receptors on T cells. *Nature*, **345**, 250–3.

Shimizu, Y., Van Seventer, G.A., Horgan, K.J., and Shaw, S. (1990 *b*). Roles of adhesion molecules in T-cell recognition: fundamental similarities between four integrins on resting human T cells (LFA-1,VLA-4,VLA-5,VLA-6) in expression, binding, and co-stimulation. *Immunological Reviews*, **114**, 110–43.

Shimizu, Y. *et al*. (1991). Activation-independent binding of human memory T cells to adhesion molecule ELAM-1. *Nature*, **349**, 799–802.

Shimizu, Y., Van Seventer, G.A., Ennis, E., Newman, W., Horgan, K.J., and Shaw, S. (1992). Crosslinking of the T cell-specific accessory molecules CD7 and CD28 modulates T cell adhesion. *Journal of Experimental Medicine*, **175**, 577–82.

Sligh, J.E., Jr. *et al*. (1993). Inflammatory and immune responses are impaired in mice deficient in intercellular adhesion molecule 1. *Proceedings of the National Academy of Sciences* (*USA*), **90**,

8529–33.

Smith, J.B., McIntosh, G.H., and Morris, B. (1970). The migration of cells through chronically inflamed tissues. *Journal of Pathology*, **100**, 21–29.

Smith, M.E. and Ford, W.L. (1983). The recirculating lymphocyte pool in the rat: a systematic description of the migratory behaviour of recirculating lymphocytes. *Immunology*, **49**, 83–94.

Spertini, O. *et al*. (1991). Leukocyte adhesion molecule-1 (LAM-1, L-selectin) interacts with an inducible endothelial cell ligand to support leukocyte adhesion. *Journal of Immunology*, **147**, 2565–73.

Springer, T.A. (1990). Adhesion receptors of the immune system. *Nature*, **346**, 425–34.

Springer, T.A. (1994). Traffic signals for lymphocyte recirculation and leukocyte emigration: the multistep paradigm. *Cell*, **76**, 301–14.

Stamper, H.B. and Woodruff, J.J. (1976). Lymphocyte homing into lymph nodes: *in vitro* demonstration of the selective affinity of recirculating lymphocytes for high-endothelial venules. *Journal of Experimental Medicine*, **114**, 828–33.

Staunton, D.E., Dustin, M.L., and Springer, T.A. (1989). Functional cloning of ICAM-2, a cell adhesion ligand for LFA-1 homologous to ICAM-1. *Nature*, **339**, 61–4.

Staunton, D.E., Dustin, M.L., Erickson, H.P., and Springer, T.A. (1990). The arrangement of the immunoglobulin-like domains of ICAM-1 and the binding sites for LFA-1. *Cell*, **61**, 243–54.

Steegmaler, M. *et al*. (1995). The E-selectin-ligand ESL-1 is a variant of a receptor for fibroblast growth factor. *Nature*, **373**, 615–20.

Streeter, P.R., Berg, E.L., Rouse, B.T.N., Bargatze, R.F., and Butcher, E.C. (1988). A tissue-specific endothelial cell molecule involved in lymphocyte homing. *Nature*, **334**, 41–6.

Subramaniam, M., Saffaripour, S., Watson, S.R., Mayadas, T.N., Hynes, R.O., and Wagner, D.D. (1995). Reduced recruitment of inflammatory cells in a contact hypersensitivity response in P-selectin-deficient mice. *Journal of Experimental Medicine*, **181**, 2277–82.

Szekanecz, Z. *et al*. (1994). Differential distribution of intercellular adhesion molecules (ICAM-1, ICAM-2, and ICAM-3) and the MS-1 antigen in normal and diseased human synovia. *Arthritis and Rheumatism*, **37**, 221–31.

Takahashi, H., Söderström, K., Nilsson, E., Kiessling, R., and Patarroyo, M. (1992). Integrins and other adhesion molecules on lymphocytes from synovial fluid and peripheral blood of rheumatoid arthritis patients. *European Journal of Immunology*, **22**, 2879–85.

Tanaka, Y. *et al*. (1992). CD31 expressed on distinctive T cell subsets is a preferential amplifier of b $_1$ integrin-mediated adhesion. *Journal of Experimental Medicine*, **176**, 245–53.

Tanaka, Y., Adams, D.H., Hubscher, S., Hirano, H., Siebenlist, U., and Shaw, S. (1993). T-cell adhesion induced by proteoglycan-immobilized cytokine MIP-1b. *Nature*, **361**, 79–82.

Taub, D.D. *et al*. (1993). Recombinant human interferon-inducible protein 10 is a chemoattractant for human monocytes and T lymphocytes and promotes T cell adhesion to endothelial cells. *Journal of Experimental Medicine*, **177**, 1809–14.

Taub, D.D. *et al*. (1995). Monocyte chemotactic protein-1 (MCP-1), -2, and -3 are chemotactic for human T lymphocytes. *Journal of Clinical Investigation*, **95**, 1370–6.

Tedder, T.F., Steeber, D.A., and Pizcueta, P. (1995). L-selectin deficient mice have impaired leukocyte recuitment into inflammatory sites. *Journal of Experimental Medicine*, **181**, 2259–64.

Thornhill, M.H. and Haskard, D.O. (1990). IL-4 regulates endothelial activation by IL-1, tumor necrosis factor or IFNg. *Journal of Immunology*, **145**, 865–72.

van Dinther-Janssen, A.C.H.M. *et al*. (1991). The VLA-4/VCAM-1 pathway is involved in lymphocyte adhesion to endothelium in rheumatoid synovium. *Journal of Immunology*, **147**, 4207–10.

van Dinther-Janssen, A.C.H.M., van Maarsseveen, A.C.M.T., de Groot, J., and Scheper, R.J. (1983). Comparative migration of T- and B-lymphocyte subpopulations into skin inflammatory lesions. *Immunology*, **48**, 519–27.

Van Epps, D.E., Potter, J., Vachula, M., Smith, C.W., and Anderson, D.C. (1989). Suppression of human lymphocyte chemotaxis and transendothelial migration by anti-LFA-1 antibody. *Journal of Immunology*, **143**, 3207–10.

Vaporciyan, A.A. *et al*. (1993). Involvement of platelet-endothelial cell adhesion molecule-1 in neutrophil recruitment *in vivo. Science*, **262**, 1580–2.

Vazeux, R. *et al*. (1992). Cloning and characterization of a new intercellular adhesion molecule ICAM-R. *Nature*, **360**, 485–8.

Von Andrian, U.H. *et al*. (1993). *In vivo* behavior of neutrophils from two patients with distinct inherited leukocyte adhesion deficiency syndromes. *Journal of Clinical Investigation*, **91**, 2893–7.

Wahl, S.M. *et al*. (1983). Leukapheresis in rheumatoid arthritis. *Arthritis and Rheumatism*, **26**, 1076–84.

Ward, P.A. (1995). Adhesion molecule knock-outs: one step forward and one step backward. *Journal of Clinical Investigation*, **95**, 1425.

Webb, L.M.C., Ehrengruber, M.U., Clark-Lewis, I., Baggiolini, M., and Rot, A. (1993). Binding to heparan sulfate or heparin enhances neutrophil responses to interleukin 8. *Proceedings of the National Academy of Sciences* (*USA*), **90**, 7158–62.

Weyrich, A.S., McIntyre, T.M., McEver, R.P., Prescott, S.M., and Zimmerman, G. (1995). Monocyte tethering by P-selectin regulates monocyte chemotactic protein-1 and tumour necrosis factor-a secretion. *Journal of Clinical Investigation*, **95**, 2297–2303.

Whyte, A., Haskard, D.O., and Binns, R.M. (1994). Infiltrating gamma delta T cells and selectin endothelial ligands in the cutaneous phytohaemagglutinin-induced inflammatory reaction in the pig. *Vetinerary Immunology and Immunopatholology*, **41**, 31–40.

Wiener, J., Lattes, R.G., and Spiro, D. (1967). An electron microscopic study of leukocyte emigration and vascular permeability in tuberculin sensitivity. *American Journal of Pathology*, **50**, 485–521.

Wilkinson, L.S., Poston, R.N., Edwards, J., and Haskard, D.O. (1993). Expression of vascular cell adhesion molecule-1 (VCAM-1) in normal and inflamed synovium. *Laboratory Investigation*, **68**, 82–8.

Wilkinson, P.C. and Liew, F.Y. (1995). Chemoattraction of human blood T lymphocytes by interleukin-15. *Journal of Experimental Medicine*, **181**, 1255–9.

Wilkinson, P.C. and Newman, I. (1994). Chemoattractant activity of IL-2 for human lymphocytes: a requirement for IL-2 receptor beta-chain. *Immunology*, **82**, 134–9.

Williams, A.F. and Barclay, A.N. (1988). The immunoglobulin superfamily-domains for cell surface recognition. *Annual Review of Immunology*, **6**, 381–405.

Wilson, R.W. *et al*. (1993). Gene targeting yields a CD18-mutant mouse for study of inflammation. *Journal of Immunology*, **151**, 1571–8.

Yednock, T.A. and Rosen, S.D. (1989). Lymphocyte homing. *Advances in Immunology*, **44**, 313–78.

Zimmerman, G.A., McIntyre, T.M., Mehra, M., and Prescott, S.M. (1990). Endothelial cell-associated platelet-activating factor: a novel mechanism for signaling intercellular adhesion. *Journal of Cell Biology*, **110**, 529–40.

3.3 Specific immune responses

B. P. Wordsworth, S. J. Bowman, and John I. Bell

A complex series of immunological events has evolved in man as a response to the threat posed by a plethora of potential pathogens. Many of these responses are of a non-specific nature (innate immunity), while others are targeted more specifically (adaptive immunity). Activation of the immune system typically leads to tissue inflammation, clearance of the invading organism, and healing of damaged tissues, often with fibrosis. In addition to infection by bacteria, viruses, and fungi, abnormal host cells such as tumours may serve as targets of the immune response. In the course of such reactions there may also be substantial damage to the host tissues and, in certain circumstances, the immune system may be directed aberrantly against components of the host tissues. At least some rheumatological disorders appear to belong to this group of 'autoimmune diseases' and, without doubt, immunological mechanisms are involved in their pathogenesis. A clearer knowledge of the basic physiology of these immunological responses may therefore improve our understanding of the pathogenesis of these conditions.

Innate immunity

There are many non-specific barriers to infection, including the skin, mucosal surfaces, and their secretions. In addition to the potential bactericidal effects of pH or proteolytic enzymes in some secretions, such as gastric digestive juices, tissue secretions may contain an enzyme, lysozyme, that disrupts the proteoglycan cell walls of many pathogens.

Several cell types, originating from bone marrow precursors, have the capacity to phagocytose organisms that penetrate these barriers. In the bloodstream these include neutrophil polymorphs and monocytes, both of which can migrate into the tissues in response to appropriate stimuli such as infection or inflammation. Some leucocytes have the capacity to recognize virus-infected cells; these natural killer cells can bind to such cells and kill them after activation by interferon produced by the host tissues. Furthermore, interferon induces resistance to viral replication in neighbouring cells and constitutes an important component of innate immunity to viral infection.

Clearance of pathogens may be enhanced greatly by the provision of opsonization, in which the target is coated by a variety of mediators that increase the efficiency of phagocytosis by macrophages and polymorphs. Examples of opsonins include complement components and acute-phase proteins (e.g. C-reactive protein) from the innate immune system and immunoglobulin antibodies from the adaptive immune system. Activation of the complement cascade (see Chapter 3.1) can lead to such diverse activities as opsonization (C3b), chemotaxis (C5a), and lysis of micro-organisms by the membrane attack complex. Many organisms activate complement through the 'alternat(iv)e pathway' which leads to their opsonization with C3b. This enhances phagocytosis because monocyte macrophages and polymorphs have specific receptors for C3b on their surface. Further activation of the complement cascade causes chemotaxis of phagocytes and other white cell types to the focus of inflammation by the accumulation of C5a. Cell surface determinants on some organisms allow their recognition by phagocytes without the need for complement, while others can be opsonized by acute-phase proteins (see Chapter 4.1). For instance, C-reactive protein is synthesized and rapidly released in many inflammatory disorders by the action of interleukin 6 on hepatocytes. C-reactive protein can bind to many organisms, activates complement *in situ* through the alternative pathway, and thereby facilitates phagocytosis through the C3b receptors on phagocytes. Consequently, hereditary deficiency of C3 may cause recurrent infections and septicaemia (Morgan and Walport 1991). During the process of acute inflammation that accompanies an immune response, there is an increase in blood supply to the affected areas and an increase in vascular permeability. Migration of leucocytes through the postcapillary venules is enhanced by the local accumulation of C5a in the tissue. Concurrently, a number of plasma enzyme systems are activated. These include the complement cascade, the kinin system (bradykinin release), the clotting system (through Hageman factor XII), the fibrinolytic system (through tissue plasminogen activator), and the release of histamine from mast cells, triggered by complement. Many of these chemicals are vasoactive and chemoattractant to neutrophils, and they lead to further increases in local tissue concentrations of other mediators of the immune response.

Adaptive immunity

Many components of the innate immune system can be activated non-specifically. However, not all pathogens stimulate such responses, because they lack the appropriate ligands for binding to phagocytes, complement, or acute-phase proteins and alternative mechanisms, involving lymphocytes, are required to initiate targeted responses. Recognition of antigens by two distinct lymphocyte lineages (B and T cells) is central to adaptive immunity. First, the production of antibodies by B lymphocytes provides considerable specificity and flexibility in adaptive interactions with antigens and can lead to activation of the immune system, including the complement cascade through the classical pathway. Second, T lymphocytes can recognize antigens directly through their cell-surface antigen receptors. Activated in this way they can either kill cells bearing non-self antigens (cytotoxic T cells) or provide 'help' to B cells to produce antibodies (helper/inducer T cells).

B lymphocytes are capable of producing antibodies that bind to the corresponding antigen with a high degree of specificity. Individual B cells are capable of synthesizing antibodies of only one specificity and carry membrane-bound copies of the antibody as cell surface receptors. Only when the appropriate antigen is encountered by the B-cell antigen receptor can that cell potentially proceed to activation and maturation towards an antibody-producing plasma cell. However, additional signals are required for this to occur, particularly from 'helper cells'. Antibodies provide a particularly flexible form of opsonin, with one end binding to specific antigenic determinants on specific microbial pathogens (Fab portion) and the other end (Fc portion) to complement or Fc receptors on the surface of some cell populations (Fig. 1).

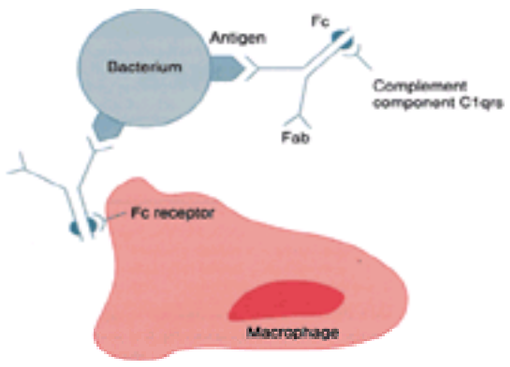


Fig. 1 Antibodies can activate complement or act as opsonins by binding through their Fc portions either to C1qrs or the Fc receptors on certain cells such as macrophages. The Fab portion recognizes specific antigenic determinants such as bacterial cell wall structures.

Specific immune responses

Immune activation is central to the pathogenesis of many of the inflammatory arthritides. In particular, rheumatoid arthritis and reactive arthritis are associated with many of the features that result from antigen-specific stimulation of the immune system. Many of the local features of these disorders within the joint can be attributed to lymphocyte activation, cytokine release, and activation of antigen-presenting cells or macrophages. Common to all these processes are the initial events associated with the recognition of peptide antigen in the context of major histocompatibility complex (**MHC**) molecules by T cells bearing the appropriate recognition elements. In the case of class-I-restricted MHC presentation this leads to the activation of cytoxic T-cell (CD8+) populations, and in the case of class II-restricted recognition, to the activation of helper T cells (CD4+). It is now widely believed that responses mediated by CD4 T cells can be separated into those mediated by a T_{H1} subset (enhanced cellular immunity) and a T_{H2} subset (facilitating B-cell humoral respones). The former responses result in the liberation of the cytokines interferon-a and interleukin 2 (**IL-2**), while the latter are mediated through the release of IL4 and IL-10 (reviewed in Paul and Seder 1994). Cascade events involving cell activation and cytokine release within the joint are undoubtedly extremely complex but include the activation of B cells and macrophages which in turn release other potent inflammatory mediators such as IL-1 and tumour necrosis factor.

The best opportunity for analysing these processes is to start with the simplest events surrounding antigen-specific, T-cell activation. There are three primary elements involved in initiating a specific immune response; these are an HLA class I or class II molecule, an antigen (in the form of peptide) that associates in the antigen-binding site of the HLA molecule, and the T-cell antigen receptor found on the surface of all T-cell populations (Fig. 2).

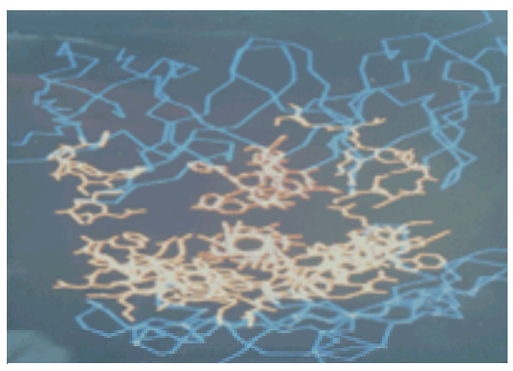


Fig. 2 Three-dimensional representation of the interaction between the MHC–peptide complex (below) and the T-cell receptor (TCR) above. The CDR1 and CDR2 of the TCR are shown in yellow and CDR3, which interacts most closely with the peptide antigen, in red. The a-helices of the MHC molecule interacting with CDR1 and CDR2 are shown in yellow and the peptide interacting with CDR3 is shown in red (Davis and Bjorkman 1988).

The interaction of these three molecules (often called the ternary complex) is assisted by the associated T-cell molecules CD4, CD8, and intercellular adhesion molecule 1, as well as by molecules such as lymphocyte function-associated antigens 1 and 2 present on the antigen-presenting receptor. This in turn leads to T-cell activation. Because of the central role of these elements of the ternary complex, this chapter concentrates on the molecular details of the components of this complex and discusses events leading to T-cell activation.

Major histocompatibility complex

Many of the genes implicated in susceptibility to autoimmune diseases, including many rheumatic conditions, map to the MHC at the centromeric end of the short arm of chromosome 6 (Stastny *et al*. 1983). Several genes in this region are essential to the mounting of normal cellular and humoral inflammatory responses and correspond to the immune response genes first described 30 years ago in the mouse. The HLA genes, for instance, are crucial to the determination of the host response not only to infections but also to transplanted tissues and they were first recognized as 'transplantation antigens'. Subsequently, much has been elucidated about their crucial role in the presentation of foreign antigens during immune responses. However, the HLA genes are but a small component of a region spanning over 3 million base pairs and containing over 100 genes. The MHC can be divided into three regions, classes I to III, each containing a series of homologous loci. (Fig. 3).

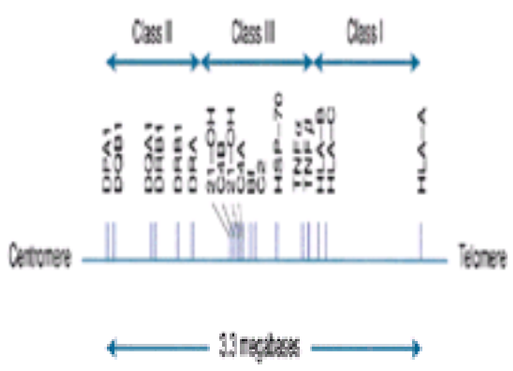


Fig. 3 The approximate genomic organization of the major histocompatibility complex on the short arm of chromosome 6.

Class I region

Encoded within this region are the polymorphic HLA-A, -B, and -C loci, encoding the corresponding 45-kDa heavy chains of HLA class I molecules, which are members of the immunoglobulin superfamily. At least 50 HLA-A, 90 HLA-B, and 30 HLA-C alleles have been identified (Bodmer *et al*. 1994), although new variants are continually being described. Each heavy chain has three extracellular domains (a_1, a_2, and a_3) a transmembrane portion, and a cytoplasmic tail. It associates with

a 12-kDa light chain (b_2-microglobulin) before expression on the surface of most nucleated cells. These molecules have long been recognized as transplantation antigens capable of triggering the rejection of tissue grafted from an HLA non-identical donor.

HLA class I molecules play an essential part in the presentation of peptide antigens to cytotoxic T cells in an HLA-restricted manner. The determination of the three-dimensional crystallographic structure of the class I molecules HLA-A2, -A68, and -B27 (Bjorkman *et al*. 1987; Garrett *et al*. 1989; Madden *et al*. 1991) has served to illustrate the probable mechanism of this process (Fig. 4). Antigens, typically peptides with eight or nine amino acids, are bound in a cleft on the surface of the HLA molecule (Jardetzky *et al*. 1991). The combined HLA–antigen complex can then interact with T lymphocytes bearing the appropriate antigen receptors to trigger an immune response. Something of the specificity of this system can be appreciated by considering T-cell clones raised against particular peptide antigens presented by target cells of known HLA type. The same clone will fail to recognize the same antigen if it is presented by target cells of a different HLA type. Likewise, even single amino acid substitutions in the peptide antigen may be sufficient to abolish recognition (Townsend and Bodmer 1989).

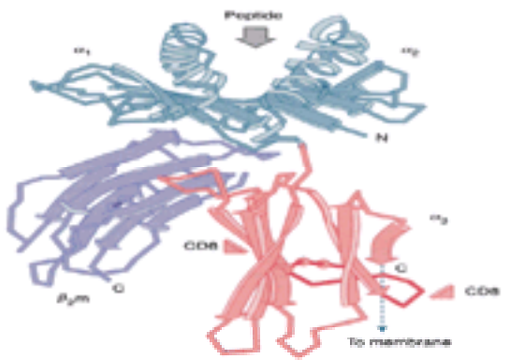


Fig. 4 Ribbon diagram of the HLA-A2 class I molecule, indicating the antigen-binding cleft. Polymorphic residues are concentrated around this structure pointing towards the peptide antigen from the floor and sides. The sites of interaction with the CD8 molecule and b_2-microglobulin are indicated.

Nucleotide sequencing of class I genes reveals that the polymorphism within these molecules lies largely in the a_1- and a_2-domains that form the sides and floor of the antigen-binding cleft (Fig. 4). Furthermore, the polymorphic residues are those that are orientated to be involved in interactions either with the peptide antigen or the T-cell receptor. Single amino acid substitutions may induce profound effects on both peptide binding and T-cell recognition. In general, class I molecules are concerned with the presentation of endogenously derived antigens to cytotoxic (CD8+) T cells and are therefore intimately concerned in the host response to tumour cells and virus-infected cells.

Class II region

Encoded within this region are the classical class II genes, HLA-DR, -DQ, and -DP as illustrated in Fig. 5. Class II HLA molecules are also members of the immunoglobulin superfamily and have considerable structural and functional homology with class I molecules, but a more limited tissue distribution. Each class II molecule is a transmembrane glycoprotein heterodimer, consisting of a non-covalently bound a- and b-chain. The amino acid sequences of these molecules share similarities with class I molecules, and the membrane distal portion of the a- and b-chains of class II molecules combine to form a similar antigen-binding cleft with a floor of eight antiparallel b-pleated sheets, and two antiparallel a-helices on either side (Brown *et al*. 1993). Class II molecules bind peptides of more variable lengths of around 14 amino acids, and this difference from class I molecules reflects a more open structure at both ends of the peptide-binding groove (Stern *et al*. 1994). Each a-chain is approximately 34 kDa and each b-chain approximately 29 kDa. Polymorphism of the various class II HLA molecules is contributed by both the a- and b-chains of DQ and DP, whereas the DR a-chain is non-polymorphic, and allelic variation is entirely confined to the DR b-chains, encoded at several distinct loci (Bell *et al*. 1989). In addition to the DRB1 locus on all haplotypes, there are supertypic specificities on many haplotypes (DR52 encoded at the DRB3 locus on DR3, -5, and -6 haplotypes; DR53 encoded at the DRB4 locus on DR4, -7, and -9 haplotypes; and DR51 encoded at the DRB5 locus on DR2 haplotypes). The DR and DQ loci exhibit extreme linkage disequilibrium but there is weaker linkage between DQ and DP. Consequently, particular DQ alleles are almost invariably associated with certain DR alleles within ethnic groups, although there may be differences in these linkage units between different races.

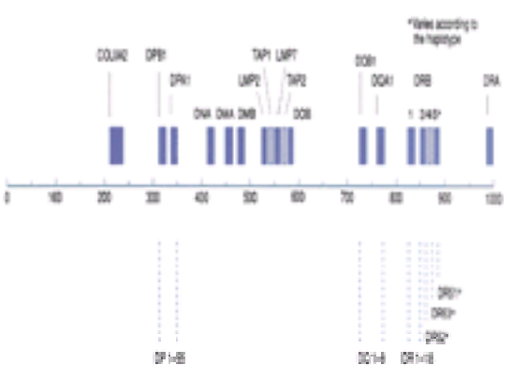


Fig. 5 The *HLA* class II region spanning approximately 800 kb which include several HLA loci as well as other loci encoding components of protesomes (LMP7, LMP2) and membrane transporters (TAP1, TAP2).

Class III region

This region, lying between the class I and class II region, encodes a number of the components of the classical complement cascade, including C2 and C4, as well as properdin factor B (Bf) of the alternate pathway. These genes are also polymorphic and may play important parts in immunological reactions during the host response, for instance in the clearance of immune complexes and pathogens. In addition, a number of other genes are found in this region, including those for 21-hydroxylase, the hsp70 heat-shock protein, and the genes encoding the cytokines tumour necrosis factor-a and -b (Spies *et al*. 1986; Sargent *et al*. 1989a). Since tumour necrosis factors are important cytokines involved in inflammatory processes (including the synovium in inflammatory arthritis), allelic variations of these genes may be involved in the predisposition to the autoimmune disease. Each of these exist as twin loci, not all of which are functional, highlighting the frequency of the gene duplication events that have given rise to many homologous loci throughout the MHC.

Definition of polymorphism of HLA molecules

Classically, serological reagents have been used to determine both HLA class I and class II specificities, either by the use of alloantisera or monoclonal antibodies raised against antigenic determinants on these molecules. Some of these may be 'public' specificities shared by a related series of molecules, while others may be more specific or 'private'. Biochemical methods have also been used to refine these definitions, but techniques such as two-dimensional gel electrophoresis or isoelectric focusing are too laborious for routine use. Initially, the mixed lymphocyte reaction has been used to refine HLA class II specificities in combination with homozygous-typing cell lines. Although this method consistently allows greater definition than do serological methods, it is also too labour intensive for routine use. In recent years direct analysis of the genomic DNA has provided a number of methods of typing for class II, and more recently class I alleles. The first of these involves analysis of the restriction fragment length polymorphisms associated with particular HLA types (Bell *et al*. 1987; MacMurray *et al*. 1987). A second method, based on nucleotide sequence analysis using sequence-specific oligonucleotides to probe HLA genes amplified to high copy number by the polymerase chain reaction, is both robust and highly sensitive and allowed a plethora of new alleles to be described (Vaughan *et al*. 1990; Wordsworth *et al*. 1990; Hill *et al*. 1991). A similar approach, also using the polymerase chain reaction, uses series of sequence-specific primers to determine the HLA class II genotype (Olerup and Zetterquist 1992).

HLA polymorphism in relation to disease

There is evidence of strong selection pressure for the maintenance of polymorphism within HLA class I and class II loci as a result of environmental factors. For example, a common class I allele in the Gambia (HLA-B53) is protective against the development of cerebral malaria. Similarly, a particular class II haplotype (DR13) protects those with malaria from severe anaemia (Hill 1991). A similar pattern is found in leprosy, in which the DR3 haplotype is associated with a strong cell-mediated response to *Mycobacterium leprae* characteristic of the tuberculoid form of this disease (Serjeantson 1983). Whether the existence of certain extended HLA haplotypes within some populations, such as A1, B8, DR3, DQ2 in north-west Europeans, can be explained in terms of protection against particular pathogens remains conjectural.

Recently described genes within the MHC

There are probably more than 120 genes within this region of the genome, most of which have not yet been characterized fully and some of which undoubtedly have important immunological functions (Sargent *et al*. 1989b). In particular, the region between the DQ and DP loci in the class II region contains further a and b class II genes (DM, DN, and DO loci) (Fig. 5). Other genes of interest in this region are those encoding for proteins involved in the class I MHC antigen-processing pathway. TAP1 and TAP2 (transporter associated with antigen processing) gene products transport peptides destined for presentation by class I molecules into the lumen of the endoplasmic reticulum. LMP2 and LMP7 (large multifunctional protease/low molecular mass polypeptide) genes encode components of a proteolytic enzyme complex known as a protesome, involved in processing complex antigens into peptides suitable for presentation by class I molecules (Monaco 1992). Since these genes are polymorphic (especially TAP2) they could contribute to disease susceptibility, although no such association has been proven to date (Monaco 1992; Trowsdale and Campbell 1992). Intriguingly, some defective antigen processing and presentation by HLA class I molecules map to the HLA class II region (see below).

Pulsed-field gel electrophoresis of the MHC has enabled a partial physical map of the region to be established. Physical mapping of the class II region has revealed that different haplotypes show variations in size. For instance, DR4 haplotypes contain an insertion of 110 kb of DNA between DR and DQ relative to most other haplotypes, while DR2 has a similar insert of 30 kb (Hardy *et al*. 1986). As both these haplotypes have been implicated in autoimmune diseases (rheumatoid arthritis with DR4 and multiple sclerosis with DR2), these inserts could be important if they contain expressed genes relevant to the disease susceptibility.

Recognition of antigens by the immune system

The mechanisms by which foreign and self antigens are recognized differ fundamentally between T and B lymphocytes. B cells carry surface immunoglobulin receptors capable of binding hydrophilic determinants on the surface of intact complex antigens. Binding of an antigen to this receptor may in principle lead to activation of the B cell and its differentiation to an antibody-producing plasma cell if the appropriate signals, such as interleukin 2, are also received from helper T lymphocytes. In contrast, T lymphocytes recognize antigenic determinants in the form of small peptides derived from more complex antigens. These peptides are typically processed fragments, which may come from parts of the intact antigen not normally exposed to immune surveillance by B cells. As these peptides are not necessarily on the surface of the molecule they may be hydrophilic or hydrophobic. The mounting of a primary antibody response is dependent on two separate systems of antigen recognition and, therefore, the chance of unwanted autoreactive antibody formation is minimized. However, it is not too difficult to envisage mechanisms by which this system could break down leading to the generation of humorally mediated autoimmunity (Fig. 6).

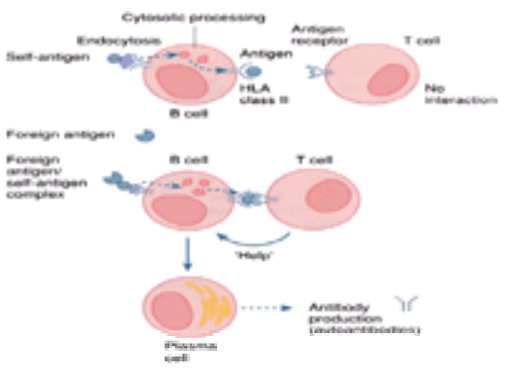


Fig. 6 One mechanism by which self-reactive antibodies might be presented. A self-antigen interacts with the surface immunoglobulin receptor of a B cell (top); as it does not receive T-cell help, it does not produce antibodies since the T-cell antigen receptor will not recognize self-antigens. However, if the self-antigen is complexed to a foreign antigen, as could happen with immunoglobulin bound to a foreign antigen, peptides from the latter could be expressed with HLA class II molecules in the B cell and elicit the necessary T-cell help to produce autoantibodies.

Processing and presentation of antigens to T lymphocytes

Antigenic peptides become associated with HLA class I and class II molecules before their presentation to lymphocytes, but the mechanisms involved are dependent on the nature of the antigen, the type of T cell involved, and the nature of the accessory molecules present on the cell surface (CD4 or CD8). Cytotoxic T lymphocytes (CD8+) generally interact with endogenously sythesized antigens bound to class I molecules. This system, common to all nucleated cells, is therefore particularly important in the immune surveillance of tumours and virus-infected cells. In contrast, helper T lymphocytes (CD4+) most commonly interact with antigens associated with HLA class II molecules, cell surface expression of which is restricted to specialized antigen-preserving cells (dendritic cells, macrophages, and B lymphocytes). However, in disease states, class II expression may be induced by the action of cytokines, such as g-interferon, on cell types including endothelium and fibroblasts that do not normally express these molecules. This may have implications for the pathogenesis of autoimmunity if it leads to the aberrant presentation of autoantigens in these tissues (Bottazzo *et al*. 1983). The processes of class I- and class II-mediated antigen presentation are not completely mutually exclusive; exogenous antigen may be presented by class I molecules, endogenous antigens by class II molecules, and some CD4+ cells are cytotoxic.

A consensus of the likely physiological process by which antigens and HLA molecules reach the cell surface has been derived from the study of certain mutant mouse and human cell lines (Monaco 1992; Neefjes *et al*. 1992). The presentation of antigens by HLA class I molecules requires the association of peptide with nascent class I heavy chains and b_2-microglobulin in the endoplasmic reticulum, and subsequent transport of the mature complex to the cell surface where it is exposed to T-cell surveillance. The peptide antigen itself is probably important in stabilizing with class I molecules. These peptide fragments may be available in the endoplasmic reticulum from *de novo* translation or else they may originate from the breakdown of proteins in the cytosol by proteases (Fig. 7). However, as these fragments lack hydrophobic leader-peptide sequences, they require active transmembrance transport into the endoplasmic reticulum if they are to associate with class I molecules.

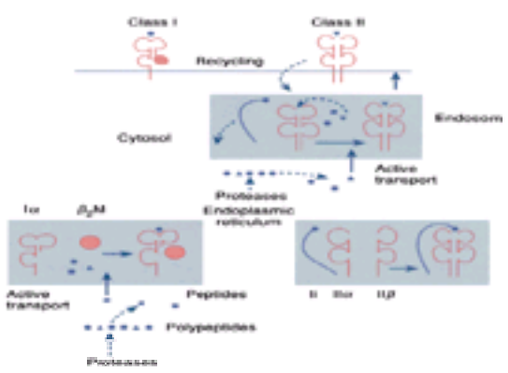


Fig. 7 Class I and class II pathways of antigen processing and presentation. The newly translated class I heavy chain (I) associated with b_2-microglobulin (b_2-m) and peptide in the endoplasmic reticulum. Peptides cannot associate with the class II molecules until the invariant chain (Ii) has been cleaved in the endosome. Active

transport of peptides from the cytosol occurs into both the endoplasmic reticulum and endosomes, and class II molecules may be recycled.

HLA class II molecules do not normally associate with endogenously synthesized peptides in the endoplasic reticulum but rather with a non-polymorphic protein known as invariant chain. This is thought to have both a signalling function, in directing the next stages of class II transport, and also to prevent endogenous peptides from binding to the class II peptide-binding groove. The invariant chain is removed by chemical cleavage in the post-Golgi system and replaced with peptide derived from endocytosed antigen that has been processed in the endosome. These complex processes involve multiple proteins, defects in any of which could potentially lead to defective antigen presentation and consequently to disease. Intriguingly, the TAP and LMP genes, whose products are involved in the class I processing pathway, are encoded within the MHC class II region (see above).

T-cell antigen receptor

The antigen receptor of circulating T lymphocytes consists of two immunoglobulin-like glycoproteins that associate as heterodimers on the cell surface in conjunction with five other invariant molecules referred to as the CD3 complex (reviewed by Moss *et al*. 1991). The cell surface heterodimers are of two major types, an a/b pair or a g/d pair. Each of these four chains is encoded as an immunoglobulin-like complex in the genome, with variable (V) diversity (D) joining (J) and constant (C) region elements which undergo germ-line rearrangement and, in association with N-region diversity, generate the extreme variation that is necessary to provide the range of antigen-presenting molecules on the T-cell surface. In peripheral blood the a/b T-cell receptor heterodimer is present on more than 95 per cent of circulating T cells. The rarer population of g/d T-cell receptors may also be involved in antigen-specific recognition, although their role is considerably less clear. They are commonly found in epithelial tissues and may have a 'sentinel' function. Recent evidence suggests that they may not require their ligand to be presented by MHC molecules and in this sense may have more in common with immunoglobulins (Davis and Chien 1995). The CD3 molecules are essential for the expression of the T-cell receptor heterodimer on the cell surface. They are also involved in transduction of signals to the T cell after the receptor interacts with peptide and MHC.

T-cell receptor genes

Germ-line organization of T-cell a, b, g, and d loci is similar to that of the immunoglobulin genes with a large number of gene segments comprising V, D, J, and C elements, which recombine resulting in a single VDJC transcript. Initial events result in the rearrangement of V, D, and J elements to produce a single exon. This exon and the exon encoding the C region are spliced together at the level of mRNA. A large number of individual V, D, and J elements for each of the major regions results in a very large amount of diversity in potential combinations. This diversity can be extended further by the use of N-region addition of nucleotides into the V, D, J and the JC junctions. In a, b, g, and d heterodimers there may be as many as 10^{15} possible combinations utilizing all the elements involved in the make-up of the receptor. There is no evidence that somatic hypermutation, as seen in immunoglobulin, occurs in these genes.

The V regions provide much of the diversity seen in a- and b-chains. Thirty-three distinct families exist for the a-chain and 25 for the b-chain. Each of these families may be subdivided further into subfamilies, representing both duplicated loci and allelic variation at the same locus. There are thought be 51 Vb (**TCRBV**) gene loci on chromosome 7, as well as a non-functional cluster on chromosome 9 (Robinson *et al*. 1993). There may be more than 100 different Va (**TCRAV**) chains in man. The g-chain has fewer V segments (14), although 4 are pseudogenes, and two clusters of J-C regions.

The d-chain lies within the a-chain complex and hence some V regions are utilized both by a- and d-chains. Multiple D elements can be used in a single transcript and extensive N-region diversity is seen in transcripts, considerably increasing the combinational diversity.

T-cell receptor repertoire

As T cells mature in the thymus, they undergo both positive and negative selection. As many as 95 per cent of T cells are deleted in the thymus, eliminating large populations of self-reactive receptors. In addition, particular receptors are positively selected in the thymus and directed to the periphery. MHC molecules are intimately involved in this positive and negative selection, and it forms one major mechanism of T-cell tolerance (Kappler and Roehm 1987). In the mouse, there is evidence of retrovirus-mediated negative selection of T-cell receptor TCRBV elements occurring intrathymically. This accounts for the profound effect on T-cell repertoire of the minor lymphocyte stimulating (**MLS**) antigen (see below). To date no equivalent process is recognized to exist in man. However, considerable diversity exists for both TCRAV and TCRBV usage in man. It is likely that both genetic (e.g. HLA type) and environmental factors contribute to this.

Allelic variation in the T-cell receptor

Despite the large range of allelic polymorphism in MHC molecules, less allelic variation has been documented in the V regions of the T-cell receptor. In part, this may be due to the difficulty in correlating polymorphism found in sequencing studies back to a single locus where multiple alleles may occur. In the absence of this formal genetic mapping, most sequence variation has been defined as a subfamily and more comprehensive studies are needed to define the extent of allelic variation. Several studies have suggested that susceptibility to some autoimmune disorders (e.g. multiple sclerosis) may be correlated with restriction fragment length polymorphisms within the germ line of the T-cell receptor complex (Oksenberg *et al*. 1989; Seboun *et al*. 1989). Major deletion/insertion polymorphisms within the TCRBV gene complex have also been described (Zhao *et al*. 1994). If this is true, allelic polymorphism may well have a role in determining an individual's susceptibility to autoimmunity (see below), and also may be important in dictating an immune response.

T-cell receptor/MHC interaction—a classical antigen recognition

Major interaction between the T-cell receptor and the MHC involves the recognition of polymorphic MHC determinants and peptides bound in the antigen-binding groove of MHC molecules by the T-cell heterodimer. Models for this classical interaction indicate that complementary-determining regions 1 and 2 of the T-cell receptor interact with the a-helices of the MHC molecule, while the highly polymorphic VDJ junctional regions interact with bound peptide (Davis and Bjorkman 1988; Stern *et al*. 1994). This is consistent with the extensive variation predicted from multiple peptides bound within a single antigen-binding site and relatively limited diversity seen within the a-helices.

Individual peptide–MHC combinations may be recognized by T-cell receptors that differ only slightly in their antigen-binding domains. This conservation of sequence suggests that genetically driven differences in T-cell repertoire may be important in determining susceptibility to autoimmune disorders where a single peptide antigen and major MHC restriction element are involved in the initiating stages.

A second form of interaction between T-cell receptor and MHC involves a class of molecules referred to as superantigens, which are encoded by either bacterial or viral pathogens and bind directly to V-region determinant and MHC elements. When expressed endogenously in the mouse by endogenous retroviruses (e.g. mouse mammary tumour virus) these MLS antigens can lead to the deletion of populations of T cells bearing particular receptor TCRVb chains (Dyson *et al*. 1991; Marrack *et al*. 1991). Similar molecules may also be expressed by environmental pathogens: the most common example of such exogenous superantigens are the b bacterial exotoxins such as staphylococcal enterotoxin and toxic shock syndrome toxin 1. These toxins can lead to the proliferation and activation of specific T-cell populations, events that may ultimately be responsible for some of the clinical features seen in these diseases.

T-cell receptor contributions to rheumatic disease

Differences between individuals in the receptor repertoire of T cells in peripheral blood may influence susceptibility to autoimmunity (see Chapter 3.1). This variation could be generated in two ways; first, as a result of germ-line polymorphism; second, as a consequence of selection processes dependent on the MHC in the thymus.

Polymorphism may exist at the level of germ-line sequences within T-cell receptor complexes. Of interest, one such polymorphism maps to the putative superantigen-binding site at residue 71 of the TCRVb chain. However, some evidence does exist from genetic studies that variation either within coding sequences or regulatory sequences may be present in T-cell receptor genomic DNA, and further, that these may contribute to particular autoimmune diseases, such as diabetes and multiple sclerosis. Studies of germ-line T-cell receptor polymorphisms in rheumatoid arthritis have produced interesting but as yet uncertain results (Funkhauser *et al*. 1992; McDermott *et al*. 1995).

The functional unit of the T-cell receptor on the cell surface has undergone extensive selection and modification during development in the thymus. This repertoire selection is both negative and positive in nature, and is in part determined by MHC products expressed within the thymus. The peripheral T-cell repertoire is more similar in identical twins than in unrelated individuals, suggesting that these variations may be one mechanism by which immune-response gene phenomena are determined (Loveridge *et al*. 1991). Analysis of the T-cell receptor repertoire in man has been possible only recently with the development of antibodies and polymerase chain reaction primers specific to individual V regions. Characterization of the repertoire in disease states has lagged even further behind. However, several studies have hinted that in autoimmune diseases, the lymphocytes infiltrating sites of tissue damage may show a biased T-cell receptor representation. This has been described in multiple sclerosis (Oksenberg *et al*. 1993) and also rheumatoid synovium (Howell *et al*. 1991; Paliard 1991). These studies all have methodological problems that make true quantification of V-region frequencies difficult, and hence their results can only be considered, at this time, speculative. In particular, the TCRBV14 predominance reported in rheumatoid joints has not been reproducible in other laboratories (Lunardi *et al*. 1992; Jenkins *et al*. 1993). For a detailed summary of the reported literature on T-cell receptors in rheumatoid arthritis to date see Struyk *et al*. (1995).

Expansion of discrete T-cell populations by superantigens

One appealing hypothesis has evolved that would help to explain the very strong associations between the presence of particular pathogens (e.g. salmonella, chlamydia) and HLA B27 in reactive arthritis. Putative superantigens from these organisms could in theory cause expansion of a particular population of T cells with the potential for autoreactivity. To date these organisms are not actually known to produce superantigens, but by inference from other bacteria and viruses, it seems quite likely that they may do so. Such expanded T-cell populations may be a subgroup with the potential to react with specific self-antigens, or alternatively, the superantigens may themselves alter the state of anergy of circulating T cells. Because of the established role of infectious pathogens in reactive arthritis, variations of repertoire may prove to be critically important in susceptibility. To date, little is known about the exact nature and alterations in peripheral T-cell repertoire to determine whether less variations might be important in dictating susceptibility to a range of inflammatory arthritides. It is likely, however, because of their critical role in the initial events in an antigen-driven immune response and because of their likely individual variation that important revelations may be forthcoming on the role of specific sequences in the initiation of the autoimmune process. Because of the interaction of the receptor with the peptide–MHC complex, such sequence conservation may be difficult to identify as it lies within the complementary-determining region 3 of the receptor. It is unlikely that V region-specific differences alone could account for the initiating events in disease and at present there is no evidence that the large V-region deletions mediated by endogenous superantigens in the mouse are present within the human. Identification of such effects may be difficult.

Both class I and class II MHC molecules have been implicated in susceptibility to certain rheumatic diseases. Characterization of the residues around the antigen-binding site suggests that the ability of individual MHC molecules to bind peptide may be central to the mechanism of susceptibility to these disorders. The following examples of MHC contributions to rheumatic disease share many similarities, but also indicate the complexities involved in assessing HLA-linked susceptibility. Not only may the primary susceptibility map either to the HLA class I or class II regions, but there may also be either primary or secondary effects arising from other loci that can influence susceptibility or the severity of these conditions.

HLA-B27 and the spondylarthritides

The spondylarthritides are characterized to varying degrees by inflammation of the spine and peripheral joints. They range from ankylosing spondylitis through the arthropathies associated with psoriasis and inflammatory bowel disease to Reiter's syndrome, in which there is clear evidence of an infective trigger. All of these conditions share similar patterns of joint involvement, are strongly associated with HLA-B27, and show at least circumstantial evidence of trigging by bacterial antigens.

Not only was the association between ankylosing spondylitis and HLA-B27 the first HLA association to be defined, but it remains to this day the strongest evidence linking susceptibility to a rheumatic disease with genes in the MHC. As at least 90 per cent of individuals with this condition are HLA-B27 positive, compared with only 7 per cent of the general population of the United Kingdom, the relative risk attributable to this antigen is approximately 120.

It is now known that there are at least 10 *HLA-B27* alleles, each differing by a limited number of amino acids around the antigen binding-site (Lopez de Castro 1995). Although the majority of these are associated with ankylosing spondylitis there is some evidence that HLA- *B*2703* in the Gambia and *HLA-B*2706* in Thailand may not be (reviewed in Khan 1995). In the United Kingdom both *HLA-B*2705* and *B*2702* are positively associated and in addition the class I allele *B*60* increases the risk threefold of ankylosing spondylitis in both B27-positive and B27-negative individuals (Brown *et al*. 1996).

Ankylosing spondylitis represents the tip of a clinical iceberg of spondylarthritides in which there are common clinical and histopathological features. Familial aggregation of these conditions is relatively common but does not necessarily breed true (i.e. some individuals in the family have ankylosing spondylitis while others have psoriasis or reactive arthritis). This can be explained by the association of each of these conditions with HLA-B27, although the strength of this association is variable. For example, HLA-B27 is only associated with those forms of enteropathic arthropathy in which sacroiliitis occurs, and even then only 50 per cent of the time, compared with over 90 per cent in ankylosing spondylitis. As triggering agents can be identified in a proportion (reactive arthritis following *Chlamydia trachomatis* infection of the genital tract or *Shigella flexneri* infection of the bowel), it is reasonable to infer that the arthropathy represents an unusual host immunological response to bacterial infection. Several ideas have been proposed to account for these observations (McMichael and Bell 1991), three of which are described below.

Linked gene

By this proposal, HLA-B27 itself is not directly involved in the disease process but is merely a marker for another closely linked gene. This could explain why the association with HLA-B27 is not absolute and why the association varies to some extent between racial groups. In some populations, such as black Africans, in which the frequency of HLA-B27 is very low (less than 1 per cent), the prevalence of ankylosing spondylitis is also low. However, the affected patients tend to be negative for HLA-B27.

There are, however, at least three major arguments against the linked gene hypothesis. First, the observation that the prevalence of the disease throughout the world generally mirrors the individual population frequencies of HLA-B27 suggests that HLA-B27 itself is involved directly. Second, the lack of families in which the disease can be seen cosegregating with a particular haplotype that is not HLA-B27 constitutes important evidence against the linked gene hypothesis. Third, additional evidence comes from the study of *HLA-B27*-transgenic rats. These animals spontaneously develop severe arthritis as well as other features sometimes associated with human spondylarthropathies, such as aortitis, uveitis, and psoriasis. The development of arthritis appears partly related to the number of copies of the *HLA-B27* gene incorporated into the genome of the transgenic animals (Hammer *et al*. 1990). In stark contrast, control animals transfected with *HLA-A2* do not develop arthritis, but the transgenic controls, for example with *HLA-A2*, are not perfect. Interestingly in this model, *HLA-B27*-transgenic rats reared in a germ-free state do not develop gut or joint abnormalities (Taurog *et al*. 1994), suggesting the involvement of both genetic (*HLA-B27*) and environmental factors, which may be of relevance to human ankylosing spondylitis.

Molecular mimicry

This theory gained considerable popularity in the late 1970s when it was demonstrated that antibodies raised against certain Gram-negative bacteria (*Klebsiella aeruginosa*) cross-reacted with HLA-B27-positive lymphocytes. Subsequently, faecal carriage of Klebsiella was shown in some studies, but not others, to relate to flares of the disease, in particular iridocyclitis and symptoms in peripheral joints. Circumstantial evidence to support this contention also came from the observation that serum IgA levels were elevated in active ankylosing spondylitis, indicating a mucosal response to bowel infection. More recently a molecular explanation for such a cross-reactive immune response has been suggested: an amino acid sequence from nitrogenase of *K. aeruginosa* shows marked homology with part of the HLA-B*2705 molecule, the most common variant of HLA-B27 in Caucasoids (Schwimmbeck *et al*. 1987). An epitope such as this could be the target for both pathogen and self-directed immune responses. Indeed, most of the amino acid substitutions corresponding to the HLA-B27 subtypes are clustered between positions 74 to 81 in the a-domain of the HLA heavy chain, precisely the region of proposed homology with Klebsiella nitrogenase.

The lesson from the reactive arthropathies is quite clearly that such organisms play a central part. Several reports have suggested the presence of bacterial antigens within the synovium of patients with both the postvenereal and postenteric form of reactive arthritis, although there is no evidence of the presence of intact, viable organisms within these joints (Howell *et al*. 1991; Hughes *et al*. 1991; Virtanen *et al*. 1991). Sensitization of the immune system to bacterial-derived antigens could in theory occur either within the joints or at a site distant from them, with subsequent traffic of activated T cells to the synovium. However, the precise mechanisms involved remain to be clarified.

Arthritogenic peptides

Arthritis can be induced in animals by a variety of materials under different sites of administration and adjuncts. Furthermore, adoptive transfer experiments have demonstrated that the disease can be transferred to other animals by T cells reactive to components of the inoculum (reviewed by van Eden 1991). Thus in adjutant arthritis, T cells recognizing the mycobacterial heat-shock protein hsp65 can pass on the disease to other animals not previously immunized. As heat-shock proteins are highly conserved between species and widely expressed in the tissue, cross-reactive immune responses could be responsible.

An immune response initially generated against a foreign immunogen may under certain circumstances by redirected against host self-antigens. For example, transgenic mice expressing the lymphocytic choriomeningitis virus (**LCV**) glycoprotein within the pancreas can develop diabetes mellitus in this way. The mice remain healthy until infected in later life by the LCV, whereupon an immune response directed against transgenic 'self' LCV glycoprotein in the pancreas leads to the development of tissue destruction and diabetes (Ohashi 1991).

Because it is now known that antigens are presented to the immune system in the form of short peptide fragments, it may be more appropriate to think in terms of 'arthritogenic peptides' as the triggering agents of these arthropathies rather than complex antigens or intact organisms. In animal models of immune-mediated disease, such as experimental autoimmune encephalomyelitis and collagen-induced arthritis, this approach has been extended to the investigation of the T-cell receptor usage in animals with these conditions. Thus TCRBV8 is commonly used by clones from defined mice strains with experimental autoimmune encephalomyelitis responding to particular myelin peptides (Urban *et al*. 1988) and anti-TCRVb8 and anti-TCRVb5 antibodies can ameliorate the disease in mice with collagen-induced arthritis (Chiocchia *et al*. 1991). Particular peptides may be common to a large number of different organisms and this could explain the wide variety of organisms that have been incriminated in the spondylarthritides. Although this type of response could account for some types of autoimmunity, it has not been described to date in man.

Definition of the genetic component of susceptibility to rheumatoid arthritis mapping to the MHC

Studies of identical and non-identical twins indicate that both genetic and non-genetic factors are involved in rheumatoid arthritis. At least 70 per cent of the identical twins of probands with rheumatoid arthritis do not develop the disease, clearly implicating non-genetic factors. However, concordance rates amongst non-identical twins (4 to 6 per cent) are approximately four to five times lower than for identical twins (15 to 30 per cent) (Lawrence 1970; Silman *et al*. 1993), from which a significant genetic contribution can also be inferred.

Most interest in the genetic contribution to rheumatoid arthritis has so far centred on the MHC. Studies of HLA haplotype sharing among affected siblings in families shows a significant trend for the sharing of at least one HLA haplotype by the affected siblings (Table 1). This observation links a gene (or genes) within the MHC to the development of rheumatoid arthritis (Payami *et al*. 1986). Further strong evidence of the importance of genes in this region comes from the study of outbred populations. Numerous studies in many racial groups have revealed associations with the HLA class II antigens DR4 (Wordsworth and Bell 1992). Furthermore, in some of these studies a weaker association with HLA-DR1 is evident, while in others there may be an association with HLA-DR1 or -DR10 (reviewed in Wordsworth *et al*. 1991). These data clearly implicate genes towards the centromeric end of the MHC in susceptibility to rheumatoid arthritis. However, they do not necessarily indicate a primary role for HLA-DR4 or -DR1 because these could equally well be genetic markers for another linked gene involved in the process of susceptibility.

	Shared haplotypes		
	2	1	0
Observed	112	147	52
Expected	78	156	78

Sharing ⩾ one haplotype χ^2=6.08; p<0.02.
The expected values are based on the prediction that 25 per cent would share two haplotypes, 50 per cent would share one, and 25 per cent would share no haplotypes. There is significant distortion of haplotypes shared by affected siblings, indicating that gene(s) within the MHC must be involved.

Table 1 Sharing of HLA haplotypes, identical by descent, between sibling pairs both of whom are affected by rheumatoid arthritis

Fine mapping of the precise susceptibility locus within the class II region has been achieved by the study of naturally occurring mutants that have given rise to a series of subtypes of HLA-DR, each producing a different predisposition to rheumatoid arthritis. Although the majority of HLA-DR specificities can be separated fairly easily on the basis of their associated DNA restriction fragment length polymorphisms, this is not true for the subtypes of HLA-DR4 (Nichlas *et al*. 1985). The sequence changes responsible for the different (formerly Dw-) subtypes of HLA-DR4 are restricted entirely to a few amino acids in the third allelic hypervariable region of the DR b-chain encoded at the DRB1 locus, which have probably arisen as a result of gene conversion events (Gregersen *et al*. 1986). In rheumatoid arthritis only subtypes *DRB1*0401* (Dw4), *DRB1*0404/8* (Dw14), and *DRB1*0405* (Dw15) are associated with predisposition to the disease, whereas *DRB1*0402* (Dw10) and *DRB1*0403/7* (Dw13) are not (Wordsworth *et al*. 1989; Gao *et al*. 1990; Nelson *et al*. 1991). The differential association of rheumatoid arthritis with particular subtypes of HLA-DR4 clearly pinpoints the DRB1 locus itself as the major determinant of susceptibility within the class II region, as the flanking DNA on all the DR4 haplotypes is the same.

Comparison of the amino acid sequences of the HLA-DR4 subtypes associated with rheumatoid arthritis and those unassociated with the disease reveals the presence of charged amino acids substitution within the third allelic hypervariable region of the DR b-chain (Table 2). These residues point into the antigen-binding site of the HLA-DR molecule (Fig. 8). They are therefore likely to have profound implications for peptide binding and interactions with T-cell receptors. Figure 9 shows an electron density of a peptide lying in the antigen-binding groove of an HLA-DR1 molecule (Brown *et al*. 1993). It is possible that the weaker associations between rheumatoid arthritis and HLA-DR1 and -DR10 may also be explicable on the basis of this third hypervariable region sequence. Both these molecules exhibit considerable similarities to the rheumatoid-associated HLA-DR4 subtypes, particularly in this region (Table 2).

			67			70	71			74		
Susceptible	DR4	DRB1*0408	Asp	Leu	Leu	Glu	Glu	Arg	Arg	Ala	Ala	Val
Susceptible	DR4	DRB1*0405	-	-	-	-	-	-	-	-	-	-
Susceptible	DR4	DRB1*0401	-	-	-	-	-	Lys	-	-	-	-
Susceptible	DR1		-	-	-	-	-	-	-	-	-	-
Susceptible	DR10		-	-	-	-	Arg	-	-	-	-	-
Susceptible	DR6	DRB1*1402	-	-	-	-	-	-	-	-	-	-
Not susceptible	DR4	DRB1*0402	-	Ile	-	-	Asp	Gln	-	-	-	-
Not susceptible	DR4	DRB1*0403?	-	-	-	-	-	-	-	-	Glu	-
Not susceptible	DR1		-	Phe	-	-	Asp	-	-	-	-	-

Table 2 Comparision of the amino acid sequences found in the third allelic hypervariable region of DR b-chains associated with rheumatoid arthritis and those that are not

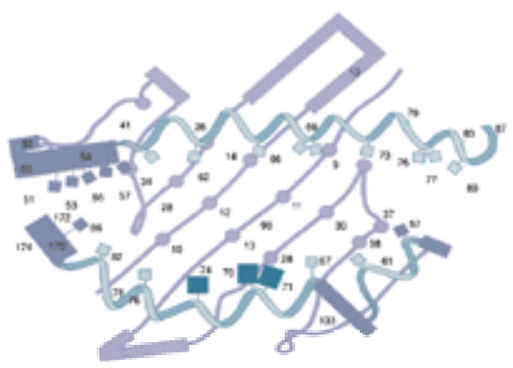


Fig. 8 Diagrammatic representation of the HLA-DR molecules, illustrating the binding site for antigens. Substitution of charged amino acids at positions pointing towards the binding site alter the susceptibility to rheumatoid arthritis. These positions (70, 71, 74) are shown as solid symbols on the DR b-chain a-helix. Relatively conservative substitutions at either end of the a-helix (positions 57, 86) do not affect susceptibility.

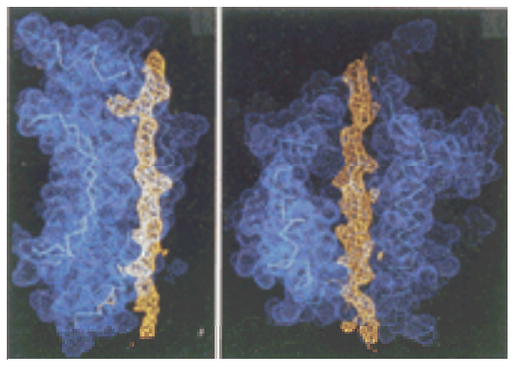


Fig. 9 Electron density map of a peptide lying within the antigen-binding groove of an HLA DR1 molecule (Brown *et al*. 1993).

Potential role of other MHC genes in autoimmune disorders

In many other HLA-associated autoimmune diseases there is evidence of a heterozygote effect or an additive effect from other HLA genes. Examples of these effects can be seen in insulin-dependent diabetes mellitus and coeliac disease. In the former, both HLA-DR3 and -DR4 are positively associated with susceptibility to diabetes but the DR3, DR4 heterozygote state greatly enhances this risk (Thomson 1988).

In northern Europeans, coeliac disease is typically associated with the extended HLA haplotype A1, B8, DR3, DQ2, but in some southern European populations an association with DR5/DR7 heterozygotes has been observed (Sollid *et al*. 1989). This observation has been interpreted as indicting that the primary restriction element in coeliac disease is DQ2, as the same DQ a-chain (*DQA1*0501*) and DQ b-chain (*DQB1*0201*) combinations can occur in both these situations (Fig. 10). In addition there appears to be a further effect in coeliac disease arising from another locus on the extended DR3 haplotype. Although DR3 and DQ2 occur on both the B8, DR3, DQ2 and B18, DR3, DQ2 extended haplotypes in the normal population of the United Kingdom, only the B8, DR3, DQ2 haplotype is found in patients with coeliac disease (Table 3). Therefore another gene in this haplotype must also be involved in determining susceptibility to this disease, although its identity is currently unknown (Rosenberg *et al*. 1989).

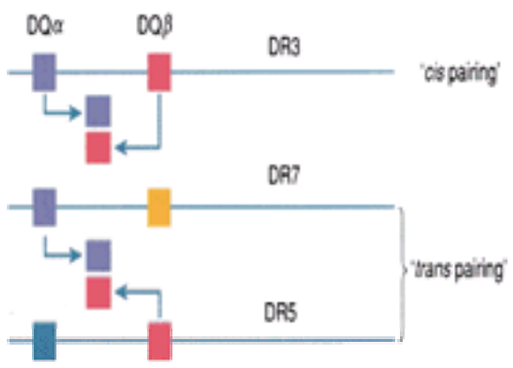


Fig. 10 The same DQ heterodimer may be assembled from the DQa and DQb paired in *cis* formation from a single DR3 haplotype or from the DQb of a DR7 haplotype paired in *trans* with the DQa from a DR5 haplotype.

	B8, DR3, DQ2	Non-B8, DR3, DQ2
DR3 coeliac patients (n=29)	29/29	0
DR3 healthy controls (n=23)	18/23	5/23

DR3-positive patients account for 90 per cent of all cases but DQ2 occurs in 95 per cent, pinpointing the latter as the primary susceptibility allele. However, a second locus on the B8, DR3, DQ2 extended haplotype must also be involved.

Table 3 The association of coeliac disease with an extended HLA-B8, DR3, DQ2 haplotype

Although rheumatoid arthritis is strongly associated with HLA-DR4, this is invariably linked to *DQB1*03* (both defined serologically) (Lanchbury *et al*. 1991). Several subtypes of *DQB1*03* have been defined, and it has been proposed that in rheumatoid arthritis, one of these subtypes, *DQB1*0301*, might predispose to more severe forms of the disease, such as Felty's syndrome. In this condition the association with HLA-DR4 is close to 95 per cent, and on these haplotypes there is an increase in the *DQB1*0301* subtype of *DQB1*03* at the expense of the *DQB1*0302* subtype. However, further analysis reveals that the primary association is with *DRB1*0401* (Lanchbury *et al*. 1991); the association with *DQB1*0301* is secondary, entirely explicable by the fact that *DRB1*0401* is more commonly linked to *DQB1*0301* than *DQB1*0302* (Table 4).

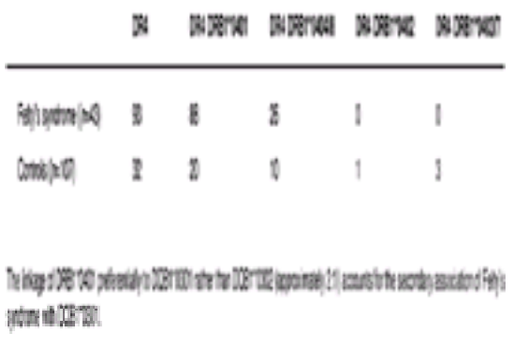


Table 4 The very strong association of DR4 and its Dw4 (*DRB1*0401*) subtype with susceptibility to Felty's syndrome

Nevertheless, in Felty's syndrome and seropositive erosive rheumatoid arthritis, there is evidence of more complex HLA-DR interactions. Several studies indicate that the HLA-DR association with rheumatoid arthritis is strongest with the most severe forms of the disease (Westedt *et al*. 1986). In the most severe forms of rheumatoid arthritis, HLA-DR4 homozygotes are also more common that would be expected by chance. Somewhat surprisingly, the majority of these DR4/DR4 homozygotes are actually *DRB1*0401/DRB1*0404/8* (Dw4/Dw14) compound heterozygotes (Table 5).

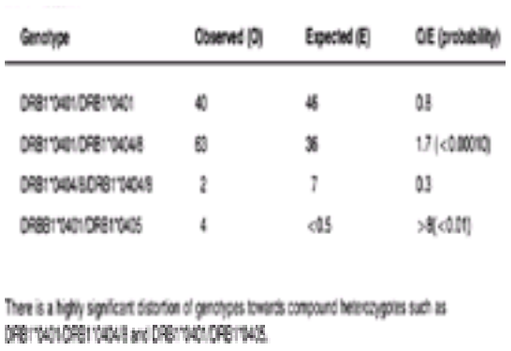


Table 5 A comparision of the subtype frequencies of HLA-DR4 in patients with severe rheumatoid arthritis or Felty's syndrome who are DR4/DR4 homozygotes

Overall these results can be interpreted in at least two ways. First there could be additional contributions from genes on particular HLA haplotypes carrying rheumatoid arthritis-associated HLA-DR alleles (particularly DR4 subtypes, DR1 and DR10). Second, these HLA-DR molecules themselves have the ability to bind and present a limited array of potentially arthritogenic peptides within the synovium to T cells. However, if this is the case, there is clearly potential synergy between these closely related alleles, in particular the compound heterozygotes, which needs further explanation. Perhaps these combinations allow more efficient binding and presentation of a range of peptides within the joint. Alternatively, the combination of two similar HLA-DR alleles may enhance the selection of a particular T-cell receptor repertoire during the thymic maturation of T lymphocytes (Wordsworth and Bell 1992).

Finally, other genes outside the MHC are undoubtedly important in determining susceptibility to rheumatoid arthritis. These could be interactive with HLA genes or provide independent susceptibility. An estimate of the importance of HLA genes in determining susceptibility can be gained from analysing haplotype sharing in sibling pairs both of whom are affected by rheumatoid arthritis. In a single-gene model of susceptibility, one would expect almost all affected siblings to share at least one HLA haplotype, but this is clearly not the case. Combined analysis of genes from over 300 families collected worldwide indicates that as many as 17 per cent of sibling pairs do not share HLA haplotypes (see Table 1). An estimate of the genetic contribution to rheumatoid arthritis arising from the MHC has been derived from a comparison of the recurrence risk in genetically identical monozygotic twins compared with HLA-identical siblings which suggests that no more than 30 per cent of the total is HLA linked (Deighton *et al*. 1989).

In insulin-dependent diabetes mellitus, another complex polygenic disorder, a genome-wide screen of a large number of families with affected sib pairs using 350 microsatellite marker loci has identified a number of susceptibility-associated loci outside the MHC (Davies *et al*. 1994). This approach is likely to yield similar findings in rheumatoid arthritis. A comprehensive understanding of the immunological responses underlying many of these disorders may require a fuller elucidation of these other genes involved in determining susceptibility.

Chapter References

Bell, J.I. *et al*. (1987). Molecular mapping of class II polymorphisms in the human major histocompatibility complex. I. DRb. *Journal of Immunology*, **139**, 562–73.

Bell, J.I., Todd, J.A., and McDevitt, H.O. (1989). The molecular basis of HLA-disease association. *Advances in Human Genetics*, **18**, 1–41.

Bjorkman, P.J. *et al*. (1987). The foreign antigen binding site and T cell recognition regions of class I histocompatibility antigens. *Nature*, **329**, 512–18.

Bodmer, J.G. *et al*. (1994). Nomenclature for factors of the HLA system 1994. *Tissue Antigens*, **44**, 1–18.

Bottazzo, G.F., Pujol-Borrell, R., Hanafusa, T., and Feldman, M. (1983). Role of aberrant HLA-DR expression and antigen presentation in induction of endocrine autoimmunity. *Lancet*, **ii**, 115–18.

Brown, J.H. *et al*. (1993). Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. *Nature*, **364**, 33–9.

Brown, M.A. *et al*. (1996). HLA class I associations of ankylosing spondylitis in Caucasians in the United Kingdom. *Annals of the Rheumatic Diseases*, **55**, 268–70.

Chiocchia, G., Boissier, M.-C., and Fournier, C. (1991). Therapy against murine collagen-induced arthritis with T cell receptor Vb specific antibodies. *European Journal of Immunology*, **21**, 2899–905.

Davies, J.L. *et al*. (1994). A genome-wide search for human type I diabetes susceptibility genes. *Nature*, **371**, 130–7.

Davis, M.M. and Bjorkman, P.J. (1988). T cell antigen receptor genes and T cell recognition. *Nature*, **334**, 395–9.

Davis, M.M. and Chien, Y-H. (1995). Issues concerning the nature of antigen recognition by ab and gd T-cell receptors. *Immunology Today*, **16**, 316–18.

Deighton, C.M., Walker, D.J., Griffiths, I.D., and Roberts, D.F. (1989). The contribution of HLA to rheumatoid arthritis. *Clinical Genetics*, **36**, 178–82.

Dyson, P.J., Knight, A.M., Fairchild, S., Simpson, E., and Tomonari, K. (1991). Genes encoding ligands for deletion of VbII T cells cosegregate with mammary tumour virus genomes. *Nature*, **349**, 551.

Funkhauser, S.W., Concannon, P., Charmley, P., Vredevoe, D.L., and Hood, L. (1992). Differences in T cell receptor restriction fragment length polymorphisms in patients with rheumatoid arthritis. *Arthritis and Rheumatism*, **35**, 465–71.

Gao, X., Olsen, N.J., Pincus, T., and Stastny, P. (1990). HLA-DR alleles with naturally occurring amino acid substitutions and risk for development of rheumatoid arthritis. *Arthritis and Rheumatism*, **33**, 939–46.

Garrett, T.P.J., Saper, J.A., Bjorkman, P.J., Strominger, J.L., and Wiley, D.C. (1989). Specificity pockets for the side chains of peptide antigens in HLA-Aw68. *Nature*, **342**, 692–6.

Gregersen, P.K. *et al*. (1986). Molecular diversity of HLA-DR4 haplotypes. *Proceedings of the National Academy of Sciences* (*USA*), **83**, 2642–6.

Hammer, R.E., Maika, S.D., Richardson, J.A., Tang, J., and D.T.J. (1990). Spontaneous inflammatary disease in transgenic rats expressing HLA-B27 and human b2m: an animal model of HLA-B27 associated human disorders. *Cell*, **63**, 1099–112.

Hardy, D.A., Bell, J.I., Long, E.O., Lindsten, T., and McDevitt, H.O. (1986). Mapping of the class II region of the human major histocompatibility complex by pulsed-field gel electrophoresis. *Nature*, **323**, 453–5.

Hill, A.V. (1991). Common West African HLA antigens are associated with protection from severe malaria. *Nature*, **352**, 596–600.

Hill, A.V. *et al*. (1991). HLA class I typing by PCR: HLA-B27 and an African B27 subtype. *Lancet*, **337**, 640–3.

Howell, M.D. *et al*. (1991). Limited T-cell receptor b-chain heterogeneity among interleukin 2 receptor-positive synovial T cells suggests a role for superantigen in rheumatoid arthritis. *Proceedings of the National Academy of Sciences* (*USA*), **88**, 10921–5.

Hughes, R.A., Hyder, E., Treharne, J.D., and Keat, A.C.S. (1991). Intra-articular chlamydia antigens and inflammatory arthritis. *Quarterly Journal of Medicine*, **80**, 575–88.

Jardetzky, T.S., Lane, W.S., Robinson, R.A., Madden, D.R., and Wiley, D.C. (1991). Identification of self peptides bound to purified HLA-B27. *Nature*, **353**, 326–9.

Jenkins, R.N., Nikaein, A., Zimmermann, A., Meek, K., and Lipsky, P.E. (1993). T cell receptor Vb gene bias in rheumatoid arthritis. *Journal of Clinical Investigation*, **92**, 2688–701.

Kappler, J.W. and Roehm, N. (1987). T cell tolerance by clonal elimination in the thymus. *Cell*, **49**, 273–9.

Khan, M.A. (1995). HLA-B27 and its subtypes in world populations. *Current Opinion in Rheumatology*, **7**, 263–9.

Lanchbury, J.S. *et al*. (1991). Strong primary selection for the Dw4 subtype of DR4 accounts for the HLA-DQw7 association with Felty's syndrome. *Human Immunology*, **32**, 56–64.

Lawrence, J. (1970). Rheumatoid arthritis—nature or nurture? *Annals of the Rheumatic Diseases*, **29**, 357.

Lopez de Castro, J.A. (1995). Structural polymorphism and function of HLA-B27. *Current Opinion in Rheumatology*, **7**, 270–8.

Loveridge, J.A., Rosenberg, W.M., Kirkwood, T.B., and Bell, J.I. (1991). The genetic contribution to human T-cell receptor repertoire. *Immunology*, **74**, 246–50.

Lunardi, C., Marguerie, C., and So, A.K. (1992). An altered repertoire of T cell receptor V gene expression by rheumatoid synovial fluid T lymphocytes. *Clinical and Experimental Immunology*, **90**, 440–6.

MacMurray, A.J. *et al*. (1987). Molecular mapping class II polymorphisms in the human major histocompatibility complex. II. DQb. *Journal of Immunology*, **139**, 574–86.

Madden, D.R., Gorga, J.C., Strominger, J.L., and Wiley, D.C. (1991). The structure of the HLA-B27 reveals nonamer self-peptides bound in an exended conformation. *Nature*, **353**, 321–5.

Marrack, P., Kushnir, E., and Kappler, J. (1991). A maternally inherited superantigen encoded by a mammary tumour virus. *Nature*, **349**, 524.

McDermott, M. *et al*. (1995). The role of T cell receptor b chain genes in susceptibility to rheumatiod arthritis. *Arthritis and Rheumatism*, **38**, 91–5.

McMichael, A. and Bell, J. (1991). HLA-B27: a disease related immune response gene. *Research in Immunology*, **142**, 475–82.

Monaco, J.J. (1992). A molecular model of MHC class-1 restricted antigen processing. *Immunology Today*, **13**, 173–9.

Morgan, B.P. and Walport, M.J. (1991). Complementary deficiency and disease. *Immunology Today*, **12**, 301–6.

Moss, P.A. *et al*. (1991). Extensive conservation of a and b chains of the human T-cell antigen receptor recognizing HLA-A2 and influenza A matrix peptide. *Proceedings of the National Academy of Sciences* (*USA*), **88**, 8987–90.

Neefjes, J. and Ploegh, H.L. (1992). Intracellular transport of MHC class II molecules. *Immunology Today*, **13**, 179–84.

Nelson, J.L. *et al*. (1991). Dw14 (DRB1*0404) is a Dw4-dependent risk factor for rheumatoid arthritis. *Tissue Antigens*, **38**, 145–51.

Nichlas, J.A., Noreen, H.J., Segall, M., and Back, F.H. (1985). Southern analysis of DNA polymorphism among Dw subtypes of DR4. *Human Immunology*, **13**, 95.

Ohashi, P.H. (1991). Ablation of 'tolerance' and induction of diabetes by virus infection in viral antigen transgenic mice. *Cell*, **65**, 305–17.

Oksenberg, J.R. *et al*. (1989). T-cell receptor Va and Ca alleles associated with multiple and myasthenia gravis. *Proceedings of the National Academy of Sciences* (*USA*), **86**, 988–92.

Oksenberg, J.R. *et al*. (1993). Selection for T-cell receptor Vb-Db-Jb-gene rearrangements with specificity for a myelin basic protein peptide in brain lesions of multiple sclerosis. *Nature*, **362**, 68–73.

Olerup, O. and Zetterquist, H. (1992). HLA-DR typing by PCR amplification with sequence-specific primers (PCR-SSP) in 2 h: an alternative to serological DR typing in clinical practice including donor-recipient matching in cadaveric transplantation. *Tissue Antigens*, **39**, 225–35.

Paliard, X. (1991). Evidence for the effects of a superantigen in rheumatoid arthritis. *Science*, **253**, 325–8.

Paul, W E. and Seder, R.A. (1994). Lymphocyte responses and cytokines. *Cell*, **76**, 241–51.

Payami, H. *et al*. (1986). Genetics of rheumatoid arthritis. *Tissue Antigens*, **27**, 57–63.

Robinson, M.A. *et al*. (1993). Organization of human T cell receptor b-chain genes: clusters of Vb genes are present on chromosomes 7 and 9. *Proceedings of the National Academy of Sciences* (*USA*), **90**, 2433–7.

Rosenberg, W.M.C., Wordsworth, B.P., Jewell, D.P., and Bell, J.I. (1989). A locus telomeric to HLA-DPB encodes susceptibility to coelic disease. *Immunogenetics*, **30**, 307–10.

Sargent, C.A., Dunham, I., Trowsdale, J., and Campbell, R.D. (1989 *a*). Human major histocompatibility complex contains genes for the major heat shock protein HSP70. *Proceedings of the National Academy of Sciences* (*USA*), **86**, 1968–72.

Sargent, C.A., Dunham, I., and Campbell, R.D. (1989 *b*). Identification of multiple HTF-island associated genes in the human major histocompatibility complex class III region. *Embo Journal*, **8**, 2305–12.

Schwimmbeck, P.L., Yu, D.T.Y., and Oldstone, M.B.A. (1987). Autoantibodies to HLA-B27 in patients with ankylosing spondylitis and Reiter's syndrome. Molecular minicry with *Klebsiella pneunomiae* as potential mechanisms of autoimmune disease. *Journal of Experimental Medicine*, **166**, 173–81.

Seboun, E. *et al*. (1989). A susceptible locus for multiple sclerosis is linked to the T-cell receptor b chain complex. *Cell*, **57**, 1095–100.

Serjeantson, S.W. (1983). HLA and susceptibility to leprosy. *Immunological Reviews*, **70**, 89–112.

Silman, A.J. *et al*. (1993). Twin concordance rates for rheumatoid arthritis: results from a nationwide study. *British Journal of Rheumatology*, **32**, 903–7.

Sollid, L.M., Markussen, G., Ek, J., Gjerde, H., and Thjorsby, E. (1989). Evidence for a primary association of coeliac disease to a particular HLA-DQ a/b heterodimer. *Journal of Experimental Medicine*, **169**, 345–50.

Spies, T. *et al*. (1986). Genes for the tumor necrosis factors-a and -b are linked to the human major histocompatibility complex. *Proceedings of the National Academy of Sciences* (*USA*), **83**, 8699–702.

Stastny, P., Ball, E.J., Dry, P.J., and Nunez, G. (1983). The human immune response region (HLA-D) and disease susceptibility. *Immunological Reviews*, **70**, 115–54.

Stern, L. *et al*. (1994). Crystal structure of the human class 2 MHC protein HLA-DR1 complexed with an influenza virus peptide. *Nature*, **368**, 215–21.

Struyk, L., Hawes, G.E., Chatila, M.K., Breedveld, F.C., Kurnick, J.T., and van den Elsen, P.J. (1995). T cell receptors in rheumatiod arthritis. *Arthitis and Rheumatism*, **38**, 577–89.

Taurog, J.D. *et al*. (1994). The germfree state prevents development of gut and joint inflammatory disease in HLA-B27 transgenic rats. *Journal of Experimental Medicine*, **180**, 2359–64.

Thomson, G. (1988). Genetic heterogeneity, modes of inheritance and risk estimates from a joint study of Caucasians with insulin-dependent diabetes mellitus. *American Journal of Human Genetics*, **43**, 799–816.

Townsend, A. and Bodmer, H. (1989). Antigen recognition by class I restricted T lymphocytes. *Annual Review of Immunology*, **7**, 601–24.

Trowsdale, J. and Campbell, R.D. (1992). Complexity in the major histocompatibility complex. *European Journal of Immunogenetics*, **19**, 43–55.

Urban, J.L. *et al*. (1988). Restricted use of T cell receptor V genes in murine autoimmune encephalomyelitis raises possibilities for antibody therapy. *Cell*, **54**, 577–92.

van Eden, W. (1991). Heat shock proteins as immunogenic bacterial antigens with the potential to induce and regulate autoimmune arthritis. *Immunological Reviews*, **121**, 5–28.

Vaughan, R.W. *et al*. (1990). The application of oligonucleotide probes to HLA class Ii typing of the DRB sub-region. *Tissue Antigens*, **36**, 149–55.

Virtanen, A.M., Arstila, T.P., Lahesmaa, R., Granfors, K., Skurnik, M., and Toivanen, P. (1991). Application of the polymerase chain reaction and immunoflurescence technique to the detection of bacteria in *Yersinia*-triggered reactive arthritis. *Arthritis and Rheumatism*, **34**, 89–96.

Westedt, M.L. *et al*. (1986). Immunogenetic heterogeneity of rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **45**, 534–8.

Wordsworth, B.P. and Bell, J.I. (1992). The immunogenetics of rheumatoid arthritis. *Springer Seminars in Immunopathology*, **14**, 59–78.

Wordsworth, B.P. *et al*. (1989). HLA-DR4 subtype frequencies in rheumatoid arthritis indicate that DRB1 is the major susceptibility locus within the HLA class II region. *Proceedings of the National Academy of Sciences* (*USA*), **86**, 10049–53.

Wordsworth, B.P., Allsopp, C.E.M., Young, R.P., and Bell, J.I. (1990). HLA-DR typing using DNA amplification by the polymerase chain reaction and sequential hybridisation to sequence-specific oligonucleotide probes. *Immunogenetics*, **32**, 413–18.

Wordsworth, B.P., Stedeford, J., Rosenberg, W.M.C., and Bell, J.I. (1991). Limited heterogeneity of the HLA class II contribution to susceptibility to rheumatoid arthritis is suggested by positive associations with HLA-DR1 and DRw10. *British Journal of Rheumatology*, **30**, 178–80.

Zhao, T.M., Whitaker, S.E., and Robinson, M.A. (1994). A genetically determined insertion/deletion related polymorphism in human T cell receptor b chain (TCRB) includes functional variable gene segments. *Journal of Experimental Medicine*, **180**, 1405–14.

3.4 Animal models of arthritis

W. B. van den Berg

Introduction

A major research goal in the field of arthritis is to unravel the pathogenesis of chronic arthritis and the concomitant joint destruction. A second, more practical goal, is to discover a treatment that selectively inhibits the progression of destructive arthritis yet leaves the host defence mechanisms virtually intact. This requires a critical understanding of the cells and mediators involved in the destruction and in the initiation, maintenance, and remission of the arthritic process. Studies in human arthritis are hampered by the fact that the precise time of onset is unknown, whereas lesional tissue is often obtained from endstage disease at the moment of joint replacement. The latter holds true in particular for cartilage specimens, since it is widely accepted that this tissue has a limited capacity for repair. In the past, synovial biopsies and synovial fluid were only taken from patients with an exacerbation of symptoms that necessitated arthroscopy or the withdrawal of an effusion, inherently representing extremes of the natural history of the disease. Moreover, specimens come from patients who have been receiving various drugs, and the shortage of control tissue is obvious. Numerous arthritis units have now started to run early arthritis clinics, in which early biopsies are taken more often and even the taking of small biopsies of articular cartilage surfaces from select areas have been considered. Another way to circumvent the problems is to resort to the use of experimental models of arthritis.

Although not ideal in the eyes of many in terms of precise mimicry of human arthritic disease, experimental models of arthritis do reflect key aspects of their human counterparts. Their time course, the easy access to tissue samples, and the facility for experimental therapeutic manipulation offer a useful approach to further understanding of the pathogenesis of arthritis. Models may also provide valuable insights into biological approaches to arthritis therapy, ranging from cytokines, cytokine inhibitors, and 'tolerizing' antigens to monoclonal antibodies against cellular receptors and vaccination directed against effector cells. Potential therapies need to be evaluated in the human diseases, indirectly approving the predictive value of findings in particular models. It is true that no single animal model of arthritis completely represents the human disease. In fact, the wide variety of agents that can induce an experimental arthritis with clinical and histopathological features close to those of the human arthritides supports the hypothesis that rheumatoid arthritis may have a variety of causes and that the characteristic features reflect common endpoints. Analysis of aspects peculiar to an individual model may be of limited value and the emphasis should be on general validity and common concepts in various models. In the following sections the models most widely used in the study of rheumatoid arthritis will be summarized and their value will be illustrated with some recent research findings. Since the questions to be answered in models must arise from elements of the human disease, current concepts will be briefly addressed first.

Concepts and features of rheumatoid arthritis

Rheumatoid arthritis is characterized by chronic inflammation in the joints and progressive destruction of bone and cartilage. Its pathogenesis is unknown, but the disease is often considered as an autoimmune process. The articular cartilage is an intriguing tissue, since it may function both as the trigger as well as the victim of this disease (van den Berg 1994). Two major events may underlie the chronic synovial inflammation: persistent stimulation of T cells by as yet unknown (auto-) antigens, or direct activation of the synovial cells by non-antigenic triggers. The latter may imply continuous stimulation by bacterial or viral triggers, but may also reflect deranged behaviour and tumour-like growth. Figure 1 depicts potential cascades in the synovial inflammation and concomitant cartilage destruction. It illustrates the various levels of potential therapeutic interference, ranging from antigen presentation to T cells, control by regulatory T cells, activation of synovial cells, and release of destructive mediators acting on the articular cartilage. It is, of course, attractive to target research in models at the level of the arthritogen, but, in the absence of an established antigen, to seek further understanding of effector mechanisms characteristic of joint inflammation and destruction seems at least as valuable. Recent research approaches in models are summarized in Table 1, which for obvious reasons cannot all be covered in detail. A recent overview by Wooley (1991) is recommended.

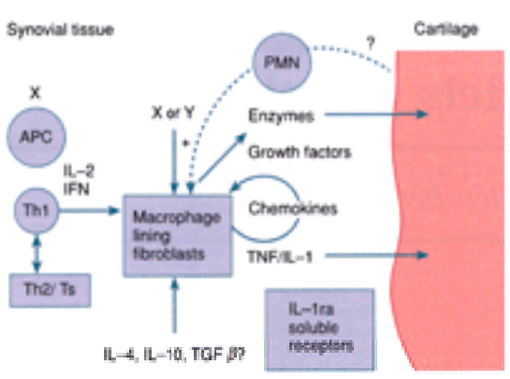


Fig. 1 Cytokine cascades in the arthritic joint.

Application of MHC blockade
Manipulation of defined T-cell subsets
Selective T-cell receptor usage/blockade
Potency/limits of T-cell vaccination
Regulatory role of heat-shock proteins
Control by bacterial flora
Preventive induction of tolerance
Immunomodulation in established arthritis?
Induction of bystander suppression
Neutralization of key cytokines
Up-regulation of cytokine inhibitors
Blockade of adhesion molecules
Identification of destructive cells/enzymes
Disease course in transgenic/knock-out mice

MHC, major histocompatibility complex.

Table 1 Recent approaches in arthritis models

Characteristic histopathological features in the rheumatoid arthritic joint include immune complexes in the articular cartilage layers and variable amounts of macrophages and T cells in the synovium, often accompanied by fibrosis and synovial hyperplasia. The antigens in the immune complexes in the cartilage are still poorly defined and, although candidates for a role in the synovitis, may even reflect an epiphenomenon of immune-complex deposition and retention in damaged areas. Considerable variation is found between patients and at various stages in one patient, making it difficult or even impossible to define critical correlations in animal models.

Models of arthritis

Historically, models of arthritis have been used to understand the stimuli provoking chronic arthritis and the mechanisms regulating chronicity and tissue destruction. When a model is established, the potential of a given stimulus is proven. The next step would be to obtain evidence that such reactions do occur in human arthritides, for instance T-cell reactivity against a cartilage antigen. If this is the case, it has still to be proved that this reactivity is of pathogenetic importance and not an epiphenomenon, for example by showing that antigen-specific immunomodulation really affects the course of the disease. Up to now such research has not yielded a clue about the definitive trigger of rheumatoid arthritis and continuing investigations can be divided into those that attempt further understanding of principles in established models and those that are still looking for new, putative triggers, and novel models and concepts.

In line with the historical concepts in rheumatoid arthritis the models most widely studied in the past decades are those of adjuvant arthritis, collagen-induced arthritis, antigen-induced arthritis, and streptococcal cell-wall arthritis. T cells play a dominant part in all of these models (Table 2). The second common principle is the presence of a chronic stimulus, either in the form of a persistent antigen or an autoantigen akin to joint structures. Persisting antigens are non-degradable bacterial cell walls in the synovial tissue, or antigen trapped in the collagenous reservoirs such as ligaments and articular cartilage (antigen-induced arthritis). Both conditions reflect escape from proper clearance by the phagocytic system. A second category of persistent stimuli is formed by autoantigens from the articular cartilage, such as collagen type II and proteoglycans. It cannot be excluded that other, yet unknown cartilage antigens may fulfil a similar role. In adjuvant and streptococcal cell-wall arthritides the cartilage could as well function as an autoantigen, related to structural mimicry between bacterial peptidoglycans and cartilage proteoglycans. However, ultimate proof that these cross-reactive responses really contribute to the arthritis has still to be provided. Of interest, destructive forms of rheumatoid arthritis tend to decline at the moment that the cartilage is fully destroyed. Moreover, total joint replacement often results in a complete remission of the arthritis in that particular joint, without the need for concomitant synovectomy. These are possible arguments for a direct role of cartilage antigens in the pathogenesis or an indirect role of cartilage components in the maintenance of the inflammatory process in the joint.

Model	Species	Immunization	Induction	Pers. Ag*	Auto Ag	T cells
Adjuvant arthritis	Rat	CFA		Bact?	?	+
Collagen-induced arthritis	Rat, mouse, primate	CII/CFA			Collagen II	+
Proteoglycan arthritis	BALB/c mouse	PG/CFA			Proteoglycan	+
Streptococcal cell-wall arthritis	Rat		i.p. SCW	SCW		+
Antigen-induced arthritis	Rabbit, rat, mouse	Ag/CFA	i.v. Ag	Antigen		+
CII-induced arthritis	Rat, mouse	CII			?	+
MRL/lpr/lpr	MRL mouse			Virus?		-

Ag, antigen; CFA, complete Freund's adjuvant; CII, collagen type II; PG, proteoglycan; SCW, streptococcal cell wall; i.p., intraperitoneal; i.v., intravenous.
*Relates to persistent antigen in the joints.

Table 2 Common models of arthritis

The third lesson from models is that microbial components, particularly cell-wall fragments from enteric organisms, are potential causative agents in humans. Apart from their ability to induce arthritis by direct localization to joint tissues, they may induce arthritis remotely as a result of molecular mimicry with joint structures. There is ample evidence from clinical observations that bacterial infections and development of arthritis may somehow be related.

It is almost a matter of choice whether persistent exogenous antigens or autoantigens should be considered as different entities or as reflecting integral parts of the body that need tight regulation of suppression and/or tolerance. Certainly, it is as yet unclear whether the regulation of these forms of immunity is the same.

In line with the increasing doubt in various research groups about the critical role of a T-cell-driven reaction to a defined (auto-) antigen, increasing attention is nowadays focused on less defined models, such as the spontaneous arthritis in MRL/lpr mice and the arthritis induced in susceptible strains with mineral oils. In addition, plain overexpression of cytokines [tumour necrosis factor (**TNF**)-a] or H_2-c-*fos* in transgenic animals provides further evidence that the expression of arthritis may occur in the absence of a clear involvement of T cells. In the following sections some of the commonly used models will be critically evaluated. Comparative information is given in Table 3 and Table 4. Technical details can be found in some overviews (Billingham 1983; Bliven and Otterness 1985; Greenwald and Diamond 1988; Hunneyball *et al*. 1989).

	AA	CIA	SCW-A	AIA
Main site of expression	Ankle	Peripheral	Ankle	Knee*
Bone marrow inflammation	++	-	+	-
Local plasma cells	-	±	-	++
Immune complexes in cartilage	-	++	-	++
Cartilage destruction	±	+++	+	++
Effect of NSAIDs†	+	++	+	-
Dominant feature	Periostitis	Destructive	Fibrosis	Destructive

*To be chosen by intra-articular injection; †effect on ongoing destruction.
AA, adjuvant arthritis; CIA, collagen-induced arthritis; SCW-A, streptococcal cell-wall arthritis; AIA, antigen-induced arthritis; NSAID, non-steroidal anti-inflammatory drug.

Table 3 Characteristic features of the arthritis models

Adjuvant arthritis
- Self-limiting arthritis
- Spontaneous tolerance
- Non-specific immunomodulation
- Dependence on bacterial flora

Collagen arthritis
- Self-limiting arthritis
- Genetic restriction
- Epitope-specific immunomodulation
- Sensitive to modulatory cytokines

Streptococcal arthritis
- Spontaneous resurrences
- Tumour-like synovial growth
- Cross-reactive flares*
- Dependence on bacterial flora

Antigen-induced arthritis
- Controlled onset, cartilage destruction
- Antigen retention in collagenous tissues
- Antigen-specific local hyperreactivity
- Chronicity by repeated flares

*Inducible in the unilateral model.

Table 4 Models used to study pathogenic mechanisms

Adjuvant arthritis

This is the first and most extensively studied model of polyarthritis, which was discovered some 40 years ago by Stoerk *et al*. (1954) as an experimental accident. Whilst trying to produce immunity to spleen extracts emulsified in Freund's adjuvant, an arthritis developed in the immunized rats. Pearson (1956) demonstrated that the arthritis was due simply to the bacterial component. The model has since been extensively used for the screening of drugs to be used in rheumatoid arthritis.

Nowadays the classical model is induced by intradermal injection of Freund's complete adjuvant containing heat-killed mycobacteria and the arthritis develops within 2 weeks in susceptible rat strains. In general, the model is induced in Lewis rats. The volume, type of oil, and composition of the suspension are critical variables that determine the incidence and severity of the arthritis. The active component in the bacteria is the cell-wall peptidoglycan and the disease can be induced with various bacteria.

The histopathological features of adjuvant arthritis mainly reflect a periarthritis, with marked periostitis instead of a synovitis, and massive inflammation in the bone marrow. Immune-complex deposition in the cartilage is not a characteristic feature and cartilage destruction is limited in the early disease (Fig. 2). The highly destructive attack on the cartilage seen in the autoimmune collagen type-II arthritis argues against cross-reactivity with cartilaginous autoantigens in adjuvant arthritis as the driving principle (van de Langerijt *et al*. 1994a).

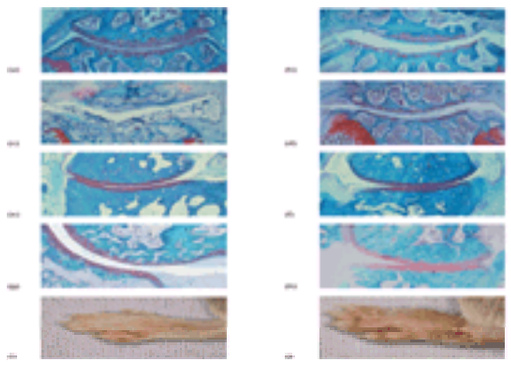


Fig. 2 Light-microscopic analysis of cartilage proteoglycan (PG) depletion in safranin O-stained knee (a–d) and ankle (e–h) joints. (a) Normal knee joint; (b) destructive collagen arthritis (CIA), day 14 after onset; (c) complete destruction in CIA, day 28; (d) antigen-induced arthritis, day 28, showing superficial loss of PG and bone erosion as well as outgrowth at the edges; (e) normal ankle; (f) adjuvant arthritis, day 17; (g) adjuvant arthritis, day 28; (h) streptococcal cell-wall, arthritis, day 17. Macroscopic appearance of collagen arthritis: (i) first onset in one toe; (j) full-blown expression.

Adjuvant arthritis is T-cell dependent and the strongest argument for an autoimmune process is the induction of arthritis by passive transfer of T cells from diseased animals. The joint inflammation may reflect the generation of a T-cell reaction to bacterial epitopes cross-reacting with endogenous bacterial fragments continuously present in synovial tissues or with cartilaginous antigens. It may also be based on non-specific immunomodulation, reflecting the adjuvant properties of the bacterium in oil preparations and the generation of a disregulated expression of autoimmunity to whatever autoimmune epitope is involved. The fact that non-antigenic adjuvants, such as the oil preparation avridin (CP 20961) and other mineral oils, can induce an arthritis indistinguishable from adjuvant arthritis may be in support of the latter hypothesis.

A bacterium-specific pathogenesis is likely since conventionally bred rats are generally resistant to adjuvant arthritis, whereas germ-free Fisher or Wistar rats are susceptible (van de Langerijt *et al*. 1994b). The germ-free rats are lacking early contact with bacteria and are therefore not 'tolerized'; colonization with bacteria before the induction of adjuvant arthritis prevented susceptibility (Kohashi *et al*. 1986). Susceptibility may also relate to an impaired steroid feedback response. Lewis rats are generally susceptible to numerous autoimmune processes and are low steroid responders to a large range of stimuli, whereas the non-susceptible Fisher rats are high responders. However, germ-free Fishers are also high steroid responders yet are susceptible to adjuvant arthritis, although the disease expression is less pronounced. This suggests that the regulation of the antibacterial T-cell responses is of prime importance and that steroid responsiveness of a rat strain or a particular individual is an element modulating the severity of the disease.

The most intriguing observation in the adjuvant arthritis model is the occurrence of spontaneous remission and the lack of susceptibility to reinduction. This resistance is antigen-specific. T-cell lines and clones were isolated that can induce disease, but when attenuated could also induce protective responses (Holoshitz *et al*. 1983). Attempts to repeat these studies were mainly unsuccessful, suggesting that the regulation is rather complex. Theoretically, it involves regulatory T cells that recognize antigen-specific receptors on the T cells driving the arthritis and are able to block the activity of those cells. The regulatory cells are not arthritogenic themselves. This principle, called T-cell vaccination, has raised a major interest in the further characterization of T-cell receptor usage in human arthritis, but so far this research has not yielded a defined receptor usage enabling therapeutic applications.

Treatment with antibodies against CD4, mainly present on T cells and macrophages, is of interest because, apart from the suppression of adjuvant arthritis, such treatment induced tolerance to the initiating arthritogen. Subsequently, it was found that rats undergoing anti-CD4 treatment before the induction of arthritis with either streptococci or non-antigenic adjuvants (mineral oils, CP20961) displayed a shared tolerance to the reinduction of both models, implying that there might be common bacterial epitopes in the regulation of adjuvant arthritides [for further details, see the review by Billingham (1994)].

The identification of epitopes on bacterial heat-shock proteins and the recognition of cross-reactive, highly conserved, endogenous heat-shock proteins in eukaryotes, has implicated these proteins as the target antigen in adjuvant arthritis (van Eden *et al*. 1989; van Eden 1991). However, subsequent research demonstrated a role for heat-shock proteins in other autoimmune models also, and a regulatory role in inflammation rather than as a critical antigen in arthritis seems more likely.

In industry, adjuvant arthritis is often the model of first choice for screening new therapeutic agents for anti-arthritic efficacy. This is mainly based on its ease of induction and the simple macroscopic observation of arthritis in the paws. The fact that non-steroidal anti-inflammatory drugs are effective inhibitors of cartilage and bone destruction in this model, in clear contrast to observations in human rheumatoid arthritis, puts some doubt on its applicability.

Collagen-induced arthritis

The model of collagen arthritis in rats was first described in 1977 by Trentham and colleagues, again as a coincidental finding in protocols to induce antibodies to purified collagen preparations (Trentham *et al*. 1977). The initial observation indicated that arthritis was confined to sensitization with native collagen type II, a major component of articular cartilage. Denatured collagen type II or other collagen types were not arthritogenic. Nowadays, data are accumulating that minor collagen types from articular cartilage may also function as arthritogens, for instance collagen type IX and XI (Holmdahl *et al*. 1993; Cremer *et al*. 1994).

The crucial element in the pathogenesis of this arthritis is the induction of immunity to foreign collagen type II, subsequently cross-reacting with homologous collagen type II. Plain immunization with homologous collagen type II can also be used, but then much stronger immunization regimens are needed to override natural tolerance. The disease can easily be induced in rats, with full-blown expression within 14 days, whereas expression in mice seems to follow tight genetic restriction (Wooley *et al*. 1981) Moreover, disease expression in mice is more gradual, starting after 3 to 4 weeks in some, whereas a 100 per cent incidence commonly takes 8 to 10 weeks. Of interest, collagen arthritis can also be induced in non-human primates. Most rhesus monkeys were susceptible (Bakker *et al*. 1991) and instead of the precipitation of a susceptibility gene, linkage studies on the major histocompatibility complex (**MHC**) revealed the presence of a gene regulating resistance.

Unlike adjuvant arthritis, collagen arthritis is not a systemic illness, but is localized to the peripheral joints and spares the spine. However, the ears may be affected, and this feature is mainly found at late stages in rats. This may suggest a role for this type of reactivity in polychondritis. In murine collagen-induced arthritis, marked disease was also found in knee joints, in addition to the paws, ankle, and wrist. Histopathology of collagen arthritis shows a distinct, acute synovitis with numerous

granulocytes, and bone erosions as well as periosteal new bone formation. Involvement of the bone marrow is limited. A characteristic feature is the direct attack by granulocytes at the cartilage surface. In contrast to findings in other models a complete loss of articular cartilage is often seen within 2 weeks (Fig. 2) and the arthritis results in ankylosis, with limited inflammation. The lack of sustaining antigen is probably the main reason for the remission of the arthritis. In addition, regulating T cells have been demonstrated in late-phase disease, which demonstrates that one should be careful with anti-T-cell therapy in human arthritis such treatment can either improve the condition or make it worse (Maeda *et al*. 1994).

The mechanism of collagen-induced arthritis is based on two principles: the presence of anticollagen antibodies and the generation of anticollagen type II T-cell immunity. Although antibodies alone are able to induce arthritis after passive administration to naive recipient animals, high concentrations are needed and, at best, a transient arthritis occurs. Passive transfer with bulk T cells or clones also produce mild or transient arthritis. Antibodies may be needed to bind to the cartilage surface and to release further collagen epitopes upon complement fixation and the attraction of leucocytes, including granulocytes and lymphocytes. A subsequent influx of anticollagen type II-specific T cells will than further drive the arthritic process.

It is convincingly demonstrated that the expression of collagen arthritis can be enhanced by extra anticollagen type II antibodies, non-specific inflammatory stimuli such as lipopolysaccharide or yeast particles (zymosan), or the simple addition of single inflammatory mediators such as interleukin (**IL**)-1, TNF-a, or transforming growth factor-b (**TGF-b**) (Hom *et al*. 1992; Thorbecke *et al*. 1992; Joosten *et al*. 1994). These observations suggest that slumbering autoimmune arthritis comes to full expression by a combination of potentiating elements. Recent studies not only implicate potentiating cytokines, but also the temporary control by modulators such as IL-4 and IL-10. This is further addressed under the heading 'Cytokines' below.

Much research has been focused on oral tolerance and intriguing progress has been made in identifying arthritogenic and tolerogenic epitopes on the collagen type II protein (Thompson *et al*. 1993; Myers *et al*. 1993a,b). Fragments containing tolerogenic epitopes may perhaps be safely applied to down-modulate anticollagen type II autoimmunity and to suppress established arthritis, without the risk of exacerbation. This collagen-induced arthritis model, using a defined autoantigen, is also suitable for analysing the restricted usage of T-cell receptors (**TCR**) and the possibility of suppressing arthritis by blocking a particular receptor. Treatment with antibodies against Vb8, a predominant TCR b-chain used in the immune response against collagen type II, prevented collagen arthritis (Moder *et al*. 1993) However, although specific TCR-Vb expression was noted in early synovial infiltrates at day 2 after the onset of collagen arthritis, this specificity was lost after only 1 week (Erlandsson *et al*. 1994). This particularly questions the usefulness of investigating TCR-Vb usage in the rheumatoid synovium.

Clear T-cell reactivity to collagen type II cannot easily be detected in patients with rheumatoid arthritis, casting some doubt on the relevance of this potential arthritogen in humans, but it is hoped that refined methods will provide further insight into relevant epitopes. A recent study of Trentham *et al*. (1993) indicated the usefulness of collagen type II in suppressing arthritis when given by the tolerizing oral route. Whether this really implies the possibility of antigen-specific down-modulation of collagen type II T-cell reactivity, indirectly proving the relevance of this response, or merely reflects non-specific 'bystander' suppression through the generation of suppressive cytokines (IL-10, TGF-b), remains to be seen. In fact, the anticollagen type II T-cell responses analysed so far were already low in patients before the treatment.

A worrying finding in terms of comparison with human arthritis is the highly destructive character of collagen arthritis and the marked sensitivity to non-steroidal anti-inflammatory drugs. Indomethacin is a very potent suppressor of both the inflammation and the joint destruction; steroids are also highly effective. This complicates experimental studies and stress can also profoundly disturb the expression of the arthritis. In contrast to the female preponderance in rheumatoid arthritis, male mice are more susceptible to collagen arthritis.

Proteoglycan arthritis

This is a rather new model and a logical extension of the collagen arthritis, since both collagen type II and proteoglycan are major components of articular cartilage. Yet again the discovery of the model was coincidental, following the immunization of mice to prepare antibodies. Repeated boosting was needed to induce consistent arthritis after 8 weeks, implicating poor antigenicity or tolerance. Importantly, arthritis was only noticed in inbred female BALB/c mice upon immunization with human fetal articular proteoglycan, stripped of chondroitin sulphate (Glant *et al*. 1987; Mickecz *et al*. 1990). The arthritogenic epitope resides in the core protein. The mechanism of induction of the arthritis is probably quite similar to that in collagen-induced arthritis: the induction of immunity to fetal human proteoglycan, subsequently cross-reacting with murine proteoglycans.

The proteoglycan model shows a polyarthritis, with severe cartilage erosion and marked ankylosis. In addition, involvement of the lumbar spine and disc regions was found, making it a model for spondylitis also. Like collagen arthritis, the most severe expression of proteoglycan arthritis is found in the presence of both antibodies and antiproteoglycan T-cell immunity. Of interest, antiproteoglycan antibodies on their own were capable of causing marked loss of proteoglycan from the cartilage in the absence of distinct synovitis.

Unfortunately, screening for the occurrence of such antiproteoglycan immunity in patients with rheumatoid arthritis did not yield unequivocal data in support of a role in human arthritis. It is to be expected that the further characterization of proteoglycan subtypes and epitopes may lead to better defined antigens, warranting thorough analysis of particular T-cell responses and antibodies in patients with arthritis.

Streptococcal cell-wall arthritis

This model was originally described by Cromartie *et al*. (1977). It was induced in Lewis rats by the systemic injection of cell-wall fragments of group A streptococci, which are highly resistant to biodegradation. Later on a similar disease was induced with cell-wall fragments from other bacteria, such as *Lactobacillus casei* or *Eubacterium aerofaciens*. The common principle resides in the poor degradability of the fragments, thereby creating a persistent stimulus. The lactobacillus and eubacterial models are of particular interest for the human disease, since these bacteria are part of the normal gastrointestinal flora (Stimpson *et al*. 1986; Hazenberg *et al*. 1992), which implies that an enormous load of potential arthritogenic stimuli is continuously present in the normal gastrointestinal tract.

Within 24 h of the administration of cell-wall fragments, acute inflammation develops in peripheral joints, coincident with the dissemination of cell-wall fragments in the blood vessels of the synovium and in the subchondral bone marrow. This acute, complement-dependent inflammation subsides over the next week and is followed within 2 weeks by a chronic, erosive polyarthritis, which involves mainly peripheral joints. In contrast with the acute phase, the chronic joint inflammation develops in only a limited number of rat strains, with the highest incidence in Lewis rats. The chronic phase often shows waxing and waning of arthritis, which brings it close to human rheumatoid arthritis. Mice strains studied so far are not susceptible to this form of arthritis. In general, female rats show a more severe arthritis than males.

Although macrophages become stimulated by the persistent bacterial fragments, cogent evidence now exists that the chronic phase is dependent on T cells. The chronic phase was not inducible in nude Lewis rats (no T cells) and cyclosporin A effectively inhibited this phase. In addition, chronic arthritis can be prevented with antibodies against the ab TCR (Yoshino *et al*. 1991). Moreover, streptococcal cell wall-specific T-cell responses were found in arthritis-susceptible Lewis rats, whereas resistant Fisher rats did not mount this immune reaction. Finally, germ-free Fishers rats were susceptible and did show streptococcal cell wall-specific T-cell reactivity. This set of data suggests that the chronic arthritis is driven by a streptococcal cell wall-specific T-cell reaction to persistent bacteria. Normal animals and most individuals are strongly tolerant of threatening arthritogenic reactions to bacterial cell walls, whereas Lewis rats and similar individuals display weak tolerance and are easily overloaded so as to lose tolerogenic control. Lewis rats display a disturbance in the hypothalamic–pituitary–adrenal axis, reflected in a low feedback response to endogenous corticosteroid (Sternberg *et al*. 1989). Such a defect is also noted in patients with rheumatoid arthritis, which is not attributable to chronic inflammation itself. As already discussed in the section on adjuvant arthritis, the defect in the hypothalamic–pituitary–adrenal axis is probably not a decisive but more a regulatory principle.

In addition to streptococcal cell wall-specific T-cell reactions, cross-reactive autoimmunity to cartilage proteoglycans may contribute to the chronic arthritis. However, it is unlikely that this is a major factor in its onset (van de Langerijt *et al*. 1994a). In fact, early histopathological appearances are those of a strong, mononuclear synovitis, with a sparse exudate in the joint space and limited loss of proteoglycan from the articular cartilage (Fig. 2). Only at later stages of the chronic arthritis were marked pannus formation and severe erosions of underlying cartilage and bone observed frequently.

In line with the involvement of growth factors and the tumour-like behaviour of synovial cells in patients with rheumatoid arthritis, similar characteristics have been found in synoviocytes from streptococcal cell-wall arthritic rats (Case *et al*. 1989). Probably due to the persistent bacterial stimuli, the cells do show continued proliferation *in vitro*, with apparent paracrine and autocrine regulation by growth factors (Lafyatis *et al*. 1989). This observation further delineates that macrophage–fibroblast activation may be a perpetuating principle, but it leaves unanswered why *in vivo* the observed T-cell dependence is still a critical, controlling

factor.

The model has rarely been used for drug studies, which is a pity. Cyclosporin shows efficacy and it has been claimed that non-steroidal anti-inflammatory drugs and steroids are suppressive as well. Gold and penicillamine were without effect, whereas methotrexate showed moderate activity. The main reason for limited studies in this model is the difficulty of preparing arthritogenic bacterial cell walls.

Antigen-induced arthritis

Since human rheumatoid arthritis is believed to be an immunologically mediated disease, albeit with an unknown inciting antigen, a model based on local antigenic challenge in a primed host appeared logical. Such a model was first developed by Dumonde and Glynn (1962) in rabbits. In principle it can be induced in any species, provided that proper immunity to a particular antigen can be mounted, and extensions have since been developed in mice, rats, and guinea-pigs. In contrast to the polyarthritis models described so far, this type of arthritis remains confined to the injected joint, enabling comparison with a contralateral joint of the same animal.

Commonly used antigens were ovalbumin, bovine serum albumin, and fibrin. Preimmunization is performed with antigen in complete Freund's adjuvant to induce strong humoral as well as cell-mediated immunity. Arthritis is usually induced 3 weeks later by a local injection in the knee joint of a large amount of antigen. Initially an Arthus type of reaction develops, followed by a T-cell-mediated chronic inflammation. In the rabbit, chronicity may last for years. The histopathological appearances are of a granulocyte-rich exudate in the joint space, thickening of the synovial lining layer, and, at later stages, a predominantly mononuclear infiltrate in the synovium, which includes numerous T cells and clusters of plasma cells. Interestingly, a large proportion (50 per cent) of these plasma cells are still making antibodies to the inciting antigen, providing evidence that the retained antigen might still be the driving force in the chronic arthritis. Intense immune-complex formation is seen in the superficial layers of the articular cartilage, which may contribute to the process of cartilage destruction. Early loss of proteoglycan, followed by pannus formation and cartilage and bone erosion, is a common finding. Of the models described so far these characteristics are the closest to those found in human rheumatoid arthritis.

Two important principles emerged from studies on antigen-induced arthritis: first, chronicity is only found in the presence of sufficient antigen retention in joint tissues, in combination with proper T-cell-mediated delayed hypersensitivity; second, joints contain numerous, avascular collagenous tissues such as cartilage, ligaments, and tendons, which allow for prolonged antigen retention by antibody-mediated trapping and charge-mediated binding (Cooke *et al*. 1972; van den Berg *et al*. 1984; Fig. 3). A key finding was the observation that antigen injected in the skin produced transient inflammation, whereas a similar dose in the joints caused chronic arthritis. The chronicity of arthritis in the standard model is probably related to the generation of local hyper-reactivity. Antigen trapped in the collagenous tissues will be released in tiny amounts to sustain low-grade chronic arthritis. As a consequence the local T-cell infiltrate will gain specificity, since retention of specific T cells is enhanced by homologous antigen. This means that small amounts of antigen are sufficient to sustain arthritis, whereas relatively large amounts are needed to induce it. This is a pivotal finding and forms the basis for exacerbations of arthritis described below (van den Berg *et al*. 1982).

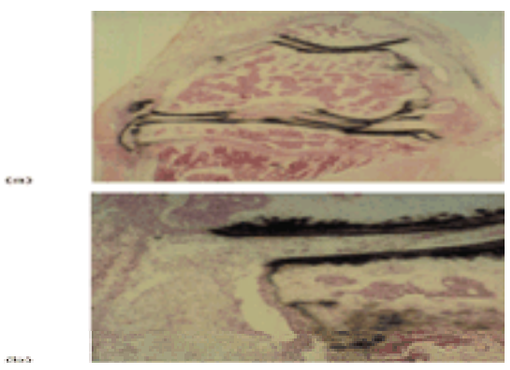


Fig. 3 Autoradiograph at day 2 after intra-articular injection of [^{125}I]-labelled bovine serum albumin (BSA) in BSA/Freund's complete adjuvant immunized mice: (a) whole knee joint, (b) detail of patella and femur; note the deep penetration of antigen in cartilaginous structures.

The finding in the mouse that cationic antigens are proper arthritogens in this model, owing to their ability to stick to the negatively charged collagenous structures of the joint, has led to a search for putative natural antigens with similar properties. Interestingly, cationic components of bacteria such as streptococci appeared to be stimuli in the model (Mertz *et al*. 1991), but subsequent searches for such T-cell reactivities in human rheumatoid arthritis were negative. However, similar reactivity was recently demonstrated in reactive arthritis (Probst *et al*. 1993), and antigen-induced arthritis, using a cationic antigen, probably makes a good model for reactive arthritis.

The model of antigen-induced arthritis is most suited to studies into the mechanism of cartilage destruction, which are facilitated by knowledge of the exact time of onset, the accessibility of the knee joint (as compared with ankles), and the presence of a contralateral control joint. Moreover, the model can be adequately used to test the feasibility of approaches to therapeutic immunomodulation. Comparison with autoimmune models may provide insight into similarities or dissimilarities in regulation of immunity to chosen protein antigens or cartilage autoantigens.

Antigen-induced arthritis is insensitive to non-steroidal anti-inflammatory drugs (de Vries and van den Berg 1989; Hunneyball *et al*. 1989), like the human condition. Steroids are highly effective, cytotoxic drugs are potent suppressants, and gold compounds were shown to be effective in the rabbit model.

Flares of arthritis

In comparison with the chronic process of human rheumatoid arthritis, a general shortcoming of most models is the relatively short duration of a severe and rapidly destructive inflammation. In that respect, models of repeated flares of arthritis, with slower development of lesions, provide a valuable extension. As stated earlier, an arthritic joint bearing a chronic T-cell infiltrate may display a state of local hyper-reactivity against the retained antigen, contributing to chronicity. However, it is intriguing that this is not restricted to retained antigen but also applies to new antigen entering the sensitized joint. In the model of bovine serum albumin-induced arthritis, flares of the smouldering arthritis are easily induced with as little as 10 ng of antigen. The flare is a T-cell-dependent process and can be completely blocked by *in vivo* antibody treatment against MHC (Ia) or CD4 (Lens *et al*. 1984). The flare can be induced by local rechallenge, but also by intravenous or even oral antigen administration. Higher dosages are, of course, needed for intravenous or oral challenges, and access to the joint is dependent on systemic antibodies and the physicochemical properties of the antigen. A model of repeated flares is probably more akin to the human state than is a model showing severe inflammation for some weeks, followed by rapid waning of arthritis. In a considerable proportion of patients with rheumatoid arthritis the disease course is characterized by exacerbations and remissions.

An important extension of this model was found by comparative dosing of IL-1 to naive joints and joints bearing a chronic infiltrate (van de Loo *et al*. 1992b). The infiltrated joint was much more sensitive to IL-1, and the reactivity seemed to reside in the macrophage infiltrate. Most importantly, the IL-1-induced flare was more highly destructive to the articular cartilage than the initial insult, and further understanding of the underlying mechanisms of such destruction is much needed.

In addition to models of flares of arthritis based on protein antigen, similar models have been developed in rats and mice using bacterial cell-wall constituents. In contrast to small protein antigens, which are only inflammatory in the context of an immune response, bacterial fragments may function as an antigen as well as a phlogistic irritant in their own right, and the ensuing reactions are a mixture of T-cell- and macrophage-driven processes. The generation of local hyper-reactivity requires large, persistent bacterial peptidoglycan–polysaccharide components, but the recurrence may happen with a variety of stimulants ranging from fragments, to lipopolysaccharide, to cytokines like IL-1 (Stimpson *et al*. 1987). The strongest flares occur in the presence of T-cell immunity and a correlation was found between the potential of fragments to induce an exacerbation and to elicit cell wall-specific T-cell proliferation *in vitro* (van den Broek *et al*. 1988; van den Broek *et al*. 1990; van den Berg *et al*. 1991). One other important aspect of these flares with cell-wall fragments resides in the presence of considerable cross-reactivity between cell walls from different bacterial origins. Flares can be induced with homologous as well as heterologous fragments (Esser *et al*. 1985) and this may even extend to cross-reactive autoantigens from the cartilage. Table 5 summarizes key aspects of such flares. These principles open up a wide range of putative stimuli involved in

exacerbations, simultaneously complicating the search for arthritogenic antigen in humans.

Local hyperreactivity in chronic infiltrate
Tiny dosages of homologous antigen sufficient
Contribution of cross-reactive antigens/irritants
Both T-cell and macrophage mediated
Hypersensitivity to TNF-α, IL-1, TGF-β

Table 5 Pathophysiology of flare reactions

Other models

In line with the trend of increasing concern about a particular T-cell-driven pathogenesis in rheumatoid arthritis, models not based on a specific antigenic trigger have received major attention in recent years. These include arthritis with various types of mineral oils, spontaneous models displaying arthritis amongst other changes, models based on superantigens or viral antigens exacerbating established models, and transgenic models based on overexpression of cytokines or mediators involved in cellular activation like c- *fos* or c-*jun*. These models reflect the hyper-reactivity of synovial cells or a general disturbance of the control of autoimmunity, either spontaneous or caused by compounds showing distinct adjuvant properties. In addition, interest has been raised in the **SCID** (severe combined immunodeficiency disease) mouse. This immunocompromised animal permits the *in vivo* study of the pathological potential of cells from animal models or patients with rheumatoid arthritis. For this purpose, cells or pieces of synovial tissue are transferred to the SCID mouse and pathological changes analysed (Williams *et al*. 1992b; Elkon and Ashany 1993). An interesting design is the combination of cells or tissue with cartilage as a target tissue, to obtain further insight into mechanisms of cartilage destruction. Detailed discussion of this model goes beyond the scope of this chapter. Some of the features of the other models will be addressed briefly.

Adjuvant oils can induce a symmetrical, destructive polyarthritis when injected intradermally in DA rats (Kleinau *et al*. 1991). Expression of the arthritis occurs between days 11 and 14, is found in 100 per cent of the rats, and lasts for 6 weeks. As in the classical adjuvant arthritis, readministration of oil to rats that had recovered from oil-induced arthritis fails to induce arthritis a second time. This points to an immunological background and indeed the arthritis could be transferred with concanavalin A-activated T cells from arthritic rats to irradiated recipients. A seemingly similar disease could be induced with adjuvant oil in certain strains of mice and was termed pristane arthritis (Wooley *et al*. 1989; Thompson and Elson 1993). The pristane disease, however, has proved difficult to characterize, due to late onset, variable penetrance, and difficulty of transfer. Moreover, in late disease numerous types of autoantibodies were noted, including rheumatoid factor, which may contribute to the expression of the arthritis and make this model less clearly T-cell driven. In clear contrast to findings in adjuvant and streptococcal cell-wall arthritides, pristane arthritis was suppressed in germ-free mice, implying involvement of bacterial flora and a bacterium-specific pathogenesis. Of interest, spontaneous arthritis may occur in DBA/1 mice, the strain that is highly susceptible to collagen-induced arthritis (Nordling *et al*. 1992). The spontaneous model seems to reflect aspects of the collagen arthritis since improvement could be achieved with anti-idiotypic antibodies to anticollagen antibodies. This is in sharp contrast with the adjuvant arthritis, which has been demonstrated to have a pathogenetic pathway different from that of collagen arthritis (Cremer *et al*. 1990; Holmdahl and Kvick 1992).

Spontaneous arthritis is also described in MRL–lpr/lpr mice (Hang *et al*. 1982; Koopman and Gay 1988). These animals develop a severe autoimmune disease, mainly characterized by massive lymphadenopathy, arteritis, immune complex-mediated glomerulonephritis, and chronic arthritis. The serological abnormalities in these animals include antibodies against native DNA, rheumatoid factors, and circulating immune complexes. This strain of mice may thus be regarded as a model of both systemic lupus erythematosus and rheumatoid arthritis. The presence of rheumatoid factors, which is lacking in most of the induced models, makes this model of potential interest. However, the incidence of arthritis is much lower than the incidence of the systemic lupus-like syndrome and is much more variable in presentation. Moreover, upon standard breeding it is often noted that the incidence of arthritis is further diminished, due to preferential survival and breeding potential of the more healthy individuals. The arthritis is characterized by synovial and mesenchymal cell hyperplasia, late T-cell infiltration, and cartilage destruction. The first signs are synovial cells with a transformed appearance and invasion of these cells into cartilage and bone, resulting in a rheumatoid arthritis-like pannus. Significant arthritis occurs only in aged mice, and signs are mild or absent before the age of 5 months. A viral cause has been suggested. Expression can be enhanced with an injection of Freund's complete antigen (Ratkay *et al*. 1993) or superantigen, the latter probably by induction of T-cell anergy (Mountz *et al*. 1994). In line with the autoimmune character, immunosuppressive drugs such as cyclophosphamide and leflunomide are effective in this model (Bartlett *et al*. 1988). In addition, chloroquine and gold were antiarthritic.

Transgenic mice expressing the human TNF transgene develop chronic polyarthritis with a 100 per cent incidence (Keffer *et al*. 1991). Hyperplasia of the synovium, inflammatory infiltrates in the joint space, pannus formation, and cartilage destruction were observed, and the expression could be completely blocked with anti-TNF-a antibodies. This model illustrates that systemic overexpression of TNF-a may lead to the precipitation of inflammatory processes in the joints.

The potential involvement of retroviral antigens in chronic arthritis was further underlined by the occurrence of arthritis after 2 to 3 months in mice transgenic for human T-cell leukaemia virus (Iwakura *et al*. 1991; Yamamoto *et al*. 1993). Recently, transgenic mice overexpressing c-*fos* were used in model studies. Overexpression of c-*fos* in synovial cells did not lead to arthritis. However, the eliciting of antigen-induced or collagen arthritides in these c- *fos* mice yielded more severe and more destructive arthritis. Remarkably, the cellular infiltrate in these mice contained hardly any lymphocytes, yet marked cartilage destruction was found, stressing the role of mesenchymal cells in that damage (Shiozawa *et al*. 1992). The expression of c-*fos* coincides with the enhanced expression of stromelysin and collagenase.

Involvement of cytokines

The synovial inflammation in patients with rheumatoid arthritis is characterized by an abundance of macrophage- and fibroblast-derived mediators and a relative lack of the T-cell factors IL-2 and interferon-g (Firestein 1991). It is encouraging to note that a considerable hierarchy exists in the plethora of cytokines and growth factors, putting major emphasis on cytokines TNF-a and IL-1. Characterization of the temporary involvement of these cytokines in the various models may shed some light on the (lack of) resemblance of the model arthritic processes to the human disease and may indirectly yield further understanding of the rheumatoid arthritic process. Before addressing the involvement of TNF/IL-1 it must be remembered that early studies into the synovial expression of IL-2 and interferon-g in proven, T-cell-driven experimental models of arthritis encountered major difficulties in the demonstration of these mediators. This should be taken as a warning that the relative absence of IL-2 and interferon-g in rheumatoid synovia must not be overinterpreted to mean exclusion of a T-cell-driven origin. Approaches to obtaining evidence of the involvement of TNF and IL-1 are the use of neutralizing antibodies, specific binding proteins or receptor antagonists in the various models (Henderson and Blake 1992).

Although adjuvant arthritis is still the model most widely used in drug discovery, information on TNF/IL-1 involvement is scant. Treatment with antibodies to the IL-1 receptor, starting directly after immunization, resulted in a marked reduction of arthritis. Such an early treatment may interfere with the generation of autoimmune responses in this model. (Schorlemmer *et al*. 1993). Anti-IL-1 treatment hardly affected late joint swelling and cell influx, but anti-TNF blocked the influx of lymphocytes and granulocytes in the arthritic joint and reduced swelling (Issekutz *et al*. 1994).

In murine collagen arthritis, anti-IL-1 and anti-TNF have been used before onset, shortly after onset, and in the established phase. TNF plays an important part in the onset of collagen-induced arthritis but is less dominant in late arthritis (Williams *et al*. 1992a; Wooley *et al*. 1993a). In contrast, IL-1 is a pivotal mediator in both early and established collagen-induced arthritis (Fig. 4). The elimination of IL-1 greatly suppressed the arthritis and yielded marked protection against cartilage destruction (van den Berg *et al*. 1994). The protection could be demonstrated using either neutralizing antibodies or IL-1 receptor antagonist, provided that large amounts (1 mg/day per mouse) of the antagonist were continuously supplied in osmotic minipumps. Given the poor pharmacokinetics and the need for sustained receptor

blocking, it must be stressed that any therapeutic application in patients with rheumatoid arthritis would need extremely high dosages.

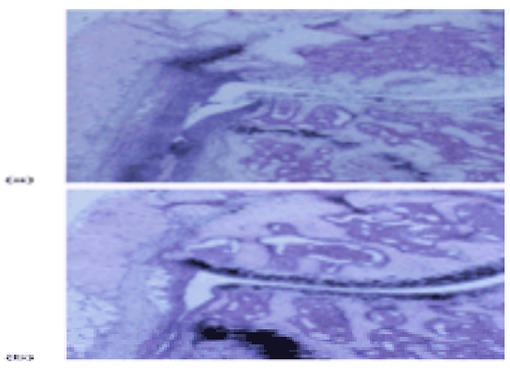


Fig. 4 Autoradiograph of a knee joint (patella/femur) at day 35 after induction of collagen arthritis: (a) arthritic control, (b) anti-IL-1ab treated (from day 28); note the restoration of ^{35}S incorporation in the chondrocytes.

In the flare model of streptococcal cell-wall arthritis in the rat both TNF-a and IL-1 appeared to be important (Schwab *et al*. 1991). A similar study in the mouse revealed a dominant role for TNF in the joint inflammation, but still a pivotal role for IL-1 in cartilage destruction. Flares induced with superantigen were IL-1/TNF independent (Schwab *et al*. 1993).

Finally, in antigen-induced arthritis in the rabbit and the mouse, elimination of both TNF-a and IL-1 was not very effective in suppressing joint inflammation, pointing to substantial 'overkill' by other mediators (Wooley *et al*. 1993b; Lewthwaite *et al*. 1994). However, elimination of IL-1 did yield protection against cartilage destruction (van de Loo *et al*. 1992a) and this was even more striking in the antigen-induced flare (van de Loo *et al*. 1995).

In summary, the dominance of IL-1 and TNF in joint inflammation in the different models is variable, implying that the stimulus, the type of process, and probably also the phase of the arthritis are important determinants. Given the apparent dominance of TNF-a in the cytokine cascade in rheumatoid synovial tissue, the model of streptococcal cell-wall arthritis is perhaps the closest to the human disease. However, all phases studied so far are rather more acute than the human disease and further study is needed in the late stages or repetitive flares of the models. Moreover, future studies are needed in arthritis models to dissect any critical dependence on the cytokine profiles of distinct T-cell or macrophage-driven processes. A consistent and important finding is the potential uncoupling of joint inflammation and cartilage destruction. Although IL-1 is not a dominant inflammatory cytokine in all models, it is certainly the pivotal cytokine in the inhibition of chondrocyte proteoglycan synthesis in all models studied so far and the blocking of IL-1 has a great impact on net cartilage destruction.

Apart from the cytokines TNF-a and IL-1, modulatory cytokines such as IL-4, IL-10, and TGF-b, and specific endogenous inhibitors like shed receptors or IL-1 receptor antagonist are of prime importance. Consistent findings have been made in collagen-induced arthritis and streptococcal cell-wall arthritis. Local TGF-b is a crucial proinflammatory mediator, yet becomes anti-inflammatory at later stages. TGF-b, when applied systemically, is generally anti-inflammatory (Santambrogio *et al*. 1993; Wahl *et al*. 1993; Wahl 1994). A similar dichotomy seems to hold for IL-4 and IL-10, which complicates any therapeutic application. These mediators may interfere at various levels of the arthritic process including skewing of T-helper cells, inhibition of macrophage cytokine production and direct action on chondrocytes (Fig. 1 and Fig. 5) The up-regulation of IL-1 receptor antagonist by IL-4 contributes to the suppression of streptococcal cell-wall arthritis (Allen *et al*. 1993). It is expected that the study of experimental arthritis in the recently developed cytokine 'knock-out' mice will further identify crucial roles of the various cytokines.

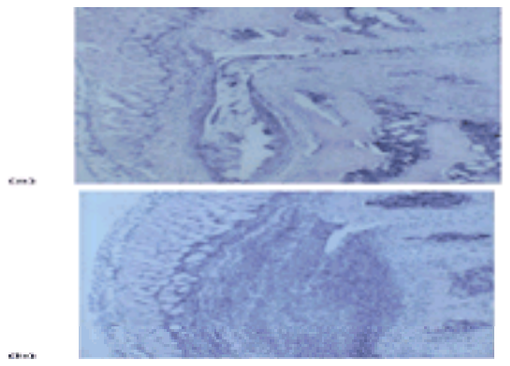


Fig. 5 Haematoxylin and eosin-stained section of a murine knee joint after triple injections of IL-1 (a) or IL-1 and transforming growth factor-b (b); note the markedly enhanced cell influx and fibrosis with the combination.

Uncoupling of cartilage destruction/joint inflammation

Apart from studies on the involvement of cytokines and on immunotherapeutic approaches, models of arthritis are valuable tools for further identifying subpopulations of infiltrating leucocytes or synovial cells involved in cartilage destruction (van Lent *et al*. 1994). The damage observed in the different models ranged from a selective loss of matrix in cartilage underlying pannus tissue to an overall loss of proteoglycans and, later on, of collagen, or even to the killing of numerous chondrocytes and the complete loss of the superficial and middle cartilage layer. This underlines that the arthritic processes can be more or less destructive, dependent on the underlying process. Large variation in progressive destruction is also noted in populations of patients with rheumatoid arthritis, which may indicate separate pathogenetic pathways. Enhanced degradation of matrix and inhibited synthesis of proteoglycans by the chondrocyte are general findings in all models.

In antigen-induced arthritis, clear uncoupling of joint swelling from cartilage destruction is a common observation. Non-steroidal anti-inflammatory drugs are effective in reducing the swelling but leave the destructive process untouched. It is even suggested that most non-steroidal anti-inflammatory drugs are harmful, by inhibition of prostaglandin production (Pettipher *et al*. 1989). Prostaglandin E_2 is a potent feedback regulator of IL-1 production and concentrations of this cytokine are higher in the presence of non-steroidal anti-inflammatory drugs with cyclo-oxygenase inhibitory activity.

The contribution of neutrophils to cartilage destruction is still unclear. Although the enzymes from neutrophils, such as elastase, can be highly destructive *in vitro*, neutrophils also contain TGF-b and IL-1 receptor antagonist and can be protective as well. Normally, neutrophils do not attach to the cartilage surface and released enzymes will be scavenged by enzyme inhibitors of the synovial fluid. Depletion of neutrophils in antigen-induced arthritis did not influence cartilage destruction and damage was similar in elastase-deficient mice (Schalkwijk *et al*. 1990). However, in the presence of dense immune complexes in the superficial cartilage layers, marked sticking of neutrophils is found in antigen- and in collagen-induced arthritis in particular. This attachment is potentiated by immobilization of the joint (Fig. 6). The ruffled cartilage surface under those conditions indicates direct destruction by the attached cells (van Lent *et al*. 1990).

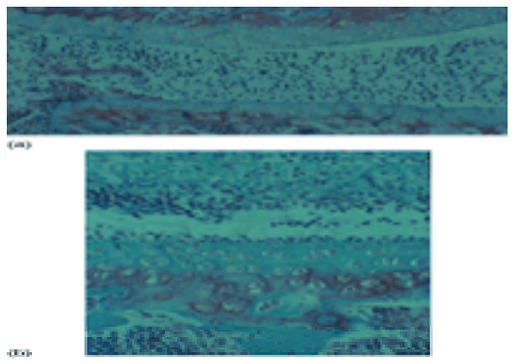


Fig. 6 Section of knee joint in antigen-induced arthritis (day 7) under normal conditions (a) or after joint immobilization (b); note the heavy sticking of granulocytes with immobilization.

Although neutrophils may be destructive under limited conditions, these cells are certainly not essential to the destruction and neither are lymphocytes. Observations in MRL lpr/lpr mice and the H_2-c-*fos* transgenic mice indicate that macrophage-rich infiltrates can be highly destructive without the presence of neutrophils and lymphocytes. Similarly, macrophage but not lymphocyte numbers in rheumatoid synovial tissue correlate with the radiological progression of joint destruction (Yanni *et al*. 1994). The critical enzymes involved in destruction in arthritis models and human rheumatoid arthritis are still far from understood. Detailed discussion of the various approaches to identify key enzymes in the models is beyond the scope of this chapter.

A general lesson that may be deduced from observations in most models is that continuing, irreversible destruction can occur under conditions that will be hardly considered as inflammatory. Symptomatic relief by anti-inflammatory therapy is promising but the main challenge remains to interrupt joint destruction. It is intriguing that combined, repeated injections of IL-1 and TGF-b give a much more profound synovitis than IL-1 alone, yet cartilage destruction is less (van Beuningen *et al*. 1994).

Final remarks

The various models have provided insight into aspects of chronic inflammation and joint destruction. Possibilities and impossibilities of targeted (immuno)therapy can be critically analysed in the models, which offers a framework for future application when the underlying process of human rheumatoid arthritis is unravelled. In recent years there has been a tendency for researchers to skip detailed analysis in complex models and to rely on models driven by T-cell clones and lines, or to jump directly to application in human rheumatoid arthritis. It is certain that the possibilities and/or limitations of, for instance, T-cell vaccination, TCR-Vb targeted therapy or immunomodulation in established disease still need evaluation in convincing animal models. Although cartilage autoantigens are attractive as putative arthritogens, final evidence for their role in human rheumatoid arthritis has yet to be provided, and future research is likely to focus on collagen type II and proteoglycan epitopes, and on novel cartilage antigens. An interesting approach is the use of non-arthritogenic antigens to provoke 'bystander' suppression to control arthritis of unknown aetiology. A still neglected area is the basis for preferential expression of the first inflammatory signs in either peripheral joints, whole paws, ankles or knees in the various models. This is of pathogenetic interest, since it is still not understood why the expression of human arthritides starts in particular joints and initially spares others.

A seemingly contradictory finding in all models is the relative abundance of periostitis in addition to synovitis, and the formation of osteophytes apart from bone and cartilage erosions. It is important that most studies are done in relatively young animals and that juvenile arthritis is also characterized by osteophytes. Moreover, periosteal lining layers and epiphyseal growth plates remain active until late in life in rodents, in contrast to the inactivity in those layers in aged humans.

Finally, the severity of the arthritis is often greater and the disease course commonly shorter in most models compared with rheumatoid arthritis. This complicates the screening of antiarthritic drugs, an issue that has not been addressed in great detail here. Reviews that selectively address drug therapy are recommended for further details (Bliven and Otterness 1985; Hunneyball *et al*. 1989). Suffice it to say that therapeutic studies on the late stages of chronic disease and on repeated flares are warranted.

A clear message from the animal models is that cartilage destruction is often dissimilar from joint inflammation. In some models, early anti-inflammatory treatment can completely prevent the onset of arthritis and under those conditions cartilage destruction is prevented too. However, in other models, cartilage destruction is still noted in the absence of marked inflammation, making it essential to define destructive pathways or mediators. It is promising that the model studies have yielded unequivocal evidence for a pivotal role of IL-1 in inhibiting proteoglycan synthesis by chondrocytes. However, the critical enzymes involved in cartilage degradation are still not fully elucidated. An intriguing approach is now offered by the analysis of enzyme-specific breakdown epitopes of proteoglycans and collagens in cartilage from human arthritides and various models, using recently developed monoclonal antibodies. Moreover, the development of arthritis in 'knock-out' mice, deficient in candidate enzymes, will provide evidence for their involvement. It is expected from the studies on current models that full protection against cartilage destruction will require a combination therapy of IL-1 blockade and neutralization of as yet unidentified, IL-1-independent enzymes.

Chapter References

Allen, J. B., Wong, H. L., Costa, G. L., Bienkowski, M. J., and Wahl, S. M. (1993). Suppression of monocyte function and differential regulation of IL-1 and IL-1ra by IL-4 contribute to resolution of experimental arthritis. *Journal of Immunology*, **151**, 4344–51.

Bakker, N. P. M., van Erck, M. G. M., Botman, C. A. D., Jonker, M., and 't Hart, B. (1991). Collagen-induced arthritis in an outbred group of rhesus monkeys comprising responder and nonresponder animals. *Arthritis and Rheumatism*, **34**, 616–24.

Bartlett, R. R., Popovic, S., and Raiss, R. X. (1988). Development of autoimmunity in MRL/lpr mice and the effects of drugs on this murine disease. *Scandinavian Journal of Rheumatology*, **75** (Suppl.), 290–9.

Billingham, M. E. J. (1983). Models of arthritis and the search for anti-arthritic drugs. *Pharmacological Therapy*, **21**, 389–428.

Billingham, M. E. J. (1994). Monoclonal antibody therapy of experimental arthritis: comparison with cyclosporin A for elucidating cellular and molecular disease mechanisms. In *Immunopharmacology of joints and connective tissue* (ed. M. E. Davies and J. T. Dingle), pp. 65–86. Academic Press, New York.

Bliven, M. and Otterness, I. (1985). Laboratory models for testing nonsteroidal antiinflammatory drugs. In *Nonsteroidal antiinflammatory drugs* (ed. J. G. Lombardino). Wiley, New York.

Case, J. P., Sano, H., Lafyatis, R., Remmers, E. F., Kumkumian, G. K., and Wilder, R. L. (1989). Transin/stromelysin expression in the synovium of rats with experimental erosive arthritis. *In situ* localization and kinetics of expression of the transformation-associated metalloproteinase in euthymic and athymic Lewis rats. *Journal of Clinical Investigation*, **84**, 1731–40.

Cooke, T. D. V., Hird, E. R., Ziff, M., and Jasin, H. E. (1972). The pathogenesis of chronic inflammation in experimental antigen induced arthritis. *Journal of Experimental Medicine*, **135**, 323–38.

Cremer, M. A., Townes, A. S., and Kang, A. H. (1990). Adjuvant-induced arthritis in rats. Evidence that autoimmunity to homologous collagens types I, II, IX and XI is not involved in the pathogenesis of arthritis. *Clinical and Experimental Immunology*, **82**, 307–12.

Cremer, M. A., Ye, X. J., Terato, K., Owens, S. W., Seyer, J. M., and Kang, A. H. (1994). Type XI collagen-induced arthritis in the Lewis rat. Characterization of cellular and humoral immune responses to native types XI, V, and II collagen and constituent a-chains. *Journal of Immunology*, **153**, 824–32.

Cromartie, W. L. *et al*. (1977). Arthritis in rats after systemic injection of streptococcal cell walls. *Journal of Experimental Medicine*, **146**, 1485–602.

de Vries, B. J. and van den Berg, W. B. (1989). Impact of NSAIDs on murine antigen induced arthritis. I. An investigation of antiinflammatory and chondroprotective effects. *Journal of Rheumatology*, **18** (Suppl. 16), 10–18.

Dumonde, D. C. and Glynn, L. E. (1962). The production of arthritis in rabbits by an immunological reaction to fibrin. *British Journal of Experimental Pathology*, **43**, 373–83.

Elkon, K. B. and Ashany, D. (1993). The SCID mouse as a vehicle to study autoimmunity. *British Journal of Rheumatology*, **32**, 4–12.

Erlandsson, H., Mussener, A., Klareskog, L., and Gold, D. P. (1994). Restricted T cell receptor usage in DA rats during early collagen induced arthritis. *European Journal of Immunology*, **24**, 1929.

Esser, R. E., Stimpson, S. A., Cromartie, W. J., and Schwab, J. H. (1985). Reactivation of streptococcal cell wall-induced arthritis by homologous cell wall polymers. *Arthritis and Rheumatism*, **28**, 1402.

Firestein, G. S. (1991). The immunopathogenesis of rheumatoid arthritis. *Current Opinion in Rheumatology*, **3**, 398–406.

Glant, T. T., Mikecz, K., Arzoumanian, A., and Poole, A. R. (1987). Proteoglycan-induced arthritis in BALB/c mice. Clinical features and histopathology. *Arthritis and Rheumatism*, **30**, 201–12.

Greenwald, R. A. and Diamond, H. S. (1988). *Handbook of animal models for the rheumatic diseases*, Vol. I. CRC Press, Boca Raton FA.

Hang, L., Theofilopoulos, A. N., and Dixon, F. J. (1982). A spontaneous rheumatoid arthritis-like disease in MRL/1 mice. *Journal of Experimental Medicine*, **155**, 1690–701.

Hazenberg, M. P., Klasen, I. S., Kool, J., Ruseler-van Embden, J. G. H., and Severijnen, A. J. (1992). Are intestinal bacteria involved in the etiology of rheumatoid arthritis? *APMIS*, **100**, 1–9.

Henderson, B. and Blake, S. (1992). Therapeutic potential of cytokine manipulation. *Trends in Physical Sciences*, **13**, 145–12.

Holmdahl, R. and Kvick, C. (1992). Vaccination and genetic experiments demonstrate that adjuvant-oil-induced arthritis and homologous type II collagen-induced arthritis in the same rat strain are different diseases. *Clinical and Experimental Immunology*, **88**, 96–100.

Holmdahl, R., Malmstrom, V., and Vuorio, E. (1993). Autoimmune recognition of cartilage collagens. *Annals of Medicine*, **25**, 251–64.

Holoshitz, J., Naparstek, Y., Ben-Num, A., and Cohen, I. R. (1983). Lines of T lymphocytes induce or vaccinate against autoimmune arthritis. *Science*, **219**, 56–8.

Hom, J. T., Cole, H., Estridge, T., and Gliszczynski, V. L. (1992). Interleukin-1 enhances the development of type II collagen-induced arthritis only in susceptible and not in resistant mice. *Clinical Immunology and Immunopathology*, **62**, 56–65.

Hunneyball, I. M., Billingham, M. E. J., and Rainsford, K. D. (1989). Animal models of arthritic disease: influence of novel compared with classical antirheumatic agents. In *New developments in antirheumatic therapy* (ed. K. D. Rainsford), pp. 93–132. Kluwer, Dordrecht.

Issekutz, A. C., Meager, A., Otterness, I., and Issekutz, T. B. (1994). The role of tumour necrosis factor-alpha and IL-1 in polymorphonuclear leukocyte and T lymphocyte recruitment to joint inflammation in adjuvant arthritis. *Clinical and Experimental Immunology*, **97**, 26–32.

Iwakura, Y. *et al*. (1991). Induction of inflammatory arthropathy resembling rheumatoid arthritis in mice transgenic for HTLV-I. *Science*, **253**, 1026–8.

Joosten, L. A. B., Helsen, M. M. A., and van den Berg, W. B. (1994). Accelerated onset of collagen-induced arthritis by remote inflammation. *Clinical and Experimental Immunology*, **97**, 204–11.

Keffer, J. *et al*. (1991). Transgenic mice expressing human tumor necrosis factor: a predictive genetic model of arthritis. *EMBO Journal*, **13**, 4025–31.

Kleinau, S., Erlandsson, H., Holmdahl, R., and Klareskog, L. (1991). Adjuvant oils induce arthritis in the DA rat. I. Characterization of the disease and evidence for an immunological involvement. *Journal of Autoimmunity*, **4**, 871–80.

Kohashi, O., Kohashi, Y., Takahashi, T., Ozawa, A., and Shigematsu, N. (1986). Suppressive effect of *Escherichia coli* on adjuvant-induced arthritis in germ-free rats. *Arthritis and Rheumatism*, **29**, 547–53.

Koopman, W. J., and Gay, S. (1988). The MRL-lpr/lpr mouse. A model for the study of rheumatoid arthritis. *Scandinavian Journal of Rheumatology*, **75** (Suppl.), 284–9.

Lafyatis, R. *et al*. (1989). Transforming growth factor-b production by synovial tissues from rheumatoid patients and streptococcal cell wall arthritic rats. *Journal of Immunology*, **143**, 1142–8.

Lens, J. W., van den Berg, W. B., and van de Putte, L. B. A. (1984). Flare-up of antigen-induced arthritis in mice after challenge with intravenous antigen. Studies on the characteristics of and mechanisms involved in the reaction. *Clinical and Experimental Immunology*, **55**, 287–94.

Lewthwaite, J., Blake, S. M., Hardingham, T. E., Warden, P. J., and Henderson, B. (1994). The effect of recombinant human IL-1 receptor antagonist on the induction phase of antigen induced arthritis in the rabbit. *Journal of Rheumatology*, **21**, 467–72.

Maeda, T. *et al*. (1994). Exacerbation of established collagen-induced arthritis in mice treated with an anti-T cell receptor antibody. *Arthritis and Rheumatism*, **37**, 406–13.

Mertz, A. K. H., Batsford, S. R., Curschellas, E., Kist, M. J., and Gondolf, K. B. (1991). Cationic Yersinia antigen-induced chronic allergic arthritis in rats. A model for reactive arthritis in humans. *Journal of Clinical Investigation*, **87**, 632–42.

Mikecz, K., Glant, T. T., Buzas, E., and Poole, A. R. (1990). Proteoglycan induced polyarthritis and spondylitis adoptively transferred to naive BALB/c mice. *Arthritis and Rheumatism*, **33**, 866–76.

Moder, K. G., Luthra, H. S., Griffiths, M., and David, C. S. (1993). Prevention of collagen induced arthritis in mice by deletion of T cell receptor Vb8 bearing T cells with monoclonal antibodies. *British Journal of Rheumatology*, **32**, 26–30.

Mountz, J. D., Zhou, T., Long, R. E., Bluethmann, H. M., Koopman, W. J., and Edwards, C. K. (1994). T cell influence on superantigen-induced arthritis in MRL-lpr/lpr mice. *Arthritis and Rheumatism*, **37**, 113–24.

Myers, L. K., Rosloniec, E. F., Seyer, J. M., Stuart, J. M., and Kang, A. H. (1993 *a*). A synthetic peptide analogue of a determinant of type II collagen prevents the onset of collagen-induced arthritis. *Journal of Immunology*, **150**, 4652–8.

Myers, L. K., Seyer, J. M., Stuart, J. M., Terato, K., David, C. S., and Kang, A. H. (1993 *b*). T cell epitopes of type II collagen that regulate murine collagen-induced arthritis. *Journal of Immunology*, **151**, 500–5.

Nordling, C., Karlsson-Parra, A., Jansson, L., Holmdahl, R., and Klareskog, L. (1992). Characterization of a spontaneously occurring arthritis in male DBA/1 mice. *Arthritis and Rheumatism*, **35**, 717–22.

Pearson, C. M. (1956). Development of arthritis, periarthritis and periostitis in rats given adjuvants. *Proceedings of the Society for Experimental Biology (New York)*, **91**, 95–101.

Pettipher, E. R., Henderson, B., Edwards, J. C. W., and Higgs, G. A. (1989). Effect of indomethacin on swelling, lymphocyte influx, and cartilage proteoglycan depletion in experimental arthritis. *Annals of the Rheumatic Diseases*, **48**, 623–7.

Probst, P., Hermann, E., Meyer zum Buschenfelde, K. H., and Fleischer, B. (1993). Identification of the *Yersinia* subunits as a target antigen for human synovial T lymphocytes in reactive arthritis. *Infection and Immunity*, **61**, 4507–9.

Ratkay, L. G., Zhang, L., Tonzetich, J., and Waterfield, J. D. (1993). Complete Freund's adjuvant induces an earlier and more severe arthritis in MRL-lpr mice. *Journal of Immunology*, **151**, 5081–7.

Santambrogio, G. M., Hochwald, G. M., Leu, C. H., and Thorbecke, G. J. (1993). Antagonistic effects of endogenous and exogenous TGFb and TNF on auto-immune diseases in mice. *Immunopharmacology and Immunotoxicology*, **15**, 461–78.

Schalkwijk, J., Joosten, L. A. B., van den Berg, W. B., and van de Putte, L. B. A. (1990). Antigen-induced arthritis in beige (Chediak–Higashi) mice. *Annals of the Rheumatic Diseases*, **49**, 607–10.

Schorlemmer, H. U., Kanzy, E. J., Langner, K. D., and Kurrle R. (1993). Immunomodulatory activity of recombinant IL-1 receptor on models of experimental rheumatoid arthritis. *Agents and Actions*, **39S**, C113–16.

Schwab, J. H., Anderle, S. K., Brown, R. R., Dalldorf, F. G., and Thompson R. C. (1991). Pro- and anti-inflammatory roles of interleukin-1 in recurrence of bacterial cell wall-induced arthritis in rats. *Infection and Immunity*, **59**, 4436–42.

Schwab, J. H., Brown, R. R., Anderle, S. K., and Schlievert, P. M. (1993). Superantigen can reactivate bacterial cell wall induced arthritis. *Journal of Immunology*, **150**, 4151–9.

Shiozawa, S., Tanka, Y., Fujita, T., and Tokuhisa, T. (1992). Destructive arthritis without lymphocyte infiltration in H_2-c-*fos* transgenic mice. *Journal of Immunology*, **148**, 3100–4.

Sternberg, E. M. *et al*. (1989). Inflammatory mediator-induced hypothalamic-pituitary-adrenal axis activation is defective in streptococcal cell wall arthritis-susceptible Lewis rats. *Proceedings of the*

National Academy of Sciences (USA), **86**, 2374–8.

Stimpson, S. A. *et al*. (1986). Arthropathic properties of peptidoglycan-polysaccharide polymers from normal flora bacteria. *Infection and Immunity*, **51**, 240.

Stimpson, S. A., Esser, R. E., Carter, P. B., Sartor, R. B., Cromartie, W. J., and Schwab, J. H. (1987). Lipopolysaccharide induces recurrence of arthritis in rat joints previously injured by peptidoglycan-polysaccharide. *Journal of Experimental Medicine*, **165**, 1688–702.

Thompson, S. J. and Elson, C. J. (1993). Susceptibility to pristane-induced arthritis is altered with changes in bowel flora. *Immunology Letters*, **36**, 227–32.

Thompson, H. S. G., Harper, N., Bevan, D. J., and Staines, N. A. (1993). Suppression of collagen induced arthritis by oral administration of type II collagen: changes in immune and arthritic responses mediated by active peripheral suppression. *Autoimmunity*, **16**, 189–99.

Thorbecke, G. J., Shah, R., Leu, C. H., Kuruvilla, A. P., Hardison, A. M., and Palladino, M. A. (1992). Involvement of endogenous tumor necrosis factor a and transforming growth factor b during induction of collagen type II arthritis in mice. *Proceedings of the National Academy of Sciences (USA)*, **89**, 7375–9.

Trentham, D. E., Townes, A. S., and Kang, A. H. (1977). Autoimmunity to type II collagen: an experimental model of arthritis. *Journal of Experimental Medicine*, **146**, 857–68.

Trentham, D. E., Dynesius-Trentham, R. A., Hafler, D. A., and Weiner, H. L. (1993). Effects of oral administration of type II collagen on rheumatoid arthritis. *Science*, **261**, 1727–30.

van Beuningen, H. M., van der Kraan, P. M., Arntz, O. J., and van den Berg, W. B. (1994). Transforming growth factor-b $_1$ stimulates articular chondrocyte proteoglycan synthesis and induces osteophyte formation in the murine knee joint. *Laboratory Investigation*, **71**, 279–90.

van de Langerijt, A. G. M., Kingston, A. E., van Lent, P. L. E. M., Billingham, M. E. J., and van den Berg, W. B. (1994 *a*). Cross-reactivity to proteoglycans in bacterial arthritis: lack of evidence for *in vivo* role in induction of disease. *Clinical Immunology and Immunopathology*, **71**, 273–80.

van de Langerijt, A. G. M., van Lent, P. L. E. M., Hermus, A. R. M. M., Sweep, C. G. J., Cools, A. R., and van den Berg, W. B. (1994 *b*). Susceptibility to adjuvant arthritis: relative importance of adrenal activity and bacterial flora. *Clinical and Experimental Immunology*, **97**, 33–8.

van de Loo, A. A. J., Arntz, O. J., Otterness, I. G., and van den Berg, W. B. (1992 *a*). Protection against cartilage proteoglycan synthesis inhibition by anti-interleukin 1 antibodies in experimental arthritis. *Journal of Rheumatology*, **19**, 348–56.

van de Loo, A. A. J., Arntz, O. J., and van den Berg, W. B. (1992 *b*). Flare-up of experimental arthritis in mice with murine recombinant IL-1. *Clinical and Experimental Immunology*, **87**, 196–202.

van de Loo, A. A. J., Arntz, O. J., Bakker, A. C., van Lent, P. L. E. M., Jacobs, M. J. M., and van den Berg, W. B. (1995). Role of interleukin-1 in antigen-induced exacerbations of murine arthritis. *American Journal of Pathology*, **146**, 239–49.

van den Berg, W. B. (1994). Articular cartilage; cause or victim in arthritis. *Scandinavian Journal of Rheumatology*, **101** (Suppl. 23), 25–31.

van den Berg, W. B., Lens, J. W., van de Putte, L. B. A., and van Beusekom, H. J. (1982). Antigen-induced arthritis: antigen handling and chronicity of joint inflammation. *Agents and Actions*, **11** (Suppl.), 233–41.

van den Berg, W. B., van de Putte, L. B. A., Zwarts, W. A., and Joosten, L. A. B. (1984). Electrical charge of the antigen determines intraarticular antigen handling and chronicity of arthritis in mice. *Journal of Clinical Investigation*, **74**, 1850–9.

van den Berg, W. B., van den Broek, M. F., van de Putte, L. B. A., van Bruggen, M. C. J., and van Lent, P. L. E. M. (1991) Experimental arthritis: importance of T cells and antigen mimicry in chronicity and treatment. In *Monoclonal antibodies, cytokines, and arthritis* (ed. T. F. Kresina), pp. 237–52. Dekker, New York.

van den Berg, W. B., Joosten, L. A. B., Helsen, M., and van de Loo, F. A. J. (1994). Amelioration of established murine collagen-induced arthritis with anti-IL-1 treatment. *Clinical and Experimental Immunology*, **95**, 237–43.

van den Broek, M. F., van den Berg, W. B., van de Putte, L. B. A., and Severijnen, A. J. (1988). Streptococcal cell wall induced arthritis and flare-up reactions in mice induced by homologous and heterologous cell walls. *American Journal of Pathology*, **133**, 139–49.

van den Broek, M. F., van Bruggen, M. C. J., Stimpson, S. A., Severijnen, A. J., van de Putte, L. B. A., and van den Berg, W. B. (1990). Flare-up reaction of streptococcal cell wall-induced arthritis in Lewis and F344 rats: the role of T lymphocytes. *Clinical and Experimental Immunology*, **79**, 297–306.

van Eden, W. (1991). Heatshock proteins as immunogenic bacterial antigens with the potential to induce and regulate autoimmune arthritis. *Immunology Reviews*, **121**, 5.

van Eden, W., Hogervorst, E. J. M., Henssen, E. J., van der Zee, R., van Embden, J. D. A., and Cohen, I. R. (1989). A cartilage-mimicking T-cell epitope on a 65k mycobacterial heat-shock protein: adjuvant arthritis as a model for human rheumatoid arthritis. *Current Topics in Microbiology and Immunology*, **145**, 27–43.

van Lent, P. L. E. M., van den Bersselaar, L., van de Putte, L. B. A., and van den Berg, W. B. (1990). Immobilization aggravates cartilage damage during antigen-induced arthritis in mice. Attachment of polymorphonuclear leucocytes to articular cartilage. *American Journal of Pathology*, **136**, 1407–16.

van Lent, P. L. E. M., Holthuysen, A. E. M., van den Bersselaar, L., van Rooijen, N., van de Putte, L. B. A., and van den Berg, W. B. (1994). Role of macrophage-like synovial lining cells in localization and expression in experimental arthritis. *Scandinavian Journal of Rheumatology*, **23** (Suppl. 101), 75–81.

Wahl, S. M. (1994). TGFb: the good, the bad, and the ugly. *Journal of Experimental Medicine*, **180**, 1587.

Wahl, S. M., Allen, J. B., Costa, G. L., Wong, H. L., and Dasch, J. R. (1993). Reversal of acute and chronic synovial inflammation by anti-transforming growth factor b. *Journal of Experimental Medicine*, **177**, 225–30.

Williams, R. O., Feldmann, M., and Maini, R. N. (1992 *a*). Anti-tumor necrosis factor ameliorates joint disease in murine collagen-induced arthritis. *Proceedings of the National Academy of Sciences (USA)*, **89**, 9784–8.

Williams, R. O., Plater-Zyberk, C., Williams, D. G., and Maini, R. N. (1992 *b*). Successful transfer of collagen-induced arthritis to severe combined immunodeficient (SCID) mice. *Clinical and Experimental Immunology*, **88**, 455–60.

Wooley, P. H. (1991). Animal models of rheumatoid arthritis. *Current Opinion in Rheumatology*, **3**, 407–20.

Wooley, P. H., Luthra, H. S., Stuart, J. M., and David, C. S. (1981). Type II collagen-induced arthritis in mice. I. MHC (I-region) linkage and antibody correlates. *Journal of Experimental Medicine*, **154**, 688–700.

Wooley, P. H., Seibold, J. R., Whalen, J. D., and Chapdelaine, J. M. (1989). Pristane induced arthritis. The immunologic and genetic features of an experimental murine model of autoimmune disease. *Arthritis and Rheumatism*, **32**, 1022–30.

Wooley, P. H., Dutcher, J., Widmer, M. B., and Gillis, S. (1993 *a*). Influence of a recombinant human soluble tumor necrosis factor receptor FC fusion protein on type II collagen-induced arthritis in mice. *Journal of Immunology*, **151**, 6602–7.

Wooley, P. H. *et al*. (1993*b*). The effect of an interleukin-1 receptor antagonist protein on type II collagen-induced arthritis and antigen-induced arthritis in mice. *Arthritis and Rheumatism*, **36**, 1305–14.

Yamamoto, H., Sekiguchi, T., Itagaki, K., Saijo, S., and Iwakura, Y. (1993). Inflammatory polyarthritis in mice transgenic for human T cell leukemia virus type I. *Arthritis and Rheumatism*, **36**, 1612–20.

Yanni, G., Whelan, A., Feighery, C., and Bresnihan, B. (1994). Synovial tissue macrophages and joint erosion in rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **53**, 39–44.

Yoshino, S., Cleland, L. G., Mayrhofer, G., Brown, R. R., and Schwab, J. H. (1991). Prevention of chronic erosive streptococcal cell wall-induced arthritis in rats by treatment with a monoclonal antibody against the T cell antigen receptor ab^1. *Journal of Immunology*, **146**, 4187–9.

3.5.1 Non-steroidal anti-inflammatory drugs

Peter Brooks

Introduction

Inflammation in the rheumatic diseases may be acute or chronic and involves a whole variety of cells and mediators interacting together to form a final common pathway to joint destruction (see previous chapters in this section). The cells involved in the inflammatory response include polymorphonuclear leucocytes, macrophages, and lymphocytes as well as the fibroblasts, platelets, and endothelial cells lining the blood vessels. Mediators of inflammation include the prostaglandins and leukotrienes, cytokines, platelet-activating factors, and kinins. In diseases such as osteoarthritis the inflammation may be mild but in the inflammatory arthropathies a number of different inflammatory pathways are activated, leading eventually to joint destruction. These are summarized in Fig. 1.

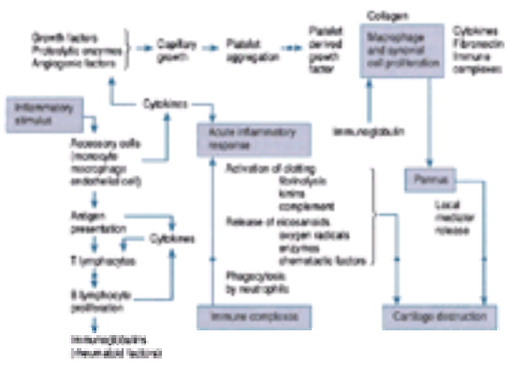


Fig. 1 Sites at which non-steroidal anti-inflammatory drugs may act on the inflammatory process. The inflammatory process comprises many detector, communication, and effector systems. These systems may be chemical or cellular, and there are clearly multiple interactions between them. · Possible action of NSAIDs. (Modified from Forrest and Brooks (1988) and reproduced with permission.)

The pharmacological agents used in the rheumatic diseases can interfere with various aspects of the inflammatory response. The mechanism of action, major adverse reactions, interactions, and the practicable prescribing of these agents will be discussed in the following chapters. The endpoint of inflammation in the rheumatic diseases is the destruction of cartilage, with its subsequent pain and loss of function. It would seem reasonable, therefore, to attempt to control inflammation as early as possible in an effort to reduce joint and cartilage damage to a minimum. A number of published studies support the view of Brook and Corbett (1977) that, if erosions are going to occur in rheumatoid arthritis, they do so within 2 years of disease onset. If we accept this premise, then it would seem important to use drugs which have antirheumatic activity at as early a stage as possible. In the past the traditional approach to the management of inflammatory arthritis has been to begin with the simple agents, such as analgesics and non-steroidal anti-inflammatory drugs (**NSAIDs**), and progress slowly towards the more potent agents. Even with this approach, the long-term outlook for rheumatoid arthritis is not good. Recently Scott *et al*. (1987) have shown that after 20 years of disease, over 50 per cent of patients are either severely disabled or dead. Significant functional disability also occurs in these patients and approximately half are work disabled 10 years after the onset of rheumatoid arthritis (Yelin *et al*. 1987). Wilske and Healey (1989) have suggested recently that the conservative 'pyramidal' approach to the management of rheumatoid arthritis, where NSAIDs are prescribed initially alone, be replaced by the more intensive and earlier use of corticosteroids, methotrexate, and other suppressive agents. The therapeutic goal in the management of inflammatory joint disease should be not just to suppress symptoms and signs of inflammation but also to try to normalize acute-phase reactants and attempt to slow and prevent bony erosion. If we are to do this, then the data support early treatment of rheumatoid arthritis with disease-modifying drugs as well as use of NSAIDs.

NSAIDs in the management of joint inflammation

NSAIDs make up one of the most commonly prescribed groups of drugs, with over 100 million prescriptions written worldwide in 1989. Over 20 per cent of patients admitted to hospitals in the United Kingdom and Australia are taking NSAIDs. There are over 50 different NSAIDs either marketed or on trial throughout the world, though most countries have between 7 and 17 available for use. Just how many NSAIDs are necessary for optimal management of patients with rheumatic disorders is not known. However, a study carried out in a number of private rheumatological practices in the United States found that 14 NSAIDs were required to control symptoms in 90 per cent of patients (Pincus and Callaghan 1989). Patients do seem to respond individually to NSAIDs, both in terms of pain relief and adverse reactions, but the extent of this variability and the reasons for it are still not clear (Brooks and Day 1991).

Mechanism of action

The major activities of NSAIDs are summarized as follows and include modulation of:

1. cyclo-oxygenase;
2. leukotriene synthesis;
3. superoxide production;
4. superoxide scavenging;
5. lysosomal enzyme release;
6. neutrophil aggregation and adhesion;
7. lymphocyte function;
8. rheumatoid factor production;
9. cartilage metabolism;
10. cell membrane activities, including:

a. enzymes (NADPH oxidase, phospholipase C),
b. transmembrane anion transport,
c. uptake of prostanoid precursor.

Prostaglandins, leukotrienes, and the cellular elements of acute and chronic inflammation might all be influenced by NSAIDs, as reviewed by Forrest and Brooks (1988). As our understanding of the way in which prostaglandins and other mediators influence the immune response increases, it is obvious that the NSAIDs do have some mild immunomodulatory activity. The differentiation between the varying classes of antirheumatic drugs into NSAIDs, slow-acting rheumatic drugs, and immunosuppressive drugs is thus becoming blurred.

The variability in patients' responses to NSAIDs is well recognized by rheumatologists (Brooks and Day 1991). The factors which might influence this variability are:

1. pharmacokinetics:
   a. absorption,
   b. protein binding,
   c. metabolism,
   d. renal clearance;
2. pharmacodynamics;
3. dosage selection and timing;
4. variability in disease state;
5. variation in adverse reaction profile.

From a practical point of view it is important that the prescriber of NSAIDs appreciates some of these variables and tries to find the most appropriate NSAID to use in a particular clinical situation.

Pharmacokinetics

The pharmacokinetics of a non-steroidal anti-inflammatory drug depend on its absorption and bioavailability, its distribution, and its half-life. Most NSAIDs are relatively lipid soluble, are weak acids, and almost all are completely absorbed from the gastrointestinal tract. The rate of absorption of NSAIDs is variable and is usually slower when they are taken with meals. Most NSAIDs are now provided as enteric-coated or sustained-release preparations and this can significantly alter the pharmacokinetics. NSAIDs are usually highly bound to plasma proteins and thus the amount of free drug (the active component) is relatively small. NSAID binding to plasma proteins might be reduced in various states such as renal and hepatic disease and in subjects with hypoalbuminanaemia. For most of the NSAIDs the unbound fraction is not altered by the plasma concentration. However, with naproxen, phenylbutazone, and salicylate, the unbound concentrations increase proportionally with dose. The binding of some other NSAIDs decreases in elderly patients and in patients with rheumatoid arthritis, emphasizing the importance of correlating unbound concentrations, rather than the total amount of drug in plasma, with clinical effect. NSAIDs are weak acids and, therefore, they tend to be trapped more effectively by cells in acid environments such as the inflamed joint, the stomach, and the renal medulla. This is of relevance since these are the sites of action and major adverse effects. Lipid solubility of NSAIDs allows them to pass easily through cell membranes, and cognitive changes, headache, and confusion are well reported as adverse effects (O'Brien and Bagby 1985). The half-life of elimination of a drug ($t_{1/2}$) is the time taken for the concentration of that drug to fall from any given concentration to 50 per cent of that concentration. NSAIDs have a wide range of plasma half-lives and can be divided into two broad categories on the basis of this variable—those with plasma half-lives of less than 6 h and those with half-lives in excess of 12 h (Day *et al*. 1988) (Table 1). NSAIDs with long half-lives show much smaller fluctuations between peak and trough concentrations, even on a once-daily dose regimen. Many of the shorter half-life NSAIDs have now been produced in slow or sustained release formulations, which tend to reduce the fluctuations between peak and trough levels and allow for once-daily dosing. As the major site of action of NSAIDs is the synovial membrane and synovial fluid, it is probably more important to consider synovial-fluid or tissue kinetics. These are obviously difficult studies to do but those that have been done show clearly that the concentrations in synovial fluid are much more stable and show fewer fluctuations than plasma concentrations, even for drugs with short half-lives. This is demonstrated in Fig. 2 where it can be seen that the peak concentrations in synovial fluid appear later and are lower than the peak concentrations in the plasma. This is one of the explanations for the fact that even NSAIDs with short half-lives often provide satisfactory pain relief on a twice-daily dosing regimen.

Short (<6 h)	$t_{1/2}$	Long	$t_{1/2}$
Aspirin	0.25	Azapropazone	15
Diclofenac	1.1	Diflunisal	13
Etodolac	3.0	Fenbufen	11
Fenoprofen	2.5	Nabumetone	26
Flufenamic acid	1.4	Naproxen	14
Flurbiprofen	3.8	Oxaprozin	58
Ibuprofen	2.1	Phenylbutazone	68
Indomethacin	4.6	Piroxicam	57
Ketoprofen	1.8	Salicylate	2–15
Pirprofen	3.8	Sulindac	14
Tiaprofenic acid	3	Tenoxicam	60
Tolmetin	1		

Table 1 Half-lives of NSAIDs

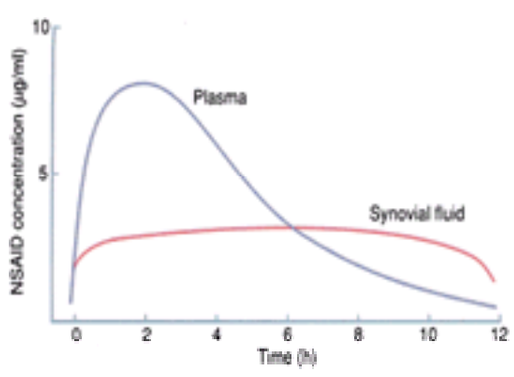


Fig. 2 Concentration in plasma and synovial fluid of a short half-life NSAID (ibuprofen) following chronic administration. The synovial fluid concentrations oscillate much less than the plasma concentrations (adapted from Day *et al*. 1988 and reproduced with permission).

The mechanism of clearance of NSAIDs is usually by hepatic metabolism with the production of inactive metabolites. Some NSAIDs, such as aspirin, nabumetone, fenbufen, and sulindac, are converted into active metabolites that produce the major anti-inflammatory pharmacological effect. Sulindac is particularly interesting because the active metabolite (sulphide) is produced in the large intestine and this process is reversible. Conversion of the active sulphide metabolite back to the inactive sulindac can occur in the kidney and has been suggested as an explanation for the reported renal-sparing effects of this drug. Some NSAIDs of the 2-arylpropionic acid class, including ibuprofen, ketoprofen, fenoprofen, naproxen, tiaprofenic acid, and flurbiprofen, exist as two optical isomers or enantiomers. Interestingly, it is only the *S*-enantiomer of these NSAIDs that inhibit cyclo-oxygenase and thus produces anti-inflammatory effects. All these compounds, except for naproxen, are produced as mixtures of the *R* and *S* forms with the inactive *R* being metabolically converted to the active *S*-enantiomer *in vivo*. Williams *et al*. (1986) have shown that the metabolism of the inactive isomer is also associated with uptake into triglycerides, which might remain in fat depots in the body for considerable periods of time.

Although only small amounts of NSAID are excreted unchanged from the urine, enal clearance does increase with increasing urinary pH, but this is only of clinical relevance for salicylate. The renal clearance of diflunisal, ketoprofen, fenoprofen, naproxen, and indomethacin is reduced in renal failure and by administration of probenecid. This is primarily due to the retention of acylglucuronide metabolites, which are then converted back to the original NSAID *in vivo*.

Side-effects of NSAIDs occur more frequently in the elderly and clearance of naproxen, ketoprofen, azapropazone, salicylate, and ibuprofen are reduced in elderly patients. In these patients it would seem sensible to commence treatment with lower than standard doses and only increase the dose in the face of an inadequate response and absence of side-effects.

Relationships between the plasma concentration of NSAID and clinical effects have been difficult to determine because of methodological problems. Linear relationships between dose and/or plasma concentration and anti-inflammatory effect have now been demonstrated for naproxen, carprofen, and ibuprofen in patients with rheumatoid arthritis. Correlations between plasma concentration and tinnitus, deafness, and gastrointestinal symptoms have also been demonstrated for some NSAIDs. Although the concept of responders and non-responders to an NSAID is generally accepted, many of the studies have flawed methods and the only study looking at reproducibility of the phenomenon has failed to demonstrate it (Preston *et al*. 1988). If responders and non-responders do exist, then it would seem an attractive hypothesis that the difference is pharmacokinetic in origin. However, it has been clearly shown that the pharmacokinetics of indomethacin and flurbiprofen do not differ between responders and non-responders and this does not seem to be the explanation for the phenomenon. Variability in pharmacokinetic parameters might play a role in individual patients; other factors that need to be considered in variability of response are listed above under mechanism of action. To these need be added genetic factors, age, gender, and the influence of a whole variety of environmental issues such as diet, smoking, and activity. Pharmacodynamic factors might also explain variability in response in that it has recently been shown that the cyclo-oxygenase enzyme system is made up of at least two components—so-called Cox-I and Cox-II. Cox-II is activated in inflammation and NSAIDs seem to have variable activities on the Cox I and II isoenzymes, thus adding to the complexity of variability of response to NSAIDs (Meade *et al*. 1993).

Adverse reactions

Adverse reactions to NSAIDs are common and have been extensively reviewed recently by O'Brien and Bagby (1985) and Simon (1995); the major side-effects are shown in Table 2. Adverse reactions to NSAIDs are the most commonly reported of any group of drugs to the Committee on Safety of Medicines in the United Kingdom. Gastrointestinal adverse events are the most frequent and usually mild (approximately 10 per cent of patients on NSAIDs) followed by renal events, skin rashes, and central nervous reactions. Although serious adverse events to NSAIDs might be extremely rare, they assume importance when up to 20 per cent of the population are ingesting these drugs on a regular basis. The real issue with serious adverse events to NSAIDs is just how frequently they occur.

Organ system	Reaction	Incidence
Gastrointestinal	Indigestion Erosions Peptic ulcer Small and large bowel ulceration	Common (>20%)
Hepatic	Hepatocellular Cholestasis	Rare (<5%)
Renal	Transient rise in serum creatinine Acute renal failure Interstitial nephritis Hyperkalaemia	
Haematological	Thrombocytopenia Neutropenia Aplasia Haemolytic anaemia	Rare
Cutaneous	Photosensitivity Erythema multiforme Toxic epidermal necrolysis	Uncommon (5–30%)
Respiratory	Bronchospasm	Rare
Central nervous system	Headache Dizziness Personality change Aseptic meningitis	Uncommon

Table 2 Adverse reactions to NSAIDs

Gastrointestinal effects (also see Chapter 1.3.6)

There has been a rapid rise in the incidence of death from gastrointestinal perforations and haemorrhage in association with increased ingestion of these drugs. The problems of risk assessment have been highlighted recently by Henry (1988) and Lichtenstein *et al*. (1995). Linkage of population databases has produced figures for relative risk of admission to hospital with upper gastrointestinal haemorrhage of approximately 1.5, with a much higher relative risk for fatal upper gastrointestinal haemorrhage in women over the age of 75 years. Controlled studies on the relationship between NSAID use and peptic ulcer complications (again summarized by Henry 1988), show odds ratios of between 1.8 and 3.8 for upper gastrointestinal haemorrhage and much higher odds ratios for ulcer complications. Until recently it has been assumed that all NSAIDs were similar in their ability to produce gastric ulceration. Recent studies, however, suggest that some NSAIDs (ibuprofen, naproxen, and diclofenac) are less frequently associated with gastric ulceration than others (Garcia-Rodriquez and Jick 1994 ; Langman *et al*. 1994).

Up to 50 per cent of patients with rheumatoid arthritis who have gastric ulcers might be asymptomatic and it is important to identify risk factors for development of peptic ulceration. The practical issues of management of NSAID-associated gastropathy are now becoming important with the advent of so-called 'gastroprotective' agents such as the prostaglandin analogues. Both peptic and duodenal ulcers will heal despite continuing use of NSAIDs, though they will heal more slowly. H_2-blockers will reduce the incidence of NSAID-associated duodenal ulcers, omeprazole will reduce the incidence of gastric ulcers, and the prostaglandin analogues will reduce the incidence of both gastric and duodenal ulcers. To date, no studies have demonstrated that use of these agents prophylactically in patients taking NSAIDs will reduce complications of peptic ulceration. This is an important issue to determine if these drugs are to be recommended widely on a prophylactic basis. Current recommendations are that prophylaxis be confined to those patients who are at particular risk of developing serious gastrointestinal side-effects, i.e. those over the age of 65 years, with a past history of peptic ulceration, on corticosteroids, or with severe rheumatoid or other system disease; the recommendations are based on risk factors demonstrated by Fries *et al*. (1989). It has been reported recently that NSAIDs also damage the small and large bowel, and oesophageal mucosal damage is also relatively frequent.

Hepatoxicity

Regular testing of liver function often shows transient elevation of liver enzymes in patients taking NSAIDs. This is reported more frequently with certain NSAIDs, such as sulindac or diclofenac. These abnormalities are usually mild and often disappear despite continuing drug therapy.

Renal effects

Many patients will have a transient rise in serum creatinine which will return to normal despite continuation of the NSAID. Prostaglandins are important modulators of renal flow, particularly in those patients with compromised renal function. It is in this group of patients that the severe problems of NSAID-associated renal dysfunction occur. Renal syndromes associated with NSAIDs include acute renal failure, interstitial nephritis, hyperkalaemia, and renal tubular damage. Patients particularly at risk from these problems include those with a past history of renal disease, those with volume depletion due to diuretic therapy or other factors such as surgery, and those with hepatocellular insufficiency. NSAIDs might also interfere with the control of hypertension in patients receiving antihypertensive agents as reviewed recently by Radack and Deck (1987). It is clear that NSAIDs interfere with blood pressure control on an individual basis, and patients with hypertension who are started on NSAIDs should have their blood pressure monitored carefully.

Haematological reactions

Iron-deficiency anaemia due to chronic blood loss is a common accompaniment of NSAID therapy, with unusual haematological adverse reactions including thrombocytopenia, agranulocytosis, aplastic anaemia, and haemotolyic anaemia. NSAIDs also affect platelet function, however all except aspirin do so in a reversible reaction, which means that platelet function is only disturbed while there are significant concentrations of NSAID in the bloodstream. This is important when calculating how long NSAIDs need to be withdrawn before interventions such as joint replacement or other surgery.

Skin

Cutaneous reactions to NSAIDs are relatively common. They are usually mild and have been reported most frequently with naproxen, sulindac, diclofenac, diflunisal, and piroxicam. The rare complications of Stevens–Johnson syndrome, which has a mortality rate of approximately 25 per cent, have also been reported with NSAIDs.

Respiratory

Bronchospasm might be precipitated in asthmatic patients after administration of NSAIDs. If asthmatics require NSAIDs, they should be monitored carefully under controlled conditions, particularly if they have a history of previous adverse reactions to such drugs.

Central nervous system

Central nervous side-effects, such as headache, are commonly described with NSAIDs such as indomethacin. Mild reactions have also been reported with a number of other NSAIDs. Ibuprofen has also been reported as producing an aseptic meningitis-like picture in patients with systemic lupus erythematosus.

Cartilage

A number of *in vitro* and animal studies have suggested that NSAIDs might interfere with cartilage integrity, leading to more rapid development of osteoarthritis. Evidence that this is an important issue in patients has not yet been obtained, as measurements of progression in osteoarthritis are still relatively crude. Whether NSAIDs do influence progression of arthritis, and in particular cartilage destruction, needs to be investigated.

Interactions

As NSAIDs are commonly taken by elderly people with other diseases, the potential for drug interaction is large. The common interactions of NSAIDs are shown in Table 3. Interactions can be seen as pharmacokinetic or pharmacodynamic, as described by Tonkin and Wing (1988). From a practical point of view, important interactions are those which might occur with oral anticoagulants, diuretics, and hypotensive agents. Although NSAIDs reduce the clearance of methotrexate, this is not a problem with the low doses used in the treatment of rheumatic arthritis.

Table 3 NSAID interactions

Practical prescribing

When prescribing NSAIDs, the most important thing is to find the right drug in terms of pain relief and lack of side-effects for each individual patient. It should only take 7 to 10 days at maximal recommended or tolerated dose to tell whether an NSAID is going to work. If pain relief does not occur and/or side-effects intervene, then that NSAID should be stopped and another one started. There does not seem to be any rationale for using combinations of NSAIDs. When using NSAIDs, the prescriber should be aware of potential drug interactions and adverse effects and should explain these appropriately to the patient.

NSAIDs can provide considerable symptom relief to patients with a variety of rheumatic diseases. Like all drugs they should be used judiciously for appropriate patients and clinical situations. Efficacy and adverse effects should be regularly reviewed. However, if these guidelines are followed, the majority of patients prescribed NSAIDs can be afforded symptom relief without adverse effects.

Chapter References

Brook, A. and Corbett, M. (1977). Radiographic change in early rheumatoid disease. *Annals of the Rheumatic Diseases*, **36**, 71–3.

Brooks, P.M. and Day, R.O. (1991). Non-steroidal anti-inflammatory drugs—differences and similarities. *New England Journal of Medicine*, **324**, 1716–25.

Day, R.O., Graham, G.G., and Williams, K.M. (1988). Pharmacokinetics of non-steroidal anti-inflammatory drugs. *Baillière's Clinical Rheumatology*, **2**, 363–93.

Forrest, M.J. and Brooks, P.M. (1988). Mechanism of action of non-steroidal anti-rheumatic drugs. *Baillière's Clinical Rheumatology*, **2**, 275–94.

Fries, J.F., Miller, S.R., Spitz, P.W., Williams, C.A., Hubert, H.B., and Bloch, D.A. (1989). Towards an epidemiology of gastropathy associated with non-steroidal anti-inflammatory drug use. *Gastroenterology*, **96**, 647–55.

Garcia-Rodriquez, L.A. and Jick, H. (1994). Risk of upper gastrointestinal bleeding and perforation associated with individual non-steroidal anti-inflammatory drugs. *The Lancet*, **343**, 769–72.

Henry, D.A. (1988). Side-effects of non-steroidal anti-inflammatory drugs. *Baillière's Clinical Rheumatology*, **2**, 425–54.

Langman, M.J-S, Veil, J., Wainwright, P., *et al*. (1994). Risks of bleeding peptic ulcer associated with individual non-steroidal anti-inflammatory drugs. *The Lancet*, **343**, 1075–8.

Lichtenstein, D.R., Syngal, S., and Wolfe, M.M. (1995). Non-steroidal anti-inflammatory drugs and the gastrointestinal tract. *Arthritis and Rheumatism*, **38**, 5–18.

Meade, E.A., Smith, W.L., and de Witt, D.L. (1993). Differential inhibition of prostaglandin endoperoxide synthase (cyclooxygenase) isoenzymes by aspirin and other non-steroidal anti-inflammatory drugs. *Journal of Biological Chemistry*, **268**, 6610–14.

O'Brien, W.M. and Bagby, G.F. (1985). Rare adverse reactions to non-steroidal anti-inflammatory drugs. *Journal of Rheumatology*, **12**, 13–20; 347–53; 562–7; 785–90.

Pincus, T. and Callaghan, L.F. (1989). Clinical use of multiple non-steroidal anti-inflammatory drug preparations within individual rheumatology private practices. *Journal of Rheumatology*, **16**, 1253–8.

Preston, S.J., Arnold, M.J., Beller, E.M., Brooks, P.M., and Buchanan, W.W. (1988). Variability in response to non-steroidal anti-inflammatory analgesics: evidence from controlled clinical therapeutic trial of flurbiprofen in rheumatoid arthritis. *British Journal of Clinical Pharmacology*, **26**, 759–64.

Radack, K. and Deck, C., (1987). Do non-steroidal anti-inflammatory drugs interfere with blood pressure control in hypertensive patients? *Journal of General Internal Medicine*, **2**, 108–22.

Scott, D.L., Symmons, D.P.M., Couton, B.L., and Popert, A.J. (1987). Long term outcome of treating rheumatoid arthritis: results after 20 years. *Lancet*, **i**, 1108–11.

Simon, L.S. (1995). Actions and toxicity of non-steroidal anti-inflammatory drugs. *Curent Opinion in Rheumatology*, **7**, 159–66.

Tonkin, A.L. and Wing, L.M.H. (1988). Interactions of non-steroidal anti-inflammatory drugs. *Baillière's Clinical Rheumatology*, **2**, 455–83.

Williams, K.M., Day, R.O., Knihinicki, R.D., and Duffield, A. (1986). The stereo selective uptake of ibuprofen enantiomers into adipose tissue. *Biochemical Pharmacology*, **35**, 3403–5.

Wilske, K. and Healey, L. (1989). Remodelling the pyramid—a concept whose time has come. *Journal of Rheumatology*, **16**, 65–7.

Yelin, E., Henke, C., and Epstein, W. (1987). The work dynamic of the person with rheumatoid arthritis. *Arthritis and Rheumatism*, **30**, 507–12.

3.5.2 Antirheumatic drugs

A. M. Denman

Introduction

The last few years have seen a change in attitude to the treatment of rheumatoid arthritis. Rheumatologists now use disease-modifying drugs early in the natural history of the disease and are less inclined to await events before resorting to potent agents (Girgis *et al*. 1994). Fuller realization of the side-effects of non-steroidal anti-inflammatory drugs has also encouraged greater reliance on other antirheumatic drugs. Awareness of the devastating effects of progressive disease on joints and general health has stimulated a less restrained attitude to trials of drug combinations. The analysis of both formal drug trials and less structured usage has also become more sophisticated so that there is more reliance on established data and less on anecdotal experience. Patients and their doctors are now better able to select a proven 'best buy'. Many of the traditional debates on the mode of action of different antirheumatic drugs have now been superseded by the realization that these agents have multiple effects on inflammatory pathways, whether immune-mediated or otherwise. A more systematic analysis of conventional antirheumatic drugs has been prompted by the acceptance that scientifically more elegant ways of specifically interfering with these events will not speedily replace other forms of treatment. This applies especially to the management of systemic connective tissue diseases where rheumatologists still rely almost exclusively on non-specific immunosuppressive and anti-inflammatory drugs, namely steroids and cytotoxic drugs.

This change in attitude means that the order in which rheumatologists resort to different antirheumatic drugs has also changed; there is less frequent recourse to gold and penicillamine and a greater dependence on methotrexate. Choice is determined partly by precedent and partly by increasing insight into the manner in which drugs function.

Much information has accrued about the pharmacokinetics, and the *in vitro* and *in vivo* effects of antirheumatic drugs. This knowledge has been gained from experimental systems and clinical observation. However, many difficulties still face the rheumatologist concerned with the intelligent application of this knowledge or involved in the quest for more effective treatment. Since the aetiology and pathogenesis of the major rheumatic diseases is unknown, the best that can be achieved is selective inhibition of key inflammatory pathways.

Improved understanding of drug actions and more sophisticated clinical trials have made drug selection more rational. As a result, a limited number of antirheumatic drugs are used predominantly and many time-honoured forms of medication are becoming obsolescent. Unfortunately, drugs in current use are still unselective in their actions and therapeutic benefit is often offset by serious toxicity, scarcely surprising given the wide range of actions of these drugs. This intolerance also reflects the long periods of administration and, quite possibly, the abnormal susceptibility of these patients to side-effects.

The great variability in the natural history of rheumatic diseases compounds the difficulties. For example the prognosis in early rheumatoid arthritis is still so imprecise that it is impossible to say whether an individual patient will enjoy early remission or will suffer inexorable disease progression of such severity that vigorous early treatment is justified. The chronicity of most rheumatic diseases poses further dilemmas; 2 years is long enough to characterize the anti-inflammatory effects and short-term toxicity of a given drug in rheumatoid arthritis, but not to evaluate its long-term benefits and risks.

Ironically, the discipline of controlled trials has given important information about the relative value of different drugs in patient populations at the expense of studying individual variation. Pharmacokinetic and genetic factors ensure that drug efficacy and tolerance vary greatly in individual patients and assuredly these factors will assume greater importance in future studies.

The order in which the antirheumatic drugs is discussed reflects their current clinical importance.

Methotrexate

After an ephemeral appearance in rheumatological practice in the 1950s, methotrexate is now universally used to treat rheumatoid arthritis because of its rapid efficacy and low toxicity. The stage in the disease at which it should be introduced is still debated, but the trend is towards early treatment. This drug is also used in systemic connective tissue diseases including systemic lupus erythematosus and Wegener's granulomatosis (Sneller *et al*. 1995).

Pharmacology

Methotrexate is aminopterin, modified by the addition of a methyl group to position 10 of the 4-aminobenzoic acid portion. It is a folic acid antagonist and in high dosage inhibits thymidylate synthetase, thereby blocking DNA synthesis (Fig. 1). However, this is not necessarily its relevant action in the much lower doses employed in rheumatological practice. Its anti-inflammatory effects are probably more important, operating by the inhibition of aminoimidazocarboxamide transformylase and leukotriene B4 production (Furst 1994).

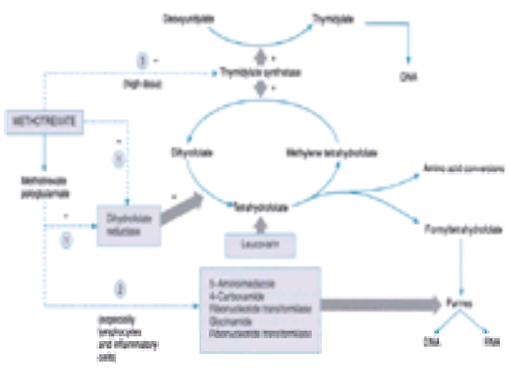


Fig. 1 Mechanism of action of methotrexate.

Since methotrexate structurally resembles pteroylglutamic (folic) acid, it is probably absorbed in the proximal jejunum. There is almost complete absorption, which is unaffected by food (Bannwarth *et al*. 1994). Bioavailability is high in doses in the range 5 to 10 mg/m ^{2} and with this dosage peak plasma concentrations are achieved as readily with oral as with parenteral administration. Subcutaneous injection is a more efficient way of sustaining high plasma levels than the intravenous route and has no disadvantages compared with intramuscular injection (Brooks *et al*. 1990). The drug is selectively concentrated in rheumatoid synovial effusions (Tishler *et al*. 1989).

Methotrexate is oxidized by aldehyde oxidase to its principal metabolite, 7-hydroxymethotrexate; both accumulate intracellularly in their polyglutamate form and have been detected in hepatocytes 1 year after drug administration. Methotrexate and its hydroxy metabolite are mainly excreted by the kidneys, but clearance of 7-hydroxymethotrexate is more prolonged. Thus methotrexate's effects are increased by impaired renal function.

Many drugs interact with methotrexate. While salicylates, most non-steroidal anti-inflammatory drugs, and sulphonamides have measurable effects on methotrexate phamacokinetics, these have not proved clinically relevant (Skeith *et al*. 1990) metabolism.

Therapeutically relevant effects (Table 1)

Table 1 Therapeutically relevant actions of methotrexate

Methotrexa. Nor has a combination of sulphonamides and methotrexate had the deleterious effects on bone marrow function predicted because of their theoretical synergistic effect on folatete affects *in vitro* and *in vivo* immune function and inflammatory reactions. Much of the experimental evidence is of dubious clinical relevance because of species differences and changes only noted with unrealistic drug doses. The most pertinent observations are those obtained by sequential observations in treated patients.

Methotrexate has proven anti-inflammatory effects. Granulocyte function is inhibited, judged by reduced chemotaxis, protease activity, leukotriene B4 production (Sperling *et al*. 1990), and superoxide radical generation (Al Balla *et al*. 1990). However, phospholipase A_2 activity is increased in peripheral blood cells of treated patients and particularly in platelets; there is some correlation with clinical efficacy (Michaels *et al*. 1994). There is also reduced granulocyte infiltration of psoriatic skin.

Methotrexate has some demonstrable immunosuppressive actions. Cell-mediated immunity is affected. A variable reduction in the numbers of circulating T cells has been reported (Wascher *et al*. 1994a); this variability is determined by drug dose and bioavailability. A reduction of blood g/ T cells has been noted in patients with juvenile chronic arthritis (Massa *et al*. 1993). However, the numbers of CD19+ T cells and activated CD25+ T cells are little changed (Lacki *et al*. 1994). Sequential synovial biopsies in treated patients showed a reduction in the numbers of infiltrating of CD4 T cells and HLA DR-positive cells (Balsa *et al*. 1993). Humoral immunity is also altered. Circulating B cell numbers are reduced (Wascher *et al*. 1994a). Rheumatoid factor titres are reduced (Spadaro *et al*. 1993) and IgA rheumatoid factor may be especially affected (Moore *et al*. 1994). Spontaneous production of rheumatoid factor by circulating B cells is also diminished (Alarcon *et al*. 1990). However,

these effects may be secondary to diminished disease activity.

Methotrexate has marginal effects on cytokine production. There is no correlation between clinical efficacy and changes in circulating cytokine levels (Kremer *et al*. 1995). Serum concentrations of interleukin 6 (**IL-6**) and tumour necrosis factor-a are unaffected by therapeutically effective doses in rheumatoid arthritis (Wascher *et al*. 1994b). It has rather unpredictable effects on the production of tumour necrosis factor-a and its soluble receptors (Barrera *et al*. 1994). It has some effects on cytokine production by peripheral blood cells and synovial cells *in vitro* including a reduction of IL-1 production, but there is no evidence that these effects are therapeutically relevant (Loetscher *et al*. 1994). Nevertheless, the expression of the gene encoding collagenase production is selectively depressed, an effect which has been attributed to an altered profile of cytokine production (Firestein *et al*. 1994).

Overall, there is little evidence that the demonstrable immunosuppressive effects of methotrexate are therapeutically relevant. Moreover, those effects which have been observed may be secondary to a reduction in disease activity through entirely different mechanisms. The most pertinent observation is that doses are commonly clinically effective which have no obvious effects on conventional markers of immune or inflammatory activity.

Clinical efficacy

Methotrexate is now an accepted drug for the treatment of rheumatoid arthritis in both hospital and community practice (Pincus *et al*. 1992). Joint inflammation is reduced and function improved in some 80 per cent of patients in the first 12 weeks of treatment and this improvement is maintained for up to 7 years in the majority of patients (Weinblatt *et al*. 1992). Controlled comparative trials have shown that methotrexate is as effective as sodium aurothiomalate and more effective than auranofin (Weinblatt *et al*. 1990). The drug is also perceived by patients to be highly effective (Tugwell *et al*. 1990). Controlled trials and assessment suggest that methotrexate may slow the appearance of radiologically detectable erosions, at least in the first 3 years of treatment (Jeurissen *et al*. 1991). There is now sufficient information about the long-term efficacy and safety of methotrexate to justify its early introduction in the treatment of rheumatoid arthritis (Furst 1994). Methotrexate may also be effective in counteracting some features of systemic lupus erythematosus and especially arthritis (Wilson and Abeles 1994), but the immunological abnormalities are not obviously affected. Controlled trials are still lacking.

Treatment regimens

Oral administration at weekly intervals is now standard treatment. There is no formally established dose schedule and clinical response and drug tolerance in individual patients are the best guidelines. The starting dose should be 5.0 mg with 2.5 mg increments at 2-monthly intervals to a maintenance dose which usually need not exceed 12.5 mg weekly. Rarely patients need higher doses of up to 20.0 mg weekly, in which circumstances closer monitoring is mandatory. A good response to higher intravenous doses has been reported in resistant disease (Gabriel *et al*. 1990). Higher doses are also not uncommonly necessary in paediatric practice. The clinical response is also the best guide to maintenance therapy. Furthermore, flares are difficult to predict from disease features in individual patients. Some patients with stable disease can be maintained on fortnightly doses (Kremer *et al*. 1995).

For use in paediatrics, see Chapter 5.6.3.

Adverse effects (Table 2)

Table 2 Adverse effects of methotrexate

With accepted dose regimens the incidence of adverse reactions necessitating drug withdrawal does not exceed 20 per cent in the first 5 years of treatment (Felson *et al*. 1990; Wolfe *et al*. 1990; Singh *et al*. 1991). Toxicity is more common in patients with impaired renal function (McKendry and Cyr 1989) and there are hints that toxicity has an immunogenetic basis (Tishler *et al*. 1988a).

Minor liver abnormalities are commonly revealed by serial blood tests and still more frequently by routine liver biopsies, but these changes are not an indication for stopping treatment and significant, irreversible damage is rare. Concomitant treatment with non-steroidal anti-inflammatory drugs increases the frequency of abnormal liver function tests, but these rarely prevent their continued, judicious use (Fries *et al*. 1993). Similarly, earlier strictures against the combined use of methotrexate and sulphonamides on pharmacokinetic grounds have been refuted (Haagsma *et al*. 1994; Nisar *et al*. 1994); clinically this combination is well tolerated. The rapid development of multiple subcutaneous nodules is a well-documented feature of patients receiving this agent (DiFrancesco *et al*. 1994).

Opportunistic infections have been reported in patients taking methotrexate. Infection by *Pneumocystis carinii* is a point of particular concern (Wollner *et al*. 1991). However, the patients were suffering from diseases associated with impaired immunity so that the contribution of methotrexate treatment is uncertain. The oncogenic risks of long-term treatment are unknown (Shiroky *et al*. 1991). Continued surveillance is essential as the numbers of patients exposed to this agent over a period of many years increases. Isolated reports of lymphomas in patients exposed to methotrexate in combination with other drugs emphasize this need (Ferraccioli *et al*. 1995).

Contrary to theoretical expectations, folic acid supplements do not reduce methotrexate's efficacy but do decrease the incidence of some adverse side-effects, notably gastrointestinal intolerance (Buckley *et al*. 1990; Morgan *et al*. 1990). It has not been established that folic acid administration prevents bone marrow suppression but serial measurements of red cell volume suggest this possibility (Weinblatt and Fraser 1989). It is now common practice to give a single oral dose of 10.0 mg of folic acid 24 h after methotrexate. In paediatric practice, a daily dose of 1.0 mg is often used to prevent nausea and the risk of bone marrow suppression.

Patient monitoring

A full blood count, serum creatinine level, chest radiograph, and liver function tests should be obtained before starting treatment. The latter should include at least serum concentrations of albumin, aspartate aminotransferase, and bilirubin. Antibodies to hepatitis B and C viruses should be sought if there is any clinical suspicion of infection by these agents. Liver biopsy is not indicated unless there is pre-existing liver disease (Kremer *et al*. 1994). Blood counts and liver function should be checked at monthly intervals for the first 3 months and an initial chest radiograph is mandatory. Thereafter 3-monthly checks are adequate in most patients. Routine liver biopsies are unnecessary and need only be performed if the cause of liver abnormalities in treated patients is in doubt. In patients with gradual increases in the blood levels of liver enzymes or alkaline phosphatase, it is usually possible to continue at a lower dose either immediately or after an intermission. Pneumonitis necessitates permanent drug withdrawal. Renal function should be monitored with particular care in patients with concomitant renal disease or who are receiving potentially nephrotoxic drugs (Seideman *et al*. 1993). In addition, impaired renal function increase the risk of toxicity (Rheumatoid Arthritis Clinical Trial Archive Group 1995).

Symptomatic discomfort such as mouth ulceration or diarrhoea may necessitate temporary withdrawal of treatment or a reduction in dose. Daily doses of folic acid (10.0 mg) may reduce the severity of symptoms. The major issue usually concerns deteriorating liver function.

Corticosteroids

The discovery of the dramatic anti-inflammatory effects of corticosteroids in the rheumatic diseases is part of medical history and has generated its own literature. The implications for current rheumatological practice are still debated. However, many pertinent questions remain unanswered.

Therapeutically relevant actions

The principal glucocorticoid hormone of the human adrenal cortex is cortisol (hydrocortisone), derived from cortisone. In common with other steroid hormones, cortisol has a basic structure of four carbon rings; the two hydroxyl groups, 11b and 17a, are important for glucocorticoid activity. Synthetic steroids are cortisol derivatives designed to increase glucocorticoid and reduce mineralocorticoid activity (Table 3). It is clearly desirable to modify the structure further so as to retain or increase the anti-inflammatory effects of these compounds at the expense of the therapeutically irrelevant or harmful actions of glucocorticoids. Deflazecort, an oxazoline analogue of prednisone, is one such compound (Gray *et al*. 1991).

Drug	Equivalent dose (mg)	Duration of action (h)
Cortisol	25	8–12
Cortisone	20	8–12
Prednisone	5	12–36
Methylprednisolone	4	12–36
Triamcinolone	4	12–36
Dexamethasone	0.75	36–72
Betamethasone	0.60	36–72

Table 3 Commonly used glucocorticoids

Free cortisol and its analogues diffuse into cells and bind to specific 95-kDa, phosphorylated protein receptors. The steroid–receptor complex moves to the nucleus where it binds reversibly to specific regulatory DNA sequences, thereby increasing or decreasing the transcription rate of adjacent genes. This in turn alters the rate of protein synthesis. In addition to these direct effects, cortisol increases the synthesis of lipocortin, which inhibits the proinflammatory enzyme phospholipase A_2 (George and Kirwan 1990). Recent evidence also suggests that steroids may interfere with cytokine-inducing transcription factors (Scheinman *et al*. 1995).

As a result, corticosteroids affect the synthesis of most major immunological and inflammatory mechanisms contributing to the pathogenesis of the rheumatic diseases (Table 4). Unfortunately, *in vitro* systems for studying these effects almost invariably discount the importance of *in vivo* regulatory pathways. Furthermore, the complexity of *in vivo* events makes it difficult to determine the relative importance of these effects. On current knowledge, it is best to attribute the therapeutic efficacy of corticosteroids to a combination of these effects.

Table 4 Therapeutically relevant actions of corticosteroids

Although synthetic steroids have enhanced glucocorticoid effects, it is difficult to see how further selectivity can be achieved by devising analogues with greater specificity in terms of target cells or therapeutically relevant actions. The only strategy currently available is to complement or replace corticosteroids with other immunosuppressive agents. Further progress depends on a better definition of the therapeutically relevant effects of corticosteroids and other immunosuppressive agents.

Indications for corticosteroid therapy

There are three main categories of rheumatic disorder in which systemic corticosteroid treatment is justified.

Systemic connective tissue diseases

The potential morbidity or mortality of many connective tissue diseases justifies corticosteroid treatment. Pertinent examples are systemic lupus erythematosus and dermatomyositis (see Section 5.7 and Section 5.9). However, the merits of other immunosuppressive drugs complicate therapeutic decisions even in these traditionally uncontroversial situations. Polyarteritis nodosa illustrates the currently contentious issues, since there is still a dearth of controlled trials indicating whether corticosteroids should complement treatment with cytotoxic drugs or be omitted entirely (see Section 5.11).

Polymyalgia rheumatica

Systemic corticosteroid treatment is still the only effective means of controlling this common disorder (see Chapter 5.11.5).

Rheumatoid arthritis

Corticosteroids now occupy an intermediate place between the euphoria which greeted their introduction and the opprobrium which followed the disclosure of gross hypercortisonism. The rehabilitation of systemic corticosteroid treatment has been aided by a truer definition of the hazards of continued treatment with non-steroidal anti-inflammatory drugs and dependence on more conservative dose regimens. Moreover, even after energetic treatment, the long-term outlook for severe rheumatoid arthritis is worse than was previously appreciated. The influence of judicious corticosteroid treatment on this progression is unknown and justifies the recent re-evaluation.

Corticosteroids are commonly used in patients over 60 years of age with rheumatoid arthritis in whom there is concern over the gastrointestinal side-effects of non-steroidal anti-inflammatory drugs and particularly the risk of haemorrhage (Schaardenburg *et al*. 1995).

Corticosteroids are also prescribed to induce a rapid remission while other drugs are taking effect. This may be given in conventional daily oral dosage, as a pulsed

oral high dose, by intramuscular injection (usually as triamcinolone acetonide), or as pulsed intravenous methylprednisolone. However, there is a risk of rebound disease activity once the steroids are withdrawn (Van Gestel *et al*. 1995).

Early corticosteroid treatment in rheumatoid arthritis has also been advocated on the grounds that this reduces the incidence of joint erosions in the first 2 years of disease activity (Kirwan *et al*. 1995). However, this is too short a period in which to judge the long-term benefits and hazards of this approach.

Adverse effects (Table 5)

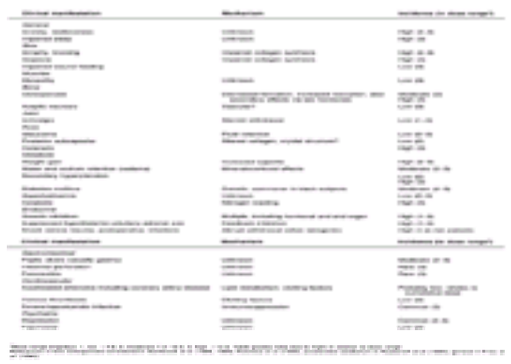


Table 5 Adverse effects of corticosteroids

The dose-related metabolic effects of systemic corticosteroid treatment have inevitable adverse consequences. The clinical manifestations depend on daily dose, duration of treatment, and genetic factors. Side-effects determined by hypersensitivity to corticosteroid analogues are rare.

Adverse effects can best be avoided by keeping maintenance doses as low as possible. Steroid treatment on alternate days is often tolerated by patients whose daily maintenance dose does not exceed 10 mg of prednisolone, but the only convincing reduction in side-effects achieved by this measure is a lower incidence of suppression of the hypothalamo–pituitary–adrenal axis.

Pulse treatment with high (1 g) or low (100 mg) doses of intravenous methylprednisolone or 1 g of oral prednisolone produce remissions of 4 to 6 weeks duration in rheumatoid arthritis. However, such approaches do not have proven short-term advantages over conventional treatment (Hansen *et al*. 1990) and there is little information about their ability to reduce the long-term adverse effects of maintenance corticosteroid treatment (Smith *et al*. 1990). Similarly, although pulse intravenous methylprednisolone is a helpful adjunct in rapidly controlling severe systemic lupus erythematosus, there is no evidence that the morbidity of maintenance corticosteroid treatment is diminished.

Sulphasalazine

Sulphasalazine was introduced in the 1930s as a rational means of treating diseases of possible infective aetiology (Svartz 1942), but it is now used pragmatically because of its established efficacy.

Pharmacology

Sulphasalazine (salicyl-azo-sulphapyridine, salazopyrine) is 5-aminosalicylic acid linked to sulphapyridine by a diazo bond. Less than 12 per cent of the intact drug is absorbed from the stomach and small intestine. It is split into its two components in the colon by bacterial action. Sulphapyridine is well absorbed and is acetylated in the liver, hydroxylated, or conjugated with glucuronic acid. The rate of acetylation is genetically determined but this does not influence efficacy or toxicity. The absorption of 5-aminosalicylic acid is poor.

Therapeutically relevant actions

Most investigational effort has centred on establishing a link between the proposed enteropathic aetiology of rheumatoid arthritis and the antibacterial effects of sulphasalazine or its moieties, but this issue is unresolved. The complete molecule and its components have different effects, which makes it difficult to identify the relevant actions. Both the parent drug and 5-aminosalicylic acid have anti-inflammatory actions by inhibiting random granulocyte mobility and chemotaxis *in vitro*; sulphapyridine has no such effects. In contrast, sulphasalazine and sulphapyridine but not 5-aminosalicylic acid inhibit *in vitro* activity of natural killer cells. Yet the parent molecule, but not its moieties, inhibits 15-hydroxydehydrogenase, thereby blocking prostaglandin inactivation by this enzyme (Porter and Capell 1990).

There are conflicting reports on the effects of sulphasalazine on immune responses. Sulphasalazine and its components have minor suppressive effects on *in vitro* lymphocyte responses to non-specific mitogens. Since the impaired *in vitro* reactivity of lymphocytes from some treated patients returns to normal with treatment, these effects might appear to have therapeutic relevance. However, the treatment dose of sulphasalazine does not affect serum concentrations of the soluble IL-2 receptor, indicating that the drug has little effect on lymphocyte activation *in vivo* (Crilly *et al*. 1993). Although the drug does not influence blood lymphocyte subpopulations, it reduces the numbers of CD3 and g/ T cells in the intestinal mucosa of treated patients (Kanerud *et al*. 1994).

Similarly, sulphasalazine suppresses B-cell activity *in vitro* (Imai *et al*. 1994), but earlier reports that it suppresses IgA production *in vivo* have not been substantiated (Jorgensen *et al*. 1993). The high incidence of IgA deficiency induced by this drug has led to speculation that immunoglobulin synthesis is more subtly affected in most recipients. Another possible effect of the drug on immunoglobulin synthesis has been proposed. There is evidence that sulphasalazine reverses the galactosylation defect in rheumatoid arthritis by suppressing autoantibodies which interfere with the enzyme galactosyltransferase in lymphocytes (Axford *et al*. 1992).

Therapeutic efficacy

Sulphasalazine reduces the intensity of rheumatoid synovitis, judged by accepted clinical and laboratory indices of disease activity. However, controlled trials show little functional improvement compared with placebo effects over a 2-year period of treatment (Capell *et al*. 1990) and 30 per cent at most benefit from long-term treatment (Porter *et al*. 1994). Hints that erosive changes may be blunted have not been confirmed.

Sulphasalazine has comparable anti-inflammatory effects in ankylosing spondylitis and its efficacy compared with placebo treatment has been confirmed in spondyloarthritis, particularly in association with psoriatic arthritis (Dougados *et al*. 1995).

Treatment regimens

The accepted regimen for instituting sulphasalazine treatment is to start with 500 mg daily, followed by weekly increases of 500 mg daily until a daily dose of 2 g is achieved after 4 weeks. Subsequent increases are dictated by tolerance and clinical response but there is no point in exceeding 40 mg/kg of body weight.

Since two-thirds of adverse effects occur within the first 3 months of treatment, clinical monitoring and blood counts are necessary at fortnightly intervals throughout this period. They should be continued at 3-monthly intervals thereafter unless there are grounds for concern. Immunoglobulins should be checked before starting treatment but need not subsequently be monitored routinely unless there is clinical suspicion of immunodeficiency.

Adverse effects (Table 6)

Reaction	Incidence (%)	Management
Discoloured urine	>50	Ignore
Oligospermia	70 (of males)	Reversible infertility Pretreatment counselling
Gastrointestinal (GI symptoms)	20	Symptomatic measures help, but 50% withdrawal
Increased mean corpuscular volume	72	No action needed
Macrocytosis	9	No action needed
Macrocytic anaemia	Rare	Rarely, headaches or haemolysis force withdrawal
Neutropenia	6	May tolerate lower dose, but 50% withdrawal
Rash (many forms from trivial to exfoliative)	5	High incidence in placebo-treated, but 50% withdrawal
Fever	2	Usually accompanies rash when significant
Stomatitis	2	Usually tolerated, may accompany severe rashes
Hepatitis	2	Withdrawal mandatory
Reduced immunoglobulin	10	Reversible
IgA deficiency	3	
Hypogammaglobulinaemia	1	

References in text.

Table 6 Adverse reactions to sulphasalazine therapy

Sulphasalazine is significantly toxic in at least 50 per cent of treated patients, forcing drug withdrawal in about 50 per cent of this number. Pre-existing sulphonamide hypersensitivity contraindicates treatment and deficiency of glucose 6-phosphate dehydrogenase enjoins extreme caution (Chalmers *et al*. 1990).

In general, these reactions are either toxic and largely dose related or attributable to hypersensitivity.

Toxicity

Orange-coloured urine is innocuous but gastrointestinal symptoms and dizziness are troublesome in up to 50 per cent of patients; symptomatic measures may avoid the need for dose reduction.

Reversible oligospermia with attendant infertility affects 70 per cent of treated males.

Red cell volumes rise in nearly 75 per cent of treated patients, amounting to frank macrocytosis in nearly 10 per cent (Farr *et al*. 1989). These changes are rarely attributable to overt folate deficiency.

Serum immunoglobulin concentrations are affected in some 10 per cent of patients; 3 per cent of treated patients develop selective IgA deficiency, 5 per cent low IgG levels, and 1 per cent panhypogammaglobulinaemia after 3 to 7 months of treatment. However, only the latter outcome is associated with a greatly increased risk of infection (Farr *et al*. 1991).

Hypersensitivity

The majority of hypersensitivity reactions can be countered by wariness and regular monitoring. Thus, while neutropenia is common and its onset often precipitous, fatal agranulocytosis is rare (Marabani *et al*. 1989). Similarly, biochemical tests often reveal evidence of asymptomatic hepatitis but clinically significant liver disease is unusual.

The overall rate of dropout for toxicity is some 20 per cent over 3 years (Felson *et al*. 1990). Concomitant steroid treatment reduces the incidence of adverse reactions (Gran and Myklebust 1993).

Immunosuppressive cytotoxic drugs

Cytotoxic drugs were introduced into rheumatological practice partly for their immunosuppressive actions and partly because of experimental evidence that they induce tolerance to specific antigens; it was reasoned that these agents would thereby restore tolerance to self-antigens in autoimmune diseases. They have indeed proved effective in rheumatoid arthritis and systemic connective tissue diseases either in isolation or as a means of reducing dependence on corticosteroids. However, it has not been established that their immunosuppressive properties account for the clinical improvement and non-specific anti-inflammatory actions may be more important. Certainly, there is no correlation between efficacy and measurable immunosuppression. The cytostatic and antiproliferative effects of cytotoxic drugs account for their therapeutic efficacy, but equally determine side-effects involving the inhibition of proliferating cells in bone marrow, gastrointestinal epithelium, and gonads. Cytotoxic drugs are also mutagenic, raising continued concern about their long-term oncogenic potential. Azathioprine is relatively safe and is still used extensively to treat rheumatoid arthritis and systemic connective tissue diseases. Alkylating agents are now reserved for life-threatening situations such as lupus nephritis and amyloidosis.

Azathioprine (Imuran)

Pharmacology

Azathioprine is a purine analogue which interferes with DNA synthesis. It is derived from 6-mercaptopurine by the addition of an imidazole group to the parent compound. The plasma half-life after oral administration is about 60 min. The drug is inactive until metabolized in liver and red cells by two pathways, one involving methylation of the sulphydryl group and later metabolites, the other oxidation of 6-mercaptopurine by xanthine oxidase.

Therapeutically relevant effects (Table 7)

Circulating lymphocyte count and subpopulation distribution	Slight non-selective fall
T cells	*In vivo* response to specific and non-specific stimuli not suppressed and may be enhanced Cytokine production unaffected Delayed hypersensitivity responses unaffected
B cells	*In vivo* response to non-specific stimulation not directly affected Minor fall in plasma immunoglobulin levels Little effect on primary and no effect on secondary antibody responses
Monocyte–macrophages	Arrival in inflammatory lesions depressed Cytokine production depressed
Natural killer cells	Circulating numbers and cytotoxicity depressed

Data relate to experiments in humans.

Table 7 In vivo immunosuppressive effects of azathioprine

The therapeutically relevant effects of azathioprine are those observed with conventional doses of 1.5 to 2.5 mg/kg of body weight. Azathioprine inhibits lymphocyte proliferation but does not affect resting lymphocytes. There is a modest, non-selective lymphopenia affecting T and B cells. Specific and non-specific *in vitro* responses of lymphocytes from treated patients are not significantly suppressed and cytokine production is also unaffected. Similarly, primary and secondary antibody responses to conventional antigens are resistant. In contrast to the unimpressive effects of azathioprine on classical immune responses, some activity of natural killer cells is inhibited (Cseuz and Panayi 1990). It is also likely that monocyte–macrophage accumulation in inflammatory lesions is blunted. Thus in the current state of knowledge, the clinical efficacy of azathioprine in inflammatory synovitis can reasonably be attributed to non-specific anti-inflammatory effects.

Therapeutic efficacy

Azathioprine is of proven benefit in rheumatoid arthritis and is as least as effective as gold and D-penicillamine. There is some indication that it limits the development of erosive changes; this protection may reflect the remarkably good, long-term tolerance of this drug enabling treatment to be continued for many years (De Silva and Hazleman 1981).

There are sufficient data from controlled studies to justify using azathioprine in patients with vasculitic diseases, polymyositis, or systemic lupus erythematosus whose disease is uncontrolled by tenable maintenance doses of corticosteroids. However, there is insufficient information about the relative benefits and disadvantages of this drug in comparison with other agents in well-documented, long-term studies. It is also unclear whether the outlook is improved by the early introduction of azathioprine or whether it should be reserved for patients with demonstrably steroid-resistant disease.

Treatment regimens

Azathioprine should not be used in a dose exceeding 2.5 mg/kg of body weight as there is no evidence that higher doses are more effective. If necessary, treatment can be continued indefinitely with appropriate monitoring.

Clinical wariness is essential in the early stages of treatment because of the risk of early intolerance masquerading as a disease flare. Monthly blood counts and liver function tests are essential. Temporary withdrawal is necessary if the neutrophil count falls below 2.5×10^{6}/l or the platelet count below 80×10^{9}/l. Treatment can usually be resumed in lower dosage once the count has recovered.

Adverse effects (Table 8)

Reaction	Incidence (%)	Management
Gastrointestinal		
Nausea	15	Many gastrointestinal symptoms controllable by drugs; often force withdrawal
Vomiting	10	
Abdominal pain	8	
Diarrhoea	5	
Mucosal ulcers	3	May remit on lower dose
Abnormal liver function	<1	May remit on lower dose
Skin		May remit on lower dose
Pruritis	2	
Rash	2	
Haematological		
Neutropenia	<1	Withdrawal; may tolerate lower dose
Thrombocytopenia	Rare	
Aplasia	Rare	Withdrawal
Acute hypersensitivity (Fever, rash, pain, often hepatotoxicity)	1	Can occur rapidly, may be confused with features of primary disease; withdrawal mandatory
Oncogenic	Slight increase	Counselling before starting treatment

Figures vary in different series. Those given are calculated on incidence in major published series.

Table 8 Adverse reactions to azathioprine

Adverse effects compel drug withdrawal in up to 30 per cent of patients in the first 6 months of treatment, but thereafter azathioprine is well tolerated (Singh I>et al. 1991). However, it is likely that this figure has subsequently been reduced owing to greater familiarity with this drug in rheumatological practice. This level of tolerance is impressive if one takes account of the disease severity and exposure to multiple drugs of patients receiving azathioprine and other cytotoxic drugs. Hypersensitivity reactions in the form of fevers, rash, and hepatotoxicity occur not uncommonly, often after only a few days of treatment (Jeurissen *et al*. 1990; Wijnands *et al*. 1990b). Gastrointestinal intolerance is common but can usually be overcome by appropriate symptomatic treatment.

Haematological problems are encountered more frequently in the elderly, in whom particular caution is essential because of the risk of marrow aplasia. Neutropenia and thrombocytopenia are usually ephemeral (Jeurissen *et al*. 1988). Macrocytosis is common and is not related to folate or vitamin B_{12} deficiency; it is not a reason for drug withdrawal.

Concern about the oncogenic potential of azathioprine has been largely assuaged in non-transplanted patients. Either no effect can be detected (Singh *et al*. 1989) or the incidence of lymphomas and related tumours is only marginally greater than that associated with rheumatoid arthritis itself (Silman *et al*. 1988).

Alkylating agents

The alkylating agents cyclophosphamide and chlorambucil (Fig. 2) have been used since the 1960s to treat immune disorders and in particular those which have proved resistant to other drugs. They have potent anti-inflammatory and immunosuppressive actions but also major adverse effects (Denman *et al*. 1992).

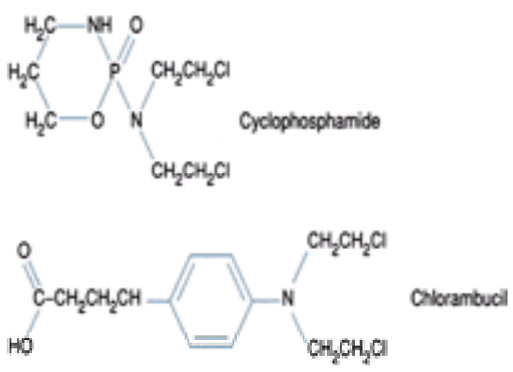


Fig. 2 Structure of alkylating agents.

Pharmacology

Alkylating agents bind covalently through their alkyl groups to other molecules, including phosphate, carboxyl, amino, sulphydryl, and imidazole groups. Binding to purine bases damages the DNA strand and interferes with mitosis and cell proliferation. The 7-nitrogen atom of guanine is particularly susceptible to covalent bonding and may be the most important target for these drugs. Damage to the DNA strand affects mitosis, thereby inhibiting cell division. However, these effects may be mutagenic rather than cytotoxic, creating long-term oncogenic risks (Lawley 1990). The damage to lymphocytes and other proinflammatory cells is immunosuppressive. In addition, the oxazaphosphorine cyclophosphamide is bifunctional (i.e. it has two active groups) so that covalent cross-linking of DNA prevents strand separation during mitosis.

Cyclophosphamide is absorbed rapidly but incompletely from the gut, with great individual variability. Maximum plasma concentrations of 1 to 2 µg/ml are reached some 2 h after an oral dose of 2 to 3 mg/kg. These levels reach 40 to 50 µg/ml after intravenous administration of larger doses (25 mg/kg). There is marked individual variation in plasma half-life irrespective of the means of administration.

Cyclophosphamide itself has neither alkylating nor cytotoxic activity but many of its metabolites possess these properties (Fig. 3). Less than 20 per cent is excreted unchanged by the kidney and the remainder is metabolized by liver microsomes. The initial oxidation product, 4-hydroxycyclophosphamide, is highly cytotoxic but also

unstable; it may also serve as a carrier molecule for transporting the final metabolite, phosphoramide mustard, across cell membranes. Renal excretion of the cytotoxic metabolites carboxyphosphamide and 4-ketocyclophosphamide accounts for some 65 per cent of the initial compound and is greatly slowed when renal function is impaired. Of clinical relevance, the intravenous administration of cyclophosphamide in high dose induces activating enzymes, which increase the rate of drug clearance and also peak plasma concentrations of its cytotoxic metabolites.

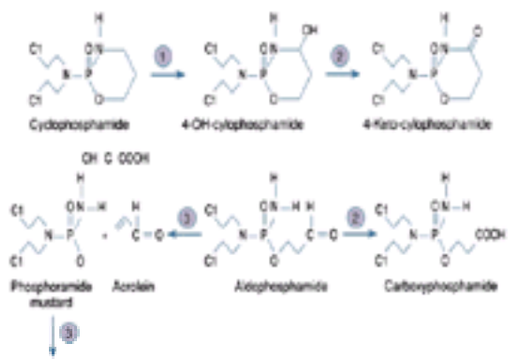


Fig. 3 Metabolic pathways of cyclophosphamide.

Therapeutically relevant effects (Table 9)

Blood lymphocytes	Non-selective lymphopenia, affecting all subpopulations; cytotoxic for resting lymphocytes and antiproliferative; marked fall after high oral dose, or pulse intravenous administration
T cells	*Suppresses:* *In vitro* activation and proliferation in response to specific and non-specific stimuli *In vitro* cytotoxicity Helper function for B cells Delayed hypersensitivity responses, if given in high or pulse dosage
B cells	Initially more sensitive than T cells *Suppresses:* *In vitro* immunoglobulin synthesis, spontaneous and induced Spontaneous *in vitro* IgM rheumatoid factor synthesis Serum immunoglobulin levels, especially with continued high dose, or pulse treatment Primary, but not secondary antibody responses to conventional antigens Autoantibody production
Monocyte–macrophages	Reduced proteolytic enzyme secretion Reduced cytokine production
Natural killer cells	Production and cytotoxicity variably affected, depending on dose

References in text, reviewed in Zeiniger *et al.* (1992).
(Only in vivo effects are considered; cyclophosphamide itself is inactive.
(Effects of chlorambucil and cyclophosphamide are similar but the latter is more effective, especially in high dose.

Table 9 Therapeutically relevant actions of alkylating agents

Alkylating agents have immunosuppressive effects of probable therapeutic importance. However, many reported observations are of dubious relevance because of species differences or contrived experimental design. For example, there is no convincing evidence that cyclophosphamide perturbs human immune responses by stimulating suppressor T-cell activity. Nor is there any data to suggest that it induces tolerance to autoantigens in a manner analogous to tolerance induced in animals to protein antigens by the correct timing of drug administration in relation to immunization.

Cyclophosphamide and chlorambucil reduce the numbers of circulating T and B lymphocytes in an essentially non-selective manner with some hint that CD4+ T cells are especially vulnerable. Primary antibody responses to test antigens are suppressed, but established antibody responses are unaffected by low-dose cyclophosphamide (0.5 to 1.0 mg/kg of body weight per day) and only slightly suppressed by higher doses given orally or as intravenous pulses. Autoantibody titres and polyclonal hypergammaglobulinaemia are reduced, but it is unclear whether these changes result from direct effects on B cells or a reduction in disease activity by other mechanisms. The activities of circulating T cells and other cells of treated patients, such as cytokine production and *in vitro* proliferation, are inconsistently affected.

Clinical efficacy

Recourse to alkylating agents is rarely justified in the treatment of rheumatoid arthritis. Their use is defensible in patients with severe systemic disease, but even in this context, there is still a lack of long-term trials comparing the benefits and risks of these drugs with those of other agents. Cyclophosphamide is of proven value in treating the nephritis and other systemic features of systemic lupus erythematosus, but the extent to which it should replace or supplement corticosteroids is controversial. Comparative data are also lacking on the relative efficacy of cyclophosphamide and newer immunosuppressive regimens. Uncontrolled experience suggests that it is more effective than corticosteroids in treating vasculitic disorders, but this has not been formally proven in comparison with other immunosuppressive agents.

Chlorambucil is established treatment for amyloidosis complicating juvenile chronic arthritis. It is now rarely necessary to use this agent in preference to azathioprine or cyclosporin to treat Behçet's syndrome and, in particular, uveitis complicating this disease.

Treatment regimens

Chlorambucil is usually given daily by mouth. The conventional starting dose is 0.1 mg/kg of body weight and subsequent doses should rarely exceed 0.2 mg/kg of body weight.

Cyclophosphamide is most commonly given orally in an initial daily dose range of 1.5 to 2.5 mg/kg of body weight. Higher, intermittent doses are often given intravenously for severe disease, but in this, as in so many situations, practice is dictated as much by local preferences as by the persuasiveness of collated data. However, there are claims that high-dose pulses of cyclophosphamide synchronized with plasma exchange can induce remissions in systemic lupus erythematosus (Euler *et al.* 1994). This is another situation in which there are unproven claims that earlier, more vigorous treatment is more effective than conventional protocols.

Clinical response and toxicity are the only guides to maintenance doses. Blood lymphocyte counts are a poor indication of efficacy or toxicity, while more sophisticated analyses of immune function have not yet proved of practical value.

Neutropenia and thrombocytopenia enforce the measures described for azathioprine treatment.

Adverse effects (Table 10)

Reaction	Incidence (%)	Management
Reproduction		
Male sterility	Virtually 100	Counselling patient or parents mandatory
Female sterility	Incomplete data, probably < 30	More data needed
Amenorrhoea	30	Hormone replacement for premature menopause unless contraindicated
Gastrointestinal		
Nausea	10	Usually responds to symptomatic measures
Vomiting	5	
Alopecia	< 1	Uncommon in dosage used in rheumatological practice
Bladder	<5	Uncommon in low oral or pulse intravenous treatment; mesna is good renal output
Infections	8	Rarely a reason for withdrawal
Haematological		
Macrocytosis	80	Not a reason for withdrawal
Neutropenia	5	Temporary withdrawal; risk of pancytopenia underestimated?
Thrombocytopenia	< 1 ?	
Pancytopenia	< 1 ?	
Lung	Rare	Withdrawal
Neoplasia	Probably justified risk/use	Keep cumulative dose to minimum
		Protect health workers exposed to drugs

References in text; see also discussion of risks of neoplasia (Fairweather *et al*. 1988), bladder complications (Stillwell *et al*. 1988), and infections (Dooley *et al*. 1990).

Table 10 Adverse actions of alkylating agents

A full explanation of the risks of treatment is even more essential before starting treatment with alkylating agents than with other antirheumatic drugs. Alkylating agents are inevitably mutagenic and clinically apparent toxicity underestimates the true extent of this process. Alopecia, stomatitis, and gastrointestinal intolerance are rarely a major problem in the doses used in rheumatological practice. Bone marrow suppression is common but almost always reversible if the drug dose is adjusted at the first sign of marrow toxicity. It is often possible to continue with a lower dose after a period of intermission. Patients with connective tissue diseases may be abnormally susceptible to bone marrow suppression (Thomas *et al*. 1983) so that especial wariness is essential. Marrow aplasia may ensue if warning evidence is ignored and may be slow to recover. Conversely, macrocytosis is very common and does not presage marrow toxicity; nor is it necessarily a marker of folate or vitamin B_{12} deficiency. It is also important to take account of low blood counts consequent on disease activity, which often improve with the response to drug treatment.

Intercurrent infections of unusual severity or by opportunistic infections are also encountered but these are not common in the dosage used in rheumatic disorders. Reversal of the immunosuppression accompanying disease may even reduce the risk of intercurrent infection.

Male sterility is almost inevitable after even low doses of alkylating agents. Amenorrhoea and sterility in females are less predictable but may still affect up to 30 per cent of those at risk (Schilsky *et al*. 1980), and serum gonadotrophin levels rise significantly after 2 months of treatment. Accurate long-term data on female fertility after treatment in childhood for rheumatic disorders are lacking.

Bladder toxicity is not usually encountered with pulsed intravenous cyclophosphamide which in general does not exceed a dose of 15 mg/kg of body weight per month. It is sufficient to ensure a good renal output thereby obviating the need for mesna. However, haematuria and aseptic cystitis are encountered in some 10 per cent of patients receiving oral cyclophosphamide for periods exceeding 12 months.

The oncogenic effects of alkylating agents are hard to assess. A long-term study has shown a small but definite increase in skin cancer and myeloproliferative diseases induced by these agents (Kinlen 1985). However, sporadic reports of unusual neoplastic diseases and bladder cancer induced by relatively short courses of cyclophosphamide suggest that the true incidence of cancer induced by alkylating agents may have been underestimated (Denman *et al*. 1992). This is especially likely to be true of bladder cancer in patients treated with long-term cyclophosphamide for lupus nephritis.

Cyclosporin

The immunosuppressive properties of cyclosporin were established in experimental and clinical transplantation. After a long period of uncontrolled observations in autoimmune diseases, its therapeutic efficacy has now been validated in controlled trials, especially in rheumatoid arthritis and uveitis. Much of the initial concern about its nephrotoxic hazards arose when high doses were employed in renal transplantation. These fears have been allayed by the low incidence of nephrotoxicity with the dose schedule used to treat autoimmune diseases; this nephrotoxicity is almost invariably reversible.

Pharmacology

Cyclosporin is a cyclic peptide derived from a fungus. Several of the amino acids are *N*-methylated and uncommon amino acids occupy the first two places in the sequence. It is usually given orally although it can also be administered by intramuscular or intravenous injection. The proven immunosuppressive dose range in clinical transplantation is 10 to 20 mg/kg of body weight per day, which produces plasma concentrations of 100 to 400 ng/ml with a half-life of 17 to 40 h. Cyclosporin is metabolized primarily in the liver and is excreted mainly via the gut; only a small fraction is recovered in the urine as the parent drug or its metabolites. The protein cyclophilin (molecular weight 17.8×10^3) is the major intracellular receptor for the agent and this binding is crucial for the drug's immunosuppressive effects. Cyclophilin consists of a single polypeptide chain with 165 residues whose structure and binding site for cyclosporin have been characterized (Braun *et al*. 1995). The complex formed between cyclosporin and cyclophilin interacts with the serine/threonine phosphatase calcineurin. This trimolecular complex interferes with factors governing the transcription of interleukin coding genes and other steps in the activation of lymphocytes and antigen-presenting cells. There is probably a family of cytophilins. A different cytophilin binds the immunosuppressive agent FK506.

Therapeutically relevant effects (Table 11)

Antigen-presenting cells (dendritic and macrophages)	*In vitro* depresses: Antigen presentation Class II molecule expression
T cells	*In vitro* depresses: Activation Proliferation Cytokine production (especially interleukin 2) Helper function and cytotoxicity *In vivo* restores: Impaired proliferative responses to *in vitro* mitogens Defective interleukin 2 receptor expression (secondary to disease activity) Impaired delayed hypersensitivity reactions No evidence that tolerance affected
B cells	*In vitro* depresses: Mitogen-induced immunoglobulin synthesis *In vivo*: Autoantibody production suppressed probably secondary to reduced disease activity

Table 11 Therapeutically relevant actions of cyclosporin A

Cyclosporin has many actions of which depressed T-cell activation is currently considered to be the most therapeutically relevant (Yocum *et al*. 1990; Yocum 1993). The problem is comparable with that posed by corticosteroids; it is not yet clear whether its efficacy in clinical immunosuppression results from a few crucially important actions or from a combination of these actions. Furthermore, the rapid relapses observed in many patients when the drug is withdrawn suggest that ephemeral anti-inflammatory effects are at least as important as T-cell suppression; the latter would be expected to be longer lasting, particularly if tolerance induction was involved. Clinical remission in rheumatoid arthritis is associated with improved *in vitro* lymphocyte responsiveness to standard antigens and reduced numbers of circulating cells expressing IL-2 receptors, findings which are as likely to reflect reduced disease activity as direct suppressive effects on T cells. Cyclosporin has little effect on conventional markers of the acute-phase response even when inflammatory activity is reduced, judged by clinical criteria.

Clinical efficacy

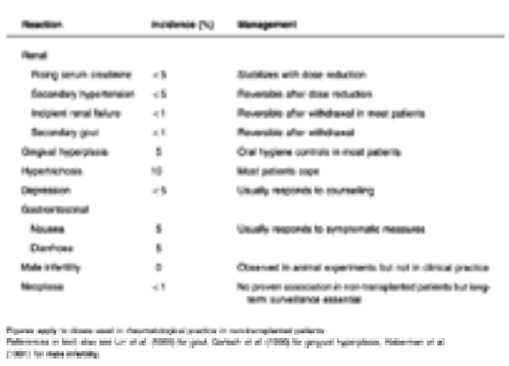
In controlled trials, cyclosporin has been shown to be effective in the treatment of severe rheumatoid arthritis in terms of reducing clinical markers of joint

inflammation. However, it has little effect on the acute-phase response and its ability to influence cartilage degradation is unproven (Dijkmans *et al*. 1993). A controlled trial has shown that cyclosporin in combination with methotrexate is more effective in the treatment of the disease than methotrexate alone (Tugwell *et al*. 1995). However, periods of follow-up have been short. So far the drug has mainly been reserved for patients whose disease has been resistant to other drug treatment. There are many uncontrolled reports of cyclosporin's efficacy in systemic autoimmune diseases. For example, anti-inflammatory activity has been shown in a short, 20-week trial of treatment in systemic lupus erythematosus (Tokuda *et al*. 1994). Its value in such situations in comparison with more conventional drug treatment has still to be established. Even in Behçet's uveitis where its short-term benefit has been proven in controlled trials, its long-term ability to suppress late disease features has not yet been demonstrated.

Treatment regimens

It is evident from controlled trials that a clinical response can be obtained and toxicity minimized if the dose of cyclosporin is kept in the range of 2.5 to 5.0 mg/kg of body weight. Renal function should be assessed before starting treatment. Blood pressure and serum creatinine should be checked at intervals of not more than 2 months. The drug should be stopped if a rising creatinine concentration or hypertension are noted, but it is often possible to resume treatment at a lower dose. For routine clinical purposes it is unnecessary to monitor blood levels of cyclosporin.

Adverse effects (Table 12)



Reaction	Incidence (%)	Management
Renal		
Rising serum creatinine	<5	Stabilizes with dose reduction
Secondary hypertension	<5	Reversible after dose reduction
Incipient renal failure	<1	Reversible after withdrawal in most patients
Secondary gout	<1	Reversible after withdrawal
Gingival hyperplasia	5	Oral hygiene controls in most patients
Hypertrichosis	10	Most patients cope
Depression	<5	Usually responds to counselling
Gastrointestinal		
Nausea	5	Usually responds to symptomatic measures
Diarrhoea	5	
Male infertility	0	Observed in animal experiments but not in clinical practice
Neoplasia	<1	No proven association in non-transplanted patients but long-term surveillance essential

Figures apply to doses used in rheumatological practice in non-transplanted patients
References in text, also see Lin et al. (2000) for gout, Gerlach et al. (1998) for gingival hyperplasia, Habermann et al. (1991) for male infertility

Table 12 Adverse actions of cyclosporin

Most of cyclosporin's adverse effects are symptomatically troublesome but do not threaten the function of vital systems. With low-dose treatment, serum creatinine concentrations do not rise significantly in the majority of patients (Landewe *et al*. 1994). Creatinine levels often stabilize with dose reduction thereby allowing treatment to be continued. Furthermore, they almost invariably return to normal once the drug has been withdrawn (Boers *et al*. 1990). The experience has been similar in large numbers of patients with psoriasis. Any associated hypertension also remits after drug withdrawal.

The impaired renal function is likely to result from reversible constriction of renal arterioles (Yougelman *et al*. 1991); this constriction may be in part secondary to selectively increased plasma concentrations of the thromboxane metabolite, 2,3-dinor thromboxane B_2 (Weinblatt *et al*. 1991). Glomerular filtration and urea excretion are also depressed (Laskow *et al*. 1990). There has also been great interest in renal interstitial fibrosis induced by cyclosporin, which may result from the enhanced transcription of procollagen coding genes (Nast *et al*. 1991). However, irreversible histological changes have been observed mainly in transplant recipients and are hard to interpret because of coexisting abnormalities associated with transplantation.

Chloroquine and hydroxychloroquine

Chloroquine and hydroxychloroquine have been used extensively to treat rheumatoid arthritis and systemic lupus erythematosus since 1943. These agents have demonstrable immunosuppressive effects *in vitro* of which the most impressive are depressed lymphocyte proliferation in response to stimulation by antigens or non-specific mitogens. This suppression is attributable to impaired antigen presentation by specialized cells, but has been analysed only in descriptive and not contemporary molecular terms. Since hydroxychloroquine is far less likely to cause retinal toxicity than chloroquine, it is more commonly used.

Pharmacology

Hydroxychloroquine and chloroquine are slowly absorbed from the upper gastrointestinal tract and a very high percentage of the drug (99.9 per cent) becomes tissue associated, with selective concentration in the liver, spleen, and heart.. These drugs bind especially avidly to melanin, a partial but incomplete explanation for their marked affinity for the iris, choroid, and skin. Drug concentration in blood white cells is also notably high. There is a delay of 3 to 4 months before steady-state plasma concentrations are reached and an equally slow fall when the drug is discontinued; excretion continues for months or even years (Tett *et al*. 1990).

Therapeutically relevant effects

Hydroxychloroquine is anti-inflammatory, primarily because it reduces the secretion of lysosomal enzymes by granulocytes and macrophages in response to stimuli such as immune complexes. It is also immunosuppressive *in vitro*, inhibiting T and B lymphocyte responses to non-specific mitogens (Van Loenen *et al*. 1990), but there is little evidence that it is immunosuppressive *in vivo* in conventional dosage. There is an intriguing suggestion that it increases pain thresholds through its effects on neurotransmitters (Middleton *et al*. 1995).

Therapeutic efficacy

Hydroxychloroquine and chloroquine measurably reduce disease activity in early rheumatoid arthritis (Tenorio and Orozco 1993; HERA Study Group 1995). However, it was not much more effective than the placebo in many reported trials (Nuver-Zwart *et al*. 1989). In terms of comparative efficacy, hydroxychloroquine is less effective than sulphasalazine judged objectively by clinicians or subjectively by patients. Any blunting of erosive changes is marginal (Van der Heijde *et al*. 1989).

Chloroquine and hydroxychloroquine are commonly used to treat systemic lupus erythematosus with the object of avoiding steroid treatment. There are uncontrolled claims that it suppresses fever, arthritis, and skin lesions, with an accompanying reduction in erythrocyte sedimentation rate and autoantibody titres. However, in a placebo-controlled trial only the joint manifestations responded (Williams *et al*. 1994). It is ineffective in combating more severe manifestations such as lupus nephritis. There is a suggestion that concomitant treatment with antimalarials may protect patients from the adverse effects of steroids on fat metabolism (Hodis *et al*. 1993).

Treatment regimens

Hydroxychloroquine is less likely than chloroquine to cause retinopathy and is therefore the more logical choice. There is good evidence that 200 mg daily is as effective as higher doses (Pavelka *et al*. 1989). Fear of ocular complications commonly restricts the duration of treatment, but the drug may be continued indefinitely with proper precautions (Carmichael 1992).

Adverse effects

Retinopathy is the complication of greatest concern because it may lead to progressive blindness. Corneal deposits disappear once the drug is stopped and are unrelated to retinal changes. A premaculopathy leads to the development of field loss to a red target between 4 and 9 degrees from fixation. The macular changes at this stage may be minimal and the visual fields may return to normal once the drug is stopped. If unchecked, the maculopathy will lead to a central, irreversible scotoma. However, this problem is hardly ever encountered in patients treated with hydroxychloroquine. It is particularly unlikely to cause ocular toxicity if the daily

dose does not exceed 6.5 mg/kg of lean body weight per day or a cumulative dose of 200 g. The eyes should be checked before starting treatment so as to establish a clinical baseline. The corrected visual acuity should be recorded, the fundi examined, and the central visual fields plotted. If any abnormalities are noted, the patient will require continued ophthalmological supervision. Patients should be warned that the drug may affect their vision and ideally they should be given a red Amsler chart with appropriate instruction on its use so that they can check their vision monthly. Any patient who notices any visual abnormality should stop the drug and obtain advice at once. An annual ophthalmological examination is not usually necessary. The prescribing physician should check vision with a red Amsler chart if any patient is unlikely to follow advice. The incidence of ocular problems necessitating drug withdrawal is now less than 1 per cent (Felson *et al*. 1990).

Gastrointestinal intolerance occurs in 10 per cent of patients and may be overlooked because of its insidious onset. It responds poorly to antacids and related drugs and is the commonest reason for premature withdrawal. Rashes appear in less than 1 per cent of patients; the drug does not cause increased photosensitivity (Seideman and Ros 1992). Hydroxychloroquine improves glucose tolerance and reduces serum cholesterol levels, changes with potential long-term but as yet uncalculated benefit (Wallace *et al*. 1990). Thrombocytopenia and other forms of bone marrow suppression are rare (Wijnands *et al*. 1990a). Toxicity overall does not exceed 20 per cent (Felson *et al*. 1990) but the drug is withdrawn in a far higher percentage of patients because of lack of efficacy (Wolfe *et al*. 1990).

D-Penicillamine

D-Penicillamine was introduced into rheumatological practice in the belief that it would dissociate immune complexes by its effects on interchain disulphide bonds in the immunoglobulin molecule. Much information has accrued about penicillamine's actions on immune and other systems but it is used in essentially pragmatic fashion.

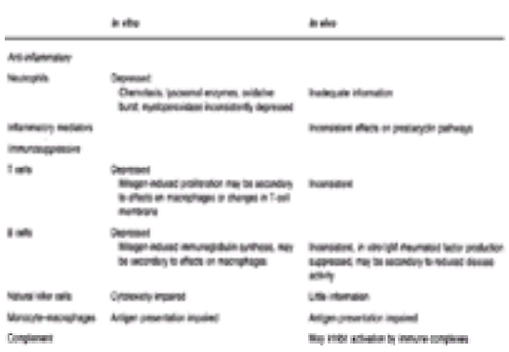
Pharmacology

D-Penicillamine (D-(–)-2-amino-3-mercapto-3-methylbutyric acid) has an a-amine, a carboxyl, and a sulphydryl functional group, conferring great complexities in terms of the bioactivity of the drug and its metabolites. There are also pharmacokinetic complexities; peak plasma concentrations are reached irregularly between 1.5 and 4 h after oral dosage because of extensive interactions with dietary proteins and iron, antacids, and gut wall proteins which favour oxidation to poorly absorbed disulphides (Joyce 1990). The drug binds to skin, connective tissue, and albumin but uptake by cells is limited.

Little D-Penicillamine is cleared by the kidneys and most is biologically transformed to disulphides or inorganic sulphate with the formation of minor products. Disulphide formation results from oxidation by interaction with an oxidant such as oxygen or through thiol–disulphide exchange and is probably dependent on the catalytic action of trace amounts of transition metals such as copper. Disulphides are not excreted renally and accumulate during treatment; thus elimination is slow when the drug is withdrawn.

The methylation product *S*-methyl-D-Penicillamine is quantitatively a minor metabolite. However it is possible, although not yet demonstrated, that this product is further metabolized to a sulphoxide. Poor sulphoxidation has been linked to an increased risk of toxicity and so this proposed pathway is of practical interest (Madhok *et al*. 1990).

Therapeutically relevant actions (Table 13)



	In vitro	*In vivo*
Anti-inflammatory		
Neutrophils	Depressed: Chemotaxis, lysosomal enzymes, oxidative burst; myeloperoxidase inconsistently depressed	Inadequate information
Inflammatory mediators		Inconsistent effects on prostacyclin pathways
Immunosuppressive		
T cells	Depressed Mitogen-induced proliferation may be secondary to effects on macrophages or changes in T-cell membrane	Inconsistent
B cells	Depressed Mitogen-induced immunoglobulin synthesis, may be secondary to effects on macrophages	Inconsistent; *in vitro* IgM rheumatoid factor production suppressed, may be secondary to reduced disease activity
Natural killer cells	Cytotoxicity impaired	Little information
Monocyte–macrophages	Antigen presentation impaired	Antigen presentation impaired
Complement		May inhibit activation by immune complexes

Table 13 Therapeutic actions of D-Penicillamine

D-Penicillamine has several actions which could contribute to its therapeutic efficacy. These results are difficult to interpret since most of the data derive from *in vitro* experiments in which, hardly surprisingly, D-Penicillamine's many functional groups perturb assay systems unprotected by *in vivo* constraints. The effects are also dose dependent, particularly where these affect *in vitro* immune function. Furthermore, while the influence of drug metabolites on macrophages is often invoked to explain the drug's immunosuppressive effects, the data relate to immunological rather than pharmacological events. Reduced reactivity of T and B cells has been documented consistently in rheumatoid arthritis, but it is not known whether these are primary drug actions or secondary to reduced disease activity. It has been proposed that the drug impairs immunoregulation by the surface sulphydryl oxidation of the lymphocyte surface membrane (Brown-Galatola and Hall 1992).

Therapeutic efficacy

Many controlled trials have shown that D-Penicillamine is more effective than placebo in reducing the symptoms and signs of rheumatoid synovitis, with a corresponding reduction in laboratory measures of disease activity such as the erythrocyte sedimentation rate (Joyce 1990). There is some indication that D-Penicillamine is more effective than sulphasalazine in maintaining long-term remission in the small percentage of patients who continue to tolerate the drug (Carroll *et al*. 1989). There is a hint that many patients may remain in remission if treatment is stopped either as a therapeutic decision or because of the patient's failure to comply (Doyle *et al*. 1993). Serum concentrations of all immunoglobulin classes, IgM rheumatoid factor, and immune complexes also decline. Nodules and vasculitic features are not influenced by D-Penicillamine treatment. It is unlikely that erosive changes are retarded in the short or long term (Scott *et al*. 1990).

There is little evidence to show that penicillamine is effective in juvenile rheumatoid arthritis and any benefit in scleroderma is marginal.

Treatment regimens

The initial daily dose of D-Penicillamine is 125 mg, increasing to 600 mg over 6 months according to tolerance and clinical response. Little further benefit can be expected with daily doses exceeding 750 mg. The drug should be taken 1 h before meals to avoid interactions with food. Clinical monitoring, urinalysis, and blood counts must be repeated at monthly intervals.

Adverse effects (Table 14)

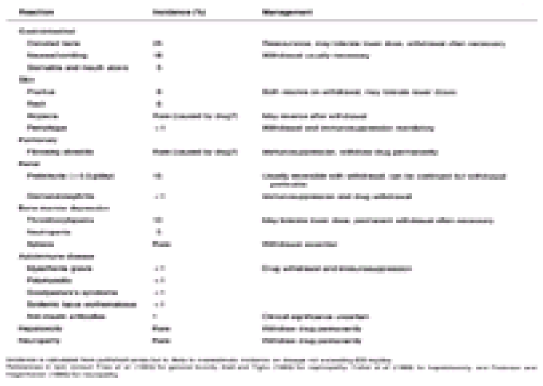


Table 14 Adverse reactions to D-Penicillamine

Adverse reactions to D-Penicillamine have been reported in some 50 per cent of patients with rheumatoid arthritis in the first 6 months of treatment; withdrawal was necessary in some 25 per cent of patients in these series in the first 5 years of treatment (Felson *et al*. 1990; Wolfe *et al*. 1990; Singh *et al*. 1991). Toxicity is associated with inheritance of a C4A null allele in the complement system, but this difference is insufficiently striking to be applicable clinically (Clarkson *et al*. 1992). Anti-Ro autoantibodies are also a possible pointer to adverse effects (Vlachoyiannopoulos *et al*. 1991). However, toxicity is dose dependent and is likely to have been greatly reduced with current prescribing practice. Proteinuria does not necessarily mean stopping treatment, but patients with this complication must be monitored carefully; the drug must be withdrawn in patients who develop nephrotic syndrome (De Silva and Eastmond 1992).

The pathological basis of this toxicity has not been elucidated. Clearly, the diversity of functional groups in the D-Penicillamine molecule gives great scope for the formation of immunogenic drug–protein complexes. Glomerular deposition of immune complexes probably accounts for the proteinuria. Similarly, thrombocytopenia is likely to follow a response to the drug acting as a hapten in combination with platelet membrane proteins. Antibody formation is probably the mechanism of penicillamine-induced myasthenia. Although myasthenia gravis is rare compared with other adverse reactions induced by D-Penicillamine, it is a particularly intriguing model because autoantibodies to the acetylcholine receptor in drug-induced and spontaneous myasthenia gravis have similar sequence and antigen-binding features (Tzartos *et al*. 1988; Morel *et al*. 1991), albeit against a different immunogenetic background. D-Penicillamine also induces the production of anti-insulin autoantibodies in some 40 per cent of treated patients with rheumatoid arthritis (Vardi *et al*. 1992), but their functional consequences are uncertain. Pemphigus has also been reported in treated patients and an autoimmune mechanism may be responsible (Bauer-Vinassac *et al*. 1992). Possibly, myasthenia and other autoimmune complications are induced by immunogenic drug–protein complexes; alternatively, the drug may induce mutations in the genes encoding variable sequences in T-cell antigen receptors or immunoglobulins.

Gold

Historically, gold was used to treat rheumatoid arthritis on the assumption that this antituberculous agent might benefit patients with a disease of presumed tuberculous aetiology. Relatively short-term trials in the 1960s showed that gold reduces disease activity and subsequently there has been much effort to lend scientific credibility to empirical usage. Gold is administered by intramuscular injection as the water-soluble thiolate, sodium aurothiomalate, or orally as the lipophilic monomeric phosphene, auranofin. However, intramuscular gold is being superseded by other drugs which are demonstrably more effective, easier to monitor, and less likely to produce severe, unforeseen toxicity. Oral gold is little used because of its gastrointestinal side-effects.

Pharmacology

Aurothiomalate is absorbed rapidly after intramuscular injection and peak plasma concentrations are reached after some 2 h (Champion *et al*. 1990). After standard weekly intramuscular doses of 50 mg of gold for 4 weeks, plasma concentrations rise to an average 300 µg/100 ml with a half-life of 3 to 5 days. With monthly maintenance doses of 50 mg, plasma levels fall to an average of 100 µg/100 ml with a half-life of 10 days. A high percentage of the injected gold is associated with cells in the circulation and in tissues so that only 40 per cent is excreted, of which 70 per cent is in the urine.

Some 25 per cent of orally administered gold is absorbed, most of the remainder being excreted in the faeces. The average plasma half-life is 15 days, with much individual variation. Although the elimination of orally administered gold is slow, it is more rapid than that of gold given by intramuscular injection.

Gold is selectively concentrated in macrophages in the synovial membrane and other tissues. This selectivity explains its widespread distribution in liver, spleen, kidney, and skin, and for its retention many years after treatment is discontinued. Only trace amounts are found in lymphocytes, granulocytes, or other cell types after *in vivo* or *in vitro* exposure. The association with phagolysosomes produces membrane-bound structures which can be readily detected in ultrastructural changes. This combination probably accounts for most of gold's anti-inflammatory and immunosuppressive effects. In addition, auranofin is taken up by polymorphonuclear granulocytes, resulting in the release of lysosomal enzymes and the inhibition of phagocytosis and the oxidative burst.

Therapeutically relevant effects (Table 15)

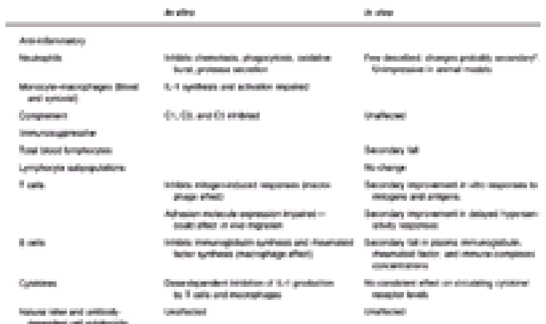


Table 15 Therapeutic actions of gold[a]

Gold suppresses non-specific inflammation in animal experiments. There is some evidence that it inhibits neutrophil accumulation by blocking the expression of the endothelial adhesion receptor ELAM-1 (Corkill *et al*. 1991). However it has little effect on immunologically induced models of chronic inflammation; nor does it alter the natural history of genetically determined, murine autoimmune disease.

Many non-specific *in vitro* immune responses are suppressed by gold. These include T-lymphocyte transformation and polyclonal immunoglobulin production by B cells in response to non-specific mitogens. This suppression can reasonably be attributed to gold's actions on monocyte–macrophages.

Gold salts also affect cytokine production. Gold sodium thiomalate and auranofin can stimulate or suppress the *in vitro* IL-1 production by monocyte–macrophages, the outcome depending on drug concentration and individual variation (Danis *et al*. 1990a). The pattern of *in vitro* IL-1 production by blood monocyte–macrophages from treated patients is highly variable, probably reflecting changes in disease activity rather than direct effects of the drug (Danis *et al*. 1990b). Monocyte accumulation in rheumatoid synovitis and cytokine production by these cells is also reduced by gold treatment (Yanni *et al*. 1994). It does not significantly affect circulating levels of soluble IL-2 receptors (Kirkham *et al*. 1993).

Delayed-type skin hypersensitivity reactions are consistently depressed in active rheumatoid arthritis, an abnormality corrected by chrysotherapy. However, this

improvement bears little relation to the clinical response (Smith *et al*. 1989). Like many antirheumatic drugs, gold reduces *in vitro* synthesis of rheumatoid factor by blood B cells from treated patients, but it is difficult to be sure if this is secondary to reduced disease activity or an independent mechanism. Gold is demonstrably anti-inflammatory through its effects on monocyte–macrophages but there is no predictable way of using this information for clinical monitoring.

Therapeutic efficacy

Sodium aurothiomalate reduces joint inflammation in rheumatoid arthritis, judged by controlled trials of up to 2-years duration. It is less certain that there is any longer-lasting benefit (Epstein *et al*. 1991); moreover, any retardation or reversal of joint destruction is modest at best. Since the natural history of rheumatoid arthritis is so variable, it is difficult to distinguish between an average, small effect on all patients with the disease and a major influence on a subset of patients which is masked when the overall benefits are analysed. Studies in the community have cast doubt on the benefit of long-term maintenance gold treatment (Epstein *et al*. 1991). Gold is of limited efficacy in psoriatic arthritis and is now rarely used in this condition.

Auranofin is of comparable efficacy to sodium aurothiomalate, judged by placebo-controlled trials of up to 2-years duration (Williams *et al*. 1988; Johnsen *et al*. 1989). The drug has also been shown to be modestly anti-inflammatory in juvenile chronic arthritis.

Dose regimens

Traditionally, sodium aurothiomalate is given in incremental weekly doses, starting at 2.5 to 5.0 mg and reaching a maximum of 50 mg. There is little evidence that the highest dose is mandatory, unless there is no clinical response to lower doses. Reduced doses are commonly tolerated by patients in whom higher doses have produced minor toxicity. Maintenance doses can be continued indefinitely in patients in remission who show no signs of intolerance.

Unfortunately, clinical response is the only satisfactory guide to dosimetry and duration of treatment. Relapses may occur when maintenance doses are discontinued and there is a suspicion, unsupported by formal evidence, that the drug may be less effective when reinstituted. Assays of drug concentrations in plasma or tissues are time-consuming and of no practical value.

Auranofin is given in a daily dose of 0.1 to 0.15 mg/kg of body weight and has been continued for up to 2 years in patients in whom it is effective and well tolerated.

Adverse effects (Table 16)

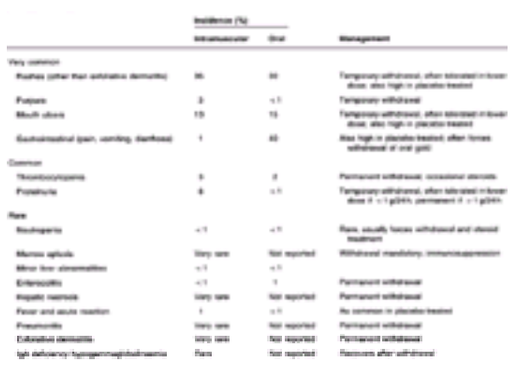


Table 16 Adverse reactions to gold therapy

Toxicity curtails sodium aurothiomalate treatment in a high percentage of patients, amounting to 30 to 45 per cent over 5 years in most reported series (Felson *et al*. 1990; Wolfe *et al*. 1990; Singh *et al*. 1991). The practical issues are to predict adverse reactions at an early stage, to decide which reactions enforce drug withdrawal, and to treat these when they occur. The theoretical problem is to elucidate the mechanisms which induce toxicity, with the possible practical benefit that patients at risk might be identified.

Over half the patients in placebo-controlled trials of auranofin develop gastrointestinal intolerance and about one-third develop rashes or stomatitis; these problems force drug withdrawal in about half the treated patients in the first 2 years of treatment (Williams *et al*. 1988; Johnsen *et al*. 1989) with a more variable cumulative drop out thereafter (Felson *et al*. 1990; Wolfe *et al*. 1990; Singh *et al*. 1991). However, many rashes which develop in patients on long-term treatment with gold are attributable to other causes and the drug is often withdrawn unnecessarily (Wilkinson *et al*. 1993).

Regular weekly clinical monitoring, blood counts, and urinalysis allow most reactions to be recognized at an early, reversible stage. Since hypersensitivity to gold is not dose-dependent, adverse reactions occur at all stages of treatment. Rashes, mouth ulcers, and proteinuria not exceeding 1.0 g/24 h necessitate temporary drug withdrawal, but treatment with lower weekly doses can often be attempted after some weeks. Bone marrow depression excludes further treatment with this agent. Marrow aplasia is rare and is the complication most likely to demand vigorous immunosuppressive treatment.

It is probable that gold toxicity is mediated by immune mechanisms. Animal experiments suggest that oxidized gold salts sensitize T lymphocytes, thereby provoking a cytotoxic reaction against gold–protein complexes (Schuhmann *et al*. 1990). Furthermore, T cells from patients with gold-induced dermatitis are specifically reactive with gold salts in proliferation assays (Verwilghen *et al*. 1992). Several studies implicate the inheritance of certain HLA coding genes in susceptibility to gold toxicity. The initial studies identified *HLA DR3* as the gene conferring an increased risk of proteinuria and thrombocytopenia; later reports suggest that this susceptibility is associated with the inheritance of an extended haplotype including the *DQA2.1* and *DQB2.1* genes on chromosome 6 (Singal *et al*. 1990; Sakkas *et al*. 1993). In contrast, the incidence of aphthous stomatitis and rashes during chrysotherapy is increased in patients inheriting the HLA-B35 antigen (Tishler *et al*. 1988b) and of enterocolitis in those with the HLA DRB1*0404 allele (Evron *et al*. 1995). These results suggest that susceptibility to gold toxicity is determined by interactions between T lymphocytes and immunogenic gold–peptide complexes presented by class II molecules. There is also a challenging contention that gold is more likely to induce remissions in patients who develop rashes (Caspi *et al*. 1989).

Thalidomide

Thalidomide (Fig. 4) has been used successfully in the treatment of lepromatous leprosy. By analogy it has also been used to treat many inflammatory rheumatic diseases in which immune complexes are thought to contribute to the immunopathological process. However, the disastrous teratogenic effects of thalidomide, which continue to be a problem in its unregulated use, and the neuropathy induced by high drug dosage deterred most rheumatologists from using this drug. However, its impressive ability to suppress orogenital ulceration in patients with Behçet's syndrome (Jenkins *et al*. 1984) led to renewed interest in the drug's therapeutic potential. Thalidomide is now regularly used in this context and also to treat similar symptoms in patients with AIDS. This interest was increased by the discovery that thalidomide is a potent inhibitor of the production of tumour necrosis factor-a by human monocytes (Sampaio *et al*. 1991). Subsequent studies have shown that it interferes with transcription of the gene encoding this cytokine.

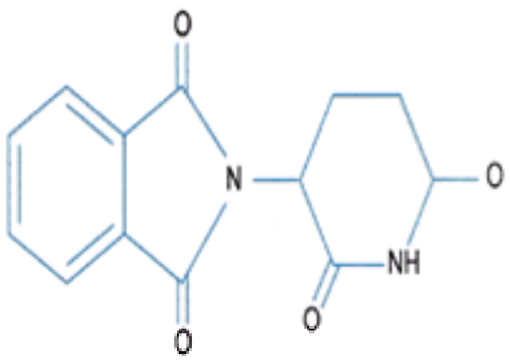


Fig. 4 Structure of thalidomide

Thalidomide may have wider immunosuppressive effects including the suppression of graft-versus-host disease in animal models and clinical practice (Heney *et al*. 1991). It now seems likely that the drug has immunoregulatory actions including the inhibition of interferon-g production by T lymphocytes because of its preferential stimulation of T_{H2} cells (McHugh *et al*. 1995).

The place of thalidomide in the treatment of other rheumatic diseases such as rheumatoid arthritis has not yet been established. When it is used, it is essential to ensure that the recipients are not at risk of pregnancy and that the drug is not taken inadvertently by pregnant women. Its use has to be carefully monitored because of its accumulation in target tissues and its slow removal from these tissues. The dose should be the minimum needed to control symptoms and should not ordinarily exceed 100 mg daily. Recurrent mouth ulcers can usually be controlled at a dose of 50 mg on alternate days.

Since thalidomide accumulates in target tissues and is slowly removed, it is essential to reduce toxicity by using low individual and cumulative doses. Drowsiness is encountered in 5 per cent of patients and its effects can be minimized by taking the drug at night. Neuropathy rarely develops in the doses employed in the treatment of rheumatic diseases. For practical purposes clinical monitoring suffices to prevent this complication, but some clinicians believe that baseline studies of nerve conduction are mandatory so that subsequent problems can be interpreted properly.

Interest in thalidomide analogues is likely to increase with the development of accurate assays for measuring plasma concentrations (Boughton *et al*. 1995) and better insight into the mechanism of thalidomide's teratogenic effects (Koch and Czejka 1986).

Antirheumatic drugs in pregnancy and lactation

There is justifiable concern about the safety of prescribing antirheumatic drugs during pregnancy and lactation and this subject is only now receiving detailed attention (Chapter 1.3.1.1 and Chapter 1.3.1.2). Most of the available data come from two sources, clinical observations and animal experiments. The interpretation of drug-induced teratogenesis in animals is complicated by species differences and dose effects. The metabolism and pharmacokinetics of cytotoxic drugs, for example, are not uniform in different species. There is also great variation in the efficiency with which potentially mutagenic DNA damage is repaired.

Until recently the clinical information was mainly anecdotal and did not evaluate other factors which may have been responsible for fetal abnormalities. However, controlled studies have now been published which take account of drug treatment, disease features, and other factors. Furthermore, there is increasing information about the genotoxic effects of antirheumatic drugs in children born to parents receiving these agents.

The available information is summarized in Table 17. It is now possible to give practical guidelines in different clinical situations. Theoretical concern over possible drug toxicity to the fetus in an established pregnancy should hardly ever be the sole grounds for termination. Techniques for monitoring fetal development provide additional safeguards. As far as possible antirheumatic drugs should be avoided or maintenance doses should be kept to the minimum during pregnancy; for example, a deterioration in lupus nephritis is a proven threat to fetal viability.

Table 17 Antirheumatic drugs in pregnancy and lactation

It is advisable to avoid starting or continuing treatment with potentially embryotoxic drugs in patients intending to become pregnant, even if in reality this decision is based on suspicion rather than firm evidence. Methotrexate and alkylating agents continue to arouse the most concern. Equally, it is important to control disease activity, particularly in systemic disorders, which pose known risks to the fetus. Higher corticosteroid doses are often the best interim solution.

Decisions about breast feeding should always be a compromise between maternal preference and medical prudence. Some drugs, notably alkylating agents, methotrexate, azathioprine, penicillamine, and maintenance corticosteroid doses above 10 mg per day contraindicate breast feeding; unavoidable treatment with other drugs is a possible reason for eschewing breast feeding but it should not be the sole grounds for influencing maternal choice.

Conclusions

There are still concerns about the long-term efficacy and safety of standard antirheumatic drugs. However, many agents have been shown to suppress inflammatory synovitis and systemic autoimmune diseases at least in the first few years of the disorder. There is also some evidence that some conventional drugs reduce joint destruction in rheumatoid arthritis. It seems likely that methotrexate, low-dose prednisolone, and cyclosporin are more effective and better tolerated than many time-honoured agents such as intramuscular gold and penicillamine. Increasing recourse to early treatment and drug combination has increased the need for accurate monitoring of drug efficacy and tolerance. Patient compliance is also an important consideration. There is still uncertainty about the mechanisms by which these agents exert their anti-inflammatory effects. One issue which will attract increasing attention in the coming years is the relative efficacy of conventional drug treatment and theoretically more selective immunosuppression with mainly biological agents. However, increasingly sophisticated analysis of the efficacy and problems of antirheumatic drugs has at least enabled rheumatologists to prescribe these potentially toxic agents with better understanding.

Chapter References

Al Balla, S., Johnston, C., and Davis, P. (1990). The *in vivo* effect of non-steroidal anti-inflammatory drugs, gold sodium thiomalate and methotrexate on neutrophil superoxide radical generation. *Clinical and Experimental Rheumatology*, **8**, 41–5.

Alarcon, G.S., Tracy, I.C., and Blackburn, W.D., Jr. (1989). Methotrexate in rheumatoid arthritis: toxic effects as the major factor in limiting long-term treatment. *Arthritis and Rheumatism*, **32**, 671–6.

Alarcon, G.S., Schrohenloher, R.E., Bartolucci, A.A., Ward, J.R., Williams, H.J., and Koopman, W.J. (1990). Suppression of rheumatoid factor production by methotrexate in patients with rheumatoid arthritis: evidence for differential influences of therapy and clinical status on IgM and IgA rheumatoid factor expression. *Arthritis and Rheumatism*, **33**, 1156–61.

Alstead, E.M., Ritchie, J.K., Lennard-Jones, J.E., Farthing, M.J., and Clark, M.L. (1990). Safety of azathioprine in pregnancy in inflammatory bowel disease. *Gastroenterology*, **99**, 443–6.

Antonelli, M.A., Moreland, L.W., and Brick, J.E. (1991). Herpes zoster in patients with rheumatoid arthritis treated with weekly, low-dose methotrexate. *American Journal of Medicine*, **90**, 295–8.

Axford, J.S. *et al*. (1992). Changes in normal glycosylation mechanisms in autoimmune rheumatic disease. *Journal of Clinical Investigation*, **89**, 1021–31.

Balsa, A., Gamallo, C., Martin-Mola, E., and Gijon-Banos, J. (1993). Histologic changes in rheumatoid synovitis induced by naproxen and methotrexate. *Journal of Rheumatology*, **20**, 1472–7.

Bannwarth, B., Labat, L., Moride, Y., and Schaeverbeke, T. (1994). Methotrexate in rheumatoid arthritis. An update. *Drugs*, **47**, 25–50.

Barrera, P., Boerbooms, A.M., Demacker, P.N., van de Putte, L.B., Gallati, H. and van der Meer, J.W. (1994). Circulating concentrations and production of cytokines and soluble receptors in rheumatoid arthritis patients: effects of a single dose methotrexate. *British Journal of Rheumatology*, **33**, 1017–24.

Bauer-Vinassac, D., Menkes, C.J., Muller, J.Y., and Escande, J.P. (1992). HLA system and penicillamine induced pemphigus in nine cases of rheumatoid arthritis. *Scandinavian Journal of Rheumatology*, **21**, 17–19.

Bennett, P.N., Humphries, S.J., Osborne, J.P., Clarke, A.K., and Taylor, A. (1990). Use of sodium aurothiomalate during lactation. *British Journal of Clinical Pharmacology*, **29**, 777–9.

Boers, M., Dijkmans, B.A., van Rijthoven, A.W., Goei-The, H.S., and Cats, A. (1990). Reversible nephrotoxicity of cyclosporine in rheumatoid arthritis. *Journal of Rheumatology*, **17**, 38–42.

Boughton, B.J. *et al*. (1995). High-performance liquid chromatographic assay of plasma thalidomide: stabilisation of specimens and determination of a tentative therapeutic range for chronic graft-versus-host disease. *Annals of Clinical Biochemistry*, **32**, 79–83.

Bourget, P., Fernandez, H., and Delouis, C. (1991). Accumulation of cyclosporine in the conceptus during the first trimester of pregnancy after liver transplantation. *Transplantation*, **51**, 1306–7.

Bradley, J.D., Brandt, K.D., and Katz, B.P. (1989). Infectious complications of cyclophosphamide treatment for vasculitis. *Arthritis and Rheumatism*, **32**, 45–53.

Braun, W., Kallen, J., Mikol, V., Walkinshaw, M.D., and Wuthrich, K. (1995). Three-dimensional structure and actions of immunosuppressants and their immunophilins. *FASEB Journal*, **9**, 63–72.

Brooks, P.J., Spruill, W.J., Parish, R.C., and Birchmore, D.A. (1990). Pharmacokinetics of methotrexate administered by intramuscular and subcutaneous injections in patients with rheumatoid arthritis. *Arthritis and Rheumatism*, **33**, 91–4.

Brown-Galatola, C.H. and Hall, N.D. (1992). Impaired suppressor cell activity due to surface sulphydryl oxidation in rheumatoid arthritis. *British Journal of Rheumatology*, **31**, 599–603.

Buchanan, N.M., Khamashta, M.A., Morton, M.E., Kerslake, S., Baguley, E.A. and Hughes, G.R. (1992). A study of 100 high risk lupus pregnancies. *American Journal of Reproductive Immunology*, **28**, 192–4.

Buckley, L.M., Vacek, P.M., and Cooper, S.M. (1990). Administration of folinic acid after low dose methotrexate in patients with rheumatoid arthritis. *Journal of Rheumatology*, **17**, 1158–61.

Caldwell, J.R. and Furst, D.E. (1991). The efficacy and safety of low-dose corticosteroids for rheumatoid arthritis. *Seminars in Arthritis and Rheumatism*, **21**, 1–11.

Capell, H.A., Marabani, M., Madhok, R., Torley, H., and Hunter, J.A. (1990). Degree and extent of response to sulphasalazine or penicillamine therapy for rheumatoid arthritis: results from a routine clinical environment over a two-year period. *Quarterly Journal of Medicine*, **75**, 335–44.

Carmichael, A.J. (1992). Hydroxychloroquine. A guide to usage. *Journal of Dermatological Treatment*, **3**, 103–6.

Carroll, G.J., Will, R.K., Breidahl, P.D., and Tinsley, L.M. (1989). Sulphasalazine versus penicillamine in the treatment of rheumatoid arthritis. *Rheumatology International*, **8**, 251–5.

Caspi, D., Tishler, M., and Yaron, M. (1989). Association betwen gold induced skin rash and remission in patients with rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **48**, 730–2.

Chalmers, I.M., Sitar, D.S., and Hunter, T. (1990). A one-year, open, prospective study of sulfasalazine in the treatment of rheumatoid arthritis: adverse reactions and clinical response in relation to laboratory variables, drug and metabolite serum levels, and acetylator status. *Journal of Rheumatology*, **17**, 764–70.

Champion, G.D., Graham, G.G., and Ziegler, J.B. (1990). The gold complexes. In *Slow-acting antirheumatic drugs and immunosuppressives* (ed. P.M. Brooks), *Baillière's Clinical Rheumatology*, Vol. 4, No. 3, pp. 491–534. Baillière Tindall, London.

Clarkson, R.W., Sanders, P.A., and Grennan, D.M. (1992). Complement C4 null alleles as a marker of gold or D-Penicillamine toxicity in the treatment of rheumatoid arthritis. *British Journal of Rheumatology*, **31**, 53–4.

Classen, J.B. and Shevach, E.M. (1991). Evidence that cyclosporine treatment during pregnancy predisposes offspring to develop autoantibodies. *Transplantation*, **51**, 1052–7.

Clegg, D.O., Furst, D.E., Tolman, K.G., and Pogue, R. (1989). Acute, reversible hepatic failure associated with methotrexate treatment of rheumatoid arthritis. *Journal of Rheumatology*, **16**, 1123–6.

Cockburn, I., Krupp, P., and Monka, C. (1989). Present experience of Sandimmun in pregnancy. *Transplantation Proceedings*, **21**, 3730.

Corkill, M.M., Kirkham, B.W., Haskard, D.O., Barbatis, C., Gibson, T., and Panayi, G.S. (1991). Gold treatment of rheumatoid arthritis decreases synovial expression of the endothelial adhesion receptor ELAM-1. *Journal of Rheumatology*, **18**, 1453–60.

Crilly, A., Madhok, R., Watson, J., and Capell, H.A. (1993). Serum concentrations of soluble interleukin 2 receptor in patients with rheumatoid arthritis: effect of second line drugs. *Annals of the Rheumatic Diseases*, **52**, 58–60.

Cseuz, R. and Panayi, G.S. (1990). The inhibition of NK cell function by azathioprine during the treatment of patients with rheumatoid arthritis. *British Journal of Rheumatology*, **29**, 358–62.

Danis, V.A., Kulesz, A.J., Nelson, D.S., and Brooks, P.M. (1990*a*). The effect of gold sodium thiomalate and auranofin on lipopolysaccharide-induced interleukin-1 production by blood monocytes *in vitro*: variation in healthy subjects and patients with arthritis. *Clinical and Experimental Immunology*, **79**, 335–40.

Danis, V.A., Kulesz, A.J., Kelly, D.E., Nelson, D.S., and Brooks, P.M. (1990 *b*). The effect of gold treatment on monocyte interleukin-l production in rheumatoid arthritis: a prospective study. *Rheumatology International*, **10**, 153–8.

Denman, A.M., Denman, D.J., and Palmer, R.G. (1992). Alkylating agents. In *Immunopharmacology in autoimmune diseases and transplantation* (ed. H.E. Rugstad, L. Endresen, and O. Forre), pp. 139–58. Plenum Press, New York.

De Sesso, J.M. and Goeringer, G.C. (1991). Amelioration by leucovorin of methotrexate developmental toxicity in rabbits. *Teratology*, **43**, 201–15.

De Silva, M. and Hazleman, B. (1981). Long-term azathioprine in rheumatoid arthritis: a double-blind study. *Annals of the Rheumatic Diseases*, **40**, 560–3.

De Silva, R.N. and Eastmond, C.J. (1992). Management of proteinuria secondary to penicillamine therapy in rheumatoid arthritis. *Clinical Rheumatology*, **11**, 216–19.

DiFrancesco, L., Miller, F., and Greenwald, R.A. (1994). Detailed immunohistologic evaluation of a methotrexate-induced nodule. *Archives of Pathology and Laboratory Medicine*, **118**, 1223–5.

Dijkmans, B.A., van Rijthoven, A.W., Goei The, H.S., Boers, M., and Cats, A. (1993). Cyyclosporine in rheumatoid arthritis. *Seminars in Arthritis and Rheumatism*, **22**, 30–6.

Doolittle, G.C., Simpson, K.M., and Lindsley, H.B. (1989). Methotrexate-associated, early-onset pancyopenia in rheumatoid arthritis. *Archives of Internal Medicine*, **149**, 1430–1.

Doris, A., Di Lenardo, L., Vario, S., Calligaro, A., Vaccaro, E., and Gambari, P.F. (1992). Cyclosporin A in a pregnant patient with systemic lupus erythematosus. *Rheumatology International*, **12**, 77–8.

Dougados, M. *et al*. (1995). Sulfasalazine in the treatment of spondyloarthropathy. A randomized, muticenter, double-blind, placebo-controlled study. *Arthritis and Rheumatism*, **38**, 618–27.

Doyle, D.V., Perrett, D., Foster, D.J., Ensor, M., and Scott, D.L. (1993). The long-term use of D-Penicillamine for treating rheumatoid arthritis: is continuous therapy necessary? *British Journal of*

Rheumatology, **32**, 614–17.

Epstein, W.V., Henke, C.J., Yelin, E.H., and Katz, P.P. (1991). Effect of parenterally administered gold therapy on the course of adult rheumatoid arthritis. *Annals of Internal Medicine*, **114**, 437–44.

Euler, H.E., Schroeder, J.O., Harten, P., Zeuner, R.A., and Gutschmidt, H.J. (1994). Treatment-free remission in severe systemic lupus erythematosus following synchronisation of plasmapheresis with subsequent pulse cyclophosphamide. *Arthritis and Rheumatism*, **37**, 1784–94.

Evron, E. *et al*. (1995). Correlation between gold-induced enterocolitis and the presence of the HLA-DRB1*0404 allele. *Arthritis and Rheumatism*, **38**, 755–9.

Farr, M., Tunn, E.J., Symmons, D.P., Scott, D.G., and Bacon, P.A. (1989). Sulphasalazine in rheumatoid arthritis: haematological problems and changes in haematological indices associated with therapy. *British Journal of Rheumatology*, **28**, 134–8.

Farr, M., Kitas, G.D., Tunn, E.J., and Bacon, P.A. (1991). Immunodeficiencies associated with sulphasalazine therapy in inflammatory arthritis. *British Journal of Rheumatology*, **30**, 413–17.

Felson, D.T., Anderson, J.T., and Meenan, R.F. (1990). The comparative efficacy and toxicity of second-line drugs in rheumatoid arthritis: results of two metaanalyses. *Arthritis and Rheumatism*, **33**, 1449–61.

Ferraccioli, G.F. *et al*. (1995). Epstein–Barr virus-associated Hodgkin's lymphoma in a rheumatoid arthritis patient treated with methotrexate and cyclosporin A. *Arthritis and Rheumatism*, **38**, 867–8.

Firestein, G.S., Paine, M.M., and Boyle, D.L. (1994). Mechanisms of methotrexate action in rheumatoid arthritis. Selective decrease in synovial collagenase gene expression. *Arthritis and Rheumatism*, **37**, 193–200.

Fishel, B., Tishler, M., Caspi, D., and Yaron, M. (1989). Fatal aplastic anaemia and liver toxicity caused by D-Penicillamine treatment of rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **48**, 609–10.

Francis, B.M. *et al*. (1990). Cyclophosphamide teratogenesis: evidence for compensatory responses to induced cellular toxicity. *Teratology*, **42**, 473–82.

Fries, J.F., Williams, C.A., Ramey, D., and Bloch, D.A. (1993). The relative toxicity of disease-modifying antirheumatic drugs. *Arthritis and Rheumatism*, **36**, 297–305.

Furst, D.E. (1994). Should methotrexate be used to treat early rheumatoid arthritis? *Seminars in Arthritis and Rheumatism*, **23** (Suppl. 2), 39–43.

Furst, D.E., Erikson, N., Clute, L., Koehnke, R., Burmeister, L.F., and Kohler, J.A. (1990). Adverse experience with methotrexate during 176 weeks of a longterm prospective trial in patients with rheumatoid arthritis. *Journal of Rheumatology*, **17**, 1628–35.

Gabriel, S., Creagan, E., O'Fallon, W.M., Jaquith, J., and Bunch, T. (1990). Treatment of rheumatoid arthritis with higher dose intravenous methotrexate. *Journal of Rheumatology*, **17**, 460–5.

Garloch, J.J., Todd, P., Thomson, J., and Kinane, D.F. (1990). Cyclosporin and gingival hyperplasia. *Clinical and Experimental Dermatology*, **15**, 466–8.

George, E. and Kirwan, J.R. (1990). Corticosteroid therapy in rheumatoid arthritis. In *Slow-acting antirheumatic drugs and immunosuppressives* (ed. P.M. Brooks), *Baillière's Clinical Rheumatology*, Vol. 4, No. 3, pp. 621–47. Baillière Tindall, London.

Girgis, L., Conaghan, P.G., and Brooks, P. (1994). Disease-modifying anti-rheumatic agents, including methotrexate, sulfasalazine, gold, antimalarials and penicillamine. *Current Opinion in Rheumatology*, **6**, 252–61.

Gran, J.T. and Myklebust, G. (1993). Toxicity of sulphasalazine in rheumatoid arthritis. Possible protective effect of rheumatoid factors and corticosteroids. *Scandinavian Journal of Rheumatology*, **22**, 229–32.

Gray, R.E., Foherty, S.M., Galloway, J., Coulton, L., de Broe, M., and Kanis, J.A. (1991). A double-blind study of deflazacort and prednisone in patients with chronic inflammatory disorders. *Arthritis and Rheumatism*, **34**, 287–95.

Haagsma, C.J. *et al*. (1994). Combination of methotrexate and sulphasalazine vs methotrexate alone: a randomized open clinical trial of rheumatoid arthritis patients resistant to sulphasalazine. *British Journal of Rheumatology*, **33**, 1049–55.

Haberman, J. *et al*. (1991). Male fertility in cyclosporine-treated renal transplant patients. *Journal of Urology*, **145**, 294–6.

Hall, C.L. and Tighe, R. (1989). The effect of continuing penicillamine and old treatment on the course of penicillamine and gold nephropathy. *British Journal of Rheumatology*, **28**, 53–7.

Hansen, T.M. *et al*. (1990). Double blind placebo controlled trial of pulse treatment with methylprednisolone combined with disease modifying drugs in rheumatoid arthritis. *British Medical Journal*, **301**, 268–70.

Hartard, C. and Kunze, K. (1994). Pregnancy in a patient with Wilson's disease treated with D-Penicillamine and zinc sulphate. A case report and review of the literature. *European Neurology*, **34**, 337–40.

Heney, D., Norfolk, D.R., Wheeldon, J., Bailey, C.C., Lewis, I.J., and Barnard, D.L. (1991). Thalidomide treatment for chronic graft-versus-host disease. *British Journal of Haematology*, **78**, 23–7.

HERA Study Group (1995). A randomised trial of hydroxychloroquine in early rheumatoid arthritis: the HERA study. *American Journal of Medicine*, **98**, 156–68.

Hodis, H.N., Quismorio, F.P., Wickham, E., and Blankenhorn, D.H. (1993). The lipid, lipoprotein, and apolipoprotein effects of hydroxychloroquine in patients with systemic lupus erythematosus. *Journal of Rheumatology*, **20**, 661–5.

Hoo, J.J., Hadro, T.A., and Von Behren, P. (1988). Possible teratogenicity of sulfasalazine. *New England Journal of Medicine*, **318**, 1128.

Imai, F., Susuki, T., Ishibashi, T., Tanaka, M., Akiyama, Y., and Dohi, Y. (1994). Effect of sulfasalazine on B cell hyperactivity in patients with rheumatoid arthritis. *Journal of Rheumatology*, **21**, 612–15.

Jenkins, J.S., Powell, R.J., Allen, B.R., Littlewood, S.M., Maurice, P.D.L., and Smith, N.J. (1984). Thalidomide in severe orogenital ulceration. *Lancet*, **ii**, 1424–6.

Jeurissen, M.E., Boerbooms, A.M., and van de Putte, L.B. (1988). Pancytopenia related to azathioprine in rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **47**, 503–5.

Jeurissen, M.E., Boerbooms, A.M., van de Putte, L.B., and Kruijsen, M.W. (1990). Azathioprine induced fever, chills, rash, and hepatotoxicity in rheumatic arthritis. *Annals of the Rheumatic Diseases*, **49**, 25–7.

Jeurissen, M.E.C., Boerbooms, A.M., van de Putte, L.B.A., Doesburg, W.H., and Lemmens, A.M. (1991). Influence of methotrexate and azathioprine on radiologic progression in rheumatoid arthritis. *Annals of Internal Medicine*, **114**, 999–1004.

Johnsen, V., Borg, G., Trang, L.E., Berg, E., and Brodin, U. (1989). Auranofin (SKandF) in early rheumatoid arthritis: results from a 24-month double-blind, placebo-controlled study. Effects on clinical and biochemical assessments. *Scandinavian Journal of Rheumatology*, **18**, 251–60.

Jorgenson, C., Bologna, C., Anaya, J.M., Reme, T., and Sany, J. (1993). Variation in the serum IgA concentration and the production of IgA *in vitro* in rheumatoid arthritis treated by sulfasalazine. *Rheumatology International*, **13**, 113–16.

Joyce, D.A. (1990). D-Penicillamine. In *Slow-acting antirheumatic drugs and immunosuppressives* (ed. P. M, Brooks), *Baillière's Clinical Rheumatology*, Vol. 4, No. 3, pp. 553–74. Baillière Tindall, London.

Kanerud, L., Scheynius, A., and Hafstrom, I. (1994). Evidence of a local immunomodulatory effect of sulfasalazine in rheumatoid arthritis. *Arthritis and Rheumatism*, **37**, 1138–45.

Kinlen, L.J. (1985). Immunosuppressive therapy and cancer. *Cancer Surveys*, **1**, 565–83.

Kirkham, B.W., Davison, S.C., Corkill, M.M., Barbatis, C., and Panayi, G.S. (1993). Serial soluble interleukin-2 receptor levels in rheumatoid arthritis: differences in response to glucocorticoid treatment and chrysotherapy. *Journal of Rheumatology*, **20**, 935–9.

Kirshon, B., Wasserstrum, N., Willis, R., German, G.E., and McCabe, E.R. (1988). Teratogenic effects of first-trimester cyclophosphamide therapy. *Obstetrics and Gynecology*, **72**, 462–4.

Kirwan, J.R. and the Arthritis and Rheumatism Council Low-Dose Glucocorticoid Study Group (1995). The effects of glucocorticoids on joint destruction in rheumatoid arthritis. *New England Journal of Medicine*, **333**, 142–6.

Koch, H.P. and Czejka, M.J. (1986). Evidence for the intercalation of thalidomide into DNA: clue to the molecular mechanism of thalidomide teratogenicity? *Zeitschrift für Naturforschung*, **41C**,

1057–61.

Kotowicz, M.A., Hall, S., Hunder, G.G., Cedel, S.L., Mann, K.G., and Riggs, B.L. (1990). Relationship of glucocorticoid dosage to serum bone Gla-protein concentration in patients with rheumatologic disorders. *Arthritis and Rhematism*, **33**, 1487–92.

Kozlowski, R.D., Steinbrunner, J.V., MacKenzie, A.H., Clough, J.D., Wilke, W.S., and Segal, A.M. (1990). Outcome of first-trimester exposure to low-dose methotrexate in eight patients with rheumatic disease. *American Journal of Medicine*, **88**, 589–92.

Kremer, J.L. *et al*. (1994). Methotrexate for rheumatoid arthritis. Suggested guidelines for monitoring liver toxicity. *Arthritis and Rheumatism*, **37**, 316–28.

Kremer, J.L. *et al*. (1995). Every-other-week methotrexate in patients with rheumatoid arthritis. *Arthritis and Rheumatism*, **38**, 601–7.

Lacki, J.K. *et al*. (1994). Immunological studies in patients with rheumatoid arthritis treated with methotrexate or cyclophosphamide. *Zeitchrift für Rheumatologie*, **53**, 76–82.

Landewe, R.B.M., Goei The, H.S., van Rijthoven, A.W.A.M., Breedveld, F.C., and Dijkmans, B.A.C. (1994). A randomized, double-blind, 24-week controlled study of low-dose cyclosporine versus chloroquine for early rheumatoid arthritis. *Arthritis and Rheumatism*, **37**, 637–43.

Laskow, D.A. *et al*. (1990). Cyclosporine-induced changes in glomerular filtration rate and urea excretion. *American Journal of Medicine*, **88**, 497–502.

Lawley, P.D. (1990). Basic concepts of carcinogenesis. In *The metabolic and molecular basis of acquired disease* (ed. R.D. Cohen, B. Lewis, K.G.M.M. Alberti, and A.M. Denman), pp. 44–73. Baillière Tindall, London.

Le Gallez, P. (1988). Teratogenesis and drugs for rheumatic disease. *Nursing Times*, **84**, 41–3.

Levy, M., Buskila, D., Gladman, D.D., Urowitz., M.B., and Koren, G. (1991). Pregnancy outcome following first trimester exposure to chloroquine. *American Journal of Perinatology*, **8**, 174–8.

Lin, H.Y., Rocher, L.L., McQuillan, M.A., Schmaltz, S., Palella, T.D., and Fox, I.H. (1989). Cyclosporine-induced hyperuricemia and gout. *New England Journal of Medicine*, **321**, 287–92.

Loetscher, P., Dewald, B., Baggiolini, M., and Seitz, M. (1994). Monocyte chemoattractant protein 1 and interleukin 8 production by rheumatoid synoviocytes. Effects of anti-rheumatic drugs. *Cytokine*, **6**, 162–70.

Madhok, R., Zoma, A., Torley, H.I., Capell, H.A., Waring, R., and Hunter, J.A. (1990). The relationship of sulfoxidation status to efficacy and toxicity of penicillamine in the treatment of rheumatoid arthritis. *Arthritis and Rheumatism*, **33**, 574–7.

Marabani, M., Madhok, R., Capell, H.A., and Hunter, J.A. (1989). Leucopenia during sulphasalazine treatment for rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **48**, 505–7.

Massa, M., de Benedetti, F., Robbioni, P., Ramenohi, B., Albani, S., and Martini, A. (1993). Association of methotrexate treatment with a decrease of double-negative (CD4-CD9-) and g/ T cell levels in patients with juvenile rheumatoid arthritis. *Journal of Rheumatology*, **20**, 1944–8.

McHugh, S.M. *et al*. (1995). The immunosuppressive drug thalidomide induces T helper cell type 2 (Th2) and concomitantly inhibits Th1 cytokine production in mitogen- and antigen-stimulated human blood peripheral mononuclear cell cultures. *Clinical and Experimental Immunology*, **99**, 160–7.

Michaels, R.M., Chang, Z.L., and Beezhold, D.H. (1994). Phospholipase A_2 activity in peripheral blood cells of rheumatoid arthritis patients treated with methotrexate: a preliminary study. *Clinical and Experimental Rheumatology*, **12**, 643–7.

Middleton, G.D., McFarlin, J.E., and Lipsky, P.E. (1995). Hydroxychloroquine and pain thresholds. *Arthritis and Rheumatism*, **38**, 445–7.

Moore, S., Ruska, K., Peters, L., and Olsen, N.J. (1994). Associations of IgA and IgA rheumatoid factor with disease features in patients with rheumatoid arthritis. *Immunological Investigations*, **23**, 355–65.

Morell, E. *et al*. (1991). Autoantibodies in D-Penicillamine-induced myasthenia gravis: a comparison with idiopathic myasthenia and rheumatoid arthritis. *Clinical Immunology and Immunopathology*, **58**, 318–30.

Morgan, S.L. *et al*. (1990). The effect of folic acid supplementation on the toxicity of low-dose methotrexate in patients with rheumatic arthritis. *Arthitis and Rheumatism*, **33**, 9–18.

Nast, C.C., Adler, S.G., Artishevsky, A., Kresser, C.T., Ahmed, K., and Anderson, P.S. (1991). Cyclosporine induces elevated procollagen a $_1$ (I) mRNA levels in the rat renal cortex. *Kidney International*, **39**, 631–8.

Needs, C.J. and Brooks, P.M. (1985). Antiarthritic medication in pregnancy. *British Journal of Rheumatology*, **24**, 282–90.

Nisar, M., Carlisle, L., and Amos, R.S. (1994). Methotrexate and sulphasalazine as combination therapy in rheumatoid arthritis. *British Journal of Rheumatology*, **33**, 651–4.

Nuver-Zwart, I.H., van Riel, P.L., van de Putte, L.B., and Gribnau, F.W. (1989). A double blind comparative study of sulphasalazine and hydroxychloroquine in rheumatoid arthritis: evidence of an earlier effect of sulphasalazine. *Annals of the Rheumatic Diseases*, **48**, 389–95.

Olshan, A.F., Mattison, D.R., and Zwanenburg, T.S. (1994). International commission for protection against environmental mutagens and carcinogens. Cyclosporine A: review of genotoxicity and potential for adverse human reproductive and developmental effects. Report of a working group on the genotoxicity of cyclosporine A. *Mutation Research*, **317**, 163–73.

Ostensen, M. (1994). Optimisation of anti-rheumatic drug treatment in pregnancy. *Clinical Pharmacokinetics*, **27**, 486–503.

Ostensen, M. and Husby, G. (1985). Antirheumatic drug treatment during pregnancy and lactation. *Scandinavian Journal of Rheumatology*, **14**, 1–7.

Parke, A.L. (1988). Antimalarial drugs, systemic lupus erythematosus and pregnancy. *Journal of Rheumatology*, **15**, 607–10.

Patapanian, H. *et al*. (1988). The oncogenicity of chlorambucil in rheumatoid arthritis. *British Journal of Rheumatology*, **27**, 44–7.

Pavelka, K., Jr., Sen, K.P., Peliskova, A., Vacha, J., and Trnavsky, K. (1989). Hydroxychloroquine sulphate in the treatment of rheumatoid arthritis: a double blind comparison of two dose regimens. *Annals of the Rheumatic Diseases*, **48**, 542–6.

Pedersen, P.B. and Hogenhaven, H. (1990). Penicillamine-induced neuropathy in rheumatoid arthritis. *Acta Neurologica Scandinavica*, **81**, 188–90.

Perhala, R.S., Wilke, W.S., Clough, J.D., and Segal, A.M. (1991). Local infectious complications following large joint replacement in rheumatoid arthritis patients treated with methotrexate versus those not treated with methotrexate. *Arthritis and Rheumatism*, **34**, 146–52.

Pilarski, L.M., Yacyshyn, B.R., and Lazarovits, A.I. (1994). Analysis of peripheral blood populations and immune function from children exposed to cyclosporine or to azathioprine *in utero*. *Transplantation*, **57**, 133–44.

Pincus, T., Marcum, S.B., and Callahan, L.F. (1992). Longterm drug treatment for rheumatoid arthritis in seven rheumatology private practices. II. Second-line drugs and prednisone. *Journal of Rheumatology*, **19**, 1885–94.

Porter, D.R. and Capell, H.A. (1990). The use of sulphasalazine as a disease-modifying antirheumatic drug. In *Slow-acting antirheumatic drugs and immunosuppressives* (ed. P. M. Brooks), *Baillière's Clinical Rheumatology*, Vol. 4, No. 3, pp. 535–51. Baillière Tindall, London.

Porter, D.R., McInnes, I., Hunter, J., and Capell, H.A. (1994). Outcome of second line therapy in rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **53**, 812–15.

Rheumatoid Arthritis Clinical Trial Archive Group (1995). The effect of age and renal function on the efficacy and toxicity of methotrexate in rheumatoid arthritis. *Journal of Rheumatology*, **22**, 218–23.

Roubenoff, R., Roubenoff, R.A., Ward, L.M., and Stevens, M.B. (1990). Catabolic effects of high-dose corticosteroids persist despite therapeutic benefit in rheumatoid arthritis. *American Journal of Clinical Nutrition*, **52**, 1113–17.

Rousseaux, C.G. and MacNabb, L.G. (1992). Oral administration of D-Penicillamine causes neonatal mortality without morphological defects in CD-1 mice. *Journal of Applied Toxicology*, **12**, 35–8.

Sakkas, L.I., Chikanza, I.C., Vaughan, R.W., Welsh, K.J., and Panayi, G.S. (1993). Gold induced nephropathy in rheumatoid arthritis and HLA class II genes. *Annals of the Rheumatic Diseases*, **52**, 300–1.

Sambrook, P.N., Eisman, J.A., Champion, G.D., and Pocock, N.A. (1988). Sex hormone status and osteoporosis in postmenopausal women with rheumatoid arthritis. *Arthritis and Rheumatism*, **31**,

973–8.

Sambrook, P.N., Cohen, M.L., Eisman, J.A., Pocock, N.A., Champion, G.D., and Yeates, M.G. (1989). Effects of low dose corticosteroids on bone mass in rheumatoid arthritis: a longitudinal study. *Annals of the Rheumatic Diseases*, **48**, 535–8.

Sampaio, E.P., Sarno, E.N., Galilly, R., Cohn, Z.A., and Kaplan, G. (1991). Thalidomide selectively inhibits tumor necrosis factor-a production by stimulated human monocytes. *Journal of Experimental Medicine*, **173**, 699–703.

Schaardenburg, D. van *et al*. (1995). Prednisone treatment of elderly-onset rheumatoid arthritis. *Arthritis and Rheumatism*, **38**, 334–42.

Scheinman, R.I., Cogswell, P.C., Lofquist, A.K., and Baldwin, A.S. (1995). Role of transcriptional activation of I_Ba in mediation of immunosuppression by glucocorticoids. *Science*, **270**, 283–90.

Schilsky, R.L., Lewis, B.J., Sherins, R.J., and Young, R.C. (1980). Gonadal dysfunction in patients receiving chemotherapy for cancer. *Annals of Internal Medicine*, **93**, 109–14.

Schuhmann, D., Kubicka-Muranyi, M., Mirtschewa, J., Gunther, J., Kind, P., and Gleichmann, E. (1990). Adverse immune reactions to gold. I. Chronic treatment with an Au(I) drug sensitizes not to Au(I) but to Au(III) and induces autoantibody formation. *Journal of Immunology*, **145**, 2132–9.

Schwartz, G.F. and Anderson, S.T. (1990). Methotrexate induced pneumonitis in a young woman with psoriasis and rheumatoid arthritis. *Journal of Rheumatology*, **17**, 980.

Scott, D.L., Greenwood, A., Davies, J., Maddison, P.J., Maddison, M.C., and Hall, N.D. (1990). Radiological progression in rheumatoid arthritis: do D-Penicillamine and hydroxychloroquine have different effects? *British Journal of Rheumatology*, **29**, 126–7.

Seideman, P. and Ros, A.M. (1992). Sensitivity to UV light during treatment with chloroquine in rheumatoid arthritis. *Scandinavian Journal of Rheumatology*, **21**, 245–7.

Seideman, P., Muller-Suur, R., and Ekman, E. (1993). Renal effects of low dose methotrexate in rheumatoid arthritis. *Journal of Rheumatology*, **20**, 1126–8.

Shiroky, J.B., Frost, A., Skelton, J.D., Haegert, D.G., Newkirk, M.M., and Neville, C. (1991). Complications of immunosuppression associated with weekly low dose methotrexate. *Journal of Rheumatology*, **18**, 1172–5.

Silman, A.J., Petrie, J., Hazleman, B., and Evans, S.J. (1988). Lymphoproliferative cancer and other malignancy in patients with rheumatoid arthritis treated with azathioprine: a 20 year follow up study. *Annals of the Rheumatic Diseases*, **47**, 988–92.

Singal, D.P., Reid, B., Green, D., D'Souza, M., Bensen, W.G., and Buchanan, W.W. (1990). Polymorphism of major histocompatibility complex extended haplotypes bearing HLA-DR3 in patients with rheumatoid arthritis with gold induced thrombocytopenia or proteinuria. *Annals of the Rheumatic Diseases*, **49**, 582–6.

Singh, G., Fries, J.F., Spitz, P., and Williams, C.A. (1989). Toxic effects of azathioprine: a national post-marketing perspective. *Arthritis and Rheumatism*, **32**, 837–43.

Singh, G., Fries, J.F., Williams, C.A., Zatarain, E., Spitz, P., and Bloch, D.A. (1991). Toxicity profiles of diseases modifying antirheumatic drugs in rheumatoid arthritis. *Journal of Rheumatology*, **18**, 188–94.

Skeith, K.J., Russell, A.S., Jamali, F., Coates, J., and Friedman, H. (1990). Lack of significant interaction between low dose methotrexate and ibuprofen or flurbiprofen in patients with arthritis. *Journal of Rheumatology*, **17**, 1008–10.

Smith, M.D., Smith, A., O'Donnell, J., Ahern, M.J., and Roberts-Thomson, P.J. (1989). Impaired delayed type cutaneous hypersensitivity in rheumatoid arthritis reversed bv chrysotherapy. *Annals of the Rheumatic Diseases*, **8**, 108–13.

Smith, M.D., Ahern, M,T., and Roberts-Thomson, P.J. (1990). Pulse methylprednisolone therapy in rheumatoid arthritis: unproved therapy, unjustified therapy, or effective adjunctive treatment? *Annals of the Rheumatic Diseases*, **49**, 265–7.

Sneller, M.C., Hoffman, G.S., Talar-Williams, C., Kerr, G.S., Hallahan, C.W., and Fauci, A.S. (1995). An analysis of forty-two Wegener's granulomatosis patients treated with methotrexate and prednisone. *Arthritis and Rheumatism*, **38**, 608–13.

Spadaro, A., Riccieri, V., Sili Scavalli, A., Taccari, E., and Zoppini, A. (1993). One year treatment with low dose methotrexate in rheumatoid arthritis: effect on class specific rheumatoid factors. *Clinical Rheumatology*, **12**, 357–60.

Sperling, R.I., Coblyn, J.S., Larkin, J.K., Benincaso, A.I., Austen, K.F., and Weinblatt, M.E. (1990). Inhibition of leukotriene B4 synthesis in neutrophils from patients with rheumatoid arthritis by a single oral dose of methotrexate. *Arthritis and Rheumatism*, **33**, 1149–55.

Stillwell, T.J., Benson, R.C., Jr., DeRemee, R.A., McDonald, T.J., and Weiland, L.H. (1988). Cyclophosphamide-induced bladder toxicity in Wegener's granulomatosis. *Arthritis and Rheumatism*, **31**, 465–70.

Svartz, N. (1942). Salazopyrin, a new sulphanilimide preparation. *Acta Medica Scandinavica*, **110**, 577–98.

Tenorio, O. and Orozco, J.A. (1993). Hydroxychloroquine compared with placebo in rheumatoid arthritis. A randomized controlled trial. *Annals of Internal Medicine*, **119**, 1067–71.

Tett, S., Cutler, D., and Day, R. (1990). Antimalarials in rheumatic diseases. In *Slow-acting antirheumatic drugs and immunosuppressives* (ed. P.M. Brooks), *Baillière's Clinical Rheumatology*, Vol. 4, No. 1, pp. 467–89. Baillière Tindall, London.

Thomas, M.R., Robinson, W.A., Boyle, D.J., Day, J.F., Entringer, M.A., and Steigerwald, J.F. (1983). Longterm effects of cyclophosphamide on granulocyte colony formation in patients with rheumatoid arthritis. *Journal of Rheumatology*, **10**, 778–83.

Tishler, M. *et al*. (1988*a*). Methotrexate in rheumatoid arthritis: a prospective study in Israeli patients with immunogenetic correlations. *Annals of the Rheumatic Diseases*, **47**, 654–9.

Tishler, M., Caspi, D., Gazit, E., and Yaron, M. (1988 *b*). Association of HLA-B35 with mucocutaneous lesions in Israeli patients with rheumatoid arthritis receiving gold treatment. *Annals of the Rheumatic Diseases*, **47**, 215–17.

Tishler, M., Caspi, D., Graff, E., Segal, R., Peretz, H., and Yaron, M. (1989). Synovial and serum levels of methotrexate during methotrexate therapy of rheumatoid arthritis. *British Journal of Rheumatology*, **28**, 422–3.

Tokuda, M. *et al*. (1994). Effect of low-dose cyclosporin on systemic lupus erythematosus disease activity. *Arthritis and Rheumatism*, **37**, 551–8.

Tugwell, P. *et al*. (1990). Methotrexate in rheumatoid arthritis: impact on quality of life assessed by traditional standard-item and individualized patient preference health status questionnaires. *Archives of Internal Medicine*, **150**, 59–62.

Tugwell, P. *et al*. (1995). Combination therapy with cyclosporine and methotrexate in severe rheumatoid arthritis. *New England Journal of Medicine*, **333**, 137–41.

Tzartos, S.J. *et al*. (1988). Fine antigenic specificities of antibodies in sera from patients with D-Penicillamine-induced myasthenia gravis. *Clinical and Experimental Immunology*, **74**, 80–6.

Uhino, M.R., Goldman, A.S., and Goto, M.P. (1993). Cyclosporin A-induced embryopathy in embryo culture is mediated through inhibition of the arachidonic pathway. *Proceedings of the Society for Experimental Biology*, **202**, 307–14.

Van der Heijde, D.M., van Riel, P.L., Nuver-Zwart, I.H., Gribnau, F.W., and van de Putte, L.B. (1989). Effects of hydroxychloroquine and sulphasalazine on progression of joint damage in rheumatoid arthritis. *Lancet*, **i**, 1036–8.

Van Gestel, A.M., Laan, R.F.J.M., Haagsma, C.J., Van de Putte, B.A., and Van Riel, P.L.C.M. (1995). Oral steroids as bridge therapy in rheumatoid arthritis patients starting with parenteral gold. A randomized double-blind placebo controlled study. *Journal of Rheumatology*, **34**, 347–51.

Van Loenen, H.J., Dijkmans, B.A., and De Vries, E. (1990). Concentration dependency of cyclosporin and chloroquine as inhibitors of cell proliferation and immunoglobulin production upon mitogen stimulation of mononuclear cells. *Clinical and Experimental Rheumatology*, **8**, 59–61.

Vardi, P., Brik, R., Barzilai, D., Lorber, M., and Scharf, Y. (1992). Frequent induction of insulin autoantibodies by D-Penicillamine in patients with rheumatoid arthritis. *Journal of Rheumatology*, **19**, 1527–30.

Verwilghen, J., Kingsley, G.H., Gambling, L., and Panayi, G.S. (1992). Activation of gold-reactive T lymphocytes in rheumatoid arthritis patients treated with gold. *Arthritis and Rheumatism*, **35**, 1413–20.

Vlachoyiannopoulos, P.G., Zerva, L.V., Skopouli, F.N., Drosos, A.A., and Moutsopoulos, H.M. (1991). D-Penicillamine toxicity in Greek patients with rheumatoid arthritis: anti-Ro (SSA) antibodies and

cryoglobulinaemia are predictive factors. *Journal of Rheumatology*, **18**, 44–9.

Wallace, D.J., Metzger, A.L., Stecher, V.J., Turnbull, B.A., and Kern, P.A. (1990). Cholesterol-lowering effect of hydroxychloroquine in patients with rheumatic disease: reversal of deleterious effects of steroids on lipids. *American Journal of Medicine*, **89**, 322–6.

Wascher, T.C. *et al*. (1994*a*). Cell-type specific response of peripheral blood lymphocytes to methotrexate in the treatment of rheumatoid arthritis. *Clinical Investigator*, **72**, 535–40.

Wascher, T.C. *et al*. (1994*b*). Serum levels of interleukin-6 and tumour necrosis factor-a are not correlated to disease activity in patients with rheumatoid arthritis after treatment with low dose methotrexate. *European Journal of Clinical Investigation*, **24**, 73–5.

Weinblatt, M.E. (1992). Long-tem prospective study of methotrexate in rheumatoid arthritis: 84 month update. *Arthritis and Rheumatism*, **35**, 129–37.

Weinblatt, M.E. and Fraser, P. (1989). Elevated mean corpuscular volume as a predictor of hematologic toxicity due to methotrexate therapy. *Arthritis and Rheumatism*, **32**, 1592–6.

Weinblatt, M.E. *et al*. (1990). Low-dose methotrexate compared with auranofin in adult rheumatoid arthritis: a thirty-six-week, double-blind trial. *Arthritis and Rheumatism*, **33**, 330–8.

Weinblatt, M. *et al*. (1991). The effects of cyclosporin A on eicosanoid excretion in patients with rheumatoid arthritis. *Arthritis and Rheumatism*, **34**, 481–5.

Wernick, R. and Smith, D.L. (1989). Central nervous system toxicity associated with weekly low-dose methotrexate treatment. *Arthritis and Rheumatism*, **32**, 770–5.

Wijnands, M.J., Perret, C.M., van Riel, P.L., and van de Putte, L.B. (1990 *a*). Generalized urticarial eruption during azathioprine treatment for rheumatoid arthritis: a case report and review of the literature. *Scandinavian Journal of Rheumatology*, **19**, 167–9.

Wijnands, M.J., Allebes, W.A., Boerbooms, A.M., van de Putte, L.B., and van Riel, P.L. (1990 *b*). Thombocytopenia due to aurothioglucose, sulphasalazine, and hydroxychloroquine. *Annals of the Rheumatic Diseases*, **49**, 798–800.

Wilkinson, S.M., Smith, A.G., Davis, M.J., Mattey, D.L., and Dawes, P.T. (1993). Suspected cutaneous drug toxicity in rheumatoid arthritis—an evaluation. *British Journal of Rheumatology*, **32**, 798–803.

Williams, H.J. *et al*. (1988). One-year experience in patients treated with auranofin following completion of a parallel, controlled trial comparing auranofin, gold sodium thiomalate, and placebo. *Arthritis and Rheumatism*, **31**, 9–14.

Williams, H.J. *et al*. (1994). Comparison of hydroxychloroquine and placebo in the treatment of arthropathy of mild systemic lupus erythematosus. *Journal of Rheumatology*, **21**, 1457–62.

Wilson, K. and Abeles, M. (1994). A 2 year, open ended trial of methotrexate in systemic lupus erythematosus. *Journal of Rheumatology*, **21**, 1674–7.

Wolfe, F., Hawley, D.J., and Cathey, M.A. (1990). Termination of slow acting antirheumatic therapy in rheumatoid arthritis: a 14-year prospective evaluation of 1017 consecutive starts. *Journal of Rheumatology*, **17**, 994–1002.

Wollner, A., Mohle-Boetani, J., Lambert, R.E., Perruquet, J.L., Raffin, T.A., and McGuire, J.L. (1991). *Pneumocystis carinii* pneumonia complicating low dose methotrexate treatment for rheumatoid arthritis. *Thorax*, **46**, 205–7.

Yanni, G., Nabil, M., Farahat, M.R., Poston, R.N., and Panayi, G.S. (1994). Intramuscular gold decreases cytokine expression and macrophage numbers in the rheumatoid synovial membrane. *Annals of the Rheumatic Diseases*, **53**, 315–22.

Yocum, D. (1993). Immunological actions of cyclosporin A in rheumatoid arthritis. *British Journal of Rheumatology*, **32** (Suppl.1), 38–41.

Yocum, D.E., Wilder, R.L., Dougherty, S., Klippel, J.H., Pillemer, S., and Wahl, S. (1990). Immunologic parameters of response in patients with rheumatoid arthritis treated with cyclosporin A. *Arthritis and Rheumatism*, **33**, 1310–16.

Youngelman, D.F., Kahng, K.U., Rosen, B.D., Dresner, L.S., and Wait, R.B. (1991). Effects of chronic cyclosporine administration on renal blood flow and intrarenal blood flow distribution. *Transplantation*, **51**, 503–9.

3.5.3 Therapeutic immunomodulation

A. M. Denman

Immunopathological basis

Immune reactions contribute to the chronic inflammation characteristic of chronic rheumatic diseases. In classic immune-mediated diseases the process is clearly initiated and maintained by an immune response to antigens. The stimulus may be an autoantigen, as in myasthenia gravis. Persistent exposure to exogenous antigens, such as drugs in drug-induced systemic lupus erythematosus or microbial antigens in untreated subacute bacterial endocarditis, also produces cumulative immune-mediated damage. It is also likely that atypical immune responses to microbial antigens in genetically predisposed individuals account for persistent inflammation in reactive arthritis. However, there are many situations in which immune reactions contribute to chronic inflammation but it is uncertain whether this is a primary or secondary event. For example, immune complex deposition in blood vessels is the main immunopathological event in systemic lupus erythematosus but the stimulus to this process is unknown. Similarly the T cells which infiltrate the rheumatoid synovial membrane undoubtedly contribute to inflammatory synovitis but it is still unproven that this process initiates the disease.

There has been major progress in elucidating the sequence of events which leads to immunopathologically mediated tissue damage. Thus details of the cytokine cascade which contributes to rheumatoid synovitis and joint erosion are largely understood. These advances have made it possible to introduce far more selective methods of immunotherapy. Whereas traditional methods of treatment were entirely pragmatic, at least the target for effective treatment can now be defined.

There are three approaches to practical immunotherapy in the rheumatic diseases. The first seeks to identify the initiating, immunologically mediated events and to block these specifically. Unfortunately there are few clinical situations in which this approach is relevant. A good experimental example is collagen-induced arthritis. Although the relevance of this model to rheumatoid arthritis is still debated, the goal of experimental immunotherapy in this model is clear, namely to block the inflammatory consequences of T-cell recognition. The second approach seeks to identify prominent immunopathological events and to block these irrespective of the process which led to their initiation. A pertinent example is rheumatoid synovitis where it is logical and feasible to interdict the T-cell component of this process or to inhibit proinflammatory cytokines such as tumour necrosis factor-a. Finally, it has become increasingly apparent that traditional antirheumatic drugs largely function through their effects on immunologically mediated events if one includes a broad range of anti-inflammatory processes in this definition. The rationale for using these drugs has often been empirical or even erroneous and only apparent in hindsight.

The immunological background

There are several stages at which immune-mediated inflammation can be blocked or blunted (Table 1; Fig. 1 and Fig. 2). These events are described in detail elsewhere in this book (see Chapter 3.1 and Chapter 3.2) and only points related to immunomanipulative strategies are considered here.

T cells
Tolerance induction by peptide administration
Tolerance induction by other methods of immunization
Receptor blocking by peptides
Antireceptor antibodies
Antireceptor immunoconjugates
Interdicting recirculation and proinflammatory traffic patterns
Antigen presentation
Peptide competition
Antireceptor antibodies
B cells and antibodies
Idiotype network manipulation
Antireceptor immunoconjugates
Intravenous immunoglobulin
Cytokine network
Cytokine administration (immunoregulatory)
Anticytokine antibodies and other agents
Anticytokine receptor antibodies and other agents
Genetic techniques — antisense oligonucleotides and gene transfection

Table 1 Strategies for selective immunosuppression

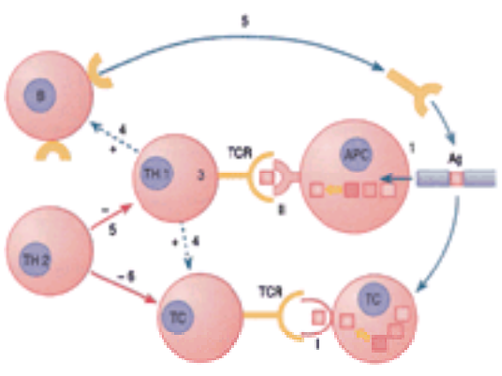


Fig. 1 Main events relevant to immunopathogenesis of rheumatic diseases. Antigen is degraded by antigen-presenting cells (APC) (1) and immunogenic peptides are presented by class II HLA molecules (II) (2). These peptide fragments are recognized by the antigen receptor of T cells (TCR) (3). T_{H1} lymphocytes are activated (+) to produce cytokines which induce a cytotoxic T-cell (T_C) response to target cells (TC) expressing the same immunogenic peptides and presented by class I HLA molecules (I) (4). T-helper cells also activate antibody-producing B cells (B) to secrete antibodies reactive with the immunizing antigen (5). Other T-cell populations of

the T_{H2} phenotype secrete cytokines which inhibit (–) many proinflammatory, antigen-driven immune responses.

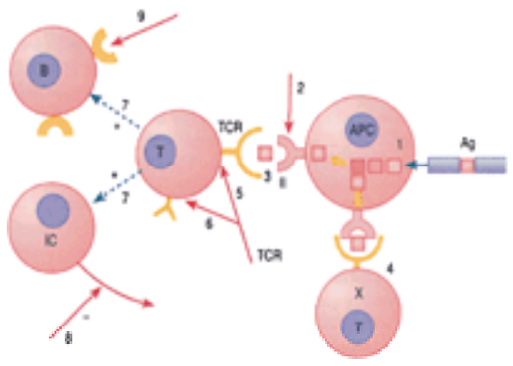


Fig. 2 Strategies for selective immunosuppression in the rheumatic diseases. Antigenic degradation into immunogenic peptides and their transport to surface class II HLA molecules may be inhibited (1). Class II molecules may be blocked (2) or their peptide-presenting groove occupied by competing peptides (3). The immunogenic response may be converted into a tolerizing event (4). Responding T cells may be inactivated by many potential monoclonal antibodies and other agents reacting with the T-cell receptor (TCR) (5) or other surface receptors essential for T-cell activation (6). Immunopotentiating cytokines may be inhibited (7) or proinflammatory cell populations (IC) may be inhibited by monoclonal antibodies and specific agents (8). Antibody-secreting B cells may be inhibited by anti-idiotypic antibodies (9).

T cells and antigen recognition

T cells recognize immunogenic peptides presented by class II histocompatibility (**HLA**) molecules that are expressed usually by specialized antigen-presenting cells but also by other cell types including B cells. The binding characteristics of the immunogenic peptides to the peptide-binding groove of the class II molecule is genetically determined and largely dictates the nature of the T-cell response. Peptide transport to the surface of specialized antigen-presenting cells is also genetically controlled. T-cell binding to peptides is specific and determined by the complementarity of the T-cell receptor. However, the rules governing T-cell binding by superantigens such as those displayed by many micro-organisms are less restrictive. T-cell interaction with antigen-presenting cells also depends on non-specific binding between ligands and their receptors expressed by both cell types. These interactions are essential for T-cell activation and also determine the outcome of the interaction between peptides and T cells. This may be activation, tolerance, or cell death (apoptosis). Interaction between the CD4 molecule and its ligands is the most crucial step in determining the outcome. It is also important in immunopathology that T cells traverse the endothelial cell barrier of small arterioles in order to reach inflammatory sites. This process also is dependent on receptors and ligands displayed by T cells and endothelial cells.

These interactions form the basis of many strategies for blocking T-cell activation and migration. The most specific interdiction involves the interaction between a given peptide and the relevant T-cell clone; the least specific depends on blocking an entire immunogenic step in the normal sequence of events.

Cytotoxic T cells

Cytotoxic T cells recognize peptides presented at the cell surface in association with class I HLA antigens. These peptides are derived from proteins that have been degraded in the cytosol and are then transported in the endoplasmic reticulum to the cell surface. The peptides bind to and take part in the assembly of glycoproteins of the class I molecule. The antigenic peptides are 8 to 20 amino acids in length and are presented to cytotoxlc T cells in the peptide-binding groove. Immunotherapeutic strategies are analogous to those employed to block T-cell recognition.

B cells and immune complex production

Normal antibody production may be qualitatively normal but quantitatively immunopathological if circulating immune complexes formed between antigens of any origin and antibodies are inadequately cleared or degraded and as a result are continually deposited in small blood vessels. A pertinent example is immune complex deposition in systemic lupus erythematosus, especially in individuals with inherited or acquired defects in the complement system. Therapeutic strategies then depend on partial suppression of antibody production, reducing antigen exposure, or correcting complement defects.

B cells and autoantibody production

It may seem self-evident that suppressing autoantibody production should be a prime object of immunosuppression. Yet two important points need to be considered. First, autoantibody production is not synonymous with abnormal B-cell behaviour, and second, autoantibodies do not invariably have immunopathological consequences.

Control of autoreactive B cells

Autoantibody production is part of the normal B-cell repertoire and autoantibodies are encoded by normal, non-mutated *VH* genes, suggesting that autoreactivity is an essential step in B-cell ontogeny. The emergence of pathogenic, autoreactive clones of B cells involves several stages. For example, systemic lupus erythematosus is likely to result from initial polyclonal B-cell proliferation followed by antigen selection of autoreactive clones. This process is aided by the failure of protective mechanisms that normally either eliminate or arrest the development of autoreactive B-cell clones.

The analysis of this regulation has been greatly assisted by transgenic experiments in which developing T cells are not exposed to selected antigens and are thereby not tolerized to these prenatally. Subsequent postnatal exposure leads to immune reactions which initiate autoimmune disease. It is evident that potentially autoreactive clones may be deleted or inactivated during tolerance induction or that loss of tolerance may be partial resulting in autoaggression restricted to certain organs. Paradoxically, experiments in which the T-cell repertoire has been selectively crippled ('knock-out' mice) show that genetically contrived immunodeficiency may lead to chronic inflammatory disease resembling human disorders. Thus increasing information about the maintenance of normal tolerance has mainly served to emphasize the complexity of restoring tolerance therapeutically.

Pathogenicity of autoantibodies

In many clinical states, autoantibodies lack pathogenetic significance and they may even be encountered in normal individuals. However, from experiments in which autoantibody-secreting B cells have been introduced into normal mice, it seems likely that autoantibodies can reproduce many of the immunopathological features of systemic lupus erythematosus and related diseases. Thus their inhibition may be a logical part of immunosuppressive strategies.

The idiotypic network

The specificity of antibodies for different antigens depends on their variable (V) region sequences. These regions are themselves novel sequences that stimulate an immune response by other lymphocyte clones. The immunogenic sequences are known as idiotypes and the set of idiotypes expressed by the complete molecule is termed its own idiotype. Idiotypes are called 'private' when they are unique to one antibody and 'public' when they are expressed by more than one antibody in the same or different individuals. Anti-idiotypic antibodies are generated as a physiological, regulatory response. They may also be created experimentally by active immunization or administered as passive immunotherapy.

The extent to which anti-idiotypic antibodies affect the antibody response partly depends on the location of the target idiotypes within the antibody molecule. Antibodies to idiotypes situated outside the antigen-combining site do not affect antigen binding by the target antibody. Antibodies to idiotypes near or contributing to

the antigen-binding site commonly interfere with antigen binding.

The same principles apply to novel sequences in the antigen receptors of T cells, which induce the production of anti-idiotypic antibodies by B cells and anti-idiotypic receptors on other T cells.

Thus, the idiotype network has long held a central place in ideas about the normal regulation of antibody production. This gave rise to the notion that autoantibody synthesis might reflect abnormal B-cell proliferation freed from the restraints of anti-idiotype regulation. This attractive idea has also stimulated schemes for the therapeutic manipulation of the network.

Cytokines

Cytokines may enhance or suppress immune responses depending on their local concentration, the nature of the target cells, and interaction with other cytokines. Cytokine production is genetically controlled and is likely to determine the duration and intensity of both normal and pathological immune responses. Cytokine production has been amply implicated in the rheumatic diseases by the direct assay of blood and inflammatory fluids and by demonstrating cytokine gene expression and protein production in lesions such as the rheumatoid synovial membrane. Since autoimmune disease may arise in experimental situations in which the control of cytokine gene expression has been artefactually disrupted, there is increasing speculation that abnormal cytokine production in disease may on occasion be the primary abnormality. These observations have stimulated a host of therapeutic ideas ranging from manipulation of cytokine-coding genes to the direct opposition of circulating cytokines with specific antibodies.

Non-selective immunomanipulation

Conventional agents

Non-specific immunosuppressive agents are still the mainstay of treating severe rheumatic disorders. Indeed the demonstrable efficacy of standard drugs such as prednisolone, methotrexate, and cyclosporin in rheumatoid arthritis has maintained interest in their mode of action. Furthermore, these drugs are simpler and cheaper to administer and monitor than many biological agents of current interest. Similarly, cyclophosphamide remains the drug of greatest proven value in controlling life-threatening features of systemic connective tissue diseases such as lupus nephritis. These agents have unpredictable effects on immune responses (Swanson and Schwartz 1967) with uncertain relevance to their clinical efficacy. Many antirheumatic drugs in standard use have been shown to affect immune-mediated inflammation. These drugs both benefit patients with rheumatic disorders and correct impaired immune responses, judged by *in vivo* and *in vitro* assays. As these effects occur simultaneously, some observers have concluded that the clinical improvement produced by cytotoxic drugs is mediated by therapeutic 'immunomodulation', but this point is far from established.

Unconventional agents

Most unconventional methods of non-specific immunosuppression are now of mainly historical interest. These were based on the premise that immunosuppression through lymphocyte depletion is the most logical means of controlling immune-mediated rheumatic diseases. Total lymphoid X-irradiation is a potent means of achieving a profound T- and B-cell lymphopenia and impressive remissions have been induced in severe rheumatoid arthritis and lupus nephritis (Strober *et al*. 1988). However, relapse is common in rheumatoid arthritis and there is a considerable incidence of illness and fatality mainly from unusually severe or opportunistic infections. While any benefit might reasonably be attributed to lymphocyte depletion, even this point has been disputed. There have been few comparative studies of the relative benefits of this approach and conventional treatment. It is probable that the side-effects of this treatment outweigh those of conventional cytotoxic drugs in patients with disease of similar severity. Thoracic duct drainage is in principle a more specific way of achieving profound depletion of the recirculating lymphocyte pool. While remissions have been recorded in severe rheumatoid arthritis resistant to other forms of treatment, this approach is no longer used. A more recent approach to the treatment of autoimmune connective tissue diseases is the extracorporeal inactivation of recirculating lymphocytes (Edelson 1991). Remissions have been claimed in open studies but controlled trials have not yet been reported. Selective removal of activated monocytes by leucapheresis has also been advocated as a means of reducing systemic inflammation (Hahn *et al*. 1993).

Strategies for selective immunosuppression

The detailed dissection of the normal immune system and its derangements in autoimmune diseases has stimulated a wealth of novel approaches for controlling immune-mediated inflammation. These are set out in Table 1. The practical problem now is to test the efficacy, safety, and relative merits of these strategies.

T-cell inactivation

Preoccupation with the proactive role of autoreactive T cells in many rheumatic disorders has stimulated interest in ways of suppressing or preferably tolerizing the abnormal clones.

Peptide administration

Impressive results have been achieved in experimental models of autoimmune diseases by administering peptides involved in the initiation of these diseases. Of obvious clinical relevance, some of these approaches suppress autoimmune disease which is already in progress. For example NOD mice with spontaneous diabetes have T-cell clones autoreactive with the 60-kDa heat-shock protein. The administration of a peptide, p277, derived from this protein both prevents and ameliorates ongoing diabetes in this model (Elias and Cohen 1994). There is particular interest in the prevention or suppression of disease by peptides given orally. Encouraging results have been achieved in the treatment of experimental allergic encephalomyelitis, collagen-induced arthritis, adjuvant arthritis, and spontaneous diabetes in NOD mice (Weiner *et al*. 1994). This principle was the basis for a trial of orally administered type II chicken collagen in patients with rheumatoid arthritis. Four out of the 28 patients remitted on this regimen (Trentham *et al*. 1993). Extended placebo-controlled trials of this principle are in progress.

Receptor competition

The factors which determine T-cell activation or tolerance after antigen exposure have been the subject of intensive investigation for many years. In addition to contact between antigenic peptides and the T-cell antigen receptor, activation requires a second signal initiated by interactions between surface receptors and their ligands. These receptors are readily accessible to experimental manipulation. One important interaction is between CD28 or CTLA-4 on T cells and B7 on antigen-presenting cells including activated B cells. T-cell activation is blocked by monoclonal antibodies to CD28 or CTLA-4. A protein constructed by genetic fusion of the extracellular domain of human CTLA-4 to an immunoglobulin C g_1-chain (termed CTLA4Ig) binds the B7 receptor and thereby blocks T-cell activation. Treatment with this novel protein inhibits the development or progression of lupus in NZB/NZW F_1 mice (Finck *et al*. 1994). Clinical trials based on this principle have not yet been reported.

Immunization by other routes

Several autoimmune diseases can be induced in experimental animals by immunization with antigens such as myelin basic protein or thyroglobulin. These disorders are mediated by T cells reactive with these autoantigens and can be transferred to other animals by autoreactive T cells or T-cell clones. On a similar theme, adjuvant arthritis is a chronic disorder of rats provoked by immunization with *Mycobacterium tuberculosis*. It is mediated by T cells reactive with a single nonapeptide epitope of the 65-kDa mycobacterial heat-shock protein. These T-cells cross-react with an autoantigen found in joints that may also prove to be a 65-kDa heat-shock protein (Karlsson-Parra *et al*. 1990).

The induction of these diseases can be prevented by vaccination with T-cell clones specifically reactive with the inducing antigen. Prerequisites for successful vaccination are attenuation of the clone by X-irradiation and its activation with the relevant antigen before vaccination. The efficacy of vaccination is greatly enhanced by the cross-linking of antigen receptors on the immunizing T cells through exposure to hydrostatic pressure or gluteraldehyde (Mor *et al*. 1990).

A novel approach has been described for achieving resistance to collagen-induced arthritis in rats and rhesus monkeys (t'Hart *et al*. 1993). The test animals inherit histocompatibility antigens which make them susceptible to this model autoimmune arthritis. Immunization with attenuated type II collagen in which arthritogenic

epitopes had been destroyed by heat inactivation conferred resistance to this disorder on subsequent challenge with the native collagen.

It is likely that several mechanisms account for the success of immunization strategies but that all of these involve anergy of potentially autoreactive T-cell clones. One intriguing possibility is that other T cells recognize idiotypic sequences in the antigen receptor of these clones and thereby inhibit their proliferation. However this strategy has not yet been applied clinically.

Monoclonal antibodies to T-cell receptors

The availability of monoclonal antibodies to T-cell receptors has prompted a spate of clinical trials. Antibodies to the CD4 receptor are of particular interest because success in experimental transplantation systems offers the prospect not simply of drug-dependent immunosuppression but of tolerizing autoreactive T-cell clones (Waldmann 1993). These hopes were increased by experience in controlling autoimmune disease in experimental models. Anti-CD4 monoclonal antibody suppresses murine lupus and the associated immunopathological abnormalities (Wofsy 1988; Ermak *et al*. 1989; Carteron *et al*. 1990). Furthermore, the disease can be suppressed without lymphocyte depletion since $F(ab)_2$ fragments of the antibody are equally effective (Carteron *et al*. 1989). In some such experiments the disease did not progress after treatment was stopped, raising the possibility that this agent had indeed induced tolerance to the inciting autoantigens. These observations led to the formulation of a strategy for treating autoimmune disease whereby the T cells mediating disease are first largely eliminated by a cytotoxic monoclonal antibody, and the remaining cells are than rendered tolerant to the inciting autoantigens by exposure to a second, non-cytotoxic monoclonal antibody.

Early impressions of anti-CD4 monoclonal antibody in rheumatoid arthritis were encouraging, with clinical improvement, minor toxicity, and a marked CD4-cell lymphopenia (Kyle *et al*. 1989). Most subsequent reports of uncontrolled trials have been similarly favourable and noted reduced joint inflammation, a decreased acute-phase response, and improved functional capacity. Different preparations of chimeric monoclonal antibody seem to have had similar effects. This amelioration of disease activity is accompanied by a CD4 lymphopenia whose intensity and duration are dose-dependent. There is an unpredictable effect on immune competence (Horneff *et al*. 1991a; Horneff *et al*. 1991b; Wendling *et al*. 1991; Wendling *et al*. 1992; Goldberg *et al*. 1991; Van der Lubbe *et al*. 1993). In addition, some patients have tolerated further courses of antibody with reported benefit (Reiter *et al*. 1991). The various hybrid antibodies are generally well tolerated; minor adverse reactions in the form of urticaria, transient hypotension, and fever have been attributed to cytokine activation, notably involving interleukin 6 (Moreland *et al*. 1993). An antibody response to the rodent hypervariable portion of the hybrid molecules which are administered has been noted in only a small percentage of the patients in these trials. This improvement probably does not simply depend on CD4 cell elimination since there is a poor correlation between circulating CD4 counts and clinical improvement (Choy *et al*. 1992; Van der Lubbe *et al*. 1994). An eventual reduction of cytokine production by circulating monocyte/macrophages has been noted (Horneff *et al*. 1993) but there is no correlation between circulating cytokine levels and clinical response.

Nevertheless the limited data from placebo-controlled trials do not as yet show any clear benefit in the treated patients (Choy *et al*. 1992). A recent trial of anti-CD4 monoclonal antibody showed no significant benefit even in early rheumatoid arthritis when agents of this kind might be expected to be most effective (Van der Lubbe *et al*. 1995). There is still uncertainty about the ability of patients to tolerate long-term or repeated courses of hybridized anti-CD4 monoclonal antibody. Despite earlier impressions, an immune response to rodent sequences or idiotypic human sequences is probably universal and may well compete for infused antibody even if the adverse effects can be combated. It also seems likely that the recirculating pool of CD4 cells is permanently depleted in some patients, particularly in those who have received potent immunosuppressive drugs (Horneff *et al*. 1992; Moreland *et al*. 1994). This raises the possibility of long-term opportunistic infections and other complications. The initial, often uncritical enthusiasm for this form of treatment has been tempered by the realization that more extensive controlled trials are needed (Burmester and Emmrich 1993). There is a still greater dearth of controlled trials of monoclonal antibodies against other T-cell surface molecules in rheumatoid arthritis.

Trials of monoclonal antibodies to T-cell surface molecules in systemic connective tissue diseases have been confined almost exclusively to anti-CD4 monoclonal antibody. The reports have involved few patients and are uncontrolled, so it is difficult to draw any conclusions about its efficacy. For example in one early report, systemic vasculitis was successfully treated by sequential administration of a monoclonal antibody that depleted CD4 cells and a non-cytotoxic anti-CD4 monoclonal antibody (Mathieson *et al*. 1990), with the intention of rendering the patient tolerant to the antigen putatively initiating the disease. Further exploitation of this idyllic principle is awaited.

Despite the theoretical advantages over previous methods of lymphocyte depletion, monoclonal antibodies against receptors expressed by all circulating T cells are still relatively non-specific. In principle, antibodies directed against the T-cell clones responsible for the autoimmune process would be the most specific form of immunosupression. Tolerizing antibody would be still more effective, thereby avoiding the need for continued treatment. This strategy depends on identifying the responsible clones by virtue of their unique antigen receptors. There is some information about the genes encoding T-cell receptors in model autoimmune diseases. Experimental allergic encephalomyelitis is a mixed inflammatory and degenerative disease of the central nervous system induced by an autoimmune attack by T cells immunized with basic myelin protein. The encephalitogenic T cells use a remarkably limited number of *V*-b genes to encode the receptor for the disease-inciting peptides. Monoclonal antibodies to these receptors prevent the induction of disease and partially suppress established disease (Kumar *et al*. 1989; Wraith *et al*. 1989). There is a similarly restricted usage of the same *V* genes in other models of organ-specific autoimmune disease, even though the T cells recognize different peptides. These observations suggest that the antigen receptors expressed by T cells responsible for autoimmune disease may have features in common (Adams *et al*. 1990; Heber-Katz 1990). Unfortunately there are two major practical problems in applying these principles to clinical practice. First, no consistent pattern of dominant T-cell clones has been identified in diseases such as rheumatoid synovitis. Some evidence of restricted T-cell receptor usage has emerged in individual patients and even in serial synovial biopsies from the same patient (Broker *et al*. 1993). However, the same dominant clone is not found in different patients so that it is as yet impossible to develop a specific reagent for general use. Second, non-specific anti-T-cell monoclonal antibodies have not induced lasting remissions so that it has been necessary to resort to repeated administration of these agents. In other words this treatment may suppress the relevant clones but there is no hint that it tolerizes them.

It would be less advantageous but still relatively helpful if diseases such as rheumatoid arthritis were mediated by T-cell clones reactive with superantigens. In these circumstances the antigen receptors would be encoded by a restricted family of *V* genes. Deleting this population of T cells would be less immunologically crippling than lymphopenia induced by entirely non-specific agents. However there is no good evidence that a mechanism of this nature is operating.

T-cell traffic

T cells accumulate at sites of inflammation by traversing the endothelial cells lining small arterioles. Recruitment by this route is largely governed by adhesion molecules and their ligands expressed on T-cell and endothelial cell surfaces. Experimental results show that it is possible to interdict this process with selective antibodies. Thus an antibody that prevents lymphocyte homing across high endothelial venules suppresses lymphadenopathy in MRL-1pr/1pr mice that develop spontaneous autoimmune disease (Mountz *et al*. 1988). Similarly, the transfer of spontaneous diabetes in mice by spleen cells can be prevented by a monoclonal antibody that interferes with macrophage adhesion in the recipient mice (Hutchings *et al*. 1990).

The interaction between the leucocyte function-associated antigen 1 and its receptor, intercellular adhesion molecule 1 (**ICAM-1**), is crucial to T-cell recruitment. One method of interdicting T-cell accumulation in rheumatoid synovitis depends on monoclonal antibodies to these adhesion molecules. Anti-ICAM-1 monoclonal antibody modifies T-cell recirculation and in an open trial produced clinical improvement in 9 of 23 treated patients (Kavanagh *et al*. 1994). The results of controlled trials are awaited.

Antigen presentation

Since T cells recognize peptide presented by class II molecules of antigen-presenting cells, several immunosuppressive strategies depend on blocking peptide presentation. Genetic susceptibility to autoimmune disease is strongly linked to the inheritance of certain class-II molecules. This linkage provides an additional rationale for this approach.

Peptide competition

In theory, autoimmune diseases could be blocked by substituting non-immunogenic peptides for the autoantigenic peptides which are presumed to break tolerance in T cells. These pathogenic peptides are probably presented in normal physiological fashion. Several technical advances have facilitated this approach. Oligopeptides can be synthesized and their binding properties to other molecules and biological structures screened on an almost industrial scale. Computer modelling has enabled their interactions with the peptide-binding groove of genetically different class I and II HLA molecules to be predicted with great accuracy. However there are still formidable problems. Some are readily foreseen, notably the difficulty in guiding immunotherapeutic peptides to the relevant receptors. Others are more unexpected.

Thus *in vitro* experiments have shown that different model peptides have a variable binding affinity for the class II molecule and that the resulting complex may be intrinsically unstable. Nevertheless, competition for antigen receptors on autoreactive T cells has been shown to operate successfully in experimental models of autoimmune disease (Fairchild *et al*. 1994).

Anti-MHC antibodies

Monoclonal antibodies to class II MHC gene products have been successfully used to treat many experimental models of autoimmune disease. Some of these models depend on immunization with tissue-specific antigens and include allergic encephalomyelitis, myasthenia gravis, autoimmune thyroiditis, and collagen-induced arthritis provoked in different species. Anti-MHC antibodies are mainly effective during the induction period. These antibodies also suppress spontaneously occurring autoimmune diseases, such as lupus in NZB/NZW mice and diabetes mellitus in BB rats (Vladutiu 1991). However, it is likely that anti-MHC antibodies do not simply block the target molecule but that they also initiate a sequence of secondary events. The overall result is a partial suppression of T-cell responses and, at least in experimental models, the activation of suppressor cell populations. Although at first sight this approach may appear to be rather non-specific, it is selectively immunosuppressive in both theory and practice. First, blocking the product of MHC alleles predisposing to autoimmune disease does not prevent the treated animal from expressing other MHC-encoded gene products that are perfectly able to engage a broad repertoire of microbial antigens. Second, microbial pathogens produce a range of antigenic determinants (epitopes) able to bind to other MHC products which T cells can recognize. Finally, class II antigens are abnormally expressed in target organs for autoimmune disease yet their local suppression does not affect systemic immune responses.

So far this principle has been little tested in clinical practice. An indirect trial of anti-class II antibodies has been conducted in rheumatoid arthritis on the basis that immunoglobulins with a high concentration of these antibodies can be obtained from the placenta and its blood supply; these antibodies may account for disease remission during pregnancy (Sany 1988). Clinical improvement has been claimed but more extensive trials are needed.

B-cell blockade

Idiotype network manipulation

The physiological importance of the network between idiotypes and anti-idiotype antibodies raised high hopes that these interactions would prove amenable to immunotherapeutic manipulation. The complexities of the idiotype network provide enormous scope for its therapeutic manipulation. For example, autoantibodies to the acetylcholine receptor contribute to the immunopathogenesis of myasthenia gravis. Antibodies to certain dextrans and to the acetylcholine receptor share idiotype determinants. Prior immunization with dextran reduces the antibody response to the acetylcholine receptor induced after subsequent challenge with this receptor (Tong and Dwyer 1990).

However, it was apparent from the outset that certain conditions would need to be met before this approach could become practicable. First, there must be common ('public') idiotypes associated with the disease to be treated. It is clearly easier to manipulate the idiotype network if certain idiotypes predominate in this condition. There are some indications that autoantibodies in patients with connective tissue diseases share common idiotypes. For example, the same idiotypes are expressed by a high proportion of anti-DNA autoantibodies in patients with systemic lupus erythematosus (Blank *et al*. 1990; Weisbart *et al*. 1990). Moreover the clones synthesizing this autoantibody appear to be stable during the course of the disease (Winkler *et al*. 1988). However, anti-DNA autoantibodies are synthesized by cells subject to considerable antigen-driven mutation so that there is the potential for considerable diversity in their variable sequence. Furthermore these idiotypes are not uniquely associated with anti-DNA autoantibodies. Similarly, while 60 per cent of monoclonal paraproteins with rheumatoid factor activity share the same idiotype determinants, there is far more idiotype diversity among rheumatoid factors synthesized by blood and synovial B cells in rheumatoid arthritis (Thompson *et al*. 1990). Thus, if selective suppression of the production of rheumatoid factor proved therapeutically desirable, it would be difficult to achieve by idiotype manipulation.

Second, the target idiotype for immunotherapy must be associated with the immunopathogenesis of the disease in question. It is not necessarily true that autoantibodies expressing a shared idiotype are pathogenic. On occasion this is clearly the case as, for example, with complement-fixing autoantibodies to formed blood elements. Occasionally the issue can be tested experimentally: for instance, it has been claimed that mice develop the immunopathological features of systemic lupus erythematosus when immunized with the idiotype 16/6 associated with anti-DNA autoantibodies (Mendlovic *et al*. 1990).

Third, the results must be therapeutically beneficial. In practice the outcome of perturbing autoantibody production in experimental systems is not easily predictable. Thus passively administered, monoclonal, anti-Sm antibodies increase or decrease the production of this autoantibody in mice with systemic lupus erythematosus, depending on the fine specificity of the administered antibody (Eisenberg *et al*. 1990).

Fourth, the benefit must be reasonably long lasting. Unfortunately, adaptive changes are likely to occur. Thus lymphomas treated with anti-idiotypes antibodies eventually escape elimination by changing their surface immunoglobulin (Rafield, *et al*. 1985). Autoantibody-synthesizing clones may behave similarly.

However, the greatest stumbling block has been the lack of practical methods for manipulating the idiotype network.

B-cell suppression

It has proved difficult to devise methods for directly inhibiting the proliferation or maturation of autoantibody-producing B cells. This is perhaps not surprising since B cells committed to conventional antibody synthesis are relatively resistant to conventional antiproliferative drugs. However, monoclonal antibodies to B-cell antigens have been used as antiproliferative agents in murine systemic lupus erythematosus (Asensi *et al*. 1989; Yakura *et al*. 1989). It is also noteworthy that anti-class II antibodies directly inhibit the spontaneous *in vitro* activation of human lupus B cells (Tanaka *et al*. 1989). The predominant attribution of autoantibody synthesis to the CD5+ subpopulation of B cells has formed the basis for a novel treatment with an immunoconjugate consisting of ricin linked to an anti-CD5 monoclonal antibody (Strand *et al*. 1993). Initial clinical improvement at least was noted in more than half the 79 patients with rheumatoid arthritis enrolled in the trial. It remains to be seen if consistent suppression of rheumatoid factor or other autoantibody production can be achieved in this fashion.

Intravenous immune gammaglobulin

There are suggestive anecdotal accounts but little controlled evidence to show that intravenous immunoglobulin benefits patients with autoimmune disorders. These include idiopathic thrombocytopenia, systemic lupus erythematosus, and rheumatoid arthritis. Many non-specific immunosuppressive effects may be operating but the most intriguing explanation is that pooled immunoglobulin from normal donors contains anti-idiotypic antibodies to idiotypes expressed by pathogenic autoantibodies (Dietrich *et al*. 1992). There is also good experimental evidence that genetically determined patterns of antibody response to micro-organisms influence autoantibody production. Polyreactive natural antibodies stimulated by infection may inhibit the production of pathogenic autoantibodies (Hentati *et al*. 1994). Thus some gammaglobulin preparations may fortuitously correct imbalances of this nature. The current impossibility of matching donor gammaglobulin preparations with postulated immunoregulatory defects in the recipient explains why this form of treatment is so hard to assess. So far, with the exception of dermatomyositis (Dalakas *et al*. 1994), there are few convincing controlled trials of this popular form of immunotherapy that have shown any benefit.

Plasma exchange was formerly a popular means of removing autoantibodies and immune complexes. Apart from its possible value as an emergency procedure in fulminating disease, it is now little used.

Cytokine manipulation

Cytokine production is a prominent feature in rheumatic diseases judged by circulating blood levels, high concentrations in inflammatory exudates, and local production in rheumatoid synovitis. Cytokines contribute to the acute-phase response characteristic of these disorders and also to joint damage in rheumatoid arthritis. There is growing evidence that cytokine production is genetically controlled and that the inherited pattern of cytokine generation influences the severity of immunopathological disorders. Detailed information about the structure and synthesis of cytokines and their receptors has led to the development of several strategies for manipulating the cytokine response in the rheumatic diseases. Many of these have already been introduced into clinical practice.

There are many possible strategies for manipulating the cytokine network for immunotherapeutic purposes. In general terms these involve administering cytokines or inhibiting cytokines or their receptors. However, the complexities of cytokine interactions mean that the administration of one cytokine may have a suppressive effect on other cytokines. Cytokine actions are pleomorphic so that inhibiting the production of a given cytokine or administering it therapeutically is likely to affect cells and systems other than the intended target. Cytokines are commonly interdependent thereby increasing the chances of unexpected actions. Transgenic models of

unopposed cytokine production have also shown that increased cytokine concentrations may directly or indirectly initiate autoimmune disorders. These considerations reinforce the need for caution over the long-term consequences of this form of treatment. The paradoxical findings in different situations also enjoin caution about superficial generalizations. Tumour necrosis factor-a (**TNF-a**) is clearly important in the pathogenesis of rheumatoid arthritis and its inhibition by monoclonal antibody is clinically beneficial. Yet spontaneous lupus nephritis in NZB × NZW F_1 mice is associated with genetically determined, low serum levels of TNF-a whose correction ameliorates the disease process (Gordon *et al*. 1989).

In the earliest attempts at cytokine therapy, patients were given interferon-g. This was a pragmatic venture since the biological effects of the interferons are unpredictable. Although interferon-g is produced by activated T lymphocytes of the kind contributing to rheumatoid synovitis, very low concentrations are encountered in rheumatoid synovitis (Firestein and Zvaifler 1990). Furthermore, blood T lymphocytes (Sadouk *et al*. 1990) and synovial macrophages (Bergroth *et al*. 1989) from patients with rheumatoid arthritis show impaired responses interferon-g. This reduction may result from the excessive production of other cytokines. Whether exogenous interferon-g can correct these defects is problematical and the therapeutic consequences are still more speculative. While one controlled trial showed some benefit in rheumatoid arthritis (Cannon *et al*. 1990), the results overall have been unimpressive and perhaps predictably so. More theoretically encouraging are cytokines whose effects are more likely to interdict other cytokines of proven immunopathogenetic relevance. For example, interleukin 10 (**IL-10**) inhibits the *in vitro* synthesis of IL-1 and TNF-a by rheumatoid synoviocytes (Katsikis *et al*. 1994) and also inhibits the induction of experimental allergic encephalomyelitis (Rott *et al*. 1994). This principle has not been tested in clinical practice but there is at least a stronger rationale than has been evident in earlier trials of cytokine manipulation.

The current phase of cytokine suppression is based on the inhibition of circulating cytokines by drugs or specific monoclonal antibodies. The proven importance of IL-1 in rheumatoid synovitis directed attention to this cytokine. Collagen-induced arthritis in mice is largely suppressed by treatment with a polyclonal anti-IL-1 antibody (Van den Berg *et al*. 1994). Tenidap has been introduced largely on the rationale that it inhibits the *in vitro* production of IL-1b and also of TNF-a and IL-6. However, it also inhibits cyclo-oxygenase production so that its anti-inflammatory effects are not entirely attributable to its effects on cytokine production. Clinical benefit has been noted in a double-blind, comparative trial of tenidap and the anti-inflammatory drug piroxicam (Littman *et al*. 1995). Tenidap was more effective than piroxicam in reducing blood levels of IL-6 and the intensity of the acute-phase response. The findings were similar in a 24-week comparative trial of tenidap and the anti-inflammatory drug diclofenac (Wylie *et al*. 1995). However, the high incidence of adverse effects and the limited duration of the reported trials makes it impossible to judge the potential of this agent on current evidence.

Arguably the most impressive results of selective cytokine inhibition have been achieved by treating rheumatoid arthritis with an anti-TNF-a monoclonal antibody. This agent also was found to be effective in earlier studies of experimental collagen-induced arthritis (Williams *et al*. 1992). Successful treatment in an open trial in 20 patients with rheumatoid arthritis (Elliott *et al*. 1993) was confirmed in a more extensive, placebo-controlled trial in 73 patients (Elliott *et al*. 1994). The clinical improvement was accompanied by a marked fall in blood C-reactive protein and IL-6 concentrations. In a similar double-blind, placebo-controlled study, a recombinant anti-TNF-a monoclonal antibody combining mouse hypervariable regions and a human IgG4 immunoglobulin reduced both joint activity and the acute-phase reaction (Rankin *et al*. 1995). However, there is as yet little information about the duration of remissions induced by this treatment, protection against joint erosion, or the patients' ability to tolerate further infusions. There is also some concern that long-term inhibition of TNF-a may make the recipients susceptible to intercurrent infections. It is still unclear how this agent achieves its anti-inflammatory effects since neutralizing the low circulating levels of this cytokine is surprisingly effective. *In vitro* and *in vivo* experiments suggest that this neutralization may have wider effects including the reversal of depressed cell-mediated immunity secondary to prolonged exposure to high TNF-a concentrations (Cope *et al*. 1994).

A potentially still more elegant method of inhibiting cytokine production depends on inhibiting the translation of messenger RNA (**mRNA**) by antisense oligonucleotides. In one experimental system an antisense nucleotide specific for IL-6 mRNA penetrated human articular cartilage *in vitro* and prevented the synthesis of this cytokine. This in turn blocked the inhibition of cartilage synthesis by IL-6 normally induced under the influence of IL-1 (Nietfeld *et al*. 1994). The success of this approach depends on the development of suitable delivery systems.

There is still only a fragmentary picture of experimental and clinical trials of methods for inhibiting cytokine receptors. Treatment with a soluble recombinant anti-IL-1 receptor antagonist reduces the intensity of collagen-induced arthritis in mice (Wooley *et al*. 1993a), but clinical trials have not yet been reported. IL-2 receptor blockade is effective in experimental autoimmune diabetes (Kelley *et al*. 1988) and there is preliminary evidence from a limited clinical trial that an intravenously administered IL-2 fusion toxin may be effective in rheumatoid arthritis (Sewell *et al*. 1993). On the same principle, a fusion protein in which the soluble TNF p55 receptor was linked to the Fc fragment of the IgG molecule efficiently blocked collagen-induced arthritis in mice (Wooley *et al*. 1993b); this agent has not yet been used in clinical practice. Similarly, a fusion protein consisting of human IL-2 and part of the diphtheria toxin selectively blocks the proliferation of activated T cells *in vitro*. In a clinical trial in psoriasis, which is T-cell dependent, this protein produced clinical improvement in four of the eight patients treated (Gottlieb *et al*. 1995).

The most fundamental way of correcting cytokine-mediated abnormalities is to devise methods of introducing genes whose transcribed proteins will block the excessive or unopposed production of proinflammatory or tissue-destructive cytokines. This could in principle be achieved locally, as in the rheumatoid synovial membrane, or systemically, in lymphoreticular tissues. Potentially therapeutic genes have been successfully transferred to the synovium. For example, a gene which encodes a protein antagonizing the IL-1 receptor was transfected to synovium and thereby protected the joint from IL-1 mediated destruction (Evans and Robbins 1994). The technology for gene transfer is well established. The main problems in the clinical application of these methods is to define cytokine expression and physiology in different diseases in precise detail. The potential hazards of gene transfer are still the subject of unproven speculation.

Dietary manipulation

There have now been intensive studies of the effects of dietary manipulation on experimental models of autoimmune disease and in clinical practice. This interest was largely stimulated by lay and medical disquiet over the failings and hazards of conventional antirheumatic drugs. It has now become a scientifically acceptable form of investigation which has revealed many interesting links between diet and inflammation (Darlington and Ramsay 1994).

In general terms there have been three main areas of study, namely the effects of dietary restriction, alterations in fat intake, and a search for dietary allergens of possible immunopathological relevance.

The most stringent form of dietary restriction is fasting and this regimen has anti-inflammatory effects in rheumatoid arthritis (Skoldstam and Magnusson 1991). In a controlled comparative trial, fasting followed by a gluten-free, lactovegetarian diet was effective for up to 1 year (Kjeldsen-Kragh *et al*. 1991) and the improvement was sustained for up to 2 years (Kjeldsen-Kragh *et al*. 1992). Interestingly, this improvement may have reflected not dietary factors but the patients attitude to treatment. The patients who responded to this form of treatment may have had confidence in a regimen under their control matched by scepticism about conventional drug treatment (Kjeldsen-Kragh *et al*. 1994a). A less draconian form of initial dietary restriction than fasting involves recourse to an elemental diet. This regimen has a beneficial effect on rheumatoid arthritis and the improvement is more likely to be attributable to protein and calorie deprivation than allergen exclusion (Kavanagh *et al*. 1995). Similar reservations were expressed after a comparable trial of this nature (Haugen *et al*. 1994a). Lactovegetarian diets also induce an improvement in rheumatoid synovitis and patients are more likely to comply with this form of diet than with more stringent ones. This diet probably has direct anti-inflammatory effects (Kjeldsen-Kragh *et al*. 1995). Gut flora is appreciably altered introducing intriguing possibilities about the cause of the clinical improvement (Peltonen *et al*. 1994). It is unlikely that the benefit can be attributed to changes in fatty acid composition (Haugen *et al*. 1994b).

The composition of dietary fat influences immune-mediated inflammation (Denman 1992). A low fat diet retards and a high fat diet accelerates genetically determined murine lupus. Furthermore, dietary fish-oil supplements also suppress disease features and increase longevity in this model. The high eicosapentaenoic acid content of fish oil is broken down into anti-inflammatory prostaglandins, possibly accounting for the protective effects.

The same dietary principles have been used to treat rheumatoid arthritis. Dietary enrichment with fish oils rich in -3 polyunsaturated acids significantly reduces joint inflammation judged by controlled trials (Nielson *et al*. 1992; Geusens *et al*. 1994; Kjeldsen-Kragh *et al*. 1994b). The likely mechanism is a relative increase in these acids in the phospholipid fraction of cell membranes at the expense of -6 polyunsaturated fatty acids. This substitution reduces the precursors of the proinflammatory 4 series leukotrienes and the 2 series prostaglandins (Tulleken *et al*. 1990). In addition to the inhibitory effects of these manoeuvres on conventional proinflammatory pathways, the production of IL-1, IL-6, and TNF-a is also inhibited (Watson *et al*. 1993).

There is little evidence that intolerance to dietary allergens contributes to disease activity in rheumatoid arthritis or systemic autoimmune diseases. This conclusion is supported by the failure to correlate changes in disease activity with the reintroduction of dietary items in patients on exclusion regimens. However, there is a suggestion that milk intolerance may be relevant in some patients and that an IgE-mediated component may contribute to the inflammation (Van de Laar and Van der Korst 1992; Van de Laar *et al*. 1992).

There is no doubt that dietary manipulation in the rheumatic diseases has achieved respectability. Treatment with fish oil or other sources of unsaturated fatty acids can be expected to achieve a 10 to15 per cent reduction in inflammatory activity in most patients. It is reasonable to encourage suitably motivated patients to pursue vegetarian diets, which generally prove nutritionally acceptable (Haugen *et al*. 1993) but nevertheless need careful supervision (Rauma *et al*. 1993).

Conclusions

Recent advances in understanding immunopathological mechanisms in the rheumatic diseases have stimulated a host of novel and ingenious methods of interdicting these pathways. Although many fundamental questions about the aetiology of most of these disorders are still unanswered, immunomanipulation is becoming increasingly rational and less pragmatic. Moreover, expectations raised by the success of these methods in experimental disease models have been gratifyingly realized in clinical studies. They usually work according to predictions, they produce at least a short-term remission, and their benefit has often been confirmed in controlled trials. Nevertheless many problems need to be settled before rheumatologists can at last move away from the hazards of traditional, non-specific immunosuppression. Not every speculation about the relative importance of different immune systems will necessarily be fulfilled. For example, the selective interdiction of CD4 T cells has not been shown to produce lasting benefit despite the central role attributed to these cells in many current ideas about the immunopathogenesis of rheumatoid synovitis. Enthusiasm over the successful suppression of rheumatoid synovitis by anticytokine monoclonal antibodies is yet to pass the real test, namely their ability to inhibit joint destruction. Only careful surveillance will test the long-term safety of even relatively selective immunomanipulation. This is especially true of the complex cytokine system whose members are interdependent and have pleomorphic effects. Simpler and cheaper drugs in routine clinical use may yet prove more cost-effective and not significantly less efficacious or more dangerous than scientifically more elegant methods.

Chapter References

Adams, S., Zordan, T., Sainis, K., and Datta, S.K. (1990). T cell V-b genes expressed by IgG anti-DNA autoantibody-inducing T cells in lupus nephritis: forbidden receptors and double-negative T cells. *European Journal of Immunology*, **20**, 1435–43.

Asensi, V., Kimeno, K., Kawamura, I., Sakumoto, M., and Nomoto, K. (1989). Treatment of autoimmune MRL/1pr mice with anti-B220 monoclonal antibody reduces the level of anti-DNA antibodies and lymphadenopathies. *Immunology*, **68**, 204–8.

Bergroth, V., Zvaifler, N., and Firestein, G.S. (1989). Cytokines in chronic inflammatory arthritis. III. Rheumatoid arthritis monocytes are not usually sensitive to g-interferon, but have defective g-interferon-mediated HLA-DQ and HLA-DR induction. *Arthritis and Rheumatism*, **32**, 1074–9.

Blank, M. *et al*. (1990). The importance of the pathogenic 16/6 idiotype in the induction of SLE in naïve mice. *Scandinavian Journal of Immunology*, **31**, 45–52.

Broker, B.M. *et al*. (1993). Biased T cell receptor usage in rheumatoid arthritis. Oligoclonal expansion of T cells expressing Va2 genes in synovial fluid but not in peripheral blood. *Arthritis and Rheumatism*, **9**, 1234–43.

Burmester, G.R. and Emmrich, F. (1993). Anti-CD4 therapy in rheumatoid arthritis. *Clinical and Experimental Rheumatology*, **11** (Suppl. 8), 139–45.

Cannon, G.W. *et al*. (1990). Prospective two-year follow-up of recombinant interferon-g in rheumatoid arthritis. *Journal of Rheumatology*, **17**, 304–10.

Carteron, N.L., Schimenti, C.L., and Wofsy, D. (l989). Treatment of murine lupus with F(ab)2 fragments of monoclonal antibody to L3T4. Suppression of autoimmunity does not depend on T helper cell depletion. *Journal of Immunology*, **142**, 1470–5.

Carteron, N.L., Wofsy, D., Schimenti, C., and Ermak, T.H. (1990). F(ab)2 anti-CD4 and intact anti-CD4 monoclonal antibodies inhibit the accumulation of CD4+ T cells, CD8+ T cells, and B cells in the kidneys of lupus-prone NZB/NZW mice. *Clinical Immunology and Immunopathology*, **56**, 373–83.

Choy, E.H., Chikanza, I.C., Kingsley, G.H., Corrigall, V., and Panayi, G.S. (1992). Treatment of rheumatoid arthritis with single dose or weekly pulses of chimaeric anti-CD4 monoclonal antibody. *Scandinavian Journal of Immunology*, **36**, 291–8.

Cope, A.P. *et al*. (1994). Chronic exposure to tumor necrosis factor (TNF) *in vitro* impairs the activation of T cells through the T cell receptor/CD3 complex: reversal *in vivo* by anti-TNF antibodies in patients with rheumatoid arhtritis. *Journal of Clinical Investigation*, **94**, 749–60.

Dalakas, M.C. *et al*. (1994). A controlled trial of high-dose intravenous immune globulin infusions as treatment for dermatomyositis. *New England Journal of Medicine*, **329**, 1993–2000.

Darlington, L.G. and Ramsey, N.W. (1994). Review of dietary therapy for rheumatoid arthritis. *Comprehensive Therapy*, **20**, 190–4.

Denman, A.M. (1992). Unsaturated fatty acids and immune disorders in unsaturated fatty acids. Nutritional and physiological significance, pp. 137–46. The British Nutrition Foundation, Chapman & Hall, London.

Dietrich, G., Kaveri, S.V., and Kazatchkine, M.D. (1992). Modulation of autoimmunity by intravenous immune globulin through interaction with the function of immune/idiotypic network. *Clinical Immunology and Immunopathology*, **62**, S73–81.

Edelson, R.L. (1991). Photophoresis: present and future pospects. *Journal of Photochemical and Photobiological Research*, **10**, 165–71.

Eisenberg, R.A., Pisetsky, D.S., Craven, S.Y., Grudier, J.P., O'Donnell, M.A., and Cohen, P.L. (1990). Regulation of the anti-Sm autoantibody response in systemic lupus erythematosus mice by monoclonal anti-Sm antibodies. *Journal of Clinical Investigation*, **85**, 86–92.

Elias, D. and Cohen, I.R. (1994). Peptide therapy for diabetes in NOD mice. *Lancet*, **343**, 704–6.

Elliott, M.J. *et al*. (1993). Treatment of rheumatoid arthritis with chimeric monoclonal antibodies to tumor necrosis factor-a. *Arthritis and Rheumatism*, **36**, 1681–9.

Elliott, M.J. *et al*. (1994). Randomised double-blind comparison of chimeric monoclonal antibody to tumour necrosis factor-a (cA2) versus placebo in rheumatoid arthritis. *Lancet*, **344**, 1105–10.

Ermak, T.H., Steiger, H.J., and Wofsy, D. (1989). Treatment of murine lupus with monoclonal antibody to L3T4. II. Effects on immunohistopathology of thymus, spleen, and lymph node. *Laboratory Investigation*, **61**, 447–56.

Evans, C. and Robbins, P.D. (1994). Prospects for treating arthritis by gene therapy. *Journal of Rheumatology*, **21**, 779–82.

Fairchild, P.J., Thorpe, C.J., Travers, P.J., and Wraith, D.C. (1994). Modulation of the immune response with T-cell epitopes: the ultimate goal for specific immunotherapy of autoimmune disease. *Immunology*, **81**, 487–96.

Finck, B.K., Linsley, P.S., and Wofsy, D. (1994). Treatment of murine lupus with CTLA4Ig. *Science*, **265**, 1225–7.

Firestein, G. and Zvaifler, N.J. (1990). How important are T cells in chronic rheumatoid synovitis? *Arthritis and Rheumatism*, **33**, 768–73.

Geusens, P., Wouters, C., Nijs, J., Jiang, Y., and Dequeker, J. (1994). Long-term effect of -3 fatty acid supplementation in active rheumatoid arthritis. A 12 month, double-blind controlled study. *Arthritis and Rheumatism*, **37**, 824–9.

Goldberg, D. *et al*. (1991). Immunological effects of high dose administration of anti-CD4 antibody in rheumatoid arhtritis patients. *Journal of Autoimmunity*, **4**, 617–30.

Gordon, C., Ranges, G.E., Greenspan, J.S., and Wofsy, D. (1989). Chronic treatment with recombinant tumor necrosis factor-a in autoimmune NZB/NZW F_1 mice. *Clinical Immunology and Immunopathology*, **52**, 421–7.

Gottlieb, S.L. *et al*. (1995). Response of psoriasis to a lymphocyte-selective toxin (DAB389IL-2) suggests a primary immune, but not keratinocyte, pathogenic basis. *Nature Medicine*, **1**, 442–7.

Hahn, G., Stuhlmeller, B., Hain, N., Kalden, J.R., Pfizenmaier, K., and Burmester, G.R. (1993). Modulation of monocyte activation in patients with rheumatoid arthritis by leukapheresis therapy. *Journal of Clinical Investigation*, **91**, 862–70.

Haugen, M.A. *et al*. (1993). The influence of fast and vegetarian diet on parameters of nutritional status in patients with rheumatoid arthritis. *Clinical Rheumatology*, **12**, 62–9.

Haugen, M., Kjeldsen-Kragh, J., and Forre, O. (1994*a*). A pilot study of the effects of an elemental diet in the management of rheumatoid arthritis. *Clinical and Experimental Rheumatology*, **12**, 275–9.

Haugen, M.A., Kjeldsen-Krogh, J., Bjerve, K.S., Hostmark, A.T., and Forre, O. (1994 *b*). Changes in plasma phospholipid fatty acids and their relationship to disease activity in rheumatoid arthritis

patients treated with a vegetarian diet. *British Journal of Nutrition*, **72**, 555–66.

Heber-Katz, E. (1990). The autoimmune T cell receptor: epitopes, idiotopes, and malatopes. *Clinical Immunology and Immunopathology*, **55**, 1–8.

Hentati, B., Notomi Sato, M., Payelle-Brogard, B., Avrameas, S., and Ternynck, T. (1994). Beneficial effects of polyclonal immunoglobulins from malaria-infected BALBc mice on the lupus-like syndrome of (NZB × NZW)F_1 mice. *European Journal of Immunology*, **24**, 8–15.

Horneff, G., Burmester, G.R., Emmrich, F., and Kalden, J.R. (1991 *a*). Treatment of rheumatoid arthritis with an anti-CD4 monoclonal antibody. *Arthritis and Rheumatism*, **34**, 129–40.

Horneff, G., Winkler, T., Kalden, J.R., Emmrich, F., and Burmester, G.R. (1991 *b*). Human anti-mouse antibody responses induced by anti-CD4 monoclonal antibody therapy in patients with rheumatoid arthritis. *Clinical Immunology and Immunopathology*, **59**, 89–103.

Horneff, G., Emmrich, F., Reiter, C., Kalden, J.R., and Burmester, G.R. (1992). Persistent depletion of CD4+ T cells and inversion of the CD4/CD8 T cell ratio induced by anti-CD4 therapy. *Journal of Rheumatology*, **19**, 1845–50.

Horneff, G., Sack, U., Kalden, J.R., Emmrich, F., and Burmester, G.R. (1993). Reduction of monocyte-macrophage activation markers upon anti-CD4 treatment. Decreased levels of IL-1, IL-6, neopterin and soluble CD14 in patients with rheumatoid arthritis. *Clinical and Experimental Immunology*, **91**, 207–13.

Hutchings, P., Rosen, H., O'Reilly, L., Simpson, E., Gordon, S., and Cooke, A. (1990). Transfer of diabetes in mice prevented by blockage of adhesion-promoting receptor on macrophages. *Nature*, **348**, 639–42.

Karlsson-Parra, A., Soderstrom, K., Ferm, M., Ivanvi, J., Kiessling, R., and Klareskog, L. (1990). Presence of human 65 kDa heat shock protein (hsp) in inflamed joints and subcutaneous nodules of RA patients. *Scandinavian Journal of Immunology*, **31**, 283–8.

Katsikis, P.D., Chu, C.-Q., Brennan, F.M., Maini, R.N., and Feldmann, M. (1994). Immunoregulatory role of interleukin 10 in rheumatoid arthritis. *Journal of Experimental Medicine*, **179**, 1517–27.

Kavanagh, A.F. *et al*. (1994). Treatment of refractory rheumatoid arthritis with a monoclonal antibody to intercellular adhesion molecule 1. *Arthritis and Rheumatism*, **37**, 992–9.

Kavanagh, S., Workman, E., Nash, P., Smith, M., Hazleman, B.L., and Hunter, J.O. (1995). The effects of elemental diet and subsequent food reintroduction on rheumatoid arthritis. *British Journal of Rheumatology*, **34**, 270–3.

Kelley, V.E., Gaulton, G.N., Hattori, M., Ikeghami, H., Eisenbarth, G., and Strom, T.B. (1988). Anti-interleukin 2 receptor antibody suppresses murine diabetic insulitis and lupus nephritis. *Journal of Immunology*, **140**, 59–61.

Kjeldsen-Kragh, J. *et al*. (1991). Controlled trial of fasting and one-year vegetarian diet in rheumatoid arthritis. *Lancet*, **338**, 899–902.

Kjeldsen-Kragh, J. *et al*. (1992). Dietary -3 fatty acid supplementation and naproxen treatment in patients with rheumatoid arthritis. *Journal of Rheumatology*, **19**, 1531–6.

Kjeldsen-Kragh, J., Haugen, M., Borchgrevink, C.F., and Forre, O. (1994 *a*). Vegetarian diet for patients with rheumatoid arhtritis—status two years after introduction of the diet. *Clinical Rheumatology*, **13**, 475–82.

Kjeldsen-Kragh, J., Haugen, M., Forre, O., Laache, H., and Malt, U.F. (1994 *b*). Vegetarian diet for patients with rheumatoid arthritis: can the clinical effects be explained by the psychological characteristics of the patients? *British Journal of Rheumatology*, **33**, 569–75.

Kjeldsen-Kragh, J. *et al*. (1995). Changes in laboratory variables in rheumatoid arthritis patients during a trial of fasting and one year vegetarian diet. *Scandinavian Journal of Rheumatology*, **24**, 85–93.

Kumar, V., Kono, D.H., Urban, J.L., and Hood, L. (1989). The T-cell receptor repertoire and autoimmune diseases. *Annual Review of Immunology*, **7**, 657–82.

Kyle, V., Coughlan, R.J., Tighe, H., Waldman, H., and Hazieman, B.L. (1989). Beneficial effect of monoclonal antibody to interleukin 2 receptor on activated T cells in rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **48**, 428–9.

Littman, B.H., Drury, C.E., Zimmerer, R.O., Stack, C.B., and Law, C.G. (1995). Rheumatoid arthritis treated with tenidap and piroxicam. *Arthritis and Rheumatism*, **38**, 29–37.

Mathieson, P.W. *et al*. (1990). Monoclonal-antibody therapy in systemic vasculitis. *New England Journal of Medicine*, **323**, 250–4.

Mendlovic, S., Brocke, S., Fricke, H., Shoenfeld, Y., Bakimer, R., and Mozes, E. (1990). The genetic regulation of the induction of experimental SLE. *Immunology*, **69**, 228–36.

Mor, F., Lohse, A.W., Karin, N., and Cohen, I.R. (1990). Clinical modeling of T cell vaccination against autoimmune disease in rats. *Journal of Clinical Investigation*, **85**, 1594–8.

Moreland, L.W., Pratt, P.W., Sanders, M.E., and Koopman, W.J. (1993). Experience with a chimeric monoclonal anti-CD4 antibody in the treatment of refractory rheumatoid arthritis. *Clinical and Experimental Rheumatology*, **11** (Suppl. 8), S153–9.

Moreland, L.W., Pratt, P.W., Bucy, R.P., Jackson, B.S., Feldman, J.W., and Koopman, W.J. (1994). Treatment of refractory rheumatoid arthritis with a chimeric anti-CD4 monoclonal antibody. Long-term followup of CD4+ T cell counts. *Arthritis and Rheumatism*, **37**, 834–8.

Mountz, J.D., Gause, W.C., Finkelman, F.D., and Steinberg, A.D. (1988). Prevention of lymphadenopathy in MRL-lpr/lpr mice by blocking peripheral lymph node homing with mcl-14 *in vivo*. *Journal of Immunology*, **140**, 2943–9.

Nielsen, G.L. *et al*. (1992). The effects of dietary supplementation with n-3 polyunsaturated fatty acids: a randomized, double-blind trial. *European Journal of Clinical Investigation*, **22**, 687–91.

Nietfeld, J.J. *et al*. (1994). Antisense nucleotides, a novel tool for the control of cytokine effects on human cartilage. *Arthritis and Rheumatism*, **37**, 1357–62.

Peltonen, R. *et al*. (1994). Changes of faecal flora in rheumatoid arthritis during fasting and one-year vegetarian diet. *Clinical and Experimental Rheumatology*, **33**, 638–43.

Raffeld, M., Neckers, L., Longo, D.L., and Cosman, J. (1985). Spontaneous alteration of idiotype in a monoclonal B-cell lymphoma. *New England Journal of Medicine*, **312**, 1653–8.

Rankin, E.C.C. *et al*. (1995). The therapeutic effects of an engineered human anti-tumour necrosis factor-a antibody (CDP 571) in rheumatoid arthritis. *British Journal of Rheumatology*, **34**, 334–42.

Rauma, A.L., Nenonem, M., Helve, T., and Hanninen, O. (1993). Effect of a strict diet on energy and nutrient intakes by Finnish rheumatoid patients. *European Journal of Clinical Nutrition*, **47**, 747–9.

Reiter, C., Kakavand, B., Rieber, E.P., Schattenkichner, M., Riethmuller, G., and Kruger, K. (1991). Treatment of rheumatoid arthritis with monoclonal CD4 antibody M-T151. Clinical results and immunopharmacologic effects in an open study, including repeated administration. *Arthritis and Rheumatism*, **34**, 525–36.

Rott, O., Fleischer, B., and Cash, E. (1994). Interleukin-10 prevents experimental allergic encephalomyelitis in rats. *European Journal of Immunology*, **24**, 1430–40.

Sadouk, M., Vaquero, C., de la Tour, B., Amor, B., and Toubert, A. (1990). Interferon-g mRNA expression upon *in vitro* T lymphocyte activation is decreased in rheumatoid arthritis patients. *Clinical Immunology and Immunopathology*, **56**, 37–45.

Sany, J. (1988). Treatment of rheumatoid arthritis by antibodies directed against class II MHC antigens. *Scandinavian Journal of Rheumatology*, **S76**, 289–95.

Sewell, K.L., Parker, K.C., Woodworth, T.G., Reuben, J., Swartz, W., and Trentham, D.E. (1993). DAB486 IL-2 fusion toxin in refractory rheumatoid arthritis. *Arthritis and Rheumatism*, **36**, 1223–33.

Skoldstam, L. and Magnusson, K.E. (1991). Fasting, intestinal permeability, and rheumatoid arthritis. *Rheumatic Disease Clinics of North America*, **17**, 363–71.

Strand, V. *et al*. (1993). Effects of administration of an anti-CD5 plus immunoconjugate in rheumatoid arthritis. Results of two phase II studies. *Arthritis and Rheumatism*, **36**, 620–30.

Strober, S. *et al*. (1988). Treatment of lupus nephritis with total lymphoid irradiation; observations during a 12–79 month follow up. *Arthritis and Rheumatism*, **31**, 850–80.

Swanson, M.A. and Schwartz, R.S. (1967). Immunosuppressive therapy: the relation between clinical response and immunologic competence. *New England Journal of Medicine*, **277**, 163–70.

t'Hart, B.A., Bakker, N.P.M., Jonker, M., and Bontrop, R.E. (1993). Resistance to collagen-induced arthritis in rats and rhesus monkeys after immunisation with attenuated type II collagen. *European Journal of Immunology*, **23**, 1588–94.

Tanaka, Y., Saito, K., Suzuki, H., Eto, S., and Yamashita, U. (1989). Inhibitory effect of anti-class II antibody on the spontaneous activation of B cells in patients with systemic lupus erythematosus.

New England Journal of Medicine, **143**, 1584–90.

Thompson, K.M. *et al*. (1990). Human monoclonal rheumatoid factors derived from the polyclonal repertoire of rheumatoid synovial membrane: incidence of cross-reactive idiotypes and expression of VH and VK groups. *European Journal of Immuology*, **20**, 863–8.

Tong, Z. and Dwyer, D.S. (1990). Alteration of the humoral immune response against muscle acetylcholine receptor by timed administration of a(1–3) dextran. *European Journal of Immunology*, **20**, 1627–34.

Trentham, D. *et al*.(1993). Effects of oral administration of type II collagen on rheumatoid arthritis. *Science*, **261**, 1727–30.

Tulleken, J.E., Limburg, P.C., Muskiet, F.A.J., and van Riiswiik, M.H. (1990). Vitamin E status during dietary fish oil supplementation in rheumatoid arthritis. *Arthritis and Rheumatism*, **33**, 1416–19.

Van de Laar, M.A. and Van der Korst, J.K. (1992). Food intolerance in rheumatoid arthritis. I. A double blind, controlled trial of the clinical effects of elimination of milk allergens and azo dyes. *Annals of the Rheumatic Diseases*, **51**, 298–302.

Van de Laar, M.A., Aalbers, M., Bruins, F.G., van Dinther-Janssen, A.C., van der Korst, J.K., and Meijer, C.J. (1992). Food intolerance in rheumatoid arthritis. II Clinical and histological aspects. *Annals of the Rheumatic Diseases*, **51**, 303–6.

Van den Berg, W.B., Joosten, L.A.B., Helsen, M., and Van de Loo, F.A.J. (1994). Amelioration of established murine collagen-induced arthritis with anti-IL-1 treatment. *Clinical and Experimental Immunology*, **95**, 237–43.

Van der Lubbe, P.A. *et al*. (1993). Chimeric CD4 monoclonal antibody cM-1412 as a therapeutic approach to rheumatoid arthritis. *Arthritis and Rheumatism*, **36**, 1375–9.

Van der Lubbe, P.A. *et al*. (1994). Treatment of rheumatoid arthritis with a chimeric CD4 monoclonal antibody (cM-1412): immunopharmacological aspects and mechanisms of action. *Scandinavian Journal of Immunology*, **39**, 286–94.

Van der Lubbe, P.A., Dijkmans, B.A.C., Markusse H.M., Nassander, U., and Breedveld, F.C. (1995). A randomised, double-blind, placebo controlled study of CD4 monoclonal antibody therapy in early rheumatoid arthritis. *Arthritis and Rheumatism*, **38**, 1097–1106.

Vladutiu, A.O. (1991). Treatment of autoimmune diseases with antibodies to class II major histocompatibility antigens. *Clinical Immunology and Immunopathology*, **61**, 1–17.

Waldmann, H. (1993). The use of monoclonal antibodies to achieve immunological tolerance. *Immunology Today*, **14**, 247–51.

Watson, J., Byars, M.L., McGill, P., and Kelman, A.W. (1993). Cytokine and prostaglandin production by monocytes of volunteers and rheumatoid arthritis patients treated with dietary supplements of blackcurrant seed oil. *British Journal of Rheumatology*, **32**, 1055–8.

Weiner, H.L. *et al*. (1994). Oral tolerance: immunologic mechanisms and treatment of animal and human organ-specific autoimmune diseases by oral administration of autoantigens. *Annual Review of Immunology*, **12**, 809–37.

Weisbart, R.H., Noritake, D.T., Wong, A.L., Chan, G., Kacena, A., and Colburn, K.K. (1990). A conserved anti-DNA antibody idiotype associated with nephritis in murine and human systemic lupus erythematosus. *Journal of Immunology*, **144**, 2653–8.

Wendling, D., Wijdenes, J., Racadot, E., and Morel-Fourrier, B. (1991). Therapeutic use of monoclonal anti-CD4 antibody in rheumatoid arthritis. *Journal of Rheumatology*, **18**, 325–7.

Wendling, D., Racodot, E., Morel-Fourrier, B., and Wijdenes, J. (1992). Treatment of rheumatoid arthritis with anti-CD4 monoclonal antibody. Open study of 25 patients with the B-F5 clone. *Clinical Rheumatology*, **11**, 542–7.

Williams, R.O., Feldmann, M., and Maini, R.N. (1992). Anti-tumour necrosis factor ameliorates joint disease in murine collagen-induced arthritis. *Proceedings of the National Academy of Sciences* (*USA*), **89**, 9784–8.

Winkler, T.H., Henschel, T.A., Kalies, I., Baenkler, H.-W., Skvaril, F., and Kalden, J.R. (1988). Constant isotype pattern of anti-dsDNA antibodies in patients with systemic lupus erythematosus. *Clinical and Experimental Immunology*, **72**, 434–9.

Wofsy, D. (1988). Treatment of autoimmune diseases with monoclonal antibodies. In *Monoclonal antibody therapy. Progress in allergy*, Vol. 45 (ed. H. Waldmann), pp.106–20. Karger, Basel.

Wooley, P.H. *et al*. (1993*a*). The effect of an interleukin-1 receptor antagonist protein type II collagen-induced arthritis and antigen-induced arthritis in mice. *Arthritis and Rheumatism*, **36**, 1305–14.

Wooley, P.H., Dutcher, J., Widmer, M.B., and Gillis, S. (1993 *b*). Influence of a recombinant human soluble tumour necrosis factor receptor Fc fusion protein on type II collagen-induced arthritis in mice. *Journal of Immunology*, **151**, 6602–7.

Wraith, D.C., McDevitt, H.O., Steinman, L., and Acha-Orbea, H. (1989). T cell recognition as the target of immune intervention in autoimmune disease. *Cell*, **57**, 709–15.

Wylie, G. *et al*. (1995). A comparative study of tenidap, a cytokine modulating anti-rheumatic drug, and diclofenac in rheumatoid arthritis: a 24-week analysis of a 1-year clinical trial. *British Journal of Rheumatology*, **34**, 554–63.

Yakura, H., Ashida, T., Kawabata, I., and Katigiri, M. (1989). Allelviation of autoimmunity in BXSB mice by monoclonal antibody to Ly-5 (CD45). *European Journal of Immunology*, **19**, 1505–8.

4.1 Acute phase response

Rosamonde E. Banks, John T. Whicher, Douglas Thompson and Howard A. Bird

What is the acute phase response and how is it assessed?

The acute phase response

The acute phase response is the term generally used to describe the concerted series of diverse systemic and local events, which accompany inflammation resulting from tissue damage. Despite the term 'acute phase', the changes may be present in both acute and chronic inflammatory situations. These events are usually considered to be protective in nature, facilitating the inflammatory process and ultimately culminating in removal of tissue debris and foreign organisms, and tissue repair. They include fever, endocrine changes, leucocytosis, muscle proteolysis, and an increase in the synthesis of specific liver-derived proteins—the acute phase proteins. It is this latter event which is often used as an indicator of the presence and extent of the inflammatory process and indeed it is often used synonymously with the term 'acute phase response'. We will also follow this convention for the purposes of this chapter.

There are approximately 30 acute phase proteins. They fall into several different structural and functional categories of plasma proteins and fulfil various roles in inflammation such as inhibition of phagocyte-derived proteases, scavenging of cellular debris, and mediation and inhibition of the inflammatory process (Table 1) (reviewed by Laurell 1985; Kushner and Mackiewicz 1987; Whicher *et al*. 1989; Schultz and Arnold 1990; Mackiewicz *et al*. 1993; Steel and Whitehead 1994). The increase in serum concentration of these proteins (which by definition must be by at least 25 per cent in the first 7 days following tissue damage (Kushner 1982) probably serves to maintain their circulating and tissue levels, which may otherwise become depleted as a result of their consumption at the site of inflammation.

Table 1 The positive acute phase proteins and their proposed functions (adapted from Whicher *et al*. 1989)

The increased synthesis of the acute phase proteins by the liver has been ascribed to several cytokines, including interleukin 1 (IL-1), tumour necrosis factor-a (TNF-a) and particularly interleukin 6 (IL-6). These cytokines are predominantly derived from activated macrophages at the site of injury although they can be produced by many other cell types such as fibroblasts and endothelial cells, and they have also been implicated in mediating other aspects of the acute response such as fever (see Chapter 3.1). From extensive experimental work, principally involving cultured hepatoma cell lines, it is apparent that the regulation of the acute phase proteins is complex and that subsets of acute phase proteins are induced by different cytokines (reviewed by Baumann and Gauldie 1994). The IL-1 group of cytokines, IL-1a and b and TNF-a and b, induce the production of so-called type-1 acute phase proteins which include a $_1$-acid glycoprotein, serum amyloid A protein, C-reactive protein, and complement component C3, whereas the IL-6 family of cytokines (IL-6, IL-11, leukaemia inhibitory factor, oncostatin M, and ciliary neurotrophic factor) induce the synthesis of type-2, IL-6 specific, acute phase proteins, fibrinogen, a $_1$-antitrypsin, a$_1$-antichymotrypsin, haptoglobin, haemopexin, and caeruloplasmin. In most cases IL-1 or TNF have no effect on the production of type-2 acute phase proteins but the IL-6 group of cytokines can induce type-1 acute phase proteins and synergistically enhance the IL-1 and TNF-mediated production of the type-1 acute phase proteins. However to a certain extent these actions depend on the species and hepatocytes being examined. In addition, two groups of factors modulate cytokine-induced acute phase protein synthesis, namely the glucocorticoids and the growth factors such as insulin, insulin-like growth factor, hepatic growth factor, fibroblast growth factor and transforming growth factor-b (Baumann *et al*. 1987; Thompson *et al*. 1991). The net result achieved is dependent on the overall interactions (Fig. 1).

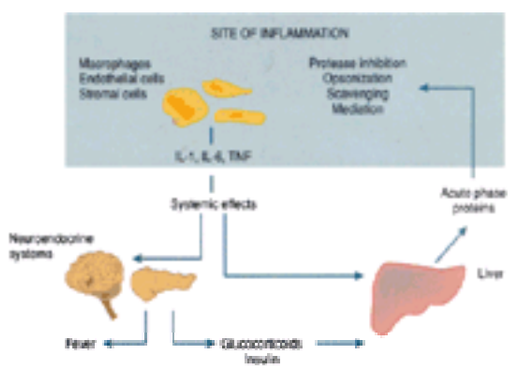


Fig. 1 The regulation of the acute phase protein response. The systemic acute phase protein response is mediated by cytokines such as IL-6, IL-1, and TNF-a released by activated macrophages, endothelial cells and stromal cells at the site of inflammation. The systemic effects of these cytokines include muscle proteolysis, fever, and the production of acute phase proteins, which exert a range of function at the site of inflammation. The synthesis of acute phase proteins can also be

modulated by hormones such as the glucocorticoids and insulin.

Most of the changes in acute phase protein production are thought to be mediated at the transcriptional level, however post-transcriptional modulation such as increases in stability of mRNA for positive acute phase proteins and altered glycosylation and hence possibly bioactivity/stability may also contribute to their regulation (reviewed by Steel and Whitehead 1994). Although the liver is generally accepted as being the site of acute phase protein synthesis, it is now recognized that extrahepatic tissues also have this capability. It has been shown for example, that cells of the monocyte–macrophage lineage can produce several complement components and a_1-antitrypsin. The significance of this and the extent to which it normally occurs in vivo, particularly at the site of inflammation, is not known.

A small group of proteins including albumin, transthyretin, retinol-binding protein and transferrin decrease in serum concentration during inflammation. These are often referred to as the 'negative acute phase proteins'. This decrease has been attributed to increased catabolism and vascular permeability although recent studies indicate that cytokines are capable of decreasing transcription of at least some of the negative acute phase protein genes (reviewed by Mackiewicz *et al*. 1993; Steel and Whitehead 1994).

Kinetics of the acute phase response in disease

The rate and extent of the increase in plasma concentration of any acute phase protein depends on factors such as molecular size, volume of distribution, rate and sensitivity to induction, and the rate of catabolism. Considerable incremental and kinetic differences are therefore seen between the various proteins. In acute inflammation, the overall pattern of the response is relatively constant, provided that the magnitude of the response is sufficient to elicit a detectable change in all the proteins. The properties of the major acute phase proteins and typical magnitudes and time courses of response are shown in Table 2. It is important to appreciate the differences in kinetics of these changes, for example while some proteins are elevated within 6 to 8 h of stimulus, there is no apparent change in others for appreciably longer. These relative changes in the acute phase proteins are exemplified in Fig. 2 which shows the changes in positive acute phase proteins in vivo following administration of IL-6.

Protein	Response time (h)	Molecular weight	Adult reference ranges
Group II: <1 × increase			
Caeruloplasmin	48–72	132 000	0.20–0.60 g/l (30–600 units)
Complement components C3	48–72	180 000	0.75–1.65 g/l (30–600 units)
C4	48–72	206 000	0.20–0.60 g/l (30–600 units)
Group II: 2–4 × increase			
α_1-Acid glycoprotein	24	41 000	0.5–1.2 g/l (male); 0.4–1.0 g/l (female); 5 to 10 years of age
α_1-Antitrypsin	10	54 000	1.3–2.1 g/l (30–600 units)
α_1-Antichymotrypsin	10	68 000	0.3–0.6 g/l
Haptoglobin	24	86 000	0.3–2.0 g/l (male); 0.4–1.6 g/l (female)
Fibrinogen	24	340 000	2.0–4.0 g/l
Group III: up to 1000 × increase			
C-reactive protein	8–10	110 000	0.008–0.3 mg/l
Serum amyloid A	8–10	180 000	

Adapted from Kushner and Mackiewicz (1987).
The normal ranges are those published by the Protein Reference Unit (1983) with the exception of fibrinogen (Blann and Lowe 1993) and C-reactive protein (Cross et al 1979). Levels of many of these proteins are affected by factors such as age and pregnancy (Kluft and Ramsoek 1979; Altman et al 1986).

Table 2 Properties of the major acute phase proteins

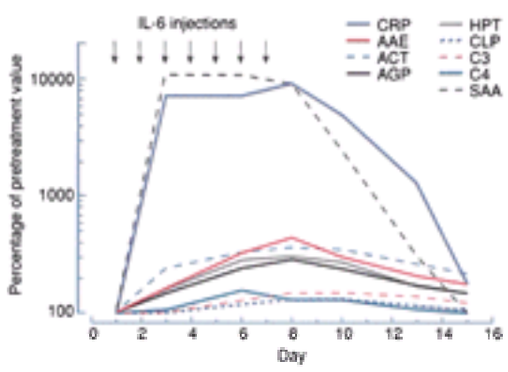


Fig. 2 Profile of the relative changes in the positive acute phase proteins following IL-6 treatment in a patient with cancer. IL-6 was administered subcutaneously once daily for 7 days (arrows) at a dose of 10 µg/kg per day (reproduced from Banks *et al*. 1995, with permission).

Generally the magnitude of the response is related quantitatively to the activity or extent of inflammation in the acute situation. In bacterial infection, an extremely large response is seen, probably as a result of systemic endotoxin-induced release of cytokines such as IL-1, IL-6, and TNF from macrophages. However, particularly in chronic inflammation, differential increases in catabolic rate may also result in altered 'profiles' for several proteins. For example when vasculitis is present, concentrations of a_1-antitrypsin are inappropriately low when compared with a group of other acute phase proteins, owing to its consumption by leucocyte enzymes. Similarly if intravascular coagulation is a complication, levels of fibrinogen may be lower than would otherwise be expected and haptoglobin is lower if haemolysis is occurring. Such differences in themselves may be useful in drawing attention to possible complications which may be present with the inflammatory process, but would contraindicate the effective use of that particular protein as an inflammatory marker in that particular situation. Often in chronic inflammation the magnitude of the overall response is lower than expected for the degree of inflammatory activity, possibly as a result of down-regulation. As will be discussed later in this chapter, some diseases show little or no acute phase response even though acute inflammation is clinically apparent.

Laboratory measurements of the acute phase response (see Chapter 4.2)

The main parameters usually measured to assess the acute phase response are the erythrocyte sedimentation rate (ESR), plasma viscosity, and the plasma concentrations of the acute phase proteins. General guidelines as to the relative merits of these tests have been drawn up by the International Committee for Standardization in Haematology (ICSH 1988). In view of the role of cytokines in the control of the acute phase response, it may be relevant in the future to include plasma levels of cytokines as an additional parameter to be used. Many cytokines can now be assayed either by bioassay or by immunoassay, but problems exist such as lack of sensitivity, poor inter-laboratory reproducibility, and unknown effects of inhibitors. However the use of cytokine assays in the clinical situation is likely to increase in the future with further research. This area has already been covered in Chapter 3.1 and will not be discussed further here.

Erythrocyte sedimentation rate (ESR) and plasma viscosity

The ESR is the oldest and probably still the most widely used index of the acute phase response (reviewed by Zlonis 1993). It is a laboratory measurement of the rate of sedimentation of erythrocytes which is dependent on their degree of aggregation and the packed cell volume (PCV). Aggregation is strongly influenced by the concentration of large asymmetric plasma proteins such as fibrinogen, a_2-macroglobulin, and immunoglobulins. The ESR is therefore a composite measurement reflecting changes in several proteins.

There are several disadvantages to this assay. It is influenced by red cell number and characteristics, abnormal immunoglobulins or complexes such as occur in myeloma or cryoglobulinaemia, age and sex, smoking, menstrual cycle, drugs, and dietary lipids therefore resulting in a circadian rhythm. In addition the ESR in the acute phase is mainly influenced by fibrinogen, an acute phase protein which is slow to increase and persists long after the inflammation subsides because of its long half-life. Its reliance on fibrinogen levels means that the ESR is often insensitive and relates poorly to the time-course of the inflammation, and is influenced by complications which result in the consumption of fibrinogen, such as the presence of intravascular coagulation. Due to the relative instability of red blood cells in terms

of their deformability and aggregation, stored samples cannot be used. Standardization and quality control are difficult although recent recommendations may improve this (ICSH 1993).

Plasma viscosity is increasingly being used in place of the ESR and has several advantages as a routine measurement (Lowe 1994). It is dependent on the concentrations of the same group of large-molecular-weight proteins but shows less interindividual variation than the ESR. This is at least partly because the plasma viscosity, unlike the ESR, is affected minimally by age, sex or pregnancy, and it is not affected by anaemia or red cell morphology. In addition it is a more rapid test to perform, easier to standardize and stored samples are suitable for use.

Despite the above reservations about the ESR and plasma viscosity, these are technically simple and inexpensive assays to perform and may provide a better screening test for the detection of many diseases than the measurement of a single acute phase protein. For example the ESR integrates the effects of anaemia, immune response and acute phase response. These tests thus still play an important role in the detection and monitoring of chronic inflammation where the hyperproteinaemia is more complex than in the acute situation, involving changes in other pathologically important proteins such as immunoglobulins.

Acute phase proteins

Individual acute phase proteins are usually assayed using specific antisera, the precise immunochemical technique being dependent on the particular laboratory. The techniques most often used are radial immunodiffusion (RID), enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), and nephelometry. Proposed functions of the acute phase proteins have been given in Table 1 and the properties of the major acute phase proteins together with their reference ranges are shown in Table 2. However it must be emphasized that reference ranges vary between laboratories and clinicians should consult their own laboratories for appropriate figures. The relative merits of the major acute phase proteins as markers of inflammation are discussed below. Detailed reviews have been published by Laurell (1985) and Thompson *et al*. (1992).

C-reactive protein (CRP)

C-reactive protein is a non-glycosylated protein composed of five identical 21 kDa non-covalently bound globular subunits, a member of the pentraxin family. It is the most useful index of the acute phase response being specific, sensitive, and rising rapidly within 6 to10 h following the inflammatory event. This time lag represents the time for complete synthesis of the molecule as no hepatic store exists. Levels reach a peak by about 48 h and, because of its short half-life of approximately 19 h in man (Vigushin *et al*. 1993), decline rapidly following the decrease in inflammatory activity. It has been shown in humans that the catabolism of CRP is independent of its plasma concentration and is not affected by the presence of inflammatory illnesses such as rheumatoid arthritis, systemic lupus erythematosus or infections (Vigushin *et al*. 1993). This, coupled with the fact that there is no significant tissue sequestration of CRP, indicates that the circulating concentration is determined solely by the its rate of synthesis and hence inflammatory stimulus. However the normal range for CRP is very wide from 0.068 to 8.0 mg/l with a median value of 0.58 mg/l (Claus *et al*. 1976), and with many routine assay methods only having a lower limit of detection of between 5 and 10 mg/l, it is apparent that CRP concentrations in some individuals could increase more than a hundredfold and still remain within normal limits. It is therefore an unsuitable protein for detecting very mild inflammation unless sensitive assays are used, serial determinations are made and the normal values are known for the individual. Typically, mild infections are associated with CRP levels of between 10 and 40 mg/l (Peltola 1982), acute inflammation and moderate bacterial infection with levels of between 40 and 200 mg/l, and severe bacterial infection can produce levels in excess of 300 mg/l (Morley and Kushner 1982). For detecting or monitoring chronic inflammation, an additional slower reacting acute phase protein such as a_1-antichymotrypsin or a_1-acid glycoprotein or the ESR should also be measured as CRP may be transiently normal during brief remission periods. A World Health Organization (WHO) international standard is available.

Serum amyloid A protein (SAA)

SAA proteins are a family of at least three different apolipoproteins with similar kinetics to C-reactive protein but possibly greater sensitivity (reviewed by Malle *et al*. 1993). It is elevated even in relatively mild conditions such as the common cold. Higher levels of SAA have been reported in diseases which carry an increased risk of amyloidosis such as rheumatoid arthritis and systemic juvenile rheumatoid arthritis (De Beer *et al*. 1982) but as SAA is not routinely measured at present, for a variety of reasons including scarcity of suitable antisera and lack of an appropriate standard, it will not be covered further here.

a_1-Antichymotrypsin (ACT)

a_1-Antichymotrypsin is a glycoprotein consisting of a single polypeptide chain, with oligosaccharide chains accounting for approximately 23 per cent of its molecular weight. Measurement of a_1-antichymotrypsin is often available routinely and normal levels lie in the range of 0.3 to 0.6 g/l. It increases almost as rapidly as C-reactive protein and SAA following tissue injury, although normally only increasing two- to fourfold. It remains elevated for longer than C-reactive protein, is easy to measure, and does not suffer from genetic variation or differential catabolism. There is however an inter-laboratory standardization problem as values obtained are dependent on the antibody and standard used. At present relatively little is known about its levels in different inflammatory states.

Haptoglobin

Haptoglobin is a member of a family of polymers of a-b-chains with the a-chain having two main genetic variants. It is much slower in response but can be measured routinely. It is generally unsuitable as an acute phase response indicator however, because of its consumption associated with haemolysis, and genetic variations giving rise to phenotypic reference anges.

a_1-Antitrypsin (AAT)

The glycoprotein a_1-antitrypsin (a_1-proteinase inhibitor) can be routinely assayed and standardized with WHO preparations. It is generally unsuitable as an indicator of the acute phase response as genetic variants give rise to decreased serum levels. Approximately 15 per cent of the Caucasian population are heterozygous and have lower normal levels than the 1.1 to 2.1 g/l found in the normal homozygous MM phenotype. In addition levels increase in pregnancy or with oestrogen treatment, and may be inappropriately lowered in vasculitis because of its consumption.

Fibrinogen

Fibrinogen is composed of three pairs of non-identical polypeptide chains, two of which are glycosylated. Fibrinogen is slow to exhibit changes in concentration during inflammation and is relatively insensitive. Additionally, it is difficult to assay because gel-based techniques such as radial immunodiffusion (RID) are unsuitable as fibrinogen has a tendency to precipitate, particularly after storage.

a_1-Acid glycoprotein (AGP)

a_1-Acid glycoprotein (orosomucoid), as its name suggests, is glycosylated with polysaccharide chains accounting for 45 per cent of the molecule. Genetic variants exist although these have no effect on normal serum ranges. Levels of a_1-acid glycoprotein are not measured routinely being less sensitive than C-reactive protein or a_1-antichymotrypsin. In addition, concentrations may be decreased in pregnancy or during oestrogen treatment and a disproportionate increase in a_1-acid glycoprotein levels may occur in chronic renal disease, resulting from a reduction in the glomerular filtration rate.

C3, C4, and caeruloplasmin

Complement components C3 and C4, and the copper-binding glycoprotein caeruloplasmin are very insensitive and too slow to increase to be of use as markers of the acute phase response, although low C3 or C4 concentrations indicating their consumption may be of use for monitoring of immunologically-based diseases.

Clinical use of the acute phase response

Although the presence of an acute phase response is associated unequivocally with inflammation, the reverse is not always true. Inflammation occurring with mild chronic tissue damage, localized disease, recurrent attacks, or certain diseases including ulcerative colitis and systemic lupus erythematosus is often found

associated with a normal ESR or CRP level. The acute phase response by nature is a non-specific response in terms of the disease and therefore must be used in conjunction with other more specific tests for diagnostic purposes as it can be elevated in many pathological conditions and indeed some acute phase proteins such as a_1-antitrypsin and caeruloplasmin also exhibit elevated levels in physiological circumstances such as pregnancy (Laurell and Rannevik 1979). Although the magnitude of the acute phase response is related generally to the mass or activity of inflammation, this is not always the case. As previously indicated, an extremely large acute phase response is seen in bacterial infections for example, presumably as a result of the systemic endotoxin-induced release of IL-1, IL-6, and TNF from macrophages.

However, together with other clinical and laboratory measurements, the assessment of the acute phase response is generally used in three main ways to assist the rheumatologist.

1. detection of organic disease;
2. assessment of extent or activity of the disease, monitoring of therapy, and indication of prognosis;
3. detection of intercurrent infection.

There is no acute phase response in patients with soft tissue rheumatism or osteoporosis. The use of the acute phase response is given below for each of the main disease groups within the arthritides.

Crystal deposition diseases

The acute phase response may be only mild or absent in crystal deposition diseases unless accompanied by superimposed infection in the joints. In gout however, an acute phase response is often seen and C-reactive protein levels may reach in excess of 100 mg/1, even in the absence of intercurrent infection (Roseff *et al*. 1987).

Osteoarthritis

Primary generalized or nodal osteoarthritis is not associated with an acute phase response, the ESR being normal or only mildly elevated in the majority of cases (reviewed by Altman and Gray 1985). This may be partly age-related. No difference is seen between erosive and non-erosive osteoarthritis and there is no practical value in attempting to determine the severity of episodes of inflammation by the measurement of the acute phase response. Plasma viscosity however has been reported to be slightly elevated in more than 80 per cent of patients examined with non-erosive osteoarthritis (Sitton *et al*. 1987).

Polymyalgia rheumatica and giant cell arteritis

Most studies report a raised ESR in 95 to 100 per cent of patients with polymyalgia rheumatica/giant cell arteritis at presentation (reviewed by Kyle 1991), and indeed an ESR of greater than 50 mm/h is now included as one of five diagnostic criteria to be used in the classification of giant cell arteritis (Hunder *et al*. 1990). This marked acute phase response is of use clinically to differentiate between polymyalgia rheumatica and other causes of myalgia such as depression, myositis and thyroid disease, due to the lack of other abnormal biochemical findings in polymyaglia rheumatica. However there are reports of patients with biopsy-proven giant cell arteritis in whom the ESR is normal (reviewed by Kyle 1991) and it has been suggested that at least in some cases this may be because of prior use of steroids (Wise *et al*. 1991). A significant correlation between plasma viscosity and ESR has been reported (Esselinckx *et al*. 1977), but in 30 out of the 112 paired readings, one parameter was abnormal in the presence of a normal value for the other. Acute phase proteins such as CRP, a_1-acid glycoprotein, a_1-antichymotrypsin, and haptoglobin are also elevated in the majority of untreated patients with polymyaglia rheumatica or giant cell arteritis (CRP may be in excess of 100 mg/1). However the elevation of a_1-acid glycoprotein by corticosteroids and the consumption of haptoglobin limit their potential usefulness in monitoring disease activity, and a_1-antichymotrypsin does not decrease with clinical remission although lower concentrations during follow-up may be indicative of a reduced risk of subsequent relapse (Pountain *et al*. 1994). There is some controversy as to whether CRP or the ESR is the better index of disease activity and both may remain normal during clinical relapse on treatment (Mallya *et al*. 1985; Kyle *et al*. 1989; Berlit 1992; Pountain *et al*. 1994). If clinical suspicion of the disease is strong yet the chosen parameter is not abnormal, it may be appropriate to measure the other. Measurement of either CRP or the ESR may be used together with clinical assessment to monitor the initial effectiveness of steroid therapy, thus minimizing the risk of arteritis, with the change in CRP more rapidly reflecting the clinical improvement and in most cases reaching normal levels within 7 to 14 days of the start of steroid administration. Neither CRP nor ESR levels can be used to predict a relapse however.

Autoimmune rheumatic disorders

The acute phase response is often moderate or absent in these disorders, even when clinically active, and if marked may suggest alternative pathology. However in systemic lupus erythematosus the ESR is elevated in nearly all patients, often falling with a decline in disease activity, although it is of little use clinically. The use of CRP levels to discriminate between intercurrent infection (with or without fever) and disease exacerbation is controversial. Most studies report generally higher CRP levels in patients with infection compared with active disease (reviewed by Pepys *et al*. 1982). It has been proposed that a serum CRP level of greater than 60 mg/1 is strongly indicative of infection, with CRP levels of less than 30 mg/1 making it unlikely that severe infection is present (Pepys *et al*. 1982). However the discriminatory power of CRP levels is not absolute and must be viewed with caution as levels in excess of 60 mg/1 have been reported in patients with systemic lupus erythematosus in the absence of infection, and patients with this disease and intercurrent infection may have normal or only moderately elevated levels of CRP (Pepys *et al*. 1982; Mackiewicz *et al*. 1987a). Recently it has been proposed that the discriminatory ability of CRP levels is only of value in the absence of serositis (Ter Borg *et al*. 1990). In addition, patients with Jaccoud's arthropathy or polyarthritis appear to have higher CRP concentrations (Spronk *et al*. 1992).

The reason for the relative lack of acute phase response in many of the autoimmune rheumatic disorders even in the presence of active inflammation is not clear, but the finding of identical clearance rates for CRP in patients with systemic lupus erythematosus compared with healthy controls make it unlikely that altered clearance mechanisms account for the apparent lack of response (Vigushin *et al*. 1993). It is also unlikely to be due to a straightforward lack of hepatic responsiveness as these patients can present such a response in the case of bacterial infection. However, with the complex regulation of the acute phase response by a network of cytokine interactions, it is possible that the response may be impaired by certain combinations of cytokines which are present in this disease, either at the level of the hepatocyte or because of cytokine inhibitors (reviewed by Gordon and Emery 1993). This is discussed further in Chapter 3.1.

Seronegative spondylarthritides

Ankylosing spondylitis is variably associated with a raised ESR, plasma viscosity and CRP. Reports regarding correlation with disease activity are conflicting, possibly due at least in part to differences in study design and definitions of active disease. Generally CRP levels are between 10 and 30 mg/1 although levels in excess of 100 mg/1 have been reported in active disease. CRP or plasma viscosity have been advocated as being more sensitive markers of active disease than the ESR (Dixon *et al*. 1981) and significantly elevated CRP levels have been found in patients with active disease compared with inactive disease (Nashel *et al*. 1986). However Cowling *et al*. (1980) found CRP and ESR correlated well, both with each other and with disease activity. In subsequent studies, although the ESR was found to correlate with levels of CRP and other acute phase proteins, none of these parameters correlated with disease activity (Laurent and Panayi 1983; Sheehan *et al*. 1986). More recently Reynolds *et al*. (1991) have shown a significant correlation between disease activity and CRP levels only in HLA-B27 positive patients.

CRP levels in patients with Behçet's syndrome do not usually exceed 20 mg/1. Most patients with psoriatic arthritis have an increased plasma viscosity compared with controls (Sitton *et al*. 1987). Both psoriatic arthritis and Reiter's syndrome are associated generally with normal or only mildly elevated levels of CRP (Sitton *et al*. 1987), although levels as high as 40 mg/1 have been reported in patients with active Reiter's syndrome (Nashel *et al*. 1986). The few reports concerning the use of CRP or ESR to monitor disease activity in these illnesses are contradictory and hence no conclusion can be drawn.

Rheumatoid arthritis

Rheumatoid arthritis is associated generally with a marked acute phase response although fever is usually absent. Plasma viscosity, ESR and CRP levels are nearly always elevated and in many cases, although not all, reflect disease activity (reviewed by Van Leeuwen and Van Rijswijk 1994). CRP levels are generally in the range of 30 to 40 mg/1 in patients with moderate disease activity, and may reach in excess of 100 mg/1 in severe disease, although in a small group of patients CRP levels lie within the normal range despite clinical evidence of active disease. CRP concentrations in excess of 150 mg/1 may be indicative of an intercurrent bacterial infection. The question of which of these parameters is the best index of disease activity is open to debate (McKenna 1988). a_1-Antichymotrypsin has also been reported to reflect disease activity in rheumatoid arthritis, being elevated even in those patients in whom CRP levels are normal despite clinical evidence of active

disease (Chard *et al*. 1988), and merits further investigation.

Studies examining the prognostic implications of the acute phase response in rheumatoid arthritis have recently been reviewed (McKenna 1988, Van Leeuwen and Van Rijswijk 1994). In general it appears that isolated measurements of the acute response are of little use in predicting outcome of the disease. However, persistently higher ESR and CRP measurements have been associated with a higher rate of radiological progression (Amos *et al*. 1977; Dawes *et al*. 1986). More recently, time-integrated values of CRP and ESR have been found to be correlated significantly with radiological progression both in patients with early disease and in those with long-standing disease (Hassell *et al*. 1993; Van Leeuwen *et al*. 1993; Van Leeuwen *et al*. 1994). However with the considerable interindividual variation in CRP values found, it would be inappropriate to attempt to use a single cut-off value for prognostic purposes but rather the longitudinal findings need to be interpreted for each individual patient (reviewed by Van Leeuwen and Van Rijswijk 1994).

Following an examination of the literature, the American College of Rheumatology have recommended that the ESR or CRP should be one of the core set of disease activity measures in clinical trials (Felson *et al*. 1993). The slow-acting disease-modifying drugs such as gold, sulphasalazine, D-penicillamine, and hydroxychloroquine are usually associated with a decline in the acute phase protein response (reviewed by Kvien and Husby 1992; Menkes 1993). An exception is cyclosporin which has no effect on the ESR. A recent European collaborative study found ESR to be a more useful measure of disease activity than CRP, with CRP presenting standardization problems between countries and laboratories (Scott *et al*. 1992). Although a vast majority of studies have failed to find any effect of non-steroidal anti-inflammatory drugs on the acute phase response, a study which differentiated between clinical 'responders' and 'non-responders' in their analysis, did suggest a decline in ESR and CRP levels in the 'responders', correlating with improvement in clinical variables (Cush *et al*. 1990). However the finding that CRP decreases even in some clinical 'non-responders' suggests that this decline in acute phase protein response may not be secondary to a suppression of joint inflammation and cytokine production by the drugs, but rather to some other mechanism. Obviously under these circumstances, the use of the acute phase reactants to monitor drug efficacy would be inappropriate.

Conflicting reports exist regarding a possible association between a $_1$-antitrypsin phenotype and rheumatoid arthritis but the evidence overall does not currently support any such association (Kahl *et al*. 1989).

Potential future developments

Glycosylation of acute phase proteins

Many serum glycoproteins exhibit microheterogeneity due to variations in their *N*-linked heteroglycan sidechains. Such variation is often assessed using crossed immunoaffinity electrophoresis, on the basis of the degree of reactivity with the lectin concanavalin A, which reacts with biantennary glycans. In patients with active polymyalgia rheumatica the glycan microheterogeneity of a $_1$-antichymotrypsin is altered, with a decrease in concanavalin A reactivity (Hachulla *et al*. 1990). This was reported to allow a discrimination between active and inactive disease with a sensitivity of 97 per cent and a specificity of 91 per cent. This decrease in concanavalin A reactivity has been confirmed (Pawlowski *et al*. 1990) both for a$_1$-antichymotrypsin and a$_1$-acid glycoprotein in patients with active polymyalgia rheumatica or polymyalgia rheumatica/giant cell arteritis and allows the differentiation of these illnesses from polymyositis. Although only studied in five patients, the concanavalin A reactivity was shown to increase towards normal following steroid therapy.

A decrease in concanavalin A reactivity of a $_1$-acid glycoprotein has been observed in patients with ankylosing spondylitis (Mackiewicz *et al*. 1989a) and in patients with active rheumatoid arthritis (Mackiewicz *et al*. 1987b) compared with healthy controls. However more recently it has been suggested that the microheterogeneity of a$_1$-acid glycoprotein in rheumatoid arthritis is dependent on disease duration, with increased and decreased concanavalin A reactivity being characteristic of early and established disease respectively, emulating the change in reactivity seen in the transition between acute and chronic bacterial infection (Hrycaj *et al*. 1993). Although not significantly different between patients with systemic lupus erythematosus of varying activity and healthy controls, the glycosylation pattern of a $_1$-acid glycoprotein has also been reported to be a much more sensitive indicator of intercurrent infection than CRP, with an increase in concanavalin A reactivity in infected patients (Mackiewicz *et al*. 1987a).

Such findings warrant further clinical studies. The exact significance of such changes is not yet known although recent studies using primary human hepatocyte cultures indicate that such changes may reflect cytokine-induced alterations in intrahepatocellular processes (Mackiewicz *et al*. 1989b; Pos *et al*. 1989). The lack of correlation in the various reports between changes in level of acute phase proteins and their glycosylation patterns indicates that the regulatory mechanisms responsible for these two events are probably different.

Drug management

Little is known about the mode of action of most of the disease-modifying drugs currently available. All were introduced for other indications or for reasons now clearly felt to lack relevance to our current understanding of the pathogenesis of rheumatoid arthritis. Already we are seeing the introduction and evaluation of new families of drugs specifically designed for this task, including specific blockers of certain cytokines which may work relatively early in pathogenesis as well as collagenase inhibitors which, by preventing final joint degradation, may act at a later stage. Close monitoring of acute phase reactants in parallel with drug administration, together with clinical and radiological assessment, is likely to be invaluable in unravelling the many different sites at which the rheumatoid process may be manipulated pharmacologically.

Chapter References

Altman, R.D. and Gray, R. (1985). Inflammation in osteoarthritis. *Clinics in the Rheumatic Diseases*, **11**, 353–65.

Amos, R.S., Constable, T. J., Crockson, R.A., and McConkey, B. (1977). Rheumatoid arthritis: relation of serum C-reactive protein and erythrocyte sedimentation rates to radiographic changes. *British Medical Journal*, **1**, 195–7.

Banks, R.E., *et al*. (1995). The acute phase protein response in patients receiving subcutaneous IL-6. *Clinical and Experimental Immunology*, **102**, 217–23.

Baumann, H. and Gauldie, J. (1994). The acute phase response. *Immunology Today*, **15**, 74–80.

Baumann, H., Richards, C., and Gauldie, J. (1987). Interaction among hepatocyte stimulating factors, interleukin 1 and glucocorticoids for regulation of acute phase plasma proteins in human hepatoma (HepG2) cells. *Journal of Immunology*, **139**, 4122–8.

Berlit, P. (1992). Clinical and laboratory findings with giant cell arteritis. *Journal of the Neurological Sciences*, **111**, 1–12.

Chard, M.D., Calvin, J., Price, C.P., Cawston, T.E., and Hazleman, B.L. (1988). Serum a $_1$-antichymotrypsin concentration as a marker of disease activity in rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **47**, 665–71.

Claus, D.R., Osmand, A.P., and Gewurz, H. (1976). Radioimmunoassay of human C-reactive protein and levels in normal sera. *Journal of Clinical and Laboratory Medicine*, **87**, 120–28.

Cowling, P., Ebringer, R., Cawdell, D., Ishii, M., and Ebringer, A. (1980). C-reactive protein, ESR, and klebsiella in ankylosing spondylitis. *Annals of the Rheumatic Diseases*, **39**, 45–9.

Cush, J.J., Lipsky, P.E., Postlethwaite, A.E., Schrohenloher, R.E., Saway, A., and Koopman, W.J. (1990). Correlation of serologic indicators of inflammation with effectiveness of non-steroidal anti-inflammatory drug therapy in rheumatoid arthritis. *Arthritis and Rheumatism*, **33**, 19–28.

Dacie, J.V. and Lewis, S.M. (1985). *Practical haematology*, 6th edn. Churchill Livingstone, London.

Dawes, P.T., Fowler, P.D., Clarke, S., Fisher, J., Lawton, A., and Shadforth, M.F. (1986). Rheumatoid arthritis: treatment which controls the C-reactive protein and erythrocyte sedimentation rate reduces radiological progression. *British Journal of Rheumatology*, **25**, 44–9.

De Beer, F.C., Mallya, R.K., Fagan, E.A., Lanham, J.G., Hughes, G.R., and Pepys, M.B. (1982). Serum amyloid-A protein concentration in inflammatory diseases and its relationship to the incidence of reactive systemic amyloidosis. *Lancet*, **ii**, 231–4.

Dixon, J.S., Bird, H.A., and Wright, V. (1981). A comparison of serum biochemistry in ankylosing spondylitis, seronegative and seropositive rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **40**, 404–8.

Esselinckx, W., Bucknall, R.C., and Dixon, A. St J. (1977). Polymyalgia rheumatica. Assessment of disease activity using erythrocyte sedimentation rate and plasma viscosity. *Annals of the Rheumatic Diseases*, **36**, 560–62.

Felson, D.T. *et al*. (1993). The American College of Rheumatology preliminary core set of disease activity measures for rheumatoid arthritis clinical trials. *Arthritis and Rheumatism*, **36**, 729–40.

Gordon, C. and Emery, P. (1993). Cytokines and the acute phase response in SLE. *Lupus*, **2**, 345–7.

Hachulla, E., Laine, A., and Hayem, A. (1990). Microheterogeneity of a $_1$-antichymotrypsin in the management of giant-cell arteritis and polymyalgia rheumatica. *Clinical Science*, **78**, 557–64.

Hassell, A.B. *et al*. (1993). The relationship between serial measures of disease activity and outcome in rheumatoid arthritis. *Quarterly Journal of Medicine*, **86**, 601–7.

Hunder, G.G. *et al*.(1990). The American College of Rheumatology 1990 criteria for the classification of giant cell arteritis. *Arthritis and Rheumatism*, **33**, 1122–8.

Hrycaj, P., Sobieska, M., Mackiewicz, S., and Muller, W. (1993). Microheterogeneity of alpha 1 acid glycoprotein in rheumatoid arthritis: Dependent on disease duration? *Annals of the Rheumatic Diseases*, **52**, 138–41.

ICSH (International Committee for Standardization in Haematology, Expert Panel on Blood Rheology) (1988). Guidelines on selection of laboratory tests for monitoring the acute phase response. *Journal of Clinical Pathology*, **41**, 1203–12.

ICSH (International Committee for Standardization in Haematology, Expert Panel on Blood Rheology) (1993). ICSH recommendations for measurement of erythrocyte sedimentation rate. *Journal of Clinical Pathology*, **46**, 198–203.

Kahl, L.E., Kamboh, M.I., Decroo, S., and Ferrell, R.E. (1989). Alpha-1-antitrypsin (Pl) and vitamin-D binding globulin (GC) phenotypes in rheumatoid arthritis: absence of an association. *Disease Markers*, **7**, 71–8.

Kushner, I. (1982). The phenomenon of the acute phase response. *Annals of the New York Academy of Sciences*, **389**, 39–48.

Kushner, I. and Mackiewicz, A. (1987). Acute phase proteins as disease markers. *Disease Markers*, **5**, 1–11.

Kvien, T.K. and Husby, G. (1992). Disease modification in rheumatoid arthritis with special reference to cyclosporin A. *Scandinavian Journal of Rheumatology*, **21** (Suppl. 95), 19–28.

Kyle, V. (1991). Laboratory investigations including liver in polymyalgia rheumatica/giant cell arteritis. In *Baillière's clinical rheumatology*, Vol. 5, No. 3, (ed. B. Hazleman and B.-A. Bengtsson), pp. 475–84. Baillière Tindall, London

Kyle, V., Cawston, T.E., and Hazleman, B.L. (1989). Erythrocyte sedimentation rate and C-reactive protein in the assessment of polymyalgia rheumatica/giant cell arteritis on presentation and during follow-up. *Annals of the Rheumatic Diseases*, **48**, 667–71.

Laurell, C.-B. (1985). Acute phase proteins—a group of protective proteins. In *Recent advances in clinical biochemistry*, No. 3 (ed. C.P. Price and K.G.M.M. Alberti), pp. 103–24.

Laurell, C.-B. and Rannevik, G. (1979). A comparison of plasma protein changes induced by Danazol, pregnancy and estrogens. *Journal of Clinical Endocrinology and Metabolism*, **49**, 719–25.

Laurent, M.R. and Panayi, G.S. (1983). Acute phase proteins and serum immunoglobulins in ankylosing spondylitis. *Annals of the Rheumatic Diseases*, **42**, 524–8.

Lowe, G.D.O. (1994). Should plasma viscosity replace the ESR? *British Journal of Haematology*, **86**, 6–11.

Mackiewicz, A., Marcinkowska-Pieta, R., Ballou, S., Mackiewicz, S., and Kushner, I. (1987 *a*). Microheterogeneity of alpha$_1$-acid glycoprotein in the detection of intercurrent infection in systemic lupus erythematosus. *Arthritis and Rheumatism*, **30**, 513–18.

Mackiewicz, A., Pawlowski, T., Mackiewicz-Pawlowska, A., Wiktorowicz, K., and Mackiewicz S. (1987 *b*). Microheterogeneity forms of alpha$_1$-acid glycoprotein as indicators of rheumatoid arthritis activity. *Clinica Chimica Acta*, **163**, 185–90.

Mackiewicz, A., Khan, M.A., Reynolds, T.L., Van Der Linden, S., and Kushner, I. (1989 *a*). Serum IgA, acute phase proteins, and glycosylation of a $_1$-acid glycoprotein in ankylosing spondylitis. *Annals of the Rheumatic Diseases*, **48**, 99–103.

Mackiewicz, A., Schultz, D., Mathison, J., Ganapathi, M., and Kushner, I. (1989 *b*). Effect of cytokines on glycosylation of acute phase proteins in human hepatoma cell lines. *Clinical and Experimental Immunology*, **74**, 70–75.

Mackiewicz, A., Kushner, I., and Baumann, H. (ed.), (1993). Acute phase proteins: molecular biology, biochemistry and clinical applications. CRC Press, London.

Malle, E., Steinmetz, A., and Raynes, J.G. (1993). Serum amyloid A (SAA): an acute phase protein and apolipoprotein. *Atherosclerosis*, **102**, 131–46.

Mallya, R.K., Hind, C.R.K., Berry, H., and Pepys, M.B. (1985). Serum C-reactive protein in polymyalgia rheumatica. *Arthritis and Rheumatism*, **28**, 383–7.

McKenna, F. (1988). Clinical and laboratory assessment of outcome in rheumatoid arthritis. *British Journal of Rheumatology*, **27** (Suppl. 1), 12–20.

Menkes, C.-J. (1993). Effects of disease-modifying anti-rheumatic drugs, steroids and non-steroidal anti-inflammatory drugs on acute-phase proteins in rheumatoid arthritis. *British Journal of Rheumatology*, **32** (Suppl. 3), 14–18.

Milman, N., Graudal, N., and Andersen, H.C. (1988). Acute phase reactants in the elderly. *Clinica Chimica Acta*, **176**, 59–62.

Morley, J.J. and Kushner, I. (1982). Serum C-reactive protein levels in disease. *Annals of the New York Academy of Science*, **389**, 406–18.

Nashel, D.J., Petrone, D.L., Ulmer, C.C., and Sliwinski, A.J. (1986). C-reactive protein: a marker for disease activity in ankylosing spondylitis and Reiter's syndrome. *Journal of Rheumatology*, **13**, 364–7.

Pawlowski, T., Aeschlimann, A., Kahn, M.F., Vaith, P., Mackiewicz, S.H., and Mueller, W. (1990). Microheterogeneity of acute phase proteins in the differentiation of polymyalgia rheumatica from polymyositis. *Journal of Rheumatology*, **17**, 1187–92.

Pepys, M.B., Lanham, J.G., and De Beer, F.C. (1982). C-reactive protein in SLE. *Clinics in Rheumatic Diseases*, **8**, 91–103.

Peltola, H.O. (1982). C-reactive protein for rapid monitoring of infections of the central nervous system. *Lancet*, **i**, 980–83.

Pos, O. *et al*. (1989). Effects of monocytic products, recombinant interleukin-1 and recombinant interleukin-6 on glycosyiation of a $_1$-acid glycoprotein. *Inflammation*, **13**, 415–27.

Pountain, G.D., Calvin, J., and Hazleman, B.L. (1994). a1-antichymotrypsin, C-reactive protein and erythrocyte sedimentation rate in polymyalgia rheumatica and giant cell arteritis. *British Journal of Rheumatology*, **33**, 550–54.

Protein Reference Unit. (1993). Reference ranges. In *PRU handbook of clinical immunochemistry*, (4th edn) (ed. A. Milford Ward, P.G. Riches, R. Fifield, and A.M. Smith), pp. 277–89. PRU Publications, Sheffield.

Reynolds, T.L, Khan, M.A., van der Linden, S., and Cleveland, R.P. (1991). Differences in HLA-B27 positive and negative patients with ankylosing spondylitis: study of clinical disease activity and concentrations of serum IgA, C-reactive protein, and haptoglobin. *Annals of the Rheumatic Diseases*, **50**, 154–7.

Roseff, R., Wohlgethan, J.R., Sipe, J.D., and Canoso, J.J. (1987). The acute phase response in gout. *Journal of Rheumatology*, **14**, 974–7.

Schultz, D.R. and Arnold, P.I. (1990). Properties of four acute phase proteins: C-reactive protein, serum amyloid A protein, a $_1$-acid glycoprotein, and fibrinogen. *Seminars in Arthritis and Rheumatism*, **20**, 129–47.

Scott, D.L., Panayi, G.S., Van Riel, P.L.C.M., Smolen, J., Van De Putte, L.B.A., and the Consensus Study Group of the European Workshop for Rheumatology (1992). Disease activity in rheumatoid arthritis: preliminary report of the Consensus Study Group of the European Workshop for Rheumatology Research. *Clinical and Experimental Rheumatology*, **10**, 521–5.

Sheehan, N.J., Slavin, B.M., Donovan, M.P., Mount, J.N., and Mathews, J.A. (1986). Lack of correlation between clinical disease activity and erythrocyte sedimentation rate, acute phase proteins or protease inhibitors in ankylosing spondylitis. *British Journal of Rheumatology*, **25**, 171–4.

Sitton, N.G., Dixon, J.S., Bird, H.A., and Wright, V. (1987). Serum biochemistry in rheumatoid arthritis, seronegative arthropathies, osteoarthritis, SLE and normal subjects. *British Journal of Rheumatology*, **26**, 131–5.

Spronk, P.E., Terborg, E.J., and Kallenberg, C.G.M. (1992). Patients with systemic lupus-erythematosus and Jaccoud's arthropathy—a clinical subset with an increased C-reactive protein response. *Annals of the Rheumatic Diseases*, **51**, 358–61.

Steel, D.M. and Whitehead, A.S. (1994). The major acute phase reactants: C-reactive protein, serum amyloid P component and serum amyloid A protein. *Immunology Today*, **15**, 81–8.

Ter Borg, E.J., Horst, G., Limburg, P.C., van Rijswijk, M.H., and Kallenberg, C.G.M. (1990). C-reactive protein levels during disease exacerbations and infections in systemic lupus erythematosus: a prospective longitudinal study. *Journal of Rheumatology*, **17**, 1642–8.

Thompson, D., Harrison, S.P., Evans, S.W., and Whicher, J.T. (1991). Insulin modulation of acute-phase protein production in a human hepatoma cell line. *Cytokine*, **3**, 619–26.

Thompson, D., Milford-Ward, A., and Whicher, J.T. (1992). The value of acute phase protein measurements in clinical practice. *Annals of Clinical Biochemistry*, **29**, 123–31.

Van Leeuwen, M.A. and Van Rijswijk, M.H. (1994). Acute phase proteins in the monitoring of inflammatory disorders. In *Baillière's clinical rheumatology*, Vol. 8, No. 3, (ed. G. Husby), pp. 531–52.

Van Leeuwen, M.A. *et al*. (1993). The acute-phase response in relation to radiographic progression in early rheumatoid arthritis: a prospective study during the first three years of the disease. *British Journal of Rheumatology*, **32** (Suppl. 3), 9–13.

Van Leeuwen, M. A. *et al*. (1994). Interrelationship of outcome measures and process variables in early rheumatoid arthritis. A comparison of radiologic damage, physical disability, joint counts, and acute phase reactants. *Journal of Rheumatology*, **21**, 425–9.

Vigushin, D.M., Pepys, M.B., and Hawkins, P.N. (1993). Metabolic and scintigraphic studies of radioiodinated human C-reactive protein in health and disease. *Journal of Clinical Investigation*, **91**, 1351–7.

Whicher, J.T., Thompson, D., Billingham, M.E.J., and Kitchen, E.A. (1989). Acute phase proteins. In *Pharmacological methods in the control of inflammation*, (ed. J.Y. Chang and A.J. Lewis), pp. 101–28. Alan R. Liss Inc., New York.

Wise, C.M., Agudelo, C.A., Chmelewski, W.L., and McKnight, K.M. (1991). Temporal arteritis with low erythrocyte sedimentation rate: a review of five cases. *Arthritis and Rheumatism*, **34**, 1571–4.

Zlonis, M. (1993). The mystique of the erythrocyte sedimentation rate. A reappraisal of one of the oldest laboratory tests still in use. *Clinics in Laboratory Medicine*, **13**, 787–800.

4.2 Haematology

Charles Richardson, Bridget Griffiths, and Paul Emery

The anaemia of chronic disease

Introduction

The anaemia of chronic disease develops as a consequence of chronic activation of the immune system (hence the alternative name of the anaemia of chronic inflammation). It is associated with autoimmune rheumatic diseases, chronic bacterial infections and malignancy. The aetiology of the anaemia of chronic disease is multifactorial but is characterized by the suppression of erythropoiesis, a blunting of erythropoietin production in response to anaemia and abnormal utilization of iron stores. The availability of recombinant erythropoietin provides an effective but expensive treatment for the anaemia of chronic disease without resorting to blood transfusion.

The anaemia of chronic disease is exemplified by the anaemia which develops in rheumatoid arthritis within 1 to 2 months of the onset of systemic disease. It is seen only after the onset of systemic symptoms but represents the most common extra-articular manifestation of the disease and is an important clinical problem. Rheumatoid arthritis has therefore been chosen as the model condition to describe the features of the anaemia of chronic disease. However, these features apply equally well to other chronic inflammatory disorders.

The severity of the anaemia in rheumatoid arthritis fluctuates with the disease course and correlates with the magnitude of the acute phase response. The prevalence is dependent on the definition used, although it has been estimated that 60 per cent of inpatients and 25 per cent of outpatients with rheumatoid arthritis are clinically anaemic, with 5 per cent of these being severely so.

The causes of the anaemia in rheumatoid arthritis are multiple (Table 1), although the most important cause is the anaemia of chronic disease itself (Baer *et al*. 1990; Vreugdenhill *et al*. 1990). The pro-inflammatory cytokines are central to the pathogenesis of the anaemia of chronic disease and their role is discussed further below. In practice the problem most commonly faced is how to differentiate the anaemia of chronic disease from iron deficiency. There are many shared features and the two conditions often coexist. The anaemia of chronic disease is a diagnosis in its own right, with characteristic abnormalities in the biochemical indices of iron metabolism in the presence of an appropriate clinical setting. It does not require further investigation whilst iron deficiency may well do so. To understand the investigations performed to distinguish between the anaemia of chronic disease and iron deficiency a knowledge of normal iron metabolism is required.

Major causes	
	The anaemia of chronic disease
	Iron deficiency anaemia
	Bone marrow suppression secondary to disease modifying drugs
Other significant causes	
	Vitamin B_{12} or folate deficiency
	Haemolysis
	Hypersplenism secondary to Felty's syndrome
	Chronic renal failure
Causes of the rapid development of anaemia	
	Bleeding, usually gastrointestinal
	Septicaemia
	Haemolysis
	Sudden increase in disease activity*

*Abnormal but can occur. It is difficult to distinguish from other causes

Table 1 The causes of anaemia in rheumatoid arthritis

Normal iron metabolism

The body iron content of normal adult males and females is 50 and 35 mg/kg of body weight, respectively. Most of the body's iron is incorporated into the haem proteins, haemoglobin and myoglobin (70 per cent). The remainder is in the storage forms ferritin and haemosiderin (30 per cent). Only a minute portion (0.1 per cent) is present in the plasma, where it is carried by the transport protein transferrin. It is worth noting that it is only this 0.1 per cent that is available for routine measurement of iron status (Fairbanks and Beutler 1983; Lee 1993).

Iron balance

In normal subjects the amount of iron present remains within narrow limits. The physiological control of body iron stores is unique in that it is mediated by control of absorption. Iron is not excreted in the usual sense, the amount lost in urine and sweat being negligible. The average daily loss of iron in adult men and non-menstruating women is 1 mg. The normal Western diet contains 10 to 30 mg per day, therefore only 5 to 10 per cent needs to be absorbed to maintain balance. The non-absorbed iron is bound to the iron-binding protein, ferritin, in the epithelial cells of the small intestine and is lost subsequently into the gut when the cells are

desquamated. The amount of iron binding to ferritin provides the physiological control of iron absorption. The system is saturated easily, and during oral iron therapy there may be passive entry of iron into the plasma. Iron entering the plasma is bound to the transport protein, transferrin. The rate of iron absorption is determined by (i) the total body iron stores and (ii) the rate of erythropoiesis. It is enhanced during iron deficiency and periods of increased erythropoiesis.

There are two pathways for iron absorption in the upper small intestine, one for iron attached to haem and another for ferrous ions. Dietary iron must be converted to one of these two forms before absorption is possible. Haem iron is derived predominantly from haemoglobin and myoglobin in food of animal origin. Haem is released from its apoprotein in the acid environment of the stomach. Non-haem iron occurs mainly in the ferric state loosely bound to phytates, oxalates, sugars, citrate, lactate, and amino acids. The pathways of iron absorption and metabolism are illustrated in Fig. 1.

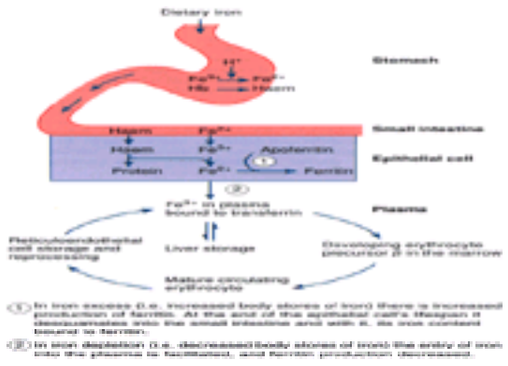


Fig. 1 The normal metabolism of iron.

The iron cycle

Iron metabolism is dominated by its role in haemoglobin synthesis. There is a continuous cycle of iron between erythrocytes, reticuloendothelial cells and the plasma (Fig. 1). The transport of iron in the plasma bound to transferrin is central to this cycling. Iron moves from the plasma to erythrocyte precursors, the mature erythrocytes enter the circulation and at the end of their lifespan are phagocytosed by macrophages and their iron extracted for reuse.

Plasma transport

Iron is transported in the plasma bound to transferrin, a b-globulin, synthesized primarily by the liver. Transferrin is distributed equally between the plasma and extravascular sites, with a half-life of approximately 10 days. Each transferrin molecule is able to bind two molecules of ferric iron.

Transferrin is measured indirectly by the total iron-binding capacity of plasma. In the normal subject the serum iron is approximately 18 mmol and the total iron binding capacity 56 mmol, indicating that only about 30 per cent of the transferrin binding sites are occupied. The rate of synthesis of transferrin is increased in iron deficiency and decreased in chronic disease. Therefore the total iron-binding capacity is increased in iron deficiency and decreased in chronic disease.

Transferrin delivers iron to the developing erythrocytes by binding to transferrin receptors found in greatest concentration on the normoblast. After binding, the entire transferrin–transferrin receptor complex forms a coated vesicle and undergoes endocytosis. The contents of the endosome are acidified with release of the iron present. Normoblast maturation is accompanied by shedding of the receptors into the plasma at a concentration that correlates with the rate of erythropoiesis. In addition it has been reported that an increase in the transferrin receptor concentration level is another indicator of tissue iron deficiency.

Ferritin is found in the plasma in small amounts (12 to 300 mg/l) and does not play an important role in iron transport (its role as a storage protein is discussed below). In a normal individual the concentration of ferritin correlates with the amount of iron stored. This may not be the case in inflammatory disease because ferritin is an acute phase respondent and its plasma concentration increases during periods of active inflammation.

Metabolism within normoblasts and macrophages

In a normal subject about 85 per cent of iron entering normoblasts is incorporated into haemoglobin, the remainder is stored within the iron storage protein ferritin.

There are two distinct populations of macrophage with respect to iron handling. The first type phagocytoses erythrocytes and stores iron, but does not return it to the circulation (e.g. the alveolar macrophage). By contrast other macrophages (e.g. the reticuloendothelial macrophages found in the liver and spleen) provide a storage pool for iron. The latter can either rapidly process iron and return it to the circulation or store it in the iron storage compounds ferritin and haemosiderin.

Iron-free ferritin (apoferritin) consists of a sphere with a central hollow cavity made from a number of subunits. Iron is deposited in the cavity in the form of a polymer of ferric hydroxide and phosphate. Under normal conditions about 2000 or less iron atoms are deposited, although in theory 4300 are possible. The synthesis of ferritin is stimulated by the exposure of cells to iron. At times of deficiency iron can be mobilized from ferritin and made available for erythropoiesis. Haemosiderin is an insoluble polymer of ferritin. It represents a more stable form of iron, but it is less readily available for utilization.

Iron is required in many other tissues besides erythrocytes for the synthesis of iron-containing enzymes. The most important tissue consumer of iron is the liver which accounts for about 5 per cent of iron leaving the plasma.

The aetiology of the anaemia of chronic disease

The aetiology of the anaemia of chronic disease is multifactorial and not fully understood. The pro-inflammatory cytokines interleukin 1 (**IL-1**), interleukin 6 (**IL-6**) and tumour necrosis factor-a (**TNF-a**) appear central in its pathogenesis. The following mechanisms are thought to operate in this anaemia:

direct inhibition of erythropoiesis

inhibition of erythropoietin production

enhanced erythrocyte phagocytosis

enhanced reticuloendothelial retention of iron

Direct inhibition of erythropoiesis

In animal studies the injection of TNF-a and IL-1 inhibits the production of mature erythrocytes by inhibition of the development of erythroid colony-forming units. It is of interest, that the colony-forming units are the most sensitive stage of erythroid development to erythropoietin, the administration of which corrects the anaemia induced by IL-1 (Schooley *et al*. 1987; Johnson *et al*. 1988; Johnson *et al*. 1989a).

Chronic exposure of nude mice to TNF-a provides a model more analogous to the situation in the anaemia of chronic disease. These animals develop an anaemia with low serum iron and normal iron stores. In addition, despite an increased concentration of erythropoietin, they have bone marrow depression of both burst-forming and colony-forming erythroid units (Johnson *et al*. 1989b). These findings may explain (i) the observation that monocytes/macrophages (which secrete pro-inflammatory cytokines) obtained from patients with anaemia of chronic disease inhibit the proliferation of erythroid colony-forming units (Roodman *et al*. 1983) and (ii) the clinical observation that suppression of the acute phase response (which is stimulated by the pro-inflammatory cytokines) produces an improvement in the

anaemia in active rheumatoid disease.

Finally, patients with rheumatoid arthritis treated with a monoclonal antibody against TNF-a increase their haemoglobin concentration (Elliot *et al*. 1994), while patients with cancer who received TNF-a (used as an experimental anticancer agent) had a fall in their haemoglobin concentration (Blick *et al*. 1987).

The inhibition of erythropoietin production

Erythropoietin is the hormone responsible for the regulation of erythropoiesis. It maintains a constant mass of red cells by stimulating the erythroid marrow to produce sufficient erythrocytes to replace the 1 per cent which reach their 120-day lifespan each day. There is a basal production of erythropoietin which can be increased one thousandfold in response to tissue hypoxia. Erythropoietin is produced predominantly by the cells of the renal mesangium, although during hypoxic stress the liver makes a significant contribution. It binds to specific receptors on erythroid progenitors in the bone marrow and increases the number of cells capable of differentiating into mature erythrocytes and augments the synthesis of haemoglobin.

The pro-inflammatory cytokines IL-1, TNF-a and IL-6 inhibit the production of erythropoietin in liver cell lines, providing a link between the acute phase response and the anaemia of chronic disease (Faquin *et al*. 1992).

For a given haemoglobin level the concentration of erythropoietin is less in patients with the anaemia of chronic disease than in patients with iron deficiency. In addition, whilst there is a significant negative correlation between the severity of iron deficiency and serum erythropoietin concentration, the same is not true for the anaemia of chronic disease (Takashina *et al*. 1990). Furthermore the ability of exogenous erythropoietin to overcome this anaemia can be taken as more evidence of an impaired erythropoietin response to anaemia. It seems likely that *in vivo* pro-inflammatory cytokines inhibit the production of erythropoietin and that this is responsible for the blunted erythropoietin response to anaemia seen in the chronic disease.

Abnormal iron utilization

The following abnormalities of iron metabolism in chronic disease have been reported:

1. an increase in the uptake of iron from the small intestine but a failure of the mucosal cells to release it into the plasma;
2. an increased uptake of iron into the reticuloendothelial cells of the liver and spleen;
3. enhanced macrophage phagocytosis of effete erythocytes with their consequent removal from the circulatory pool;
4. retention of iron within the cells of the reticuloendothelial system, making it unavailable for erythropoiesis.

It has been suggested that the following mechanisms contribute to abnormal iron utilization. The enzyme lactoferrin, released during phagocytosis, may compete with transferrin for iron binding. The lactoferrin–iron complex binds to specific receptors on liver and spleen macrophages, enters the cells and becomes incorporated into ferritin; thus preventing its use in erythropoiesis. Alternatively apoferritin (an acute phase protein) production is increased in active rheumatoid disease with consequent increase in iron binding reducing that available for erythropoiesis. However, neither of these abnormalities provides an explanation for effects seen at the intestinal mucosa (Lee 1993).

Shortened erythrocyte survival

The survival of erythrocytes in patients with active rheumatoid arthritis is reduced. Membrane-bound complement and/or immunoglobulins do not contribute to this except in special circumstances, neither is there any intrinsic defect in the erythrocyte. The shortened lifespan is probably secondary to an increase in phagocytosis of effete erythrocytes by macrophages due to a generalized activation of the immune system. If the bone marrow were functioning normally it should easily be able to overcome this slight decrease in erythrocyte survival. That it does not, implies that there is a relative failure of the marrow to respond to the degree of anaemia.

In summary

The anaemia of chronic disease has a multifactorial aetiology with the pro-inflammatory cytokines occupying a central role in the pathogenesis. Whilst there is evidence of abnormal iron metabolism (which forms the basis of the commonly used investigations), the lack of correlation between serum iron and marrow iron turnover combined with the finding that erythropoeitin can correct the anaemia all indicate that iron deficiency is not the limiting factor in the anaemia of chronic disease. The fundamental problem is a failure of the bone marrow to respond appropriately to anaemia. This is mainly due to cytokine inhibition of both erythropoiesis and erythropoietin production.

Diagnosis

Clinical findings

The diagnosis of anaemia of chronic disease should only be considered in the appropriate clinical setting. A complete history and physical examination must be performed. The anaemia of chronic disease normally takes 1 to 2 months to develop, the diagnosis is unlikely in a patient who has been unwell for only a matter of days (iron deficiency can develop in hours). A history of overt gastrointestinal bleeding should be sought and risk factors for this determined. Symptomatic anaemia indicates a rapid fall in haemoglobin; the most common cause for this is a gastrointestinal bleed. Physical examination may reveal melaena or signs of circulatory collapse, although in most instances it is unhelpful in determining the cause of anaemia.

Laboratory features

The acute phase response

The most commonly measured markers of the acute phase response are the erythrocyte sedimentation rate and the C-reactive protein. The differential diagnosis of the cause of anaemia in a patient with a known chronic disease would be simplified if a straightforward relationship between the magnitude of the acute phase response and the degree of anaemia could be formulated. Whilst it is common to observe change in anaemia fluctuating with disease course, a simple relationship has never been established. Clearly there are a large number of confounding factors which have prevented this, including the nature of the chronic disease, its duration and the drug therapy employed.

Figure 2 shows the haemoglobin concentration of over 400 patients with rheumatoid arthritis at presentation to an early arthritis clinic, plotted against increasing ranges of C-reactive protein concentration. Figure 3 shows the percentage who were anaemic. This was defined as lying 2 standard deviations below the normal population average concentration. The normal ranges for our laboratory were 13.4 to 17.4 g/dl and 11.4 to 15.4 g/dl for male and female subjects, respectively. None of these patients had received any disease-modifying therapy (including steroids). The following points are demonstrated:

1. The percentage of patients with anaemia and its severity increases with the magnitude of the acute phase response.
2. For a given level of C-reactive protein the haemoglobin is lower in females, but fewer lie outside the normal range.
3. On average the degree of anaemia is not severe until the C-reactive protein is markedly elevated.
4. A slight rise in the C-reactive protein is not associated with anaemia; in these circumstances the anaemia may warrant further investigation.

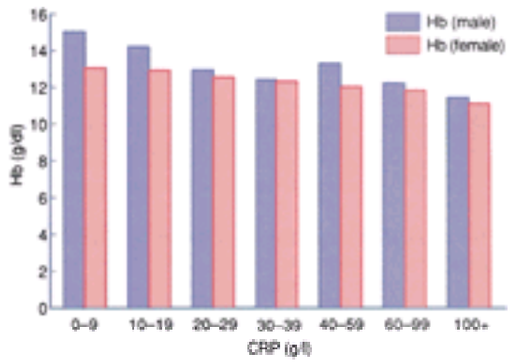


Fig. 2 The relationship between haemoglobin and C-reactive protein for patients presenting with rheumatoid arthritis.

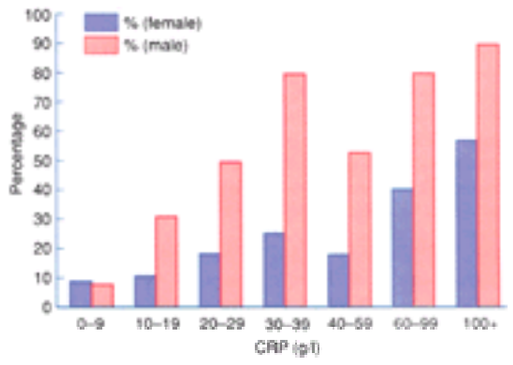


Fig. 3 The percentage of patients having anaemia at presentation with rheumatoid arthritis for increasing concentration of C-reactive protein.

Haematology investigations

In established chronic disease the haemoglobin level is rarely below 10 g/dl, a concentration less than this is more in keeping with a gastrointestinal bleed (Isenberg *et al*. 1986). Examination of blood films reveals a normochromic normocytic anaemia in 50 per cent of cases and a hypochromic microcytic picture in 30 per cent of cases. The latter is seen in the more severe chronic disease. Anisocytosis and poikilocytosis occur but not to the extent seen in iron deficiency anaemia. None of these features enable the anaemia of chronic disease to be distinguished reliably from iron deficiency. More sophisticated computer-assisted analysis of blood films can help to distinguish these conditions, although these techniques are not widely available. A high platelet count may be indicative of chronic gastrointestinal bleeding or active inflammation.

The definitive test to diagnose iron deficiency is the bone marrow aspiration. The presence of stainable iron stores in the marrow excludes iron deficiency as a cause for anaemia. The major drawback of marrow aspiration is its invasive nature.

Investigations of iron metabolism

Serum iron and total iron binding capacity

Under most clinical circumstances the serum iron is a measure of the amount of iron bound to transferrin. In an acute infection or inflammatory condition there is a rapid fall (within 24 h) in the serum iron, which persists as long as the illness does. Therefore the serum iron is low in both the anaemia of chronic disease and iron deficiency and does not distinguish between the two. Further problems with the interpretation of serum iron levels are (i) technical difficulties in measurement and (ii) the 10 to 40 per cent variation in levels seen in an individual during the course of the day.

Transferrin is measured indirectly by the total iron binding capacity of plasma. In the normal subject the serum iron is approximately 18 mmol and the total iron binding capacity 56 mmol, indicating that only about 30 per cent of the transferrin binding sites are occupied. Transferrin has a half-life of 10 days and is not subject to rapid change in its concentration, as such the total iron binding capacity remains relatively constant. It is one of the most useful tests to distinguish between chronic disease and iron deficiency. In iron deficiency the total iron binding capacity is elevated (due to increased hepatic synthesis of transferrin) whilst in chronic disease it is reduced (due to decreased hepatic synthesis of transferrin), although there is considerable overlap in the values of the two groups.

The transferrin saturation is calculated by dividing the serum iron by the total iron binding capacity and expressing the result as a percentage. The transferrin saturation is reduced in both chronic disease and iron deficiency, although to a greater extent in the latter. Again there is a wide overlap in the range of values obtained.

Ferritin

The iron storage protein, ferritin, is valuable in distinguishing between the anaemia of chronic disease and iron deficiency (Finch 1986; Porter *et al*. 1994). In patients without inflammatory disease the serum ferritin correlates with body iron stores (normal range is 20 to 500 mg /l in men and 10 to 200 mg/l in women) and therefore low ferritin indicates iron deficiency. In the rheumatoid patient the situation is more complicated because ferritin is an acute phase protein: a low serum ferritin indicates iron deficiency but a normal ferritin does not exclude deficiency in the presence of an acute phase response. Markedly high levels of ferritin are seen in particular myeloproliferative disorders and systemic juvenile chronic arthritis (see later).

Transferrin receptor

In the absence of chronic disease the concentration of the transferrin receptor in the plasma correlates with iron deficiency. The receptor concentration is increased in the anaemia of chronic disease, although not to the extent seen in iron deficiency. It is possible that this may prove to be a useful test to distinguish the two, although at present it does not confer any advantages over the conventional tests (Nielsen *et al*. 1994; Petterson *et al*. 1994).

In summary

There is no single diagnostic test for the anaemia of chronic disease. Table 2 summarizes the investigations that are possible. The diagnosis is made on the basis of an appropriate clinical picture, an elevated acute-phase response and the characteristic findings in the full blood count and biochemical indices of iron metabolism. The most useful indices are the total iron binding capacity, ferritin and serum iron. In most circumstances these will provide enough information to make the correct diagnosis.

		Anaemia of chronic disease	Iron deficiency
Clinical features	Systemic disease	present	absent
	Duration	> 2 months	variable
	Age	any	increased age
	NSAIDs	+/-	Yes
	Steroids	-	Yes
	Peptic ulcer history	-	Yes
Inflammatory markers	ESR	↑↑	N or ↑
	CRP	↑	-
Haematology indices	Haemoglobin	normally > 90 g/dl	May be < 10 g/dl
	MCV	Normal or ↓	↓
	Bone marrow	-	Absent iron stores
	Erythropoietin	↑	↑↑
	Response to iron	Poor*	Good*
Iron studies	Serum iron	↓	↓
	Ferritin	↑	↓
	TIBC	↓	↑
	Transferrin saturation	↓	↑
Biochemical indices	Serum urea	-	↑ (if recent bleed)
	Faecal occult bloods	+/-	positive

*Values marked are of greatest diagnostic significance
The following points are emphasized:
a A low serum ferritin indicates iron deficiency
b In the presence of an elevated acute-phase response a normal or slightly raised serum ferritin does not exclude iron deficiency
c In patients with acute systemic disease it is impossible not to administer some parenteral iron. Intravenous iron is normally given as iron dextran in a gastrointestinal bleed
ESR, erythrocyte sedimentation rate; MCV, mean corpuscular volume; NSAIDs, non-steroidal anti-inflammatory drugs; TIBC, total iron binding capacity

Table 2 Comparison of the typical features seen in the anaemia of chronic disease and iron deficiency

Treatment

The priorities in the treatment of the anaemia of chronic disease must be suppression of the disease activity and identification of any concomitant haematinic deficiencies. These measures are of obvious benefit to the patient and are relatively straightforward. Blood transfusion is used rarely, normally being reserved for the management of severe symptomatic anaemia or as a preoperative measure.

The availability of two recombinant forms of erythropoietin (erythropoietin-a and erythropoietin-b) provide a potential new therapy. The recombinant hormone is effective in the treatment of the anaemia of chronic renal failure, which is characterized by very low concentrations of erythropoietin. In contrast the concentration of erythropoietin is increased in the anaemia of chronic disease (with respect to the normal population), although there is a relative deficiency for the degree of anaemia. It is clear that the administration of recombinant erythropoietin can increase the haemoglobin to a normal level in the anaemia of chronic disease, the problem is the assessment of the cost–benefit of this.

Erythropoietin must be administered parenterally either intravenously or subcutaneously. The intravenous route has a higher bioavailability and peak serum concentration but a shorter half-life than that achieved by subcutaneous administration, although there are no distinct clinical advantages of one route over the other. Erythropoietin is normally commenced at a dosage of 50 to 150 U/g body weight thrice weekly, the dosage is then titrated against the response. A number of clinical studies have reported investigating the effectiveness of erythropoietin in the treatment of the anaemia of chronic disease in rheumatoid arthritis (Pincus *et al*. 1990; Salvarani *et al*. 1991; Petterson *et al*. 1993; Murphy *et al*. 1994). The following conclusions can be drawn:

1. Erythropoietin is able to correct the anaemia of chronic disease.
2. The effectiveness of erythropoietin is reduced in the presence of an elevated acute-phase response.
3. Erythropoietin increases the haemoglobin concentration by a direct effect on erythroid precursors in the marrow.
4. There is little evidence to suggest that erythropoietin has any disease suppressive activity, although in certain patients the ability to perform aspects of daily living is improved.
5. Treatment with erythropoietin can unmask latent deficiencies of haematinics (most commonly iron, but vitamin B_{12} and folate deficiency can occur). These should be monitored for, as a deficiency will limit the effectiveness of erythropoietin, and regular iron prophylaxis has been recommended.
6. Treatment with erythropoietin is associated with very few side-effects in the anaemia of chronic disease.

In summary, erythropoietin is effective in the treatment of anaemia of chronic disease and in some patients results in a symptomatic benefit. At present there is no place for its routine use in the treatment of anaemia of chronic disease in rheumatoid arthritis, except in special circumstances. It may prove useful for patients with severe symptomatic anaemia who are intolerant of conventional treatments (including blood transfusion), and in patients receiving autologous blood transfusion at the time of elective joint arthroplasties (Saikawa *et al*. 1994).

The major limiting factor preventing the use of erythropoietin is cost. Currently a 1000 U dose of erythropoietin costs £9. This translates to a cost of £8400 for a 60 kg patient receiving 300 U/kg per week for a year, and £1300 for a 1-month course at 600 U/kg per week prior to arthroplasty.

The acute phase response (see also Chapter 4.1)

The term acute phase response is used to describe the complex series of local and systemic events which accompany tissue inflammation (acute or chronic). It is normally self limiting and results in tissue repair. However if it becomes persistent, the process results in tissue damage and the systemic features of chronic inflammation.

Central to the pathogenesis of both the local inflammation and systemic effects are the pro-inflammatory cytokines interleukin 1 (IL-1), interleukin 6 (IL-6), and tumour necrosis factor-a (TNF-a). These products of macrophages, fibroblasts and endothelial cells are the mediators of the local and systemic effects of inflammation. Of particular note is their stimulation of the liver to generate the acute phase proteins whose measurement in the plasma can be used to quantify the magnitude of the acute phase response (Ganpathi *et al*. 1991; Mackiewicz *et al*. 1991; Thompson *et al*. 1992). The acute phase proteins most commonly measured are the C-reactive protein (measured directly) and fibrinogen (measured indirectly as the erythrocyte sedimentation rate and plasma viscocity). These measurements are non-specific markers of an inflammation but are often used clinically in the following situations:

1. discrimination between inflammatory and non-inflammatory conditions;
2. assessment of disease activity and response to therapy.

There are significant differences between basal concentration, incremental changes and response times in inflammatory situations of the acute phase proteins (see Chapter 4.1). The other determinant of the concentration of the acute phase reactants is the rate of catabolism. Compared with the other acute phase reactants the C-reactive protein has the distinct advantage of a constant rate of catabolism independent of other factors. In contrast, disseminated intravascular coagulation, intravascular haemolysis and complement consumption will all result in decreased levels of fibrinogen, haptoglobin, and complement and therefore give an underestimate of the magnitude of the acute phase response.

The erythrocyte sedimentation rate (ESR)

The ESR provides an indirect measure of the acute phase response. The ESR is a measure of the rate of sedimentation of erythrocytes in a vertical tube over a standard period of 1 h. The rate of sedimentation of erythrocytes is determined by the negative electrical charge on their surface which prevents aggregation. The presence of certain plasma proteins (particularly fibrinogen) reduces the negative charge on the erythrocyte surface causing an increase in aggregation and sedimentation. The relative contribution of the plasma proteins to the ESR are: fibrinogen 55 per cent, a $_2$-macroglobulin 27 per cent, immunoglobulin 11 per cent, and albumin 7 per cent (Stuart and Whicher 1988). Erythrocyte aggregation is also influenced by the number, shape, density and deformability of erythrocytes, so that the ESR increases with decreasing haematocrit. As the ESR predominantly reflects changes in the concentration of fibrinogen, there is generally a lag phase of 24 to 48 h between the onset or resolution of the inflammatory stimuli and changes in the ESR. However, as this reflects changes in the concentrations of several acute phase reactants (as well as immunoglobulins), it provides a wider screen for the detection of inflammatory disease than the measurement of a single acute phase protein. The major advantages of the ESR are its sensitivity, ease of performance, low cost and familiarity. The major disadvantages are its low specificity, the interobserver variation of its measurement and the wide range of other factors which will affect it. The relative merits of the ESR, plasma viscosity and C-reactive protein are shown in Table 3.

	ESR	Viscosity	CRP
Sample	Must be fresh	Can be stored	Can be stored
Method	Manual	Automated	Automated
Variables	Dependent on age, anaemia etc.	Independent of age and anaemia	Independent of age and anaemia
Half-life	Rises 24 to 48 h after tissue damage.	Rises 24 to 48 h after tissue damage	Rises 6 h after tissue damage
	Subsides with half-life of 4 to 6 days	Subsides with half-life of 4 to 6 days	Subsides with half-life of 48 h
Useful for monitoring	Chronic inflammatory change	Chronic inflammatory change	Acute inflammatory change

Table 3 The relative merits of the erythrocyte sedimentation rate, plasma viscosity and C-reactive protein in the monitoring of rheumatic disease

The plasma viscosity

The plasma viscosity, like the ESR, is determined predominantly by the concentration of fibrinogen. It has the same kinetics in response to inflammation as the ESR, but the advantage of being independent of age and erythrocyte concentration and morphology. In addition the test can be automated and performed on stored samples.

The C-reactive protein (CRP)

The CRP is generally accepted as the most accurate measure of the acute phase response and hence of tissue inflammation. In the normal individual the basal concentrations are extremely low or undetectable. As a consequence of this a patient's CRP concentration can increase many times from basal level and still lie within the normal population range whilst being markedly abnormal for that individual. In response to inflammation the concentration can rapidly increase up to one thousandfold. Even at these high concentrations the rate of catabolism is constant and therefore the plasma concentration is determined by the rate of synthesis. The CRP is not affected by any of the confounding factors which influence the ESR. In addition it can be quantified accurately even at very high levels of inflammation.

The rapid change in the concentration of CRP in response to changes in the degree of tissue inflammation make it the best measurement available for the monitoring of therapy in the rheumatic diseases. All three investigations are suitable for the screening of inflammatory disease when rapid changes in tissue inflammation are of less importance

Serum amyloid A protein is an acute phase reactant that shares many of the characteristics of the CRP and interest focuses upon it because it is the precursor of type AA amyloid fibrils. However, it is poorly immunogenic and there is no readily available assay for its measurement.

The clinical use of the measurement of the acute phase response

Rheumatoid arthritis

The CRP and ESR are commonly used as screening investigations for an inflammatory arthritis. In addition, the CRP can be used prognostically and to assess the contribution of disease activity to the anaemia of chronic disease (see previous section).

There is a good correlation between clinical measures of disease activity and the acute phase response measured by either ESR, plasma viscosity or CRP (Amos *et al*. 1977; Dawes *et al*. 1986; Scott *et al*. 1987). However, it should be emphasized that the kinetics of the CRP make it the only suitable test for monitoring rapid changes in the acute phase response, as occurs when systemic steroids are used. These may decrease the acute phase response by 50 per cent in 48 h, even after intra-articular injection. As a generalization the ESR is used for screening and the CRP for monitoring

In early rheumatoid arthritis, generalized osteoporosis is correlated significantly with an elevated CRP (Gough *et al*. 1994). When the disease process and secondarily the acute phase are suppressed by therapy there is a decrease in skeletal demineralization and stabilization of the functional state of the patient. If the patient's acute phase response appears elevated inappropriately for the clinical extent of disease the following should be considered:

1. Systemic amyloidosis—long-term chronic inflammation is associated with the development of secondary amyloidosis with deposition of amyloid fibrils in the heart, kidneys, liver, and gastrointestinal tract. The increased production of serum amyloid A is presumably responsible for the elevated ESR.
2. Infection—a concomitant bacterial infection will markedly increase the acute phase response. A particular concern is the development of septic arthritis that may be clinically unsuspected, with the physical signs of joint infection being attributed wrongly to a flare of rheumatoid disease. If this is considered seriously then joint aspiration and culture must be performed.
3. Malignancy—the development of occult malignancy must always be borne in mind.

The markedly elevated ESR

Modest elevations in the ESR (30 to 50 mm) are difficult to interpret and, especially in the absence of elevated CRP, a specific diagnosis may not be made. An ESR greater than 100 mm, conversely, is very specific for organic pathology and there are few conditions which present frequently to rheumatologists that are associated consistently with this level of ESR.

In polymyalgia rheumatica the ESR is markedly elevated but is not essential for the diagnosis. If raised the ESR (or CRP) can be used to monitor the disease and response to treatment. In giant cell arteritis, the raised ESR is a reliable indicator of disease activity, sufficient to be a major factor in dictating therapy.

In patients presenting with back pain and a markedly elevated ESR, myeloma, infective discitis, and osteomyelitis all need to be considered. Finally the possibility of the development of malignancy must be considered.

Elevated ESR/normal CRP

All the above conditions would also be accompanied by a markedly elevated CRP. However, there are occasions when these two are discrepant. These conditions are characterized by factors that will elevate the ESR but do not cause elevation of the acute phase proteins. Examples include anaemia, hyperlipidaemia, and hypergammaglobulinaemia. This is part of the explanation of the situation in systemic lupus erythematosus (see later) and Sjögren's syndrome.

The acute phase response in systemic lupus erythematosus (see Chapter 5.7.1)

In clinically active systemic lupus erythematosus, the ESR is elevated and may return to normal in remission. In contrast the CRP is rarely elevated in active disease, except in the presence of marked synovitis or concomitant bacterial infection (Pepys *et al*. 1982; Ter Borg *et al*. 1990). A clinical problem commonly encountered in the management of patients with systemic lupus erythematosus is to distinguish between an acute disease flare and the development of infection. The CRP can be helpful in this respect in that it is elevated in bacterial infection but not disease flares (a normal CRP makes infection an unlikely cause of an acute illness but an elevated CRP does not exclude a lupus flare). The reason why the CRP is not elevated in active lupus is not fully understood (Gordon and Emery 1993).

There is evidence to suggest that the circulating cytokines are not as active in systemic lupus erythematosus as in rheumatoid arthritis. The evidence for a hepatocyte response to TNF-a and IL-6 in active lupus disease is not consistent (Meijer *et al*. 1993). CRP, which is induced predominantly by IL-6, is elevated rarely except with serositis and then only modestly (Gabay *et al*. 1993). In contrast, a-acid glycoprotein, which is induced predominantly by TNF-a, can be elevated in flares of systemic lupus erythematosus alone. Similarly, the levels of increase of fibrinogen and depression of albumin are less in lupus than one would see in rheumatoid arthritis.

However, the failure to mount an acute phase response is not due to a deficiency in the capacity to produce CRP, because high levels are found during infection. This may be due to lower circulating levels of TNF-a and IL-6 in systemic lupus erythematosus without infection than in rheumatoid arthritis, or to the lack of another cytokine such as IL-11, which has recently been shown to promote the CRP response. In other inflammatory diseases such as rheumatoid arthritis, the acute phase response is associated not only with the hepatocyte effects but also with end-organ responses such as osteoporosis and thrombocythaemia. These features are seen to a lesser degree, if at all in systemic lupus erythematosus. This suggests that the cytokines or their soluble receptors are not elevated to the same degree or are not as active in lupus as rheumatoid arthritis. This may reflect differences in the cytokines, soluble receptors and other proteins produced in response to an increase in TNF-a or IL-6, which control the effects of these cytokines, as it is often difficult to isolate the effects of one or more cytokine when they are all part of a network of interacting proteins.

The acute phase response in adult Still's disease

Adult Still's disease is an acute systemic inflammatory condition with multiple non-specific clinical manifestations. These include fever, arthralgia and arthritis, lymphadenopathy, sore throat, and skin rashes. Elevated vitamin B_{12} levels are also found. The correct diagnosis is often difficult to make because their are no pathognomonic symptoms, physical signs or investigations. This means that a large number of investigations are performed to exclude other conditions, with consequent delay in the commencement of definitive therapy.

The characteristic haematological investigations are for normochromic normocytic anaemia (ACD), leucocytosis and thrombocytosis. There is an elevated acute-phase response with increased CRP and ESR. The abnormality of greatest interest is the serum ferritin, which is often dramatically elevated to more than 4000 mg/l (normal level is less than 300 mg/l) in active disease and may reach levels greater than 10 000 mg/l whilst the serum iron is low (Ohta *et al*. 1987a; Ohta *et al*. 1987b; Gonzales-Hernandez *et al*. 1989). Indeed in the appropriate clinical setting this may be taken as supportive evidence of the diagnosis and can be used to monitor the response to therapy (Schwaz-Eywill *et al*. 1992). There are a number of other conditions associated with a markedly raised ferritin including acute and chronic leukaemias, lymphomas, melanomas, germ cell tumours, and haemochromatosis, but in all of these it is unusual for the ferritin to be greater than 4000 mg/l. The elevation of the ferritin is due to its behaviour as an acute phase protein, but why it is elevated to such an extent in comparison with the other acute phase reactants in adult Still's disease is not known.

Haematological abnormalities in systemic lupus erythematosus (see Chapter 5.7.1)

Abnormalities of the erythrocyte

Perhaps the most common abnormality in patients with systemic lupus erythematosus is the anaemia of chronic disease. This is seen in between 30 and 50 per cent of patients. It has the features described in the previous section and will not be described further (Table 4).

Abnormalities of erythrocytes	Anaemia of chronic disease
	Autoimmune haemolytic anaemia
	Iron deficiency anaemia
	Chronic renal failure
	Microangiopathic haemolytic anaemia
	Evan's syndrome
Abnormalities of leucocytes	Lymphopenia
	Granulocytopenia
Abnormalities of platelets	Autoimmune thrombocytopenia
	Thrombotic thrombocytopenic purpura
	Acquired defects of platelet function
Clotting abnormalities	Antiphospholipid syndrome

Table 4 The haematological manifestations of systemic lupus erythematosus

A direct Coombs' test positive for autoimmune haemolytic anaemia is seen in 10 per cent of patients, and overt symptomatic haemolysis is uncommon. Haemolysis is secondary to immunoglobulin directed against antigens on the erythrocyte membrane. Complement bound to the erythrocyte membrane is attached normally to specific C3b receptors and not directed against erythrocyte antigens, and therefore does not induce haemolysis. Both warm and cold haemolysis occurs, although the former is more common. The degree of haemolysis is generally mild and does not require specific treatment. If the haemolysis is severe, immunosupression with steroids, azathioprine, and cyclophosphamide may be employed. Evan's syndrome is the combination of haemolytic anaemia and immune thrombocytopenia. Microangiopathic haemolysis is rare but can occur as a complication of the antiphospholipid syndrome (Jain *et al*. 1994; Nesher *et al*. 1994).

Iron deficiency anaemia is common and can develop secondary to drug therapy (e.g. non-steroidal anti-inflammatory drugs) or, extremely rarely, as a consequence of pulmonary haemorrhage due to pulmonary vasculitis. Anaemia will occur if chronic renal failure develops.

Abnormalities of leucocytes

Lymphopenia is a common feature of active disease. The condition is due to lymphocytotoxic antibodies (either of IgM or IgG class) directed against resting or active cells. Immunosupression of active disease produces an increase in the lymphocyte count.

Granulocytopenia is seen in active disease. It is due to a combination of central inhibition of granulocytopoiesis, splenic sequestration of granulocytes, and cytotoxic antibodies. The other major cause of granulocytopenia, which must never be forgotten, is as a consequence of drug therapy.

Abnormalities of platelets

A mild immune thrombocytopenia (100 to 150×10^9/l) occurs in up to 45 per cent of patients with active disease. Clinically significant thrombocytopenia (less than 50×10^9/l) occurs in 10 per cent of patients, who are at risk of bleeding. Up to 15 per cent of patients (especially in childhood) with lupus will present with thrombocytopenia similar to classical immune thrombocytopenia. This may precede the development of other signs of lupus by months or years. The thrombocytopenia is the result of increased peripheral destruction of platelets consequent upon autoantibodies directed against membrane antigens. Bone marrow examination shows normal or increased numbers of megakaryocytes to compensate for the increased peripheral consumption. In the antiphospholipid syndrome, anticardiolipin antibodies may have a direct role in platelet destruction.

In mild cases of immune thrombocytopenia no treatment is needed. In severe cases high doses of prednisolone (0.25 to 1 mg/kg) are required. The initial dose is maintained for 3 to 4 weeks and then titrated against the response. Steroid failures have been treated with intravenous immunoglobulin, splenectomy, vinca alkaloids, and cyclosporin.

In addition acquired defects of platelet function have been reported, including decreased aggregation and defective uptake of adenosine diphosphate and serotonin.

Clotting abnormalities

In patients with systemic lupus erythematosus there is an increased tendency to thrombosis due to a higher incidence of antiphospholipid antibodies, which predispose to the development of thrombosis. There are three serological tests commonly used to detect antiphospholipid antibodies:

1. Anticardiolipin antibodies—these can be measured quantitatively by immunoassay and correlate best with the presence of antiphospholipid antibodies. These are found in 15 to 40 per cent of patients.
2. Lupus anticoagulant—either an IgM or IgG antibody which reacts with the negatively charged phospholipid component of the prothrombin activator complex

resulting in a prolonged activated partial thromboplastin time. The name is a misnomer because it is encountered more frequently in patients without systemic lupus erythematosus and is associated with a clotting tendency. These are found in up to 2 per cent of patients with lupus.

3. False-positive VDRL test—antiphospholipid antibodies cross-react with phospholipid in the VDRL test to give a false-positive result when testing for syphilis serology.

A clinical syndrome termed the antiphospholipid syndrome (see Chapter 5.7.3) has been increasingly recognized in association with antiphospholipid antibodies (for review see Asherson and Cerevera 1992; Hughes 1993). The syndrome can be primary or secondary; it is primary if there is no evidence of a pre-existing connective tisssue disease (e.g. systemic lupus erythematosus), or secondary if the latter is present. The classical syndrome comprises widespread arterial and venous thrombosis, thrombocytopenia and recurrent abortions, although improved recognition of the condition has led to the realization that there are many more features (Table 5). These are thought to be attributable to the thrombogenic effects of the antiphospholipid antibodies. The range of possible mechanisms which results in thrombotic tendency associated with antiphospholipid antibodies include effects on platelet membranes, endothelial cells, and on clotting components such as prothrombin, protein C and protein S. The thrombotic tendency is dependent on the antibody titre, isotype (IgG), and the duration of time that the antibodies have been present

Site of thrombosis	Complications
Hepatic	Budd–Chiari syndrome
	Hepatic infarction
Adrenal	Addison's disease
Pulmonary	Thromboembolic pulmonary hypertension
Cerebral	Cognitive impairment
	Movement disorders
	Epilepsy
	Migraine
	Transverse myelitis
Cardiac	Cardiomyopathy
	Valvular disorders
Renal	Thrombotic microangiopathy
Dermatological	Livedo reticularis
	Chronic leg ulceration

Table 5 Complications of the antiphospholipid syndrome associated with thrombosis

The antiphospholipid syndrome occurs in 5 to 10 per cent of patients with systemic lupus erythematosus. It does not correlate clearly with disease activity and does not respond well to immunosuppression. The mainstay of treatment is long-term oral anticoagulation.

Leucocyte abnormalities in rheumatic disorders

Neutrophilia

The commonest causes of an increased neutrophil count (neutrophilia) in patients with rheumatic diseases are steroid therapy and infection. Other causes are shown in Table 6.

Common rheumatological causes	Steroid therapy	
	Infection	
	Active disease:	Rheumatoid arthritis
		Juvenile chronic arthritis
		Gout
Other rheumatological causes	Vasculitis:	Polyarteritis nodosa
		Giant cell arteritis
		Wegener's granulomatosis
Non-rheumatological causes	Tissue infarction	
	Metabolic:	Uraemia
		Diabetic acidosis
	Malignancy	
	Bleeding	
	Pregnancy	

Table 6 The causes of neutrophilia in patients with rheumatic diseases

Neutropenia

Neutropenia is defined as the circulatory neutrophil count being below 1.5×10^{9}/l. The most important causes are second-line agents and non-steroidal anti-inflammatory drugs used in the treatment of the rheumatic diseases. These drugs often cause marrow depression, in particular neutropenia. For some drugs, such as gold and penicillamine, this side-effect is not dose related but appears to be idiosyncratic and major histocompatibility complex (MHC) class II DR3 related. For other drugs, for example the cytotoxics, it is dose related and generally involves other cell lines as well. For other causes of neutropenia see Table 7.

Common rheumatological causes	Drugs:	Idiosyncratic
		Dose-dependent
	Active disease:	Systemic lupus erythematosus
Other rheumatological causes	Felty's syndrome	
Non-rheumatological causes	Viral infections:	Influenza
		Infectious mononucleosis
		Infectious hepatitis
	Bacterial infections:	Overwhelming sepsis*
		Typhoid
		Brucellosis
	Severe folate or vitamin B_{12} deficiency	
	Racial (common in Afro-Caribbeans)	
	Pancytopenia, any cause	

*Especially Gram-negative staphylococci.

Table 7 The causes of neutropenia in patients with rheumatic diseases

Presentation

Patients may complain of a sore throat or mouth ulceration, or may have signs and symptoms of infection. More commonly, they may be asymptomatic and the neutropenia only discovered on screening.

Management

If a drug is thought to be responsible for the neutropenia then it should be withdrawn immediately. Haemopoietic recovery normally occurs within one to two weeks. If

the patient is 'well' and afebrile, no specific treatment is necessary but they need to be monitored closely for a continuing deterioration in the neutropenia and for signs of infection. However, if the patient is unwell they require admission. The differential diagnosis includes infection and active disease. Neutropenia *per se* may produce a low-grade fever. An infection screen needs to be performed and bone marrow aspiration considered. The prognosis is particularly poor if the patient has a serious infection when a granulocytosis is first discovered. Broad spectrum antibiotics are required together with supportive therapy. The administration of recombinant human granulocyte-colony stimulating factor may help by reducing the duration of the neutropenia and thereby decreasing the incidence of sepsis. Its expense precludes its routine use at present.

Eosinophilia

Eosinophilia is defined as the number of eosinophils being greater than 0.4×10^{9}/l or greater than 5 per cent of the total white cell count, in the peripheral blood. History, examination and a few simple tests usually point to the cause of the eosinophilia. Often little is gained by further exhaustive testing to find a cause. For causes see Table 8.

- Drug reactions, e.g. gold induced rashes
- Severe rheumatoid arthritis: associated with episcleritis, subcutaneous nodules, vasculitis, pleuropericarditis, pulmonary fibrosis, high rheumatoid factor titres, raised gammaglobulin
- Vasculitis, e.g. polyarteritis nodosa, Churg-Strauss syndrome
- Sarcoidosis
- Eosinophilic fasciitis (localized infiltration of the connective tissue of the upper and lower limbs, usually sparing the digits. May lead to fibrosis and subsequent functional loss).
- Eosinophilic myalgic syndrome
- Non-rheumatological causes pulmonary disorders, e.g. bronchial asthma, tropical pulmonary eosinophilia, allergic bronchial aspergillosis, Loeffler's syndrome, hypereosinophilic syndrome; parasite infections; allergic reactions, e.g. hay fever; skin disorders, e.g. urticaria, pemphigus, eczema, dermatitis herpetiformis malignancy

Table 8 The causes of eosinophilia in patients with rheumatic diseases

Felty's syndrome

This syndrome was reported by Felty in 1924 (Felty 1924). He noted the association of rheumatoid arthritis, splenomegaly and neutropenia. One per cent of patients with rheumatoid arthritis are affected.

It is usually seen in patients with long-standing seropositive, nodular, deforming rheumatoid arthritis, with systemic involvement. It is more common in certain HLA genotypes; over 90 per cent of patients are DR4 positive with a striking proportion being homozygous for the HLA-Dw4 allele (a subtype of HLA-DR4) (Sanson *et al*. 1987; Lanchbury *et al*. 1991). Approximately one-third of patients have no active synovitis when Felty's syndrome develops. The splenomegaly is usually mild to moderate but it can be huge. Sometimes it may precede leucopenia by years. Other clinical manifestations include weight loss, lymphadenopathy, leg ulcers, skin hyperpigmentation, peripheral neuropathy, chronic indolent infections, and nodular regenerative hyperplasia of the liver (Moots *et al*. 1994). Bacterial infections are common due to the marked neutropenia and account for most deaths in this syndrome.

Laboratory investigations typically reveal high titres of rheumatoid factor and antinuclear antibodies, with hypocomplementaemia. Thrombocytopenia is seen in about 40 per cent of cases but is usually mild. The coexisting anaemia is usually consistent with the anaemia of chronic disease but with the addition of a reticulocytosis, in response to increased erythrocyte destruction in the spleen.

The neutrophil count is usually between 0.5×10^{9} and 2.5×10^{9}/l. Several mechanisms for the cause of this neutropenia have been postulated. It is generally too profound to be wholly accounted for by hypersplenism and may persist after splenectomy (Logue *et al*. 1981). Bone marrow examination has shown evidence of maturation arrest of the granulocyte cell line. Antigranulocyte antibodies and the presence of T lymphocytes in the marrow that inhibit granulopoiesis have also been demonstrated. Corticosteroid therapy often improves the neutrophil count in patients in whom these inhibitory T cells have been demonstrated.

There is no consistently effective therapy for this syndrome. Splenectomy, corticosteroids, methotrexate, gold, penicillamine, cyclophosphamide, cyclosporin, lithium salts, parenteral testosterone, and granulocyte-colony stimulating factor have all been used, with varying degrees of success in different patients (Dillon *et al*. 1986; Fiechtner *et al*. 1989; Markusse *et al*. 1990). Splenectomy does not correlate well with reduction in the risk of recurrent infections or with ulcer healing. When splenectomy is planned, it is important to protect the patient against infection by bacteria with polysaccharide coats, such as *Streptococcus pneumonaie* and *Salmonella* spp. This may be best achieved by vaccination with Pneumovax (protecting against *S. pneumoniae)* before splenectomy, or daily penicillin V if prevaccination is not possible. Corticosteroids, especially in higher doses, may make the infections worse. Sometimes treatment of the underlying disease process will lead to an improvement of the neutropenia. The use of second-line drugs is therefore not contraindicated in the presence of a leucopenia but judicious monitoring is necessary. Bone marrow aspiration is probably advisable before commencing such therapy in order to exclude marrow suppression. Methotrexate is probably the treatment of choice.

Large granular lymphocyte syndrome/granular lymphocyte expansion

This is a recently recognized syndrome of neutropenia, associated with a proliferation of large granular lymphocytes in blood and bone marrow. The lymphocytes have many azurophilic granules in the cytoplasm and tend to be a subset of T lymphocytes with CD3+ CD8+Leu7+ surface phenotype (Snowden *et al*. 1991). B-cell function is also abnormal with autoantibody production and marked hypergammaglobulinaemia. The actual white blood cell count may be normal or increased because of the lymphocytosis. Other features are thrombocytopenia, anaemia, splenomegaly, and arthritis. It occurs in 0.6 per cent of patients with rheumatic disease and tends to affect older patients early in the course of their arthritis. Generally it is a benign condition but treatment in more severe cases is with glucocorticoids and/or immunosuppressives. Intermittent chlorambucil has achieved a good clinical response in patients with weight loss, fevers, and night sweats. This treatment has also produced an increase in haemoglobin and/or neutrophils, a small fall in large granular lymphocytes, and reduction in splenomegaly (Snowden *et al*. 1991). Splenectomy is not recommended. The main reason for ill health in these patients is infection secondary to the neutropenia.

At times it can be difficult to distinguish between large granular lymphocyte expansion and Felty's syndrome. The definitive method requires consideration of lymphocyte morphology, bone marrow examination, and analysis of lymphocyte surface phenotype. The condition is contrasted with Felty's syndrome in Table 9.

	Felty's syndrome	Granular lymphocyte expansion
Neutrophil count	↓	↓
Lymphocyte count	↓	↑
Thrombocytopenia	Yes	Yes
Anaemia	Yes	Yes
Splenomegaly	Yes	Yes
Age of patient	Younger	Older
Duration of disease	Longer	Shorter
Prognosis	More severe	More benign

Table 9 Comparison of Felty's syndrome with granular lymphocyte expansion

Platelet abnormalities

Thrombocytosis

A raised platelet count is common in patients with active rheumatoid arthitis. There is a positive correlation between platelet count and disease activity. Extreme thrombocytosis is noticed with some extra-articular manifestations of the disease, particularly pulmonary involvement, peripheral neuropathy, and vasculitis. The mechanism is uncertain. Interestingly it does not predispose to an increase in thrombotic events. Other causes to be considered include acute bleeding and iron deficiency. In these cases, the platelet count will return to normal if the underlying disease process is treated.

Thrombocytopenia

Two important causes of thrombocytopenia to consider in patients with rheumatic diseases are autoimmunity and drugs.

1. Autoimmune thrombocytopenia: see lupus haematology section. It has also been reported in mixed connective tissue disease, dermatomyositis, and systemic sclerosis.
2. Drug-related thrombocytopenia: Many drugs used to treat rheumatic diseases can cause thrombocytopenia by marrow suppression. It is therefore paramount that patients are monitored closely with regular full blood counts whilst on these drugs. If the platelet count drops below 150×10^{9}/l the medication needs to be reviewed immediately and stopped. The platelet count usually recovers within the next couple of weeks, although it can take longer with some drugs, for example gold which has a long half-life. If the patient experiences severe thrombocytopenia with bleeding and marked purpura, then platelet transfusions may be necessary.
3. Other causes of thrombocytopenia are Felty's syndrome and large granular lymphocyte expansion. Infections may also be responsible and these include cytomegalovirus, hepatitis B, human immunodeficiency virus, subacute bacterial endocarditis, and Gram-negative sepsis.

Infections, in particular, need to be considered in patients with rheumatic diseases as these patients are frequently immunocompromised.

Chapter References

Asherson, R.A. and Cerevera, R. (1992). The antiphospholipid syndrome: a syndrome in evolution. *Annals of the Rheumatic Diseases*, **51**, 147–50.

Baer, A.N., Dessypris, E.N., and Krantz, S.B. (1990). The pathogenesis of anaemia in rheumatoid arthritis: a clinical and laboratory analysis. *Seminars in Arthritis and Rheumatism*, **19**, 209–23.

Blick, M., Sherwin, S.A., Rosenblum, M., and Gutterman, J. (1987). Phase 1 study of recombinant tumor necrosis factor in cancer patients. *Cancer Research*, **47**, 2986–9.

Dillon, A.M., Luthra, H.S., Conn, D.L., and Ferguson, R.H. (1986). Parenteral gold therapy in the Felty syndrome. *Medicine*, **65**, 107–12

Elliot, M.J. *et al*. (1994). Randomised double blind comparison of chimeric monoclonal antibody to tumuor necrosis factor a (cA2) versus placebo in rheumatoid arthritis. *Lancet*, **344**, 1105–10.

Fairbanks, V.F. and Beutler, E. (1983). Iron metabolism. In *Hematology*, (3rd edn) (ed. W.J. Williams, E. Beutler, A.J. Erslev, and M.A. Litchman), pp. 301–10. McGraw Hill, London.

Faquin, W.C., Schneider, T.J., and Goldberg, M.A. (1992). Effect of inflammatory cytokines on hypoxia-induced erythropoietin production. *Blood*, **79**, 1987–84.

Felty, A.R. (1924). Chronic arthritis in the adult associated with splenomegaly and leucopaenia. *Johns Hopkins Hospital Bulletin*, **35**, 16.

Fiechtner, J.J., Miller, D.R., and Starkenbaum, G. (1989). Reversal of neutropaenia with methotrexate treatment in Felty's syndrome: correlation of response with neutrophil-reactive IgG. *Arthitis and Rheumatism*, **32**, 194–200.

Finch, C.A. (1986). Plasma ferritin determination as a diagnostic tool. *Western Journal of Medicine*, **145**, 657–63.

Gabay, C., Roux-Lombard, P., de-Moerloose, P., Dayer, J.M., Vischer, T., and Guerne, P.A. (1993). Absence of correlation between interleukin 6 and C-reactive protein in systemic lupus erythematosus compared with rheumatoid arthritis. *Journal of Rheumatology*, **20**, 815–21.

Ganpath, M.K., Rzewnicki, D., Samols, D., Jiang, S.L., and Kushner, I. (1991). Effect of combination of cytokines and hormones on synthesis of serum amyloid A and C-reactive protein in Hep-3B cells. *Journal of Immunology*, **147**, 1261–5.

Gonzalez-Hernandez, T, Martin-Mola, E., Fernandez-Zamoeano, A., Balso-Criado, A., and de Miguel-Menieta, E. (1989). Serum ferritin can be useful for diagnosis in adult onset Still's disease. *Journal of Rheumatology*, **16**, 412–13.

Gough, A.K.S., Lilley, J., Eyre, S., and Emery, P. (1994). Generalised bone loss in patients with early rheumatoid arthrtis. *Lancet*, **344**, 23–7.

Hughes, G.H. (1993). The antiphospholipid syndrome: ten years on. *Lancet*, **342**, 341–4.

Isenberg, D.A., Martin, P., Hajirousou, V., Todd-Pokropek, A., Goldstone, A.M., and Snaith, M.L. (1986). Haematological re-assessment of rheumatoid arthritis. *British Journal of Rhematology*, **25**, 152–7.

Jain, R., Chartash, E., Susin, M., and Furie, R (1994). Systemic lupus erythematosus complicated by thrombotic microangiopathy. *Seminars in Arthritis and Rheumatism*, **24**, 173–82.

Johnson, C.S., Chang, M.J., and Furmanski, P. (1988). *In vitro* effects of tumor necrosis factor a in normal and erythroleukemic mice: characterisation and therapeutic applications. *Blood*, **76**, 1875–83.

Johnson, C.S., Keckler, D.C., Torrer, M.I., Braunshweiger, P.G., and Furmanski, P. (1989 *a*). *In vivo* haemopoietic effects of recombinant interleukin-1a in mice: stimulation of granulocytic, monocytic, megakaryocytic, and early erythroid progenitors, suppression of late stage erythropoiesis, and reversal of erythroid supression with erythropoietin. *Blood*, **73**, 678–83.

Johnson, R.A., Waddelow, T.A., Caro, J., Oliff, A., and Roodman, G.D. (1989 *b*). Chronic exposure to tumor necrosis factor *in vivo* preferentially inhibits erythropoiesis in nude mice. *Blood*, **74**, 130–38.

Lanchbury, J.S.S *et al*. (1991). Strong primary selection for the Dw4 subtype of DR4 accounts for the HLA-DQw7 association with Felty's syndrome. *Human Immunology*, **32**, 56.

Lee, R.G. (1993). Microcytosis and the anaemias associated with impaired haemoglobin synthesis. In *Wintrobe's clinical haematology*, (9th edn) (ed. R.G. Lee, T.C. Bitchell, J. Foerster, J.W. Athens, and J.N. Lukens), pp. 791–807. Lea and Febiger, Philadelphia.

Logue, G.L., Huang, A.T., and Shimm, D.S. (1981). Failure of splenectomy in Felty's syndrome. The role of antibodies supporting granulocyte lysis by lymphocytes. *New England Journal of Medicine*, **304**, 580–3.

Mackiewicz, A., Speroff, T., Ganapathi, M.K., and Kushner, I. (1991). Effect of cytokine combinations on acute phase protein production in two hepatoma cell lines. *Journal of Immunology*, **146**, 3032–7.

Markusse, H.M., Breedveld, F.C., Fibbe, W.E., and Stern, A.C. (1990). Granulocyte-macrophage colony-stimulating factor treatment of neutropenia associated with Felty's syndrome. *Arthitis and Rheumatism*, **33**, 1865–7.

Meijer, C., Huysen, V., Smeenk, R.T., and Swaak, A.J. (1993). Profiles of cytokines (TNF a and IL-6) and the acute phase proteins (CRP and a 1 AG) related to the disease course in patients with systemic lupus erythematosus. *Lupus*, **2**, 359–65.

Moots, R.J., Elias, E., Hubscher, S., Salmon, M., and Emery, P. (1994). Liver disease in twins with Felty's syndrome. *Annals of the Rheumatic Diseases*, **53**, 202–5.

Murphy, E.A., Bell, A.L., Wojtulewski, J., Brzeski, M., Madhok, R., and Capell, H.A. (1994). Study of erythropoietin in treatment of anaemia in patients with rheumatoid arthritis. *British Journal of Medicine*, **309**, 1337–8.

Nesher, G., Hanna, V.E., Moore, T.L., and Hersh, M. (1994). Thrombotic microangiopathic hemolytic anaemia in systemic lupus erythematosus. *Seminars in Arthritis and Rheumatism*, **24**, 165–72.

Nielsen, O.J., Anderson, L.S., Hansen, N.E., and Hansen, T.M. (1994). Transferrin receptor levels in anaemic patients with rheumatoid arthritis. *Scandinavian Journal of Clinical and Laboratory*

Investigation, **54**, 75–82.

Ohta, A., Higashi, S., Suzuki, H., and Eto, S. (1987 *a*). Increased serum ferritin levels in adult Still's disease. *Lancet*, **i**, 562–3.

Ohta, A., Yamaguchi, M., Kaneoka, H., Nagyoshi, T., and Hilda, M. (1987 *b*). Adults Still's disease: review of 228 cases from the literature. *Journal of Rheumatology*, **14**, 1139–45.

Pepys, M.B., Lanham, J.G., and De Beer, F.C. (1982). C-reactive protein in systemic lupus erythematosus. *Clinics in the Rheumatic Diseases*, **8**, 91–103.

Petterson, T., Rosenlof, K., Friman, C., and Mickos, A. (1993). Successful treatment of the anaemia of rheumatoid arthritis with subcutaneously administered recombinant human erythropoietin. Slower response in patients with more severe inflammation. *Scandinavian Journal of Rheumatology*, **22**, 188–93.

Petterson, T., Kivivuori, S.M., and Siimes, M.A. (1994). Is serum transferrin receptor useful for detecting iron deficiency in anaemic patients with chronic inflammatory diseases? *British Journal of Rheumatology*, **33**, 740–47.

Pincus, T.P. *et al*. (1990). Multicentre study of recombinant human erythropoietin in correction of anaemia in rheumatoid arthritis. *American Journal of Medicine*, **89**, 161–8.

Porter, D.R., Sturrock, R.D., and Cappell, H.A. (1994). The use of serum ferritin estimation in the investigation of anaemia in patients with rheumatoid arthritis. *Clinical and. Experimental Rheumatology*, **12**, 179–82.

Roodman, D.G., Horadam, V.W., and Wright, T.L. (1983). Inhibition of erythroid colony formation by autologous bone marrow adherent cells from patients with the anaemia of chronic disease. *Blood*, **62**, 406–12.

Saikawa, I. *et al*. (1994). Autologous blood transfusion with recombinant erythropoietin treatment. Twenty-two arthroplasties for rheumatoid arthritis. *Acta Orthopaedica Scandinavia*, **65**, 15–19.

Salvarani, C. *et al*. (1991). Recombinant human erythropoietin therapy in patients with rheumatoid arthritis with the anaemia of chronic disease. *Journal of Rheumatology*, **18**, 1168–71.

Salvarani, C. *et al*. (1991). The role of interleukin 1, erythropoietin and red cell bound immunoglobulins in the anaemia of rheumatoid arthritis. *Clinical and Experimental Rheumatology*, **9**, 241–6.

Sanson, D.M., Bidwell, J.L., Maddison, P.J., Campion, G., Klouda, P.T., and Bradley, B.A. (1987). HLA DQ alpha and DQ beta restriction fragment length polymorphisms associated with Felty's syndrome and DR4-positive rheumatoid arthritis. *Human Immunology*, **19**, 269.

Schwaz-Eywill, M., Heilig, B., Bauer, H., Breitbart, A., and Pezzutto, A. (1992). Evaluation of serum ferritin as a marker for adult Still's disease activity. *Annals of the Rheumatic Diseases*, **51**, 683–5.

Snowden, N. *et al*. (1991). Large granular lymphocytes, neutropaenia and polyarthropathy: an underdiagnosed syndrome? *Quarterly Journal of Medicine*, **78**, 65–76.

Stuart, J. and Whicher, J.T. (1988). Tests for detecting and monitoring the acute phase response. *Archives of Disease in Childhood*, **63**, 115–17.

Takashina, N., Kondo, H., and Kashiwazaki, S. (1990). Suppressed serum erythropoietin response to anaemia and the efficacy of recombinant erythropoietin in the anaemia of rheumatoid arthritis. *Journal of Rheumatology*, **17**, 885–7.

Ter Borg, E.J., Horst, G., Limburg, P.C., Van Rijswijk, M.H., and Kallenberg, C.G.M. (1990). C-reactive protein during disease exacerbations and infections in systemic lupus erythematosus: a prospective longitudinal study. *Journal of Rheumatology*, **17**, 1642–8.

Thompson, D., Whicher, J.T., and Banks, R.E. (1992). Acute phase reactants in predicting disease outcome. In *Baillère's clinical rheumatology. Management of early inflammatory arthritis*, Vol. 6, pp. 393–404. Baillière Tindall, London.

Vreugdenhil, G., Wognum, A.W., van Eijk, H.G., and Swaak, A.J.D. (1990). Anaemia in rheumatoid arthritis: the role of iron, vitamin B12, and folic acid deficiency, and erythropoietin responsiveness. *Annals of the Rheumatic Diseases*, **49**, 93–8.

4.3 Biochemistry

Rosamonde E. Banks, Douglas Thompson, John T. Whicher, and Howard A. Bird

The rheumatologists' requirements of the clinical biochemistry laboratory vary according to the disease being managed. Automation of assays has led to a wider variety of biochemical estimations becoming more easily and inexpensively available. However, the contribution of the clinical biochemist to the diagnosis and management of the rheumatic illnesses is still relatively minor when compared with many other illnesses, possibly reflecting the relative chronicity and non-life-threatening nature of many of the rheumatic illnesses and the difficulty of devising new clinical laboratory tests when the aetiology of the disease is relatively uncertain. This chapter concentrates on those biochemical assays that are likely to be available in most hospitals for the routine management of rheumatic patients.

The clinical problem in rheumatology

Rarely is a single biochemical laboratory test truly diagnostic (except in the case of inborn errors of metabolism); rather they point to a pathological process or suggest damage or dysfunction of a particular organ, or they are used to monitor the course of a disease. As in most specialties, biochemical tests serve a number of well-defined purposes in rheumatology, which may be outlined as follows.

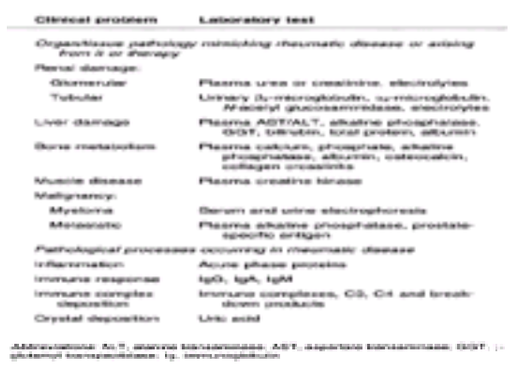
1. Indication of a pathological process and its intensity, e.g. inflammation, immune response, haemolysis, malignancy.
2. Indication of damage or dysfunction of organs, e.g. liver and kidney.

Tests may be used to:

1. Confirm or exclude a specific diagnosis. Several tests must be interpreted together with the clinical picture to create a probability of a diagnosis. Critical decisions are whether a particular organ or tissue is affected and what process (e.g. inflammation) is present.
2. Exclude other disorders which could (although deemed unlikely in the clinical context) give rise to the same symptoms. A panel of screening tests for other disorders is required, tailored to the clinical presentation. In a given situation these should always be used, thus preventing other quite different diagnoses from being overlooked. For example, malignant disease may cause myositis, myopathy or generalized symptoms suggestive of soft tissue rheumatism. However, normal test results never completely exclude a condition.
3. Assess and monitor disease activity. Prognosis and choice of therapy may depend on disease activity at presentation and control of therapy may depend on monitoring disease activity. Examples would include longitudinal measurements of acute phase proteins to assess disease activity in rheumatic diseases and the use of serum uric acid levels to monitor the effectiveness of long-term prophylactic allopurinol therapy for gout.
4. Detect a complication of the disease. This usually means detecting damage to an organ, such as the kidney, either at the time of diagnosis or during the course of the disease. Rheumatic diseases particularly associated with such complications include systemic lupus erythematosus, polyarteritis nodosa, Sjögren's disease, and gout.
5. Detect side-effects of therapy. Many of the drug therapies used in rheumatology can have unwanted side-effects, particularly in elderly people, often involving the liver or kidney. Sulphasalazine, azathioprine, and methotrexate are amongst this group of drugs.

The biochemical tests in rheumatology

Table 1 indicates the appropriate biochemical tests for assessing organ/tissue pathology and disease processes in rheumatic diseases, with an indication of the imprecision of these tests provided in Table 2. In the following section we will explain the use of these tests, suggest how they may be interpreted, and examine their clinical relevance in rheumatology. A guide to the use of these tests in the overall management of patients in rheumatology is provided in Table 3.



Clinical problem	Laboratory test
Organ/tissue pathology mimicking rheumatic disease or arising from it or therapy	
Renal damage:	
Glomerular	Plasma urea or creatinine, electrolytes
Tubular	Urinary β2-microglobulin, α1-microglobulin, *N*-acetyl glucosaminidase, electrolytes
Liver damage	Plasma AST/ALT, alkaline phosphatase, GGT, bilirubin, total protein, albumin
Bone metabolism	Plasma calcium, phosphate, alkaline phosphatase, albumin, osteocalcin, collagen crosslinks
Muscle disease	Plasma creatine kinase
Malignancy:	
Myeloma	Serum and urine electrophoresis
Metastatic	Plasma alkaline phosphatase, prostate-specific antigen
Pathological processes occurring in rheumatic disease	
Inflammation	Acute phase proteins
Immune response	IgG, IgA, IgM
Immune complex deposition	Immune complexes, C3, C4 and breakdown products
Crystal deposition	Uric acid

Abbreviations: ALT, alanine transaminase; AST, aspartate transaminase; GGT, γ-glutamyl transpeptidase; Ig, immunoglobulin

Table 1 Biochemical tests for assessing organ/tissue pathology and pathological processes in rheumatic disease

Analyte	Range	Units	Imprecision
Sodium	123–157	mmol/l	1.4
Potassium	3.7–6.9	mmol/l	0.08
Urea	3.3–17.1	mmol/l	0.32
Creatinine	149–627	µmol/l	9.8
Alkaline phosphatase	169–513	IU/l	12.6
Bilirubin	52–64	µmol/l	2.9
AST	51–189	IU/l	3.8
ALT	32–119	IU/l	3.7
GGT	40–148	IU/l	2.4
Calcium	1.80–3.55	mmol/l	0.06
Phosphate	1.09–3.09	mmol/l	0.06
Albumin	26–45	g/l	1.0
Creatine kinase	54–280	IU/l	20.3
Uric acid	0.212–0.409	mmol/l	0.012

The data are extracted, with permission, from the Wellcome Clinical Chemistry Quality Assessment Programme, End of Period Report, December 1996 (Wellcome Diagnostics, Dartford, England). Range refers to the analytical range over which the imprecision of each analyte was assessed. The imprecision shown is the mean within-laboratory imprecision of all British laboratories in the Programme for each analyte. These data are reproduced for illustrative purposes only and it is emphasized that individual laboratories should be consulted as to their reference ranges and precision data.
Abbreviations: ALT, alanine transaminase; AST, aspartate transaminase; GGT, γ-glutamyl transpeptidase.

Table 2 Imprecision of biochemical tests

Disease	At diagnosis		At follow-up
	A	B	
Soft-tissue rheumatism	N/A	Acute phase response, full profile, creatine kinase	N/A
Gout	Uric acid, acute phase response	N/A	[illegible]
Pyrophosphate deposition	N/A	Bone profile, uric acid	N/A
Polymyalgia rheumatica/temporal arteritis	Acute phase response	Creatine kinase, serum/urine electrophoresis	[illegible]
Osteoarthritis	N/A	Acute phase response, full profile	N/A
Rheumatoid arthritis	Acute phase response, rheumatoid factor	Full profile	[illegible]
Psoriatic arthritis	Acute phase response	Rheumatoid factor, full profile	Acute phase response annually
Ankylosing spondylitis	Acute phase response	Rheumatoid factor, full profile	Acute phase response annually
Systemic lupus erythematosus	Acute phase response, antinuclear antibodies	Rheumatoid factor, full profile	[illegible]
Polymyositis/dermatomyositis	Creatine kinase, acute phase response	N/A	Repeat tests and renal profile annually
Systemic sclerosis	Acute phase response	N/A	Repeat tests and renal profile annually

[illegible]

Table 3 Routine biochemical management of rheumatic patients

Organ or tissue pathology mimicking rheumatic disease or arising from it or from therapy

At presentation and during the course of the disease, the presence of various organ or tissue abnormalities which may either be the cause of the symptoms or occur as a result of complications of the disease or therapy, needs to be excluded.

Renal function

The tests available for the assessment of renal function are described below (reviewed by Turney and Cooper 1988; Waller *et al*. 1989; Flynn 1990; Duarte and Preuss 1993; Lindeman 1993; Moore, Jr. and Carome 1993) (see also Chapter 1.3.11).

Glomerular function

The two principal indicators of glomerular function are the permeability of the glomerular basement membrane and the glomerular filtration rate (**GFR**). The GFR can be assessed by measuring the clearance of substances that must fulfil the requirements of being freely filtered at the glomerulus, neither secreted nor reabsorbed by the renal tubule, be physiologically and metabolically inert, and that have a constant concentration in the plasma throughout the period of urine collection. Although the fructose polymer inulin is an ideal substance, the requirement of administration by intravenous infusion precludes its use routinely. An alternative is ^{51}Cr-labelled EDTA, the clearance of which can be assessed from the plasma radioactivity at one time-point after a single intravenous dose rather than collecting timed urine samples. However, disadvantages include cost and unsuitability of use during pregnancy. Although clearance of inulin or radioisotopes are the most accurate methods of measuring GFR, they are used less frequently than the more inaccurate methods described below, largely because of the relative complexity and cost (Duarte and Preuss 1993).

The GFR is most often assessed by measuring the creatinine clearance. This is not ideal as creatinine is secreted actively by the renal tubules, the assay of plasma creatinine is subject to interference from non-creatinine chromogens, the reference range is age dependent, and plasma concentrations may increase after exercise or after the consumption of cooked meat during the period of urine collection (24 h). Although creatinine clearance provides a reasonably accurate assessment of GFR when renal function is normal, as glomerular function becomes impaired the GFR is increasingly overestimated. This partly results from increased tubular secretion of creatinine, together with the decreasing inaccuracy of the plasma creatinine measurement as the relative contribution from the non-creatinine chromogen falls. The overall imprecision of creatinine clearance measurements can be high (up to 27 per cent coefficient of variation (CV), Flynn 1990) since it is calculated from three measurements, namely plasma, urinary creatinine, and 24 h urine volume, each having their own imprecision. The main sources of error include incomplete and inaccurately timed urine collection, and bacterial degradation of urinary creatinine. It is recommended that clearance measurements should only be undertaken when absolutely necessary (Turney and Cooper 1988).

In view of the inaccuracy and imprecision of measurements of creatinine clearance, the concentration of plasma creatinine is viewed by many to be a more useful indicator of glomerular function. Measurements of plasma creatinine are less imprecise than those of urinary creatinine, and the adult reference range for plasma creatinine is not affected by age, although it is sex dependent and affected by pregnancy and muscle mass. The plasma creatinine may be elevated falsely by drugs that block tubular secretion (e.g. cimetidine and salicylates), ketone bodies and bilirubin (which may interfere in some assays), exercise, and the consumption of meat. It is a relatively insensitive indicator of renal impairment, with no appreciable change in plasma creatinine until the GFR is reduced by about half, after which point the creatinine rises exponentially with the decline in GFR. The reciprocal of the creatinine concentration at several time-points may provide a better indicator of declining GFR (Turney and Cooper 1988). As the reciprocal of the plasma creatinine decreases linearly with progressive renal failure, the rate of progression of renal impairment can be determined from the slope of the line. Additional adverse occurrences or effects of therapy are thus apparent from a change in the slope of the line (Fig. 1).

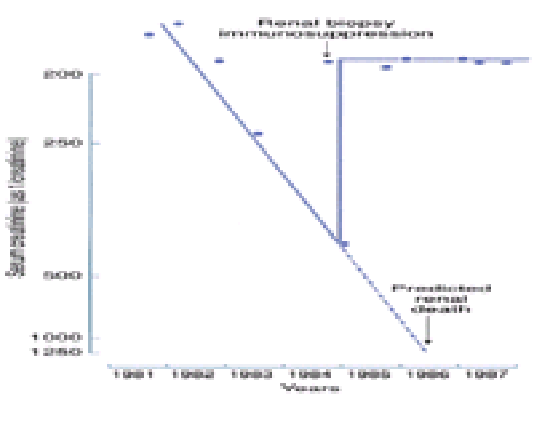


Fig. 1 The use of reciprocal plasma creatinine to monitor renal function. The reciprocal creatinine concentrations are shown for a female Caucasian (date of birth 18/8/41) who developed arthralgia, photosensitive skin rash and mildly positive serology for systemic lupus. Persistent proteinuria and a progressive rise in plasma creatinine were found. On nephrological referral 3 years later, a renal biopsy showed diffuse, proliferative glomerulonephritis with deposition of immunoglobulins and complement, and the patient was treated with prednisolone and cyclophosphamide. The graph illustrates the use of the reciprocal creatinine to monitor the progressive decline in renal function. With therapy, renal function improved and reached a new steady state, although it was permanently impaired. (Adapted from

Turney and Cooper 1988.)

The decline in GFR that occurs with increasing age is not reflected in the serum creatinine, which remains within the 'normal' range. This is due to the coincidental reduction in lean body mass and resultant decrease in endogenous creatinine production. When using the serum creatinine (S_{cr}) to predict the creatinine clearance (C_{cr}), there are several nomograms and formulae, such as the one below, which allow correction for effects of body weight and age (Cockroft and Gault 1976):

$$C_{cr} = \frac{(140 - \text{age}) \times \text{bodyweight (kg)}}{0.81 \times S_{cr}(\mu\text{mol/l})}$$

For an adult woman, the value obtained should be reduced by 15 per cent to allow for lower muscle mass. It must be emphasized that the analytical and preanalytical variability of creatinine measurements are inherent in such formulae.

Plasma urea is a less sensitive measurement than plasma creatinine and is influenced by the rate of protein catabolism, diet, fluid balance, infection, and concomitant medication such as corticosteroids (Turney and Cooper 1988; Moore, Jr. and Carome 1993). As the excretion of urea is very dependent on urine flow rate, dehydration leads to an elevated plasma urea, owing to increased reabsorption. In patients known to have chronic renal failure, intercurrent infection may drastically increase the rate of protein catabolism, with a deterioration of the biochemical abnormality and possible rapid clinical impairment. In this situation, plasma urea is a more rapid and sensitive indicator of the clinical state than is creatinine. If the patient with chronic renal failure is undergoing strict dietary control to reduce the rate of urea formation, measurement of creatinine is appropriate as the levels of urea may fall considerably as the result of the dietary measures alone. The urea:creatinine ratio can be used to indicate reductions in renal blood flow that occur as a result of either real or apparent hypovolaemia. A reduction of the normal ratio (approximately 1:20 SI units) indicates renal outflow obstruction, hypotension, increased protein intake, or dehydration.

Proteinuria can reflect either glomerular or tubular damage, and can be characterized as primarily glomerular, primarily tubular, or of mixed origin, based on the protein constituents of the urine (Waller *et al*. 1989). Haematuria together with proteinuria usually localizes the cause of the proteinuria to a glomerular pathology. Glomerular proteinuria is often associated with larger amounts of urinary protein (>2.5 g/24 h) and the urine contains proteins of moderate and high molecular weight to a much greater extent than in primary tubular proteinuria (4.4 g/24 h compared with 0.4 g/24 h). Proteinuria may also result from strenuous exercise and the possibility of Bence-Jones proteinuria should not be overlooked.

Tubular function

The proximal tubule is the most common site of renal tubular damage and substances normally reabsorbed at this site appear in the urine in increased concentrations (e.g. glucose, amino acids, proteins, phosphate, and potassium). The most sensitive test of proximal tubular reabsorption is the measurement of low-molecular-weight proteins such as b_2-microglobulin (11.8 kDa), a_1-microglobulin (27 kDa), and retinol-binding protein (21 kDa), which are filtered normally by the glomeruli and almost totally reabsorbed in the proximal tubule. Increased plasma concentrations of these proteins, such as may occur with b_2-microglobulin in lymphoproliferative disorders or when the GFR falls excessively in chronic renal failure, may also result in low-molecular-weight proteinuria as the tubular reabsorptive capacity may be exceeded. In acid urine b_2-microglobulin is unstable, thus favouring measurement of retinol-binding protein or a_1-microglobulin. However, the lack of rapid reliable assays currently limits their use.

Renal enzymes such as the lysosomal enzyme *N*-acetyl-d-glucosaminidase (NAG) and the brush-border enzyme alanine aminopeptidase, are released during damage to tubular cells and are sensitive markers, usually preceding changes in urinary protein. Such enzymes are of most use for monitoring the acute rather than the chronic state, where protein assays may be superior.

Clinical relevance in rheumatic disease

Soft tissue rheumatism and osteoarthritis are usually associated with normal renal function tests but it is important to exclude serious coincidental renal pathology. Similarly, renal function is virtually never impaired in the temporal arteritis-polymyalgia rheumatica syndrome (Sonnenblick *et al*. 1989). There is a high prevalence of fatal renal disorders in patients with rheumatoid arthritis (10 to 34 per cent; reviewed by Boers (1990)). Apart from drug toxicity, the kidney is not often clinically involved in systemic juvenile chronic arthritis or adult rheumatoid arthritis, with abnormalities mainly occurring as a result of amyloid deposition or, less frequently, vasculitis. Most patients with amyloidosis present with a marked proteinuria, and the disease runs a variable but slow course towards endstage renal failure (Turney and Cooper 1988) that can be monitored by the reciprocal of serum creatinine. It has been suggested that in view of the many renal-associated deaths in rheumatoid arthritis which are not amyloid-related, there may also be a 'rheumatoid arthritis nephropathy' (Boers 1990), which remains subclinical during the course of the disease because of insensitive screening methods.

Between 10 and 20 per cent of patients with ankylosing spondylitis have some renal abnormality (reviewed by Youssef and Russell 1990). These may be coincidental, related to therapy with non-steroidal anti-inflammatory drugs, the result of amyloid deposition that is associated with long-term ankylosing spondylitis, or alternatively due to IgA nephropathy.

Systemic lupus erythematosus, polyarteritis nodosa, and Sjögren's disease, which may complicate rheumatoid arthritis, all carry a much higher risk of renal involvement. Renal disorders, mainly mild interstitial nephritis with distal tubular acidosis, can occur in up to 30 per cent of patients with primary Sjögren's syndrome (reviewed by Boers 1990), although clinically overt renal disease is generally found in less than 5 per cent of patients. Vasculitis, which may be primary as in Wegener's granulomatosis and polyarteritis nodosa, or secondary to disorders such as rheumatoid arthritis and many of the autoimmune rheumatic disorders, may cause a glomerulonephritis with temporary or permanent glomerular dysfunction, a decrease in GFR, and ultimately renal failure.

Potentially serious renal disease may complicate systemic lupus (reviewed by Balow and Austin 1988) and regular renal assessment is essential. The onset may be insidious or abrupt, usually following a chronic course with remissions and exacerbations. The pathogenesis is complex and the renal disease may be glomerular, tubular, or vascular, with concomitant biochemical changes as described above, predominantly those of glomerulonephritis. Serial creatinine clearances have been reported to be of little use in patients with lupus nephropathy, as they do not accurately measure the direction or the magnitude of the change in GFR (Petri *et al*. 1988). Both plasma creatinine and qualitative urinary protein can predict mortality in systemic lupus (reviewed by Gladman 1990). However, there are patients with systemic lupus who have normal plasma creatinine concentrations and urinalysis but abnormal findings on renal biopsy.

Tests of renal function are essential in patients with gout, both at diagnosis and at regular intervals thereafter, in order to detect the important complication of renal damage, which may dictate the need for specific drug therapy with allopurinol. Approximately half of patients with chronic gout have proteinuria with normal plasma creatinine concentrations (Turney and Cooper 1988). The incidence of renal stones (predominantly uric acid) is in excess of 100-fold greater than in the general population. Hypertension and atherosclerosis may also result from renal damage, as well as being associated with gout.

Several drugs used in the treatment of the rheumatic diseases, particularly the disease-modifying antirheumatic drugs, may cause nephrotoxicity. If renal function is already compromised, the risks of nephrotoxicity may be greatly increased. In addition, with drugs which are excreted primarily by the kidney such as methotrexate, a decrease in renal function may lead to accumulation of the drug and increased toxicity. It is thus important that renal function is assessed prior to commencing therapy and monitored appropriately during therapy.

Liver function

The standard 'liver function tests (LFTs)', i.e. serum or plasma bilirubin, plasma aspartate aminotransferase (AST, previously known as glutamic oxaloacetic transaminase or GOT), alanine aminotransferase (ALT, previously known as glutamic pyruvic transaminase or GPT), alkaline phosphatase, g-glutamyl transferase (GGT), total protein, and albumin, are used routinely to identify the presence of and monitor the course of liver disease, although they are generally poor at identifying the precise pathology (reviewed by Johnson 1989, Rosalki and Dooley 1994). Several newer tests attempt to assess actual liver function by measuring the elimination

of substances such as caffeine or bile acids by the liver, but these are not routinely available and will not be considered here.

Serum bilirubin

Bilirubin is excreted in the bile after conjugation with glucuronide in the liver. Although less than 500 mg of bilirubin are produced each day from breakdown of haem, the liver is capable of conjugating up to three times this amount. Hyperbilirubinaemia is therefore an insensitive indicator of parenchymal liver disease. Clinical jaundice may not be apparent until the bilirubin concentration is more than twice the upper limit of normal. The major causes of hyperbilirubinaemia include haemolysis (usually accompanied by normal liver function tests), hepatitis, cirrhosis, gallstones, tumour infiltration, and drugs such as rifampicin that impair the handling of bilirubin.

Serum enzymes

The aminotransferases AST and ALT are released from damaged cells and are used as markers of hepatocellular integrity, although they are not liver specific, being released from other tissues such as skeletal muscle, heart muscle, and kidney. Very high activities (>10× the upper limit of normal) occur in acute hepatocellular disease, with even higher values (>20× the upper reference limit) strongly suggestive of acute hepatitis from drugs or viral infection. Very high values may also occur in shock, cardiac failure, or sepsis.

Total circulating alkaline phosphatase consists of several different forms (isoenzymes) with the main two in the healthy adult being derived from liver and bone. The liver isoenzyme is produced by hepatocytes adjacent to the biliary canaliculi and elevated plasma activities of alkaline phosphatase are found in extra- or intrahepatic bile-duct obstruction (cholestasis) as a result of increased synthesis. Normal or only mildly elevated levels are found in hepatitis. A high plasma alkaline phosphatase and low plasma aminotransferase indicate cholestatic jaundice, although cholestasis may be accompanied by hepatocellular damage and resultant increases in AST and ALT. Pregnancy also results in an elevated plasma alkaline phosphatase. If the elevation in alkaline phosphatase is unaccompanied by abnormalities of other liver function tests, a non-hepatic source should be considered. For example the most common cause of an isolated rise in plasma alkaline phosphatase is due to an increase in the circulating bone isoenzyme in Paget's disease. An hepatic origin of the alkaline phosphatase is indicated by the finding of a simultaneous elevation of a liver-derived enzyme such as g-glutamyl transferase, which does not occur in bone disease. This enzyme is sensitive but not tissue-specific and may be elevated either through hepatocellular damage or induction by drugs or alcohol. The finding of a normal g-glutamyl transferase result does not necessarily imply a bone origin for the alkaline phosphatase however, as an isolated increase in alkaline phosphatase is often the only abnormality in cases of liver metastases. A more specific approach is the actual determination of the alkaline phosphatase isoenzymes (reviewed by Moss 1987). Placental and intestinal alkaline phosphatases are 'true' isoenzymes, being encoded by individual structural genes. The kidney, liver, and bone isoenzymes are all encoded by the same gene but differ through tissue-specific glycosylation. Isoenzymes can be differentiated by their heat stability or more precisely through electophoretic separation, which is not always routinely available. An immunoassay measuring the bone isoenzyme has recently been developed, but it is still the subject of evaluative research.

Cirrhosis results in a modest increase of plasma aspartate aminotransferase and g-glutamyl transferase, sometimes accompanied by alkaline phosphatase. Drug-induced hepatitis may be primarily hepatocellular (often without jaundice) or the result of intrahepatic cholestasis with jaundice. The aspartate aminotransferase, g-glutamyl transferase, and alkaline phosphatase are usually raised, the last especially with cholestasis. Alkaline phosphatase is raised in many inflammatory conditions, primarily as part of the acute phase response. Sometimes inflammation elsewhere results in an inflammatory infiltration of the biliary tracts with modest jaundice and greater increases of alkaline phosphatase. The mechanism of this is not understood.

Clinical relevance in rheumatic disease

Systemic juvenile chronic arthritis may present with abnormal liver function, and abnormal transaminases are often found after salicylate therapy. In patients with rheumatoid arthritis, plasma transaminases are usually normal, although occasionally the aspartate aminotransferase may increase, suggesting low-level hepatitis. Between 7.5 per cent and 50 per cent of patients with rheumatoid arthritis have an abnormal alkaline phosphatase level, an abnormality which has been proposed to be primarily of hepatic origin, based on the finding (although variable) of simultaneous elevations of g-glutamyl transferase or 5'-nucleotidase, but it may reflect the acute phase response rather than hepatobiliary dysfunction. However, other studies attribute this increase to the bone isoenzyme and it is a question which still remains unresolved, both in patients with rheumatoid arthritis and in those with ankylosing spondylitis. Patients with rheumatoid arthritis have also been found to have elevated g-glutamyl transferase (less than 5 per cent to 77 per cent of patients depending on the study). It has been suggested that a latent or subclinical hepatobiliary pathology may exist in patients with rheumatoid arthritis, with an association between this involvement and disease activity (Aida 1993). The possibility of subclinical hepatobiliary damage is supported by the non-specific histopathological changes of liver biopsies from patients with rheumatoid arthritis (reviewed by Mills and Sturrock 1982; Aida 1993).

Serious hepatic pathology is not usually found in osteoarthritis, although a wider variation in results of liver function tests may be expected than would be encountered in a younger population.

Elevated plasma alkaline phosphatase may be found in patients with the temporal arteritis–polymyalgia rheumatica syndrome (Sonnenblick *et al*. 1989), with abnormal transaminases being infrequent and elevated serum bilirubin extremely rare. Resolution of the biochemical abnormalities occurs after successful corticosteroid therapy. Of patients undergoing liver biopsy, 73 per cent overall had positive findings, although the pathogenesis of the liver involvement in this illness is not clear.

In patients with polymyositis and dermatomyositis, lactate dehydrogenase (LDH) and aspartate aminotransferase may be elevated. An elevated serum alkaline phosphatase is a common finding in systemic sclerosis (approximately 25 per cent) and this condition is often seen in patients with chronic liver disease (reviewed by Weinblatt *et al*. 1982). Although most studies find little evidence of clinically significant hepatic disease in systemic lupus erythematosus (Weinblatt *et al*. 1982), liver disease as detected by altered liver function tests and confirmed in most cases histologically has been reported in more than 20 per cent of patients studied (Runyon *et al*. 1980). This was unrelated to aspirin-induced hepatitis, which is recognized as a potential complication in patients with systemic lupus.

Several of the drugs used in the treatment of the rheumatic diseases may be hepatotoxic, particularly sulphasalazine, azathioprine, cyclophosphamide, and methotrexate. It is essential that hepatic function is assessed before commencing therapy and at regular intervals thereafter.

Calcium and bone

An elevated plasma alkaline phosphatase, particularly in the absence of other abnormal liver function tests, may reflect an increase in the bone isoenzyme and hence indicate increased bone turnover. Appropriate measures can be used to identify the origin of the isoenzyme (as discussed previously).

Calcium is present in blood combined with plasma proteins (40 per cent, 1.0 mmol/l), complexed to anions such as citrate and phosphate (10 per cent, 0.25 mmol/l), or as the free ion (50 per cent, 1.2 mmol/l). The maintenance of the blood calcium is achieved by parathyroid hormone, which promotes calcium absorption from the gut and renal tubule and mobilization from bone, calcitonin, which has a less pronounced action tending to lower blood calcium, and vitamin D, which increases the blood calcium directly by acting on the gut and renal tubule and indirectly via a synergistic action with parathyroid hormone on bone (described in Chapter 2.4). As the secretion of parathyroid hormone is under direct feedback control by ionized calcium, it is this form of calcium that should ideally be measured. However, this is often not routinely available and total plasma calcium is usually assayed by one of a large number of available methods. Several preanalytical factors can affect the result (reviewed by Gosling 1986, Bourke and Delaney 1993), two of the most important of which are venous stasis and the posture of the patient at sample collection. Total plasma calcium should always be corrected for the plasma albumin concentration, as this corrected result shows a better correlation with ionized calcium. The simultaneous measurement of phosphate provides an indication of the pathology of any calcium abnormality, with calcium and phosphate concentrations usually changing in the same direction unless parathyroid hormone is inappropriately in excess or deficient, or unless renal failure is present.

Clinical relevance in rheumatic disease

Bone pains may be confused with pains of articular or soft-tissue origin. Bone pain in adults may be caused by osteomalacia, Paget's disease, bony metastases, myeloma, or very occasionally primary hyperparathyroidism. All these conditions, with the exception of myeloma, are associated with increased osteoblastic activity and a consequent raised plasma alkaline phosphatase due to an increase in the bone isoenzyme. A raised alkaline phosphatase of bone origin is most likely to be associated with Paget's disease. Myeloma is investigated easily by serum and urine protein electrophoresis, with the almost invariable finding of a paraprotein in the serum or Bence-Jones protein in the urine. Advanced osteomalacia may be associated with frank hypocalcaemia, while hypercalcaemia is invariably present in

primary hyperparathyroidism, and often in myeloma and bony metastases.

Serial biochemical assessments are not useful in soft-tissue rheumatism but plasma alkaline phosphatase and serum calcium should be measured to exclude osteomalacia. Similarly, clinical biochemistry plays no part in the management of calcium pyrophosphate deposition disease (CPPD), osteoarthritis, or osteoporosis with a normal calcium, phosphate and alkaline phosphatase. Occasionally, osteoporosis may be secondary to metabolic disorders such as Cushing's syndrome. As previously discussed, there is controversy as to whether elevations in the hepatic or the bone isoenzyme account for the elevated total alkaline phosphatase often seen in diseases such as rheumatoid arthritis, and to the relationship of this elevation with disease activity.

No routine biochemical markers for bone or cartilage turnover are yet available, although the use of substances such as osteocalcin, bone-derived alkaline phosphatase, hydroxyproline or the collagen-crosslinks deoxypyridinoline and pyridinoline as potential markers of matrix turnover is currently the subject of active research (reviewed by Peel *et al*. 1992, also see Chapter 2.4).

Muscle disease

Muscle pains may be due to myositis, with damage to muscle fibres resulting in the release of various enzymes. The most clinically useful is creatine kinase (CK). This is composed of two monomers designated M and B and exists in human tissues as three isoenzymes. The isoenzymes CK-1 (CK-BB), CK-2 (CK-MB), and CK-3 (CK-MM) differ in electrophoretic mobility. Skeletal muscle contains predominantly CK-MM (98 per cent) whereas CK-BB is confined mainly to the brain and thyroid. Myocardium contains both CK-MM and CK-MB, with a much greater proportion of CK-MB than skeletal muscle (30 and 2 per cent, respectively). Myocardial infarction, myopathies including those induced by alcohol and drugs, severe exercise, intramuscular injection, surgery, or hypothyroidism can all cause an elevated total creatine kinase. Abnormal levels can occur in asymptomatic female carriers of Duchenne muscular dystrophy. Most serum creatine kinase is normally CK-MM, rising with damage to skeletal muscle. Increases in CK-MB as a percentage of total creatine kinase are indicative of myocardial infarction, whereas an increase of CK-BB indicates brain damage.

Plasma electrolytes should also be measured to exclude electrolyte imbalance, particularly hypokalaemia, as a cause of muscle pain.

Clinical relevance in rheumatic disease

If soft-tissue rheumatism is suspected, measurement of creatine kinase may be required to exclude myositis, although this condition is usually obvious clinically. The myalgia accompanying polymyalgia rheumatica/giant-cell arteritis is not associated with abnormal serum activities of muscle enzymes. Other causes of myalgia, such as anxiety, depression, fibrositis, thyroid disease, and Parkinson's disease, must be excluded. Creatine kinase is the most reliable enzyme test in polymyositis and dermatomyositis, with plasma activity levels correlating with disease activity (Vignos and Goldwin 1972). Rarely, creatine kinase may be normal but an elevation usually distinguishes polymyositis from polymyalgia.

Several antirheumatic drugs including corticosteroids, d-penicillamine, aspirin, gold salts, and the non-steroidal anti-inflammatories, are known to cause myopathy (reviewed by Zuckner 1990).

Malignancy

Apart from metastatic malignancy of bone, malignant disease may cause myositis, myopathy, or generalized symptoms suggestive of soft-tissue rheumatism. Serum and urinary electrophoresis may be required to exclude myeloma, and plasma alkaline phosphatase and serum calcium analysis to exclude other neoplastic involvement. Prostate-specific antigen (PSA) in conjunction with digital rectal examination can be used to check for carcinoma of the prostate. Pain caused by such conditions affecting bone is interpreted frequently by the patient as arising in adjacent soft tissues. Further clinical tests will be required to investigate this.

Pathological processes occurring in rheumatic disease

The division between clinical chemistry and immunology laboratories is becoming blurred, particularly with the automation of many assays. The measurements used to detect and monitor the following pathological processes may be carried out in laboratories of either discipline. For completeness, they are mentioned briefly here, being covered in greater detail elsewhere.

Inflammation

The main biochemical indicator of the inflammatory episode is the acute phase response (see Chapter 4.1). Inflammatory cells release cytokines, which influence both local and systemic effects, that can be measured and may be of use as described elsewhere. In addition, enzymes and oxygen radicals are released at the inflammatory site, which may damage surrounding extracellular matrix elements. Enzymes, free radicals, and the products of free radical-mediated damage may be measured, although currently this is largely the province of the research laboratory (Holley and Cheeseman 1993).

The measurement of lysosomal and granular enzymes released from polymorphonuclear leucocytes has been used as an indicator of phagocytic activity, but has proved to be insensitive due to the short half-life of the free enzyme. The measurement of enzyme–inhibitor complexes such as elastase–a_1 proteinase inhibitor (E–a_1PI) has now been studied by several groups and may prove to be of future use, although further research is needed. A significant proportion of patients with active rheumatoid arthritis have raised concentrations of E–a_1PI complexes.

Immune response

Immune responses feature in many rheumatic diseases, being most notable early in the disease or during exacerbation. During a phase of immune stimulus, IgM is raised modestly, subsiding in 2 to 3 weeks to be replaced by IgG. Raised concentrations of IgG are found commonly, especially in the autoimmune rheumatic diseases. Infiltration of the synovium with plasma cells results in IgA production. Thus the immunoglobulin profile is complex and is related insufficiently to disease process or activity to be useful for differential diagnosis or monitoring.

Immune complex diseases

Although the presence of immune complexes is usually not in itself a very helpful marker of clinical disease or its treatment and is difficult to assay, significant deposition of complexes in the vascular system results in activation of mediator systems, notably complement, which may be measured. Complement activation implies the occurrence of immune complex mediated damage, indicating the need for therapeutic intervention. Total C3 and C4 concentrations may be measured as they are consumed by the process of activation, although their consumption may be masked by their increased synthesis as part of the acute phase response. Breakdown fragments of C3 or C4, for example C3a, are a much more sensitive indicator of complement activation. Changes in C4 are frequently more sensitive than those in C3, particularly in the case of systemic lupus erythematosus, when used in conjunction with anti-DNA antibody measurements.

Crystal deposition

Biochemical tests are mainly useful in the case of gout. The relationship between hyperuricaemia and gout has been reviewed by Becker (1988). Uric acid, the end product of the metabolism of purines, is present normally as the urate ion at physiological pH and has a low solubility. Saturation and consequently precipitation of urate crystals can occur after only minor elevations of circulating concentrations. The serum urate concentration is governed by the rate of excretion (approximately two-thirds via the kidneys and the remainder via the gut) and the rate of formation, which is dependent on the balance between dietary intake, *de novo* synthesis, and degradation of nucleotides. Serum levels tend to be higher in males, rise with age, and reference ranges vary between ethnic groups.

Gout may be secondary to conditions that cause increased formation or decreased excretion of uric acid, such as hyperparathyroidism or myeloproliferative disorders. Dietary factors and alcohol often exacerbate hyperuricaemia. Such conditions need to be excluded before appropriate treatment can be commenced. Gout is more often primary or idiopathic, being found in patients who have an elevated rate of urate synthesis (10 to 15 per cent) or inappropriately low excretion (85 to 90 per cent). The exact nature of the defects is not known. Measurement of the urinary uric acid excretion aids in the differentiation of the two mechanisms. Although the risk

of gout increases with a higher serum urate, hyperuricaemia does not necessarily imply the presence or the development of gout. Figure 2 shows clearly that while measurements of uric acid are useful, the relationship between the concentration of uric acid and clinical gout is not very clear. The definitive diagnosis of gout can only be made on the finding of crystals of monosodium urate in joint fluid. Gout may cause, or occur as a result of, renal failure, and assessment of renal function is essential in the interpretation of raised concentrations of uric acid. It is also prudent to screen at diagnosis for conditions sometimes associated, such as hypertension and hyperlipidaemia. Measurements of serum uric acid can be used to determine the efficacy of long-term prophylactic therapy with allopurinol or uricosuric agents. Hyperuricaemia has also been reported to occur in 10 to 20 per cent of patients with psoriatic arthritis (Lambert and Wright 1977), probably reflecting the increased breakdown of nucleoprotein with the rapid turnover of skin cells.

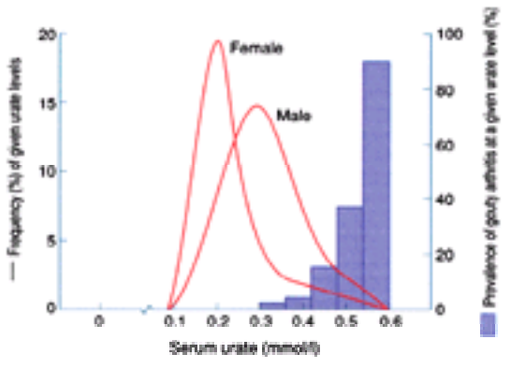


Fig. 2 Frequency distribution of serum uric acid values and the prevalence of gouty arthritis (adapted from Whicher *et al*. 1989). Serum urate concentrations are not normally distributed and the upper limit of the reference ranges (2 SD above the mean) are 0.42 mmol/l and 0.36 mmol/l for males and females respectively. Approximately 20 per cent of subjects with urate concentrations within the reference range suffer from gout and a significant percentage of individuals with hyperuricaemia are not affected by gout.

Although relatively uncommon as causes of the deposition of calcium pyrophosphate, metabolic disorders such as hyperparathyroidism, hypomagnesaemia, or hypophosphatasia should be excluded in pseudogout, using appropriate tests.

Chapter References

Aida, A. (1993). Alkaline phosphatase isoenzyme activities in rheumatoid arthritis: hepatobiliary enzyme dissociation and relation to disease activity. *Annals of the Rheumatic Diseases*, **52**, 511–16.

Balow, J.E. and Austin, H.A., III (1988). Renal disease in systemic lupus erythematosus. *Rheumatic Disease Clinics of North America*, **14**, 117–33.

Becker, M.A. (1988). Clinical aspects of monosodium urate monohydrate crystal deposition disease (gout). *Rheumatic Disease Clinics of North America*, **14**, 377–93.

Boers, M. (1990). Renal disorders in rheumatoid arthritis. *Seminars in Arthritis and Rheumatism*, **20**, 57–68.

Bourke, E. and Delaney, V. (1993). Assessment of hypocalcemia and hypercalcemia. *Clinics in Laboratory Medicine*, **13**, 157–81.

Cockroft, D.W. and Gault, M.H. (1976). Prediction of creatinine clearance from serum creatinine. *Nephron*, **16**, 31–41.

Duarte, C.G. and Preuss, H.G. (1993). Assessment of renal function—glomerular and tubular. *Clinics in Laboratory Medicine*, **13**, 33–52.

Flynn, F.V. (1990). Assessment of renal function: selected developments. *Clinical Biochemistry*, **23**, 49–54.

Gladman, D.D. (1990). Prognosis of systemic lupus erythematosus and factors that affect it. *Current Opinion in Rheumatology*, **2**, 694–702.

Gosling, P. (1986). Analytical reviews in clinical biochemistry: calcium measurement. *Annals of Clinical Biochemistry*, **23**, 146–56.

Holley, A.E. and Cheeseman, K.H. (1993). Measuring free radical reactions in vivo. *British Medical Bulletin*, **49**, 494–505.

Johnson, P.J. (1989). Role of the standard 'liver function tests' in current clinical practice. *Annals of Clinical Biochemistry*, **26**, 463–71.

Lambert, J.R. and Wright, V. (1977). Serum uric acid levels in psoriatic arthritis. *Annals of the Rheumatic Diseases*, **36**, 264–7.

Lindeman, R.D. (1993). Assessment of renal function in the old. *Clinics in Laboratory Medicine*, **13**, 269–77.

Mills, P.R. and Sturrock, R.D. (1982). Clinical associations between arthritis and liver disease. *Annals of the Rheumatic Diseases*, **41**, 295–307.

Moore, J., Jr. and Carome, M.A. (1993). *Clinics in Laboratory Medicine*, **13**, 21–31.

Moss, D.W. (1987). Diagnostic aspects of alkaline phosphatase and its isoenzymes. *Clinical Biochemistry*, **20**, 225–30.

Peel, N., Eastell, R., and Russell, G. (1992). Markers of bone and collagen breakdown in early inflammatory arthritis. *Baillière's Clinical Rheumatology*, **6**, 351–72.

Petri, M. *et al*. (1988). Serial assessment of glomerular filtration rate in lupus nephropathy. *Kidney International*, **34**, 832–9.

Rosalki, S.B. and Dooley, J.S. (1994). Liver function profiles and their interpretation. *British Journal of Hospital Medicine*, **51**, 181–6.

Runyon, B.A., LaBrecque, D.R., and Anuras, S. (1980). The spectrum of liver disease in systemic lupus erythematosus. Report of 33 histologically-proved cases and review of the literature. *American Journal of Medicine*, **69**, 187–94.

Sonnenblick, M., Nesher, G., and Rosin, A. (1989). Nonclassical organ involvement in temporal arteritis. *Seminars in Arthritis and Rheumatism*, **19**, 183–90.

Turney, J.H. and Cooper, E.H. (1988). Renal biochemistry in rheumatic disease. *Baillière's Clinical Rheumatology*, **2**, 233–58.

Vignos, P.J. and Goldwin, J. (1972). Evaluation of laboratory tests in diagnosis and management of polymyositis. *American Journal of Medicine*, **263**, 291–308.

Waller, K.V., Ward, K.M., Mahan, J.D., and Wismatt, D.K. (1989). Current concepts in proteinuria. *Clinical Chemistry*, **35**, 755–65.

Weinblatt, M.E., Tesser, J.R.P., and Gilliam, J.H., III (1982). The liver in rheumatic diseases. *Seminars in Arthritis and Rheumatism*, **11**, 399–405.

Whicher, J., Walters, G., and Henley G. (1989). Using the biochemistry laboratory. *Medical Forum*, **31**, 6–9.

Youssef, W. and Russell, A.S. (1990). Cardiac, ocular, and renal manifestations of seronegative spondyloarthropathies. *Current Opinion in Rheumatology*, **2**, 582–5.

Zuckner, J. (1990). Drug-induced myopathies. *Seminars in Arthritis and Rheumatism*, **19**, 259–68.

4.4 Microbiology and diagnostic serology

Geoffrey Scott

Optimal use of the microbiology laboratory

Many other chapters in this book are concerned with specific infections that directly or indirectly involve joints. This short chapter does not set out to iterate descriptions of these diseases but to give a brief idea of how microbiologists deal with specimens sent to their laboratories and the limitation of the tests that are applied. The aim of the microbiologist is to help the clinician make a more tangible diagnosis if infection or reactive arthritis are suspected. The clinician can help the microbiologist by ensuring that appropriate specimens are taken, that they get to the laboratory quickly, and by communication with the laboratory scientist or doctor so that all relevant tests can be done. The request form that accompanies a specimen is the principal method of communication, but some 50 per cent of request forms to this laboratory are completed inadequately. In most hospitals, microbiologists will see patients on request, examine notes, advise on prescribing of antimicrobials, and may even impose restrictions on their use. In an acute or difficult clinical setting, direct communication is worthwhile because it helps to ensure that someone in the laboratory will take a special interest in the processing of a specimen. In most laboratories, some simple tests, such as examination for acid-fast bacilli, will be made only if specially requested by the clinician or if there is some message cryptically encoded in the information given on the request form. Scientific purists would say that all tests should be done on very specimen and that they should be examined for 'all' possibilities. In practice, this would be impossibly expensive and time wasting. The interpretative microbiologist, perhaps more so than the haematologist or chemist, requires some sense of what the clinician is searching for. The more relevant (concise) history that can be given, the better.

In clinical pathology, it is important to consider the likelihood that a result is actually a true representation. Computer-printed results sometimes suggest a degree of scientific precision that, particularly in microbiology, may be unfounded. A patient must not be made to fit an unlikely diagnosis suggested by test results, especially when they are preliminary (Stokes *et al*. 1993). The sensitivity of a test is the percentage of the infected group detected as positive by the test. The specificity indicates the percentage of the negative group detected as negative by the test. These values are in themselves meaningless unless put into the context of the total number of tests done, and the expected rates of positives and negatives in a population. The figures that can be calculated to express these data are the predictive values of positive (**PVP**) and negative (**PVN**) test results, respectively (Easmon 1990). The problem in trying to calculate these values is that one needs to know how many in the population actually do or do not have the infection. PVP and PVN can be calculated for a new test when compared with a 'gold standard' test assumed to be 100 per cent sensitive and specific. An example would be the comparison of a new enzyme immunoassay for chlamydial antigen against culture of the organism (the 'gold standard'), or against amplification of chlamydial DNA. The last will inevitably detect more positives than culture, but it will be impossible to know whether they are truly positive for a variety of reasons, including difficulty in identifying specific DNA and the risk of contamination. When comparing two tests, it becomes obvious that one test can detect positives which the other cannot. Yet the more sensitive test will almost certainly have a higher rate of false positivity.

If a test gives a continuum of signals between positive and negative, a cut-off point has to be chosen to categorize its results. This is illustrated by the microbiology of urine specimens. The cut-off between positive and negative is chosen by examining large numbers of samples and selecting a point (10^5 c.f.u./ml) between the two peaks of the bimodal distribution of the concentration of bacteria present. Some of those with 'significant bacteriuria' will, by definition, have only 10^4 to 10^5 organisms per ml of urine. Rather than accepting this single result as meaningful, the microbiologist will hope to confirm or refute the suggestion of positivity by requesting a repeat test. In recently developed enzyme immunoassays, the cut-off between negative and positive is selected in exactly the same way. If the cut-off is selected at 20 units and the test has a result of 19 units, this should be reported as negative yet may well be positive. Only alternative, confirmatory tests can be of real help in these cases and the clinician must again avoid the temptation to make patients fit a dubious result. False-positive immunoassays are common in some conditions (for example, first-generation assays for antibody to human immunodeficiency virus), but repeated tests of different design and specificity will yield more reliable results.

Unless the sensitivity and specificity of a test are both 100 per cent, PVP and PVN will be markedly influenced by the prevalence of the disease in the population under study. Comparing likely isolation rates of chlamydia in patients attending a genitourinary medicine clinic (17 per cent) with those attending a rheumatology clinic (assumed 1 per cent), and using a test with sensitivity of 90 per cent (i.e., 1 in 10 positive patients give a negative result) and specificity of 95 per cent (1 in 20 tests are false positive), the PVPs in the two clinics would be 78 and 15 per cent, respectively. In the low-prevalence population this is of profound importance in trying to make a microbiological diagnosis in a patient with reactive arthritis.

Tests and their limitations

The influence of antibiotics on culture

Microbiological culture studies are relatively insensitive and will be almost certainly negative when the patient is already on antimicrobial agents. Prophylactic antibiotics are often prescribed as a matter of routine for orthopaedic operations, but when infection is suspected and the operation is partly for diagnostic purposes it is crucial that the administration of antimicrobials be delayed until after all the specimens have been taken. We advise that anaesthetists give the antimicrobials intraoperatively in infected cases.

Specimens

The rheumatologist will tend to send fluid or biopsies from joints or blood in order to diagnose bacterial infection, and special culture specimens from elsewhere (e.g. faeces, urethral fluid), or serum, to support the diagnosis of reactive arthritis. All specimens are treated in the laboratory as though a risk of infection to the staff. Contamination of the outside of a container is obviously a particular hazard, so all containers should be sent in a sealed, clear-plastic bag separate from the request form. Staples should not be used. Specimens that leak may be discarded in the laboratory because opening the bag poses a high risk to the staff. Unlabelled specimens and those accompanied by an inadequate request form should also be discarded. Laboratory staff should check with clinicians before a specimen is discarded to ensure that it can be replaced.

An outline of microbiological investigations is shown in Table 1.

Table 1 Outline of microbiology investigations for rheumatology patients

Synovial fluid (pus)

Specimens must be of the highest quality. It is crucial to take specimens using aseptic precautions. Interpretation of the relevance of a culture containing avirulent skin commensals such as coagulase-negative staphylococci, *Acinetobacter* spp., or a-haemolytic streptococci is almost impossible. Furthermore, the inoculation of any of these organisms into a damaged joint may well result in secondary infection.

In the laboratory, pus will be examined by Gram stain. Modified Giemsa is necessary to identify cell morphology. Methylene blue and acridine orange may reveal organisms that do not stain with Gram because they are damaged. Acridine orange will stain DNA and RNA, and will confirm the presence of micro-organisms, but will not reveal their classical morphology. When indicated, acid-fast bacilli are stained with auramine O and Ziehl-Nielsen.

The specimen is cultured under a range of conditions on non-selective media and, when a clue is given by the Gram stain, on selective media. It is quite common to see organisms in pus but for these not to grow, and this may reflect inhibition by host defences still present in the pus rather than insensitive culture conditions. Restoration of 'stressed' organisms is done very poorly in clinical laboratories. There is a move to apply special techniques used in food microbiology to clinical specimens. Essentially, the culture is enriched under optimal conditions for bacterial growth without any potentially suppressive selective pressure.

All fluids should be incubated alone and inoculated into enrichment broth or cooked-meat medium, then subcultured at 24 and 48 h. Dilution in broth may allow growth, even in the presence of antimicrobial agents.

It is better to send a good volume of fluid or pus rather than a swab of the pus. It is not worth looking for mycobacteria from a swab. If the specimen is aspirated with a syringe, send the capped syringe to the laboratory rather than decanting the pus into another container. Aeration of a specimen may reduce the yield of anaerobes. Anaerobes in joints are usually found only after penetrating trauma. The pus can be subjected to gas-liquid chromatography to demonstrate volatile fatty acids, which are products of anaerobic metabolism. This is a rather more technical way of identifying anaerobes than simply smelling the offensive odour of a specimen. Specific gas-liquid chromatographic patterns may, however, indicate a particular anaerobic genus, long before it has been cultured and identified.

Most fluids are cultured aerobically on enriched media (such as chocolatized blood agar), in an atmosphere of 5 to 10 per cent CO_2, and anaerobically because certain facultative anaerobes (e.g. *Streptococcus pyogenes*) show more characteristic colonial morphology under these conditions. Many true anaerobes require several days of incubation before they give visible growth.

If swabs of wounds are taken, they should be placed in transport medium supplied (e.g. Stuart or Amies). The best recovery that can be expected from a swab is less than 30 per cent of the organisms sampled. Swabs and pus from joints should in general not be refrigerated because of the risk of killing cold-sensitive organisms such as *Neisseria gonorrhoeae*.

Blood cultures

Blood cultures should be taken from any patient suspected of having pyogenic arthritis. Two sets taken by separate venepuncture at different times will increase the positive yield but it is not worth taking more. Disinfection of the skin twice with 70 per cent isopropyl alcohol allowed to dry is quite sufficient provided that the venepuncture is done deftly. Contamination of the culture with skin commensals nearly always follows poor technique or difficult venepuncture. Blood-culture bottles should be filled from a syringe before other sample containers because often these others are not sterile.

Handling the blood is dangerous. Needlestick injuries while changing needles and inoculating the bottles are very common, so extreme care must be taken. Some recommend not changing needles for this reason and because it has been shown not to reduce significantly the level of contamination. The bottles should not be overfilled with blood (easily done for bottles with a vacuum), because blood over a certain concentration is inhibitory to bacterial growth. Different style bottles require different sizes of inoculum. Most modern bottles have a cap over a sterile rubber membrane and the membrane should not be swabbed before inoculation.

Semiautomatic, vacuum ('Vacutainer', Becton Dickinson) blood-culture bottles are available but are not popular in the laboratory. Automatic systems have revolutionized the handling of blood-culture specimens but the special bottles required are, in general, not suitable for 'Vacutainer' phlebotomy.

Three main types of automatic culture detection systems are in use: detection of $^{14}CO_2$ released from labelled carbohydrate; detection of CO_2 by infrared spectrophotometry or some patented indicator system; and detection of increased pressure of gas within a bottle by a simple cap indicator. These systems do away with the routine subcultures, which are labour-intensive and lead to contamination of the cultures. They also accelerate the process of detection. It is not unusual for a blood culture taken during the night to cue positive in the morning with coliforms or *Strep. pneumoniae*, without visible signs of growth.

It is important to realize the limitation of the blood-culture system routinely used in the local laboratory. Do cultures regularly detect fastidious organisms such as neisseriae and brucellae? When in doubt, cultures should be specially subcultured on to rich medium after a week's incubation, but this will only be done by special request.

Antimicrobial sensitivity testing

In general, antibiotic tests will be done only when an organism has been isolated in pure culture. However, direct sensitivities can be done by plating the specimen and using antibiotic discs, but this depends on having seen an organism with characteristic morphology. The inoculum is inaccurate so the preliminary results are suspect until confirmed using the isolated organism in pure culture.

Antimicrobial testing is done in most laboratories by one of two methods. The principle of the first is that antimicrobials diffuse from a small filter-paper disc into agar in such a way that exponential concentration is proportional to the distance from the disc. A sensitive organism plated on to the agar is inhibited up to but not beyond the minimal inhibitory dilution of the antibiotic. The zone radius or diameter is then measured and recorded. In the United Kingdom, it is practice to compare the zone of inhibition of test organisms with a control of known sensitivity (Stokes method; Stokes and Ridgway 1987), but in the United States the size of the zone given by the unknown is compared with a database of known sensitive organisms (Kirby-Bauer method). A reduction in zone size indicates reduced sensitivity, which may be declared 'moderate' or 'resistant'. The advantages of the test are that the organism can be seen to have grown with characteristic morphology, that it is not contaminated, and that 'sensitivity' is a clear signal. The main limitations of the disc diffusion test are that it is standardized for fast-growing organisms and causes problems with difficult organisms (e.g. *Haemophilus* spp., *Neisseria* spp.). Some other specific problems are as follows:

1. Requirement for a special medium or culture conditions:
   a. enriched for *Neisseria*, *Haemophilus* etc.;
   b. thymidine-free—testing folate antagonists (sulphonamide etc.);

c. special supplements—pyridoxine for streptococci;
d. staphylococcal sensitivity to methicillin is temperature-dependent (best demonstrated at 30°C).

2. Poor diffusion of antibiotics (a small zone of inhibition is difficult to interpret) as with:
   a. glycopeptides—vancomycin and teicoplanin;
   b. polymyxin.
3. Inaccurate results:
   a. inappropriate sensitivity of inactivating enzyme producers, for example:
      i. *Enterobacter* spp. may appear sensitive to ampicillin yet always produce b-lactamase;
      ii. penicillin-resistant *Staph. aureus* often shows a good zone of inhibition to penicillin;
      iii. penicillin-resistant pneumococci may show a zone to penicillin but not to oxacillin.
   b. Organism shows heterogeneous resistance:
      i. *Staph. aureus* to methicillin.

The other method in common use involves the incorporation of antimicrobials into agar or broth at concentrations above and below the so-called breakpoint between sensitive and resistant. Unknown and control sensitive and resistant organisms are spot inoculated on to a series of plates incorporating different antimicrobials. Although this method gives more precise information about the minimal inhibitory dilution of an antibiotic against a certain organism, it also has flaws. If an organism does not grow, one cannot be certain that the organism was not dead when inoculated. The effect of penicillin against *N. gonorrhoeae* is a good example where there is a continuum of increasing resistance conferred by sequential changes in penicillin-binding proteins, permeability, and finally b-lactamase. It is impossible to know where to set the breakpoint in this situation.

Sensitivity to antibiotics of slow growers such as mycobacteria is best done by a modified agar incorporation method that has been standardized. Mycobacterial sensitivities should be ready 1 month after primary isolation but rapid sensitivity tests employing radiocarbon release as $^{14}CO_2$ from a suitable broth can be done in specialized laboratories. For the future, we may expect molecular methods for species identification and to detect important antibiotic resistance genes.

These examples illustrate the point that sensitivity testing of bacteria *in vitro* has problems whatever method is adopted. In addition, when large numbers of tests are done, there are likely to be transcriptional errors. When a patient does not appear to respond to recommended therapy, question first the appropriateness of the results of sensitivity tests, have them double-checked in the laboratory, and then consider the pharmacology of the antimicrobial selected.

The minimal inhibitory and the minimal bactericidal dilutions of an antimicrobial are tested in broth culture using highly accurate concentrations of antimicrobial. This is done for isolates causing infective endocarditis but may exceptionally be done also for a pyogenic arthritis that is failing to resolve. Antimicrobial activity can also be measured in synovial fluid as a check on pharmacokinetics. Experimentally, killing of organisms and the postantibiotic effect (inhibition of growth even after the antibiotic has been removed) may be measured; such tests are part of the development, and may influence the use, of new antimicrobials.

Antibiotic choice

Antibiotics chosen empirically for joint infections should be appropriate to the isolate or most likely organism, but also need to penetrate well. Some agents, such as chloramphenicol, co-trimoxazole, tetracycline, clindamycin, rifampicin, or metronidazole, are well absorbed from the intestine and will probably be effective when given orally. Most b-lactam antimicrobials are poorly absorbed and penetrate into bone and joint tissue poorly, so the first phase of treatment for acute pyogenic infections should be given parenterally in high doses.

Antimicrobials that penetrate well into joint fluid include tetracycline (useful for brucellae), co-trimoxazole (active against coliforms and some *Staph. aureus*), rifampicin and fucidic acid (very active against *Staph. aureus*, but should not be given alone), and chloramphenicol. Ciprofloxacin in recommended doses may not achieve sufficiently high concentration to inhibit *Pseudomonas aeruginosa*. Although absorbed from the gut and suitable for pseudomonal urinary infection, the antimicrobial should probably be given parenterally to achieve sufficiently high concentrations in the synovium. *Stenotrophomonas maltophilia*, an emerging nosocomial pathogen in immunosuppressed patients, is often resistant to ciprofloxacin and carbapenems so may be selected by the widespread use of these drugs. Nosocomial enterococcal infections are also selected by using cephalosporins.

Antigen detection

Tests for antigen detection have developed in two rather different directions. First, rapid tests, which usually depend on agglutination of latex particles, red cells, or staphylococci (binding Fc) coated with antibody, are sold in kit form to the routine laboratory for rapid detection or confirmation of common organisms found in certain specimens. An example would be *H. influenzae* type b in cerebrospinal fluid. Kits are designed so that a single test can be done economically, together with positive and negative controls. These tests have superseded countercurrent immunoelectrophoresis, which was cumbersome, slow, and insensitive.

The alternative tests for antigen are immunoassays (either radiometric or enzyme-linked), suitable for processing large numbers of specimens. These are often highly sensitive but may not be specific, so confirmatory tests of positives are required. Reference laboratories tend to develop and validate their own immunoassays. The classical application of successful immunodiagnosis of antigen is in the detection of hepatitis B virus surface antigen.

Despite antigenic diversity, it seems that, in practice, only a limited number of the antigenic shapes expressed on micro-organisms are useful for the detection of species. Monoclonal antibodies are preferred to polyclonal ones for their greater specificity, but polyclonals will usually include some antibody of great avidity, which can significantly improve a test. Sandwich immunoassays using high-avidity antibody capture and one highly specific antibody seem to be the most reliable. The antigen recognized by the antibody chosen should be common to and expressed on all of the species and subspecies that are to be searched for. Yet there should be no cross-reactivity with antigens on other species and this is unachievable. In an immunofluorescence test, where the antibody (directly or indirectly labelled) is applied to a clinical specimen and examined microscopically, it is possible for the experienced reader to detect alternative organisms or debris that appear to bind yet have a different shape from the organism under investigation. This indicates a serious problem with enzyme immunoassays in which positives are detected simply by a graded colour change. It must be assumed that certain of the results will be false-positive and that secondary confirmatory tests are required.

Specific DNA or RNA detection

Labelled, specific, single-stranded DNA probes will anneal with complementary DNA that has been cleaved into single strands at high temperature. The labelled complex is detected in the appropriate way (radioisotope detection or enzymatic colour change). This method appears to be specific (although it is often impossible to prove positives or negatives) but is not sensitive because it will detect only relatively large amounts of DNA present. Therefore, the label itself may be amplified by a chemical trick. Labelled DNA probes can also be made to anneal with ribosomal RNA, which will be present at a higher copy number than the DNA.

Alternatively, DNA segments in the test sample may be amplified using flanking sequences that define a specific segment and a heat-stable polymerase to replicate the DNA repeatedly. This is known as the polymerase chain reaction (**PCR**). Theoretically, one DNA molecule can be amplified to a detectable level but rather more than one molecule is needed for the test to work in practice. Nevertheless, the problem with DNA amplification is extreme sensitivity and false-positivity due to contamination: the test should be done away from laboratories where the organism being sought is routinely grown and handled. These tests are expensive, experimental, and not yet universally available.

Recent advances in PCR technology include 'nesting', using two pairs of primers, the second internal to the first, which increases the sensitivity of the test some hundred-fold and also the specificity. Secondly, viral and bacterial RNA can be amplified after it has been converted to complementary DNA by reverse transcription. Thirdly, the amplification product can be annealed to a specific labelled probe to confirm the specificity of the product.

Good examples of the useful application of PCR are to be found in the study of viruses that cannot (easily) be grown (e.g. hepatitis C virus, human immunodeficiency virus). A highly sensitive test such as PCR for hepatitis C virus is too difficult and expensive for screening of blood products even though seroconversion may take 3 to 6 months after infection.

These tests are, in general, experimental. When a new test is introduced, inevitably there is a exploratory period when the results are quite uninterpretable and unreliable.

Antibody detection

A single serum test for antibody almost never yields a definitive diagnosis. Even a confirmed positive antibody test for human immunodeficiency virus merely indicates exposure to, and latent infection with, the virus. It does not infer anything about the relation of the present symptoms and the stage of the disease.

Two specimens separated by at least 2 and preferably 4 weeks should be tested to make a diagnosis of a recent infection, especially when a positive test result is obtained. Only the evolution of titre can indicate how recent an infection was and then only within wide confidence limits. Some infections (e.g. legionellosis) are associated with a delayed rise in antibody, sometimes as long as 3 weeks after the onset of illness. Many infections fail to stimulate any detectable antibody response at all.

Antibody tests are the only way of making a non-invasive diagnosis in many infections that predispose to arthritis. Patients with polyclonal B-cell activation may show strong positivity in several tests and anamnestic responses are common.

The tests for antibody have developed enormous sophistication. Nevertheless, crude agglutination reactions are still made on dilutions of serum mixed with polyclonal antisera, usually as screening tests. This is the basis for screening tests for brucellae and enteric fever. Positive screening for *Brucella* agglutination in the routine laboratory is usually non-specific, cannot be confirmed in the reference laboratory by other immunoassays, and should not be taken automatically as indicating brucellosis.

With the high rate of false-positive tests in the first-generation enzyme immunoassays for human immunodeficiency virus, particularly in patients from Africa, it became conventional to confirm the specificity of antibodies by Western immunoblotting. The antigens of the organism under test are separated by chromatography and transferred to a membrane for reaction with the patient's serum. The patterns of reactivity are then compared with positive controls. Non-specific Western blot reactivity has been clearly shown not to indicate infection with human immunodeficiency virus. Western blotting may also be used to confirm the specificity of antibody in other conditions (e.g. Lyme disease) (Karlsson 1990). Enthusiasm for what was thought to be a very specific test has now waned because of difficulties in standardization and the definition of a positive test.

Specific tests have been developed over the years for certain organisms. They include the dye test for *Toxoplasma gondii*. This is still used as the 'gold standard' in reference laboratories, but routine agglutination tests have been replaced by highly sensitive latex-particle agglutination and specific IgM kits, which can be applied in routine laboratories. Some IgM kits have been found to be oversensitive.

The antistreptolysin-O titre for antibodies to *Strep. pyogenes* remains the mainstay of laboratory diagnosis of recent infection. It uses the neutralization of a standard inoculum of streptolysin by dilutions of the patient's serum. The indicator is haemolysis of sheep or human erythrocytes. However, the test is not specific and suffers false-positives and -negatives. Anti-DNAase B, streptokinase, and hyaluronidase can be measured in the specialist laboratory and may be useful additional tests.

It should be noted that the likelihood of accurate results from serological tests for antigen and antibody depend on how frequently the tests are done in the laboratory. Furthermore, commercial test kits are very expensive. Hence the reliance on reference laboratories for support and verification. Inevitably, such test results are delayed.

Acute joint infections (Fig. 1, Fig. 2, Fig. 3)

Fig. 1 Acute back pain with fever may be due to spinal osteomyelitis. Tuberculosis would be a common cause but in this case an aspirate of pus yielded *Brucella melitensis*. The patient came from Turkey.

Fig. 2 A chronic discharging sinus from the back of the finger of this man with psoriatic arthropathy surprisingly yielded *Mycobacterium tuberculosis*.

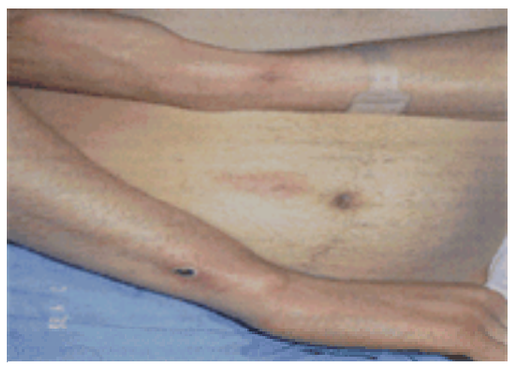


Fig. 3 This patient has acute arthritis of one knee and these skin lesions indicate bacteraemia particularly with *Neisseria gonorrhoeae*.

Acute joint infections engage a relatively small part of rheumatology practice in Europe and America. In certain parts of the world, however, brucellosis and

tuberculosis of joints are common. With increasing invasive health care, nosocomial bacteraemia, especially from intravenous-line infections, may contribute to some bone and joint infections.

Microbiological examination of fluid from an infected joint is only important because of the guidance the results ultimately give towards sensible chemotherapy. In the relatively few acutely inflamed joints that are infected, a very wide range of organisms is involved and, with very few exceptions, isolates have increasingly unpredictable sensitivity patterns to antimicrobials. Moreover, when positive, culture gives a firm diagnosis in a doubtful case.

Bacterial infections of healthy joints occur after bacteraemia with virulent organisms including *Strep. pneumoniae*, *Staph. aureus*, and *N. gonorrhoeae* but are much more likely to occur in joints damaged by other conditions such as osteoarthrosis or rheumatoid arthritis. Patients are often debilitated and on corticosteroids. In these patients, it is critical to make a positive diagnosis as soon as possible to prevent the inadvertent injection of steroids into an infected joint. Clinically, acute spontaneous pyogenic arthritis in a previously healthy individual is often restricted to one joint and is relatively easy to recognize, but may be much less evident when it supervenes on a chronic inflammatory polyarthropathy in a patient on steroids. To make a diagnosis none of the surrogate markers of infection (e.g. white blood-cell count, erythrocyte sedimentation rate, or C-reactive protein) compares favourably with direct examination of synovial fluid from inflamed joints or blood for relevant micro-organisms. It may be sufficient to take blood cultures and treat with antimicrobials on the basis of the most probable organism, in expectation of a positive culture in due course. However, it should always be the practice to consider sampling synovial fluid, which can also be examined for crystals and other important markers (see Chapter 4.6).

Routine cultures should readily yield *Br. melitensis* within a week but less certainly *Br. abortus*. Special cultures are needed for mycobacteria and negative results are not available for 8 or 12 weeks, depending on laboratory routine. Simple tests will distinguish *M. tuberculosis* from atypical species but the laboratory scientist will be reluctant to declare a species to the clinician until reference laboratory identification is completed.

The best way of monitoring the success of long-term therapy for deep *Staph. aureus* infections is to test serial sedimentation rates and C-reactive protein levels. Antibodies to staphylococcal a- and g-lysins should not be used to make a diagnosis of deep staphylococcal infection because false-positives may occur. However, if the diagnosis has been made by conventional culture, then serial assays of sera taken every 1 to 2 months or so can be done in parallel to monitor progress.

Investigation of reactive arthritis

Many infections predispose to reactive arthritis. There is considerable clinical overlap between full-blown Reiter's disease associated with HLA-B27 and so-called seronegative reactive arthritis. These conditions arise because of immune reactivity to organisms infecting particularly the genitourinary and gastrointestinal tracts. It may be possible to elicit a very specific, temporally relevant history of a symptom complex that points to a diagnosis. Common syndromes preceding arthritis are 'febrile rash' illness, diarrhoea, urethritis (sometimes with uveitis), and influenza-like illness with and without jaundice. Unless there is a preceding illness or accompanying signs, searching for the cause is likely to be fruitless.

Urethritis

It is important to search for the cause. When urethritis or prostatitis precedes arthritis, then it should be presumed to have been sexually acquired. Patients should ideally be examined in a department of genitourinary medicine, where there are facilities to take the optimal specimens and direct liaison with the laboratory. Furthermore, staff there will arrange contact tracing and may obtain circumstantial evidence for infection of the proband, by detecting chlamydiae or *N. gonorrhoeae* in an asymptomatic carrier partner.

The quality of the specimens is of great importance. For chlamydia, separate special swabs must be used for immunodiagnosis (enzyme immunoassay), immunofluorescence, DNA detection, and culture. Special transport media containing 2-sucrose phosphate and preservation at -70°C are required if the swabs are not to be inoculated directly into tissue culture. Special tissue culture is labour-intensive and does not detect 1/10 positives (by all tests) but has the advantage of being highly specific. The antigen-detection tests by immunofluorescence or by enzyme immunoassays tend to yield false-positives for a variety of technical reasons. Optimal diagnostic power in patients with low risk of chlamydial infection is obtained by doing several tests (Ridgway and Taylor-Robinson 1991). For the patient with posturethritis arthritis or Reiter's disease, diagnosis is important because eradicative antibiotic therapy is an important part of management, even though the immunological disease is usually not affected by such treatment.

Experimental methods, such as direct probes to detect chlamydial DNA or RNA, or DNA amplification, are under development. The direct probes are not particularly sensitive, but DNA amplification is oversensitive, very prone to contamination, and may yield false-positive results. Antigen-detection tests alone and DNA amplification are not yet specific enough to be absolutely sure about the meaning of a positive result.

The chlamydiae, therefore, are typical examples of organisms that commonly cause disease and are detected by rapid, simple means other than conventional staining and culture, that is, enzyme immunoassay and immunofluorescence of antigen, with DNA or RNA hybridization proving a useful experimental tool. In contrast, detection of gonorrhoea should be by culture.

The likelihood of detecting these organisms diminishes progressively with time from the acute infection. Therefore, the tests may be negative in patients with reactive arthritis. Although antibody may be detected, this does not help with the diagnosis of recent infection. A very high titre of antibody to *Chlamydia trachomatis* would be a pointer to the diagnosis.

However, the results of tests for antibody to *N. gonorrhoeae* are so totally unreliable that they should not be requested or performed.

Diarrhoea-associated reactive arthritis

When a patient gives a history of diarrhoea preceding the arthropathy then it is worth examining stools for continuing carriage of salmonellae, shigellae, and yersiniae, even though the acute illness may have occurred some time ago. All culture methods are selective, so lack sensitivity. Therefore, three separate stools should be examined to increase the value of a negative result. Intestinal biopsies may reveal organisms by immunofluorescence but this technique is very likely to be misleading because of cross-reactivity (de Koning *et al.* 1989). There is preliminary evidence that fluoroquinolones such as ciprofloxacin may be able to eradicate the carriage of certain organisms in the stool, but this action has not yet been shown to affect postdysenteric arthropathy

Yersinia enterocolitica causes a syndrome of prolonged diarrhoea with abdominal pain a particular feature (Cover and Aber 1989). Patients may have terminal ileitis and pseudoappendicitis. Five per cent of patients develop erythema nodosum or reactive arthritis or full-blown Reiter's disease. *Yersinia enterocolitica* serogroups 0:3 and 0:9 are particularly associated with postinfective arthritis in Scandinavia, in parallel with a high incidence of HLA-B27. Antibodies to *Yersinia*, particularly IgA, persist and can be measured in reference laboratories. Cross-reactivity between numerous O serogroups and other organisms (brucella, morganella, salmonella) and thyroid tissue antigens is likely to make agglutination tests for antibodies unreliable, although experimental enzyme immunoassays are under evaluation (Gronberg *et al.* 1989).

Non-specific illness

Short-lived, influenza-like illness predisposing to arthritis will probably not be diagnosed. From a survey of 10 microbiology laboratories, Waghorn (1995) discovered that the following serological tests might be performed on serum from a 30-year-old female with 'an influenza-like illness and arthralgia': (in order of frequency) rubella, parvovirus, *Mycoplasma*, influenza, *Chlamydia*, adenovirus, enterovirus, hepatitis B, Q-fever, Epstein–Barr virus, antistreptolysin titre, *Borrelia*, mumps, respiratory syncytial virus, *Yersinia*, *Legionella*, toxoplasmosis, syphilis, and cytomegalovirus.

More protracted illness would be a feature of glandular fever, toxoplasmosis, or hepatitis. It is worth checking particularly for evidence of recent anicteric hepatitis. Markers of recent infection with hepatitis B virus include antihepatitis B core antigen (IgM). Serial sera will also reveal the evolution of an acute infection or chronic carrier state. Hepatitis A virus IgM is specific to recent infection within the last 6 months because the carrier state does not occur.

Positive hepatitis C virus antibody tests occurred particularly in rheumatoid arthritis, implying a causal link until it was discovered that this reactivity was non-specific.

However, there have been a small number of case reports of reactive arthritis associated with this infection.

Rash illness

Common exanthematous infections such as rubella and parvovirus are associated with arthritis either during the latter stages of the acute infection, when the diagnosis is usually clear, or when the infection appears to have resolved. Serological tests are needed to make the diagnosis. Specific IgM tests are available for these organisms.

Lyme disease (see also Chapter 5.3.4)

The rheumatologist may be alerted to the possibility of Lyme disease as a cause of rheumatic symptoms during the secondary stage of the disease by concomitant rash, neurological signs, or myopericarditis. History of exposure, of known tick bite, and of the erythema chronicum migrans of the primary illness will also be useful. Large-joint arthritis is characteristic of the tertiary stage and there may or may not be skin and neurological manifestations.

Although the spirochaetes can sometimes be seen in biopsy material, realistically the diagnosis can only be made serologically and several methods of detecting antibody are available in kit form: indirect immunofluorescence (based on the fluorescent treponemal antibody test but using *Borrelia burgdorferi*), passive haemagglutination of red cells or latex beads coated with *B. burgdorferi*, various enzyme immunoassays, and a specific IgM capture test (Barbour 1988). Rheumatoid factor will give false-positives in the last. Many laboratories will do one of these as a screening test but will not do confirmatory tests except in an area of high prevalence (Schwartz *et al*. 1989). Most will rely on a reference laboratory to make confirmatory tests.

False-positives are common in all tests, probably because of exposure to commensal spirochaetes, and the most specific tests use a preliminary absorption step. Schwartz *et al*. (1989) found some discrepancies between test results from different laboratories. It is important that the screening test should be highly sensitive (and therefore gives some false-positives) so that positives are not missed. False-positives should be excluded by confirmatory tests but Western immunoblotting does not add much to the specificity of a good enzyme immunoassay (Karlsson 1990).

Rheumatic fever (see Chapter 5.3.12)

In the relatively rare condition of acute rheumatic fever, *Strep. pyogenes* may be obtained from the throat swab of the proband and family contacts several weeks after the acute infection. During an outbreak it is worth determining the serotype of any isolate. However, except in very unusual conditions, these are not consistently of one M-type (Kaplan *et al*. 1989). Certain M-types seem to be more commonly associated with rheumatic fever but this is probably an epiphenomenon and the diversity of types usually found suggests an epidemic cofactor, such as a coincident viral infection. No such cofactor has ever been identified. Isolation of *Strep. pyogenes* is a helpful indication but is not the mainstay of the diagnosis (Pope 1989). High titres, preferably with a demonstrated fourfold rise or fall in antistreptolysin-O and anti-DNAase B titres in serial samples, are the strongest supportive evidence for rheumatic fever. However, only 80 per cent of patients who fulfil the clinical criteria for acute rheumatic fever have a significantly raised antistreptolysin-O titre, and antibodies against other determinants may be more sensitive.

Conclusion

The rheumatologist can gain much information by using the diagnostic microbiology laboratory to the best advantage but can also expend much money and fruitless energy uncritically requesting irrelevant serology. Withholding antibiotics until all the relevant culture specimens have been taken and consultation over difficult cases may be helpful.

Chapter References

Barbour, A. G. (1988). Laboratory aspects of Lyme borreliosis. *Clinical Microbiology Reviews*, **1**, 399–413.

Cover, T. L. and Aber, R. A. (1989). *Yersinia enterocolitica. New England Journal of Medicine*, **321**, 16–24.

de Koning, J., Heeseman, J., Hoogkamp-Korstanje, J. A. A., Festen, J. J. M., Houtman, P. M., and Van Oijen, P. L. M. (1989). *Yersinia* from intestinal biopsy specimens from patients with seronegative spondyloarthropathy: correlation with specific IgA antibodies. *Journal of Infectious Diseases*, **159**, 109–12.

Easmon, C. S. F. (1990). Laboratory diagnosis of bacterial diseases. In *Principles of bacteriology, virology, and immunity*, Vol. 3, (ed. G. R. Smith and C. S. F. Easmon), pp. 1–9. Edward Arnold, London.

Gronberg, A., Fryden, A., and Kihlstrom, E. (1989). Humoral response to individual *Yersinia enterocolitica* antigens in patients with and without reactive arthritis. *Clinical and Experimental Immunology*, **76**, 361–5.

Kaplan, E. L., Johnson, D. R., and Cleary, P. P. (1989). Group A streptococcal serotypes isolated from patients and sibling contacts during the resurgence of rheumatic fever in the United States in the mid-1980s. *Journal of Infectious Diseases*, **159**, 101–3.

Karlsson, M. (1990). Western immunoblot and enzyme-linked immunosorbent assay for serodiagnosis of Lvme borreliosis. *Journal of Clinical Microbiology*, **28**, 2148–50.

Pope, R. M. (1989). Rheumatic fever in the 1980s. *Bulletin on the Rheumatic Diseases*, **38**, 1–8.

Ridgway, G. L. and Taylor-Robinson, D. (1991). Current problems in microbiology. 1. Chlamydial infections: which laboratory test? *Journal of Clinical Pathology*, **44**, 1–5.

Schwartz, B. S., Goldstein, M. D., Ribeiro, J. M. C., Schulze, T. L., and Shahied, S. I. (1989). Antibody testing in Lyme disease: a comparison of results in four laboratories. *Journal of the American Medical Association*, **262**, 3431–4.

Stokes, E. J. and Ridgway, G. L. (1987). Antibiotic sensitivity tests. In *Clinical microbiology*, pp. 204–33. Edward Arnold, London.

Stokes, E. J., Ridgway, G. L., and Wren M. (1993). The practice of clinical microbiology. In *Clinical microbiology*, pp. 1–11. Edward Arnold, London.

Waghorn, D. J. (1995). Serological testing in a microbiology laboratory of specimens from patients with suspected infectious disease. *Journal of Clinical Pathology*, **48**, 358–63.

4.5 Autoantibody profile

P. J. Maddison

Autoantibodies are a prominent feature of autoimmune rheumatic diseases and are the subject of intensive study to understand the underlying pathogenesis of these disorders. At the same time methods to detect certain of these antibodies have provided the clinician with valuable tools to help diagnose and assess patients with connective tissue disease. It is becoming increasingly clear that autoantibodies in the context of autoimmune rheumatic diseases are directed against highly selected targets and are not just the result of non-specific polyclonal B-cell activation. Of the 2000 or more mammalian intracellular proteins only relatively few are targeted by autoantibodies (Christian and Elkon 1986). In many cases, when they occur they belong primarily to the IgG1 and IgG3 subclasses of IgG, are high affinity, and occur in large amounts, showing all the features of an antigen-driven, T-cell-dependent immune response. Antigen specific T cells have now been detected for some of these autoantigens (Fenning *et al*. 1995).

Table 1 lists antibodies which have a role in diagnosis, in predicting patterns of disease expression and prognosis and in assessing disease activity. Although at first glance there is a wide range of antibody specificities, in the individual patient the autoantibody profile is much more restricted and this can be used to our clinical advantage. It is now appreciated that certain antibody profiles are associated not only with a clinical diagnosis but also with particular clinical manifestations and, as a consequence, the prognosis. Therefore, once antinuclear antibodies (**ANA**) have been detected with a screening test, it is important to determine their specificity. To be of most use to the clinician, it is important that the serology laboratory should have the facility to detect a wide range of relevant antibody specificities.

Table 1 Serological tests in the assessment of patients with connective tissue diseases

Autoantibody tests should be planned, and the results interpreted, in the light of the clinical picture. Thus, in the situation where the clinician wishes to exclude the possibility of systemic lupus erythematosus, examination of the serum for ANA by indirect immunofluorescence is sufficient as a screening test and it is not cost-effective automatically to test for anti-DNA or other antibody specificities (Clough *et al*. 1989). If the ANA test is positive, then it is important to look for antibodies reacting with DNA or nucleic acid-binding proteins. In other clinical situations particular combinations of serological tests are indicated. For example, where systemic sclerosis is suspected clinically it is worthwhile from the beginning looking for antibodies to nucleolar antigens, kinetochore-related proteins, or topoisomerase 1 (Scl 70), since these antibodies have disease specificity and one or the others will be detected in over 85 per cent of patients with systemic sclerosis.

Antinuclear antibodies

The discovery of the lupus erythematosus cell phenomenon by Hargraves and colleagues (1948) led to the first serological test to be used in the diagnosis of systemic lupus erythematosus. The next landmark was the development of indirect immunofluorescence (Holborow *et al*. 1957), which has taken over from the lupus erythematosus cell test because it is more sensitive, simpler to perform, reproducible, and easier to interpret. Indirect immunofluorescence detects a wide range of autoantibodies and, therefore, is generally used as a primary screening test for ANA.

At one time thin sections of mammalian tissues, especially mouse or rat liver or kidney, were used as the substrate to incubate with the patient's serum. Now there is a trend towards using human cell lines grown as a monolayer on the glass slide. The advantages are that the cells lie on a flat plane allowing good recognition of both nuclear and cytoplasmic structures and that actively dividing cells present antigens only expressed during certain stages of the cell cycle, which are either absent or occur only in small quantities in the resting nuclei of tissue sections. At present, HEp2 epithelial cells, derived from a human larynx carcinoma (American Type Culture Collection CCL-23), are most widely used because both nuclei and nucleoli are large and enable a wide range of staining patterns to be recognized. They can be cultured in the laboratory or slides with monolayers of cells can be purchased.

The antinuclear antibodies fixed to the nuclear antigens are visualized by adding a fluorescein-conjugated antibody to human gammaglobulin and examining the slide under a fluorescence microscope (Fig. 1). Alternative enzyme-linked conjugates are now commercially available, after addition of which the slides can be examined under an ordinary microscope. The incidence of ANA depends upon the assay system, the titres chosen to separate positive and negative sera, the ability of the technician, and the type of microscope. Uniformity of the test between different laboratories has been aided by the availability of positive standards (Feltkamp 1990).

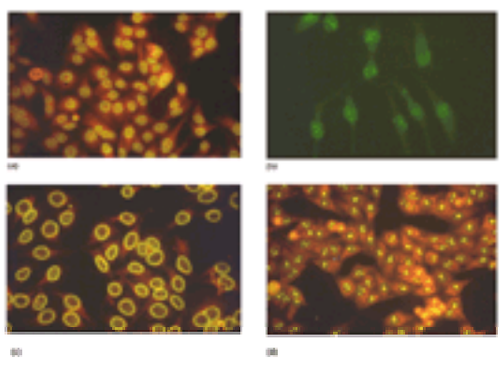


Fig. 1 Patterns of immunofluorescence using HEp2 cells: (a) speckled (pattern typical of anti-U1RNP); (b) centromere: characterized by discrete speckles in interphase nuclei which segregate on to the metaphase plate during cell division; (c) peripheral; (d) nucleolar.

The use of HEp2 cells enhances the sensitivity of this test in systemic lupus erythematosus so that antinuclear antibodies can be detected in 95 per cent of active, untreated patients. The main difference from using rodent substrates is the increased detection of patients with an immune response predominantly to Ro. However, a very small number of such patients are still ANA-negative even using HEp2 cells. As shown in Table 2, the specificity of the test for systemic lupus erythematosus is low (approximately 0.6) and antinuclear antibodies are frequently detected in other autoimmune rheumatic diseases and in a proportion of the normal population. Higher titres tend to occur in disease states than in normal subjects. Some reports have shown a fall in the frequency of antinuclear antibodies in the normal population after the age of 70 and have suggested that this is due to the earlier death of ANA-positive individuals (Hooper *et al*. 1972).

Systemic lupus erythematosus	Other autoimmune rheumatic diseases		Other diseases		Normal population
95%	Systemic sclerosis	80%	Chronic active hepatitis	100%	8%
	Sjögren's syndrome	80%	Drug-induced lupus	100%	
	Rheumatoid arthritis	50%	Myasthenia gravis	50%	
	Polymyositis	40%	Waldenström's macroglobulinaemia	25%	
	Polyarteritis nodosa	10%	Diabetes	25%	

Table 2 Antinuclear antibodies in various diseases using HEp2 cells as substrate

Different patterns of fluorescence can be observed, often reflecting the predominant antibody in the serum. It is difficult to attribute the pattern to a particular specificity and additional tests are required, but there are exceptions when the pattern of fluorescence can be clinically useful. Thus, a peripheral nuclear pattern is almost exclusively seen in systemic lupus erythematosus and often corresponds to the presence of anti-nDNA, although it only occurs in a small proportion of such patients. Similarly, antikinetochore antibodies (anticentromere, **ACA**) and high titres of antinucleolar antibodies are most commonly seen in the context of systemic sclerosis. For a comprehensive review of fluorescence patterns the reader is referred to Humbel (1993).

The titre of antinuclear antibodies frequently fluctuates but does not correspond well to clinical assessment of disease activity. The titre *per se* does not have prognostic significance.

Antibodies to DNA

Autoantibodies binding DNA have a central place in the immunology of lupus. Two broad groups of antibody specificity correspond to the major conformations of DNA: native, double-stranded DNA (**nDNA**) and denatured, single-stranded DNA (**ssDNA**). Antibodies to nDNA do not react with the free purine or pyrimidine bases but do react with some aspect of the secondary structure of the denatured molecule, generally with higher affinity. Recent evidence suggests that DNA circulating in the blood is always in the form of nucleosomes, i.e. closely associated with histones (Rumore *et al*. 1992).

In the past various techniques have been used to detect anti-DNA including immunodiffusion, haemagglutination, and complement fixation. With the availability of chemically-pure antigen it has been possible to develop sensitive and quantitative methods and now most laboratories use radioimmunoassays or enzyme-linked immunosorbent assays (**ELISA**). Either commercial kits are used or in-house assays. The Farr assay is a fluid phase radioimmunoassay in which antibodies combined to ^{125}I-labelled DNA are precipitated by 50 per cent saturated ammonium sulphate (Smeenk *et al*. 1990a). Modifications of this include the use of filters or an antihuman immunoglobulin serum. The Farr assay detects high avidity antibody. To detect antibodies to DNA by means of an ELISA, DNA has to be coated on to an ELISA plate. Being relatively negatively charged, nDNA in particular binds poorly to plastic and so the ELISA plate needs to be precoated with poly-l-lysine or, even better, protamine sulphate. Other methods have included the binding of biotinylated DNA to the plastic via streptavidin (Emlen *et al*. 1990) or to coat DNA to ultraviolet-irradiated plates (Zouali and Stollar 1986). During the assay, DNA-bound antibody is detected by adding (i) an antihuman immunoglobulin antibody conjugated to an enzyme such as horseradish peroxidase or alkaline phosphatase, and (ii) the appropriate substrate for the enzyme to produce a colour reaction, which is quantified by measuring the resulting optical density.

Different assay systems are not always comparable and are influenced by important factors such as characteristics of the DNA antigen, how the DNA is presented to antibody in the serum, and the reaction conditions. This is illustrated by a recent demonstration of the variability between commercial ELISA kits (Aviña-Zubieta *et al*. 1995).

The ELISA is more sensitive than the Farr assay because it detects low avidity antibody and can be used to distinguish its isotype. However, being polyanionic, DNA is a 'sticky' molecule with a tendency to bind serum proteins non-specifically. To minimize interlaboratory variation the chosen assay should be standardized using a freely available standard serum from the World Health Organization (Feltkamp 1990). When the Farr technique and an ELISA are compared, the former, according to some reports, is less sensitive but more specific for systemic lupus erythematosus (Smeenk *et al*. 1990b).

Antibodies reacting with nDNA have greatest relevance to the diagnosis of systemic lupus erythematosus (Aarden *et al*. 1976), but assays to detect these antibodies rely on a source of impeccably pure, double-stranded DNA. DNA extracted from *Escherichia col*, for example, needs to be treated with S1 nuclease, or purified on columns of diethylaminoethylcellulose. Radiolabelled, circular double-stranded DNA from the bacteriophage PM $_2$ also serves this purpose well or, alternatively, the synthetic polynucleotide, polydAT can be used. Indirect immunofluorescence using the haemoflagellate, *Crithidia luciliae* (Aarden and Smeenk 1981), is also an accepted technique for detecting anti-nDNA, combining strong sensitivity with high disease specificity (Smeenk *et al*. 1982). This micro-organism contains a giant mitochondrion which consists of pure nDNA and it is the fluorescence of this which constitutes a positive test. This technique is more sensitive than the Farr assay, since it detects antibodies with lower affinity; it can also be modified to identify complement-fixing antibodies.

Approximately 70 per cent of untreated patients with active systemic lupus erythematosus have anti-nDNA detected by the *C. luciliae* technique. These antibodies have a high degree of specificity (Table 3). The Farr technique, which selects out high avidity antibodies, is somewhat less sensitive and the ELISA more sensitive but less specific since low avidity antibodies are detected in other conditions. Antibodies reacting with purine and pyrimidine bases and other sites revealed by denaturing DNA are very commonly found in systemic lupus erythematosus but are also seen in a variety of other conditions (Table 3).

Disease	Anti-nDNA (%)	Anti-ssDNA (%)
Systemic lupus erythematosus	70	90
Drug-induced systemic lupus erythematosus	0	60
Chronic active hepatitis	5	60
Rheumatoid arthritis	5	25
Primary biliary cirrhosis	0	15
Normal subjects	0	5

Table 3 Incidence of antibodies to native and single-stranded DNA

There is evidence that antibodies to both nDNA and ssDNA have a role in pathogenesis. The presence of high avidity, complement-fixing antibodies to DNA are associated with an increased risk of lupus nephritis. In some patients, but not in all, a steady increase in anti-DNA levels followed by a sharp drop in titre precedes a clinical exacerbation (Smeenk *et al*. 1990b). Consequently there is value in monitoring serial serum anti-DNA levels. In studies where the relative proportion of high and low avidity anti-DNA antibodies was measured, clinical exacerbation was often heralded by an increase in high avidity antibodies (Smeenk *et al*. 1990b).

Antihistone antibodies

Antibodies to histones are detected in idiopathic and drug-induced lupus, rheumatoid arthritis, and other conditions including scleroderma-related disorders (Bernstein *et al*. 1985; Wallace *et al*. 1994). Histones are a set of basic proteins that organize chromosomal DNA in eukaryotes into nucleosomes which are the repeating units of chromatin. The nucleosome consists of the histone octamer $(H2A–H2B–H3–H4)_2$, approximately 200 base pairs of DNA and H1. Antibodies are directed to a range of epitopes expressed on individual histones, histone–histone complexes, and histone–DNA complexes (Burlingame and Rubin 1990). They can be detected by indirect immunofluorescence, typically producing an homogeneous nuclear pattern which can be abolished by acid treatment of the substrate (Fritzler and Tan 1978). These antibodies primarily react with the (H2A–H2B)–DNA complex which is exposed in chromatin. ELISA have also been reported (Burlingame and Rubin 1990) but can be affected by DNA contamination of the antigen preparation. Immunoblotting using commercial sources of extracted histones can be used to detect antibodies to all five histones, although only the reactivity to denatured individual histones is measured by this technique.

Antihistone antibodies are a striking feature of the immune response in drug-induced lupus, occuring in over 90 per cent cases. Although they lack disease specificity, the presence of autoantibodies primarily to histones in the appropriate clinical setting strongly suggests a drug-induced syndrome (Thompson *et al*. 1993). In syndromes induced by a variety of drugs including procainamide, quinidine, hydralazine, penicillamine, and sulphasalazine the antibodies are predominantly IgG against the (H2A–H2B)–DNA complex (Totoritis *et al*. 1988; Rubin *et al*. 1992; Bray *et al*. 1994).

Nucleic acid-binding proteins

The presence of antibodies reacting with certain highly conserved, nucleic acid-binding proteins present in eukaryotic cells is a very characteristic feature of autoimmune rheumatic diseases. The first observations were made over 30 years ago when antibodies to what are now called Ro and La were detected by immunodiffusion in the sera of patients with Sjögren's syndrome (Jones 1958; Anderson *et al*. 1961). Subsequent clinical interest in these systems results from observations that certain profiles of these antibodies are associated with particular patterns of disease expression within the spectrum of autoimmune rheumatic diseases.

Traditionally, these antibodies have been detected by immunodiffusion using buffered saline extracts of mammalian tissue, such as rabbit or calf thymus extract and human spleen extract. A range of prototype sera are used to detect a precipitin system. Counter-immunoelectrophoresis (Fig. 2(a)) is used frequently to screen sera but the Ouchterlony technique of passive immunodiffusion, which is slightly less sensitive, is sometimes needed to confirm comparisons with prototype sera, particularly if multiple precipitin systems are present (Fig. 2(b)). Increasingly, more sensitive methods of antibody detection are being used, such as immunoblotting (Fig. 3), protein or RNA immunoprecipitation (Fig. 4), and ELISA. With immunoblotting, specific antibodies are recognized by binding to a profile of proteins of particular molecular weight which have been separated by electrophoresis of a cell extract and transferred to nitrocellulose sheets. As a technique it is too labour intensive for the routine laboratory and suffers from the disadvantage of detecting antibodies only to epitopes on denatured proteins. However, it is a good way of demonstrating the profile of antibodies in a particular serum and determining the fine specificity of these antibodies and their isotypes. ELISAs using purified antigens have been developed and provide a sensitive, quantitative way of detecting these antibodies. Initially, immunoaffinity purified antigens were used but recombinant antigens are becoming available. Compared with immunodiffusion, these techniques increase the prevalence of these antibodies (Table 4). However, with the exception of anti-Ro which can be found in low titre in approximately 12 per cent of the normal population (Reichlin and Harley 1988), they are virtually confined to patients with autoimmune rheumatic diseases.

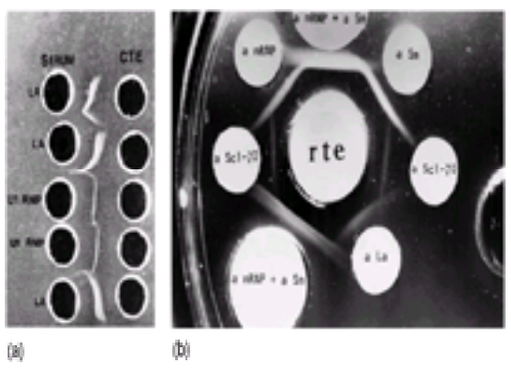


Fig. 2 Detection of nucleic acid-binding proteins by immunodiffusion. (a) Counter-immunoelectrophoresis to demonstrate reactions of identity. (b) The Ouchterlony technique used to confirm the identity of precipitins in a serum containing antibodies to U1RNP and Sm.

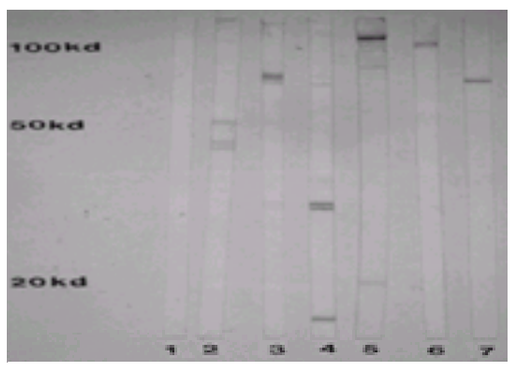


Fig. 3 An example of immunoblotting in which antibody specificities are confirmed by the specific profile of protein binding. Key: (1) normal serum, (2) anti-La(SSB), (3) anti-U1RNP (4) anti-Sm, (5) anticentromere antibody, (6) antitopoisomerase 1, (7) antimitochondrial antibody.

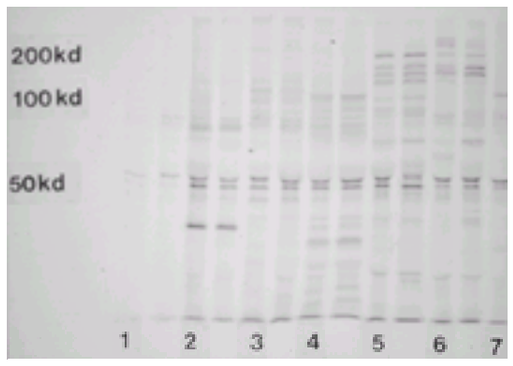


Fig. 4 The technique of protein radioimmunoprecipitation used to identify the specificity of antinucleolar antibodies demonstrated by indirect immunofluorescence. Key: (1) normal serum, (2) anti-U3RNP, (3) anti-ThRNP, (4) anti-PM-Scl, (5) anti-RNA polymerase-I/-III, (6) anti-RNA polymerase-I/-II/-III, (7) antitopoisomerase 1.

Specificity	Systemic lupus erythematosus		Normal sera	
	Immunodiffusion	ELISA	Immunodiffusion	ELISA
Anti-URNP	33	35	0.1	0.1
Anti-Sm	10	25	0.1	0.1
Anti-Ro	40	50	1	12
Anti-La	15	28	0.1	0.1

Numbers are percentages

Table 4 Frequency of antibodies to RNA-binding proteins: ELISA compared with immunodiffusion

Systemic lupus erythematosus

In systemic lupus erythematosus, antibodies react most frequently with four groups of RNA-binding proteins, namely Sm, U1RNP, Ro, and La (see Reichlin and Harley 1988 for a review). The Sm and U1RNP antigens are expressed on protein components of a series of small nuclear ribonucleoprotein (**snRNP**) particles, each with its distinctive uridine-rich RNA molecule (Van Venrooij and Sillekens 1989) (Fig. 5). Sm epitopes, for example, are expressed on proteins designated B_, B, D, E, F, and G, which are amongst core proteins found in all snRNP particles. The B and D proteins, however, are the major autoantibody targets. U1RNP epitopes are expressed on three proteins, designated 70-kDa, A, and C, uniquely found in association with U1RNA of which the 70-kDa and A proteins are the major targets. The U1RNP particle therefore expresses epitopes reacting with both anti-Sm and anti-U1RNP. Most anti-U1RNP sera react with epitopes on the 70-kDa protein of the U1RNP particle and virtually all anti-Sm sera bind simultaneously to the B', B, and D proteins, while the pattern of binding to other protein components varies from serum to serum. Sera containing anti-Ro react with a number of different Ro-bearing proteins complexed in ribonuclear protein particles with a distinctive series of small RNA species, which in humans are designated hY1–5 (Reichlin and Harley 1988; Pruijn *et al*. 1990). In contrast to Sm and U1RNP, Ro proteins have evolved through the species so that in some autoimmune rheumatic disease sera, antibodies react almost exclusively with human Ro. Autoimmune sera recognize two major proteins, a 60-kDa protein known as Ro60, which binds directly to hYRNA, and a 52-kDa protein known as Ro52 which does not bind to hYRNA (Boire *et al*. 1995) (Fig. 6). The relationship *in vivo* between Ro52 and the RoRNP complex is unknown. Another protein, calreticulin, initially reported to be the 60-kDa protein is an unrelated autoantigen (Rokeach *et al*. 1991). Antibodies to La react with a 50-kDa protein which transiently binds to transcripts of RNA polymerase III which include the Ro RNAs. The physical association of Sm and U1RNP on one hand and Ro and La on the other explains, at least in part, the linked antibody responses to these antigens.

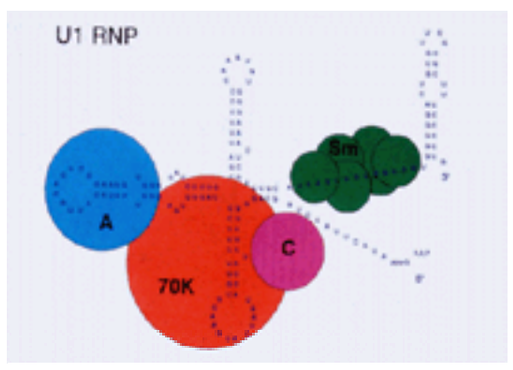


Fig. 5 The U1RNP molecule showing the relationship of U1RNP and Sm antigens to the U1RNA molecule (by courtesy of Dr G. Pruijn and Dr W. van Venrooij, University of Nijmegen, The Netherlands).

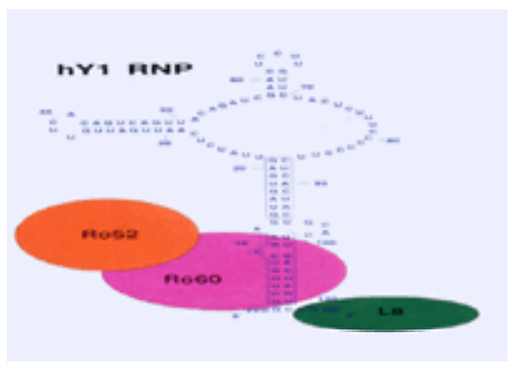


Fig. 6 The Ro/La particle showing the relationship of the Ro and La antigens to hYRNA (by courtesy of Dr G. Pruijn and Dr W. van Venrooij, University of Nijmegen, The Netherlands).

High titres of these antibodies, for example as detected by immunodiffusion, are found frequently and almost exclusively in the context of autoimmune rheumatic diseases (Table 5). Anti-Sm has the greatest specificity for systemic lupus erythematosus but there is a marked ethnic variation in the presence of these antibodies, being more commonly found in Afro-Caribbeans than in northern European Caucasoids (Arnett *et al*. 1988). These antibodies identify distinctive serological subsets within the spectrum of systemic lupus erythematosus. Antibodies to Sm frequently occur in association with anti-U1RNP and antibodies to La are virtually always accompanied by anti-Ro. It is apparent that these serological subsets are associated with certain patterns of disease expression (Table 5) in which they may have a

pathogenetic role.

Disease	Frequency of precipitins	Major specificities		Clinical association
Systemic lupus erythematosus	75	U1RNP	(30)	Raynaud's syndrome, deforming arthritis, mixed connective tissue disease.
		Sm	(10–25)	Vasculitis
		Ro	(40)	Photosensitive dermatitis, subacute cutaneous lupus erythematosus, Sjögren's syndrome, pulmonary fibrosis, neonatal lupus, congenital heart block
		La	(15)	
Sjögren's syndrome	60	Ro	(80)	Systemic disease, vasculitis, neurological disease
Systemic sclerosis	55	La	(30)	
Polymyositis	60	Scl-70	(20)	Diffuse scleroderma, cardiopulmonary disease
Non-connective tissue disease	<0.5	Jo-1	(20)	Interstitial lung disease
Normal subjects	<0.5			

Numbers are percentages.

Table 5 Antibodies to nucleic acid-binding proteins detected by immunodiffusion in connective tissue diseases

These antibodies are usually present from the beginning of the clinical presentation and are detectable throughout the course of the disease. The presence of high titres in an asymptomatic individual usually heralds the future development of an autoimmune rheumatic disease (McCune *et al*. 1987). Using an ELISA, fluctuations in antibody titre can be detected but there is an inconsistent relationship between titre measured in longitudinal studies and disease activity (de Rooij *et al*. 1990).

A variety of other antibody specificities, for example anti-PCNA (cyclin) (Asero *et al*. 1987), SL (Ki) (Sakamoto *et al*. 1989), and ribosomal P protein (Elkon *et al*. 1985), can be detected in systemic lupus erythematosus. They occur in a small proportion of sera and although clinical associations have been reported, such as anti-ribosomal P proteins and neuropsychiatric lupus, these associations require confirmation in larger, prospective studies and the role of these antibodies in 'routine' systemic lupus erythematosus serology is not yet defined.

Antibodies to ribosomal P proteins are directed predominantly to a conserved region of the C-terminus which is shared by the P proteins. They occur in approximately 10 per cent of patients with systemic lupus erythematosus but have been reported in a higher proportion with active disease with an association with diffuse neuropsychiatric manifestations (Elkon *et al*. 1994). However, this association is not apparent in all published studies (reviewed by Teh and Isenberg 1994). Reasons for discrepancies between reports include methodological differences for detecting anti-ribosomal P, a lack of uniform criteria for classifying patients with lupus involving the central nervous system, and demographic variation between the study populations. For example, anti-ribosomal P antibodies have been reported more frequently in certain ethnic groups, such as the Malaysian Chinese (Teh *et al*. 1993).

Sjögren's syndrome (see Chapter 5.10)

With sensitive techniques, antibodies to Ro and La can be detected in virtually all patients (Harley *et al*. 1986). They are a marker for Sjögren's syndrome developing in systemic lupus erythematosus, systemic sclerosis, and primary biliary cirrhosis. Several studies, including that of Pease *et al*. (1993), have shown that antibodies to Ro and La identify patients at greatest risk of developing extraglandular complications such as vasculitis. The fine specificity of anti-Ro antibodies may correspond to the clinical expression of Sjögren's syndrome as opposed to systemic lupus erythematosus (Ben-Chetrit *et al*. 1990; St Clair *et al*. 1994). In a recent study, Barakat *et al*. (1992), using synthetic peptides corresponding to regions of Ro60 in an ELISA, demonstrated that one peptide representing residues 21–41 of the protein was recognized by sera from the majority of patients with Sjögren's syndrome but only a minority of anti-Ro-positive sera from those with systemic lupus erythematosus.

Sjögren's syndrome in rheumatoid arthritis is usually not associated with anti-Ro. However, the small proportion of rheumatoid arthritis patients with anti-Ro have a distinctive clinical picture (Moutsopoulos *et al*. 1985; Tishler *et al*. 1994).

Mixed connective tissue disease

The concept of mixed connective tissue disease was predicated initially upon the identification of autoantibodies to U1RNP (Sharp *et al*. 1972). Whether or not mixed connective tissue disease is a distinctive autoimmune rheumatic disease is controversial. These antibodies produce a characteristic nuclear pattern on indirect immunofluorescence (Fig. 1(a)). Initially, the presence of these antibodies was confirmed by a haemagglutination reaction to an RNase-sensitive component of a buffered saline extract of mammalian tissue (extractable nuclear antigen, ENA). Subsequently, immunodiffusion became the standard assay in most serology laboratories. Using techniques such as immunoblotting or ELISA with recombinant protein components of U1RNP, it has been suggested that the reaction of antibodies to the 70-kDa polypeptide of U1RNP is an important serological marker for mixed connective tissue disease, whereas antibodies to U1RNP in the context of systemic lupus erythematosus are directed to epitopes on other polypeptides, particularly B'B (Habets *et al*. 1983; Reichlin and Van Venrooij 1991).

Systemic sclerosis (see Chapter 5.8)

Antibodies to kinetochore proteins (anticentromere antibodies), topoisomerase 1, and nucleolar constituents are very characteristic of systemic sclerosis and identify distinctive serological groups within the spectrum of the disease (Maddison 1988). Anticentromere antibodies can be detected in approximately 30 per cent of patients with systemic sclerosis, particularly in those with the limited cutaneous form of the disease. They include the majority falling into the so-called CREST subset, up to 90 per cent of whom have these antibodies. In contrast, antitopoisomerase 1, found in approximately 25 per cent of patients with systemic sclerosis, is seen mainly in those with diffuse scleroderma and systemic involvement, particularly of heart and lungs. Antinucleolar antibodies, detected in high titre by indirect immunofluorescence in approximately 40 per cent of patients with systemic sclerosis, react with multiple targets and can be defined further by immunodiffusion, immunoblotting, or immunoprecipitation (Reimer 1990) (Fig. 4). These include the PM-Scl system (antibodies occur in approximately 5 per cent of patients, many with associated myositis, and produce homogeneous nucleolar staining), U3RNP-associated fibrillarin (produce clumpy nucleolar staining on indirect immunofluorescence and are not associated with a particular clinical subgroup), and nucleolar 7–2 ribonucleoproteins (produce homogeneous nucleolar staining and are not associated with a particular clinical group). Autoantibodies to RNA polymerase I, II, and III (anti-RNAP) occur in approximately 15 per cent of patients in various combinations (Hirakata *et al*. 1993). Most patients have diffuse cutaneous disease and there is an association with the development of renal disease.

Dermato- and polymyositis (see Chapter 5.9.1 and Chapter 5.9.2)

Multiple antibody systems are found in polymyositis or dermatomyositis, and precipitin systems can be detected in 60 per cent of patients by immunodiffusion (Targoff 1989). Each specificity occurs in a small proportion of patients. Antibodies to Jo-1 (histidyl-tRNA synthetase) are the most common specificity, occurring in approximately 20 per cent of adult patients with polymyositis, especially those developing interstitial lung disease, but are not common in dermatomyositis or disease in children.

Antiphospholipid antibodies (seeChapter 5.7.3)

Antiphospholipid antibodies, which are responsible for the false-positive VDRL test in lupus, have been associated with a characteristic spectrum of features (Table 6) (Lockshin *et al*. 1990). A range of antibody specificities exists, including antibodies to cardiolipin which can conveniently be detected by ELISA. A less sensitive method is to detect lupus anticoagulant activity in specially esigned functional tests of clotting. Antibodies are found in a wide range of clinical settings including individuals in whom the 'antiphospholipid syndrome' occurs in the absence of other autoimmune rheumatic disease features. They are also found in a proportion of the normal population. Anticardiolipin antibodies occur in up to 40 per cent of systemic lupus erythematosus patients, but only a small proportion develop the associated clinical features. Those most at risk have high titres of IgG anticardiolipin antibody and/or a lupus anticoagulant.

Major features
- Recurrent arterial and venous thrombosis
- Recurrent pregnancy loss
- IgG anticardiolipin antibody and/or positive lupus anticoagulant

Additional features
- Thrombocytopenia
- Pre-eclampsis
- Livedo reticularis
- Multi-infarct dementia
- Chorea
- Lupoid sclerosis
- Immune thrombocytopenia
- Haemolytic anaemia
- Endocardial disease

Table 6 The 'antiphospholipid syndrome'

Antineutrophil cytoplasmic antibodies (ANCA)

A significant development in recent years has been the identification of antibodies to selected cytoplasmic targets in polymorphs and monocytes, primarily to myeloid-specific lysosomal enzymes (see Wiik 1995 for a review).

A standard method of detecting these antibodies is indirect immunofluorescence using normal neutrophils, carefully separated and fixed in ethanol. Using this technique, different patterns of fluorescence indicate different antibody specificities (Fig. 7). Many of the conflicting findings in the literature can be accounted for, at least in part, by the use of different methods of fixation.

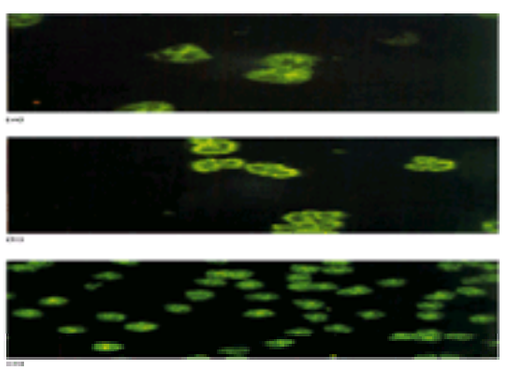


Fig. 7 Detection of ANCA using normal human polymorphonuclear leucocytes fixed in ethanol. (a) cANCA corresponding to antibodies to proteinase 3. (b) pANCA corresponding to antibodies directed to a variety of constituents, often myeloperoxidase but sometimes elastase and other lysosomal enzymes (by courtesy of Dr T. Wallington, Southmead Hospital, Bristol). (c) Atypical ANCA.

The typical pattern associated with Wegener's granulomatosis is coarse granular staining of the cytoplasm with central accentuation (**cANCA**) (Fig. 7(a)). cANCA corresponds to the presence of antibodies to a 29-kDa protein component of primary granules identified as serine proteinase 3.

The perinuclear pattern of ANCA (**pANCA**) (Fig. 7(b)) may be associated with antibodies to a variety of structures including myeloperoxidase, elastase, lactoferrin, cathepsin G, and other as yet unidentified cytoplasmic constituents. This pattern is artefactual and is due to the release of highly charged components of the primary granules on fixation of the cells, which are attracted to DNA.

The high specificity of cANCA for Wegener's granulomatosis has been demonstrated in several studies. Overall, sensitivity is 60 to 70 per cent for the disease depending on the patient population selected. cANCA are found in at least 90 per cent of patients with diffuse Wegener's granulomatosis and renal involvement but less commonly in patients with more localized disease. In some patients, measurement of the titre of cANCA provides an additional laboratory tool for assessing disease activity. However, since this is not true of all patients, changes in antibody titre have limited prognostic value (Kerr *et al*. 1993).

pANCA are found in a broader spectrum of vasculitis and in patients with idiopathic segmental necrotizing glomerulonephritis. Other patterns of ANCA are also seen. For example, a pattern distinctive from both cANCA and pANCA, and termed xANCA or atypical ANCA (Fig. 7(c)), is sometimes seen in patients with chronic inflammatory bowel disease, autoimmune hepatitis, and in a small proportion of patients with spondylarthropathies (Gross *et al*. 1993; Koh *et al*. 1995). Knowledge of the existence of these ANCAs is important in case physicians are misled towards a misdiagnosis of vasculitis.

Rheumatoid factor

Rheumatoid factors are antibodies directed against antigenic determinants on the Fc fragment of IgG. In rheumatoid arthritis, rheumatoid factor is polyclonal and reacts with a wide range of determinants on both the CH2 and CH3 domains of IgG including sites expressed by aggregated or denatured IgG, cross-reactive antigens shared by human and animal IgG, species-specific or even subclass-specific sites found only on human IgG, and genetically determined alloantigens such as Gm.

Rheumatoid factor is found in all classes of immunoglobulin. 'Classical' rheumatoid factor is a 19s IgM molecule which is still detected in many laboratories by techniques employing agglutination or flocculation of IgG-coated cells or particles. An example is the rheumatoid arthritis latex test that uses latex beads coated with human IgG which are cross-linked by rheumatoid factor to produce visible flocculation. This detects primarily IgM rheumatoid factor because of its high valency. The rheumatoid arthritis latex test also detects antiallotypic antibodies resulting from transfusion or pregnancy; thus, to increase specificity, techniques such as the Rose Waaler test have been used in which rheumatoid factor interacts with sheep cells coated with subagglutinating doses of rabbit IgG to cause haemagglutination. Since the presence of heterophile antibodies which react directly with sheep red cells can give rise to false-positive results, sera are usually tested in parallel with both coated and uncoated red cells and the difference in titres is expressed as the differential agglutination titre. Alternatively, the test serum is preincubated with uncoated sheep red cells (SCAT technique). Methods using rabbit IgG are less sensitive and give lower titres than the rheumatoid arthritis latex test because only a minor proportion of rheumatoid factors cross-react with rabbit IgG in rheumatoid arthritis. The presence of immune complexes or aggregated IgG can interfere with these techniques, sometimes to the extent that there is a false-negative result ('hidden rheumatoid factors'). Other methods for detecting rheumatoid factors are increasingly being used. These are laser nephelometry (Knight and Pritchard 1982), which is highly reproducible, and radioimmunoassays and ELISAs which are sensitive and quantitative and can be used to detect different isotypes of rheumatoid factor (Hay *et al*. 1975; Allen *et al*. 1981).

Tests for rheumatoid factor have some value as screening tests for rheumatoid arthritis, since rheumatoid factor is present in at least 75 per cent of patients with rheumatoid arthritis using a cut-off level for positivity which excludes 95 per cent of the normal population. In an early-synovitis clinic, for example, the combination of rheumatoid factor and low serum sulphydryl levels at the initial presentation with symptoms of less than 3 months' duration correctly predicted persistent disabling arthritis at 5 years in over 90 per cent of cases (Woolf *et al*. 1991). However, the clinical setting for interpreting this test is all important and the chance finding of rheumatoid factor on routine screening of a disease unlikely to be rheumatoid arthritis is of little clinical significance. Rheumatoid factor is not specific for rheumatoid arthritis (Table 7) and can be found, for example, during the course of certain infections and in approximately 15 per cent of the population over the age of 65.

1.	Rheumatoid arthritis	75%
2.	Other rheumatic diseases	Systemic lupus erythematosus, systemic sclerosis, Sjögren's syndrome, mixed connective tissue disease
3.	Acute viral infections	Mononucleosis, hepatitis, influenza etc., after vaccination
4.	Parasitic infections	Trypanosomiasis, malaria, schistosomiasis etc.
5.	Chronic inflammatory diseases	Tuberculosis, leprosy, syphilis, subacute bacterial endocarditis etc.
6.	Neoplasms	After irradiation or chemotherapy
7.	Other hyperglobulinaemic states	Hyperglobulinaemic purpura, sarcoidosis, chronic liver disease, cryoglobulinaemia

Table 7 Diseases commonly associated with rheumatoid factor

The presence of rheumatoid factor in rheumatoid arthritis indicates a poorer prognosis (Feigenbaum *et al*. 1979; Heliövaara *et al*. 1995) and a high frequency of systemic and extra-articular manifestations. The presence of high titres of IgG rheumatoid factor has been associated with necrotizing vasculitis (Allen *et al*. 1981).

Other serological tests in rheumatoid arthritis

Antiperinuclear factor can be demonstrated in up to 80 per cent of patients with rheumatoid arthritis (Berthelot *et al*. 1995). It is also found in a small proportion of psoriatic arthritis, juvenile chronic arthritis, autoimmune rheumatic diseases, and lung cancer but rarely in other clinical situations or in the normal population. There is no obvious relationship with disease activity or particular clinical manifestations. Recent studies suggest that the antigen target is related to human epidermal filaggrin and is the same as that recognized by so-called antikeratin antibodies (Sebbag *et al*. 1995).

Antibodies detected by immunoblotting to a 33-kDa protein in Hela cell extract were initially reported to be a highly specific finding in rheumatoid arthritis (Hassfeld *et al*. 1989). The RA33 protein has subsequently been shown to be homologous to the A2 protein of the heterogeneous nuclear ribonucleoprotein (hnRNP) complex (Steiner *et al*. 1992). Anti-RA33 antibodies have been found in approximately one-third of rheumatoid arthritis patients and have no relationship to rheumatoid factor, being occasionally detected in 'seronegative' rheumatoid arthritis. These antibodies are not specific for rheumatoid arthritis and are identified in other conditions such as systemic lupus erythematosus where their presence is associated with arthropathy, especially in erosive arthritis (Isenberg *et al*. 1994).

Variations in the carbohydrate content of IgG have been associated with rheumatoid arthritis and a restricted group of other diseases (Rahman and Isenberg 1995). In particular, the number of oligosaccharide chains attached to the C domains which lack the terminal galactose moiety (G0) is invariably raised in patients with active rheumatoid disease (Rademacher *et al*. 1988). In a recent study of undiagnosed patients with early-onset synovitis, a combination of a raised G0 and a positive rheumatoid factor gave a positive prediction factor of 94 per cent, significantly greater than either value alone (Young *et al*. 1991). In a similar study of 127 Dutch patients over a 6-year period, it was shown that the raised percentage of G0 correlated significantly with the number of swollen joints, number of prescribed second-line drugs, presence of rheumatoid nodules, and the radiological erosion score. Over the follow-up period, those patients whose first G0 level had been high had greater and more rapidly increasing erosion scores, a larger number of swollen joints, and were given more antirheumatic drugs (Van Zeben *et al*. 1994).

Chapter References

Aarden, L.A. and Smeenk, R. (1981). Measurements of antibodies specific for DNA. In *Immunological methods*, Vol. 2 (ed. I. Lefkovitis and B. Pernis), pp. 75–82. Academic Press, New York.

Aarden, L.A., Lakmaker, F., and Feltkamp, T.E.W. (1976). Immunology of DNA. I. The influence of reaction conditions on the Farr assay as used for the detection of anti-dsDNA. *Journal of Immunological Methods*, **10**, 39–48.

Allen, C., Elson, C.J., Scott, D.G.I., Bacon, P.A., and Bucknall, R.C. (1981). IgG antiglobulins in rheumatoid arthritis and other arthritides: relationship with clinical features and other parameters. *Annals of the Rheumatic Diseases*, **40**, 127–31.

Anderson, J.R., Gray, K.G., and Beck, J.S. (1961). Precipitating autoantibodies in Sjögren's disease. *Lancet*, **ii**, 456–60.

Arnett, F.C. *et al*. (1988). Increased frequencies of Sm and nRNP autoantibodies in American blacks compared to whites with systemic lupus erythematosus. *Journal of Rheumatology*, **15**, 1773–6.

Asero, R., Origgi, L., Crespi, S., Bertetti, E., D'Agostino, P., and Riboldi, P. (1987). Autoantibody to proliferating cell nuclear antigen (PCNA) in SLE: a clinical and serological study. *Clinical and Experimental Rheumatology*, **5**, 241–6.

Aviña-Zubieta, J.A., Galindo-Rodriguez, G., Kwan-Yeung, L., Davis, P., and Russell, A.S. (1995). Clinical evaluation of various selected ELISA kits for the detection of anti-DNA antibodies. *Lupus*, **4**, 370–74.

Barakat, S. *et al*. (1992). IgG antibodies from patients with primary Sjögren's syndrome and systemic lupus erythematosus recognise different epitopes in 60 kDa SSA/Ro protein. *Clinical and Experimental Rheumatology*, **89**, 38–45.

Ben-Chetrit, E., Fox, R.I., and Tan, E.M. (1990). Dissociation of immune responses to the SSA(Ro) 52 kDa and 60 kDa polypeptides in systemic lupus erythematosus and Sjögren's syndrome. *Arthritis and Rheumatism*, **33**, 349–55.

Bernstein, R.M., Hobbs, R.H., Lea, D.J., Ward, D.J., and Hughes, G.R.V. (1985). Patterns of antihistone antibody specificity in systemic rheumatic disease. *Arthritis and Rheumatism*, **28**, 285–92.

Berthelot, J.M., Maugars, Y., Audrain, M., Youinou, P., and Prost, A. (1995). Specificity of antiperinuclear factor for rheumatoid arthritis in rheumatoid factor-positive sera. *British Journal of Rheumatology*, **34**, 716–20.

Boire, G., Gendron, M., Monast, N., Bastin, B., and Menard, H.A. (1995). Purification of antigenically intact Ro ribonucleoproteins; biochemical and immunological evidence that the 52-kD protein is not a Ro protein. *Clinical and Experimental Immunology*, **100**, 489–98.

Bray, V.J., West, S.G., Boumpas, D.T., and Rubin, R.L. (1994). Antihistone antibody profile in sulfasalazine induced lupus. *Journal of Rheumatology*, **21**, 2157–8.

Burlingame, R.W. and Rubin, R.L. (1990). Subnucleosome structures as substrates in enzyme-linked immunoabsorbent assays. *Journal of Immunological Methods*, **134**, 187–99.

Christian, C. L. and Elkon, K.B. (1986). Autoantibodies to intracellular proteins: clinical and biological implications. *American Journal of Medicine*, **80**, 53–61.

Clough, J.D., Calabrese, L.H., and Valenzuela, R. (1989). The ANA profile; quality and cost effective laboratory utilization. *Cleveland Journal of Clinical Medicine*, **56**, 245–8.

de Rooij, D.J.R.A.M., Habets, W.J., van de Putte, L.B.A., Hoet, M.H., Verbeek, A.L., and van Venrooij, W.J. (1990). Use of recombinant RNP peptides 70K and A in an ELISA for measurement of antibodies in mixed connective tissue disease: a longitudinal follow up of 18 patients. *Annals of the Rheumatic Diseases*, **49**, 391–5.

Elkon, K.B., Parnassa, A.P., and Foster, L.L. (1985). Lupus autoantibodies target ribosomal P proteins. *Journal of Experimental Medicine*, **162**, 459–71.

Elkon, K.B., Bonfa, E., Weissbach, H., and Brot, N. (1994). Antiribosomal antibodies in SLE, infection, and following deliberate immunisation. *Advances in Experimental Medicine and Biology*, **347**, 81–92.

Emlen, W., Jarusiripipat, P., and Burdick, G. (1990). A new ELISA for the detection of double-stranded DNA antibodies. *Journal of Immunological Methods*, **132**, 91–101.

Feigenbaum, S.L., Masi, A.T., and Kaplan, S.B. (1979). Prognosis in rheumatoid arthritis. A longitudinal study of newly diagnosed younger adult patients. *American Journal of Medicine*, **66**, 377–83.

Feltkamp, T.E.W. (1990). Standards for ANA and anti DNA. *Clinical Rheumatology*, **1**, 78–81.

Fenning, S. *et al*. (1995). T cell lines recognizing the 70-kD protein of U1 small nuclear ribonucleoprotein (U1snRNP). *Clinical and Experimental Rheumatology*, **101**, 408–13.

Fritzler, M.J. and Tan, E.M. (1978). Antibodies to histones in drug-induced and idiopathic lupus erythematosus. *Journal of Clinical Investigation*, **62**, 560–67.

Gross, W.L., Schmitt, W.H., and Csernok, E. (1993). ANCA and associated diseases: immunodiagnostic and pathogenetic aspects. *Clinical and Experimental Immunology*, **91**, 1–12.

Habets, W.J., de Rooij, D.J., Salven, M.H., and van Venrooij, W.J. (1983). Antibodies against distinct nuclear matrix proteins are characterisic for mixed connective tissue disease. *Clinical and Experimental Immunology*, **54**, 267–76.

Hargraves, M.M., Richmond, H., and Morton, R. (1948). Presentation of two bone marrow elements. The 'tart' cell and 'LE' cell. *Proceedings of Staff Meetings of the Mayo Clinic*, **23**, 25–30.

Harley, J.B. *et al*. (1986). Anti-Ro (SSA) and anti-La (SSC) in patients with Sjögren's syndrome. *Arthritis and Rheumatism*, **29**, 196–201.

Hassfeld, W. *et al*. (1989). Demonstration of a new antinuclear antibody (anti-RA33) that is highly specific for rheumatoid arthritis. *Arthritis and Rheumatism*, **32**, 1515–20.

Hay, F., Nineham, L.J., and Roitt, I.M. (1975). Routine assay for detection of IgG and IgM antiglobulins in seronegative and seropositive rheumatoid arthritis. *British Medical Journal*, **3**, 203–7.

Heliövaara, M., Aho, K., Knekt, P., Aromaa, A., Maatela, J., and Reunanen, A. (1995). Rheumatoid factor, chronic arthritis and mortality. *Annals of the Rheumatic Diseases*, **54**, 811–14.

Hirakata, M. *et al*. (1993). Identification of autoantibodies to RNA polymerase II. Occurrence in systemic sclerosis and association with autoantibodies to RNA polymerases I and III. *Journal of Clinical Investigation*, **91**, 2665–72.

Holborow, E.J., Weir, D.M., and Johnson, G.D. (1957). A serum factor in lupus erythematosus with affinity for tissue nuclei. *British Medical Journal*, **2**, 732–6.

Hooper, D., Whittingham, G., Matthews, J.D., Mackay, I.R., and Curnow, D.H. (1972). Autoimmunity in a rural community. *Clinical and Experimental Immunology*, **12**, 79–84.

Humbel, R.L. (1993). Detection of antinuclear antibodies by immunofluorescence. In *Manual of biological markers of disease*, (ed. W.J. van Venrooij and R.N. Maini) **A2**, 1–16. Kluwer, Dordrecht.

Isenberg, D.A., Steiner, G., and Smolen, J.S. (1994). Clinical utility and serological connections of anti-RA33 antibodies in systemic lupus erythematosus. *Journal of Rheumatology*, **21**, 1260–3.

Jones, B.R. (1958). Lacrymal and salivary precipitating antibodies in Sjögren's syndrome. *Lancet*, **ii**, 773–7.

Kerr, G.S., Fleisher, T.A., Hallahan, C.W., Leavitt, R.Y., Fauci, A.S., and Hoffman, G.S. (1993). Limited prognostic value of changes in antineutrophil cytoplasmic antibody titer in patients with Wegener's granulomatosis. *Arthritis and Rheumatism*, **36**, 365–71.

Knight, R.K. and Pritchard, M.H. (1982). Nephelometry compared with differential antibody titre in routine rheumatoid factor measurements. *Annals of the Rheumatic Diseases*, **41**, 426–9.

Koh, W-H., Dunphy, J., Calin, A., McHugh, N.J., and Maddison, P.J. (1995). Atypical antineutrophil cytoplasmic antibodies in patients with ankylosing spondylitis. *British Journal of Rheumatology*, **34**, 695–6.

Lockshin, M.D., Qamar, T., and Levy, R.A. (1990). Anticardiolipin and related antibodies: thrombosis and fetal death. In *Pregnancy, autoimmunity, and connective tissue disorders*, (ed. J.S. Scott and H.A. Bird), pp. 185–211. Oxford University Press.

Maddison, P.J. (1988). Humoral immunity. In *Systemic sclerosis: scleroderma*, (ed. M.I.V. Jayson and C.M. Black), pp. 45–59. John Wiley, Chichester.

McCune, A.B., Weston, W.L., and Lee, L.A. (1987). Maternal and fetal outcome in neonatal lupus erythematosus. *Annals of Internal Medicine*, **106**, 518–23.

Moutsopoulos, H.M. *et al*. (1985). Anti-Ro (SSA) positive rheumatoid arthritis. A clinico-serologic group of patients with high incidence of penicillamine-D side effects. *Annals of the Rheumatic Diseases*, **44**, 215–19.

Pease, C.T., Charles, P.J., Shackles, W., Markwick, J., and Maini, R.N. (1993). Serological and immunogenetic markers of extraglandular primary Sjögren's syndrome. *British Journal of Rheumatology*, **32**, 574–7.

Pruijn, G.J.M., Slobbe, R.L., and Van Venrooij, W.J. (1990). Structure and function of La and Ro RNPs. *Molecular Biology Reports*, **14**, 43–8.

Rademacher, T. *et al*. (1988). The role of IgG glycoforms in the pathogenesis of rheumatoid arthritis. *Springer Seminars in Immunopathology*, **10**, 231–41.

Rahman, A. and Isenberg, D.A. (1995). Does it take sugar? A clinical role for measuring the glycosylation of IgG? *Annals of the Rheumatic Diseases*, **54**, 689–91.

Reichlin, M. and Harley, J.B. (1988). Immune response to the RNA protein particles in systemic lupus erythematosus. A distinctive dichotomy. *American Journal of Medicine*, **85** (Suppl. 6A), 3507.

Reichlin, M., and Van Venrooij, W.J. (1991). Autoantibodies to the U1RNP particles: relationship to clinical diagnosis and nephritis. *Clinical and Experimental Immunology*, **83**, 286–90.

Reimer, G. (1990). Autoantibodies against nuclear, nucleolar and mitochondrial antigens in systemic sclerosis (scleroderma). *Rheumatic Disease Clinics of North America*, **16**, 169–83.

Rokeach, L.A. *et al*. (1991). Characterisation of the autoantigen calreticulin. *Journal of Immunology*, **147**, 3031–9.

Rubin, R.L., Bell, S.A., and Burlingame, R.W. (1992). Autoantibodies associated with lupus induced by diverse drugs target a similar epitope in the (H2A–H2B)–DNA complex. *Journal of Clinical Investigation*, **90**, 165–73.

Rumore, P.M., Muralidhar, B., Lin, M., Lai, C., and Steinman, C.R. (1992). Haemodialysis as a model for studying endogenous plasma DNA: oligonucleosome-like structure and clearance. *Clinical and Experimental Immunology*, **90**, 56–62.

Sakamoto, M., Takasaki, Y., Yamanaka, K., Kodama, A., Hashimoto, H., and Hirose, S. (1989). Purification and characterization of Ki antigen and detection of anti-Ki antibody by enzyme-linked immunosorbent assay in patients with systemic lupus erythematosus. *Arthritis and Rheumatism*, **32**, 1554–62.

Sebbag, M. *et al*. (1995). The antiperinuclear factor and the so-called antikeratin antibodies are the same rheumatoid arthritis-specific autoantibodies. *Journal of Clinical Investigation*, **95**, 2672–9.

Sharp, G.C., Irving, W., Tan, E., Gould, G., and Holman, H. (1972). Mixed connective tissue disease: an apparantly distinct rheumatic disease syndrome associated with a specific antibody to an extractable nuclear antigen (ENA). *American Journal of Medicine*, **52**, 148–59.

Smeenk, R., van der Lelij, G., and Aarden, L.A. (1982). Measurement of low avidity anti-dsDNA by the *Crithidia luciliae* test and the PEG assay. *Clinical and Experimental Immunology*, **50**, 603–10.

Smeenk, R., Brinkman, K., Van den Brink, H., and Swaak, T. (1990 *a*). A comparison of assays used for the detection of antibodies to DNA. *Clinical Rheumatology*, **9**, 63–73.

Smeenk, R. *et al*. (1990*b*). Antibodies to DNA in patients with systemic lupus erythematosus. Their role in the diagnosis, the follow-up and the pathogenesis of the disease. *Clinical Rheumatology*, **9**, 100–10.

St Clair, E.W., Burch, J.A., and Saitta, M. (1994). Specificity of autoantibodies for recombinant 60-kd and 52-kd Ro autoantigens. *Arthritis and Rheumatism*, **37**, 1373–9.

Steiner, G. *et al*. (1992). Purification and partial sequencing of the nuclear autoantigen RA33 shows that it is indistinguishable from the A2 protein of the heterogeneous nuclear ribonucleoprotein complex. *Journal of Clinical Investigation*, **90**, 1061–6.

Targoff, I.N. (1989). Immunologic aspects of myositis. *Current Opinion in Rheumatology*, **1**, 432–42.

Teh, L.S. and Isenberg, D.A. (1994). Antiribosomal P protein antibodies in systemic lupus erythematosus. A reappraisal. *Arthritis and Rheumatism*, **37**, 307–15.

Teh, L.S. *et al*. (1993). Antiribosomal P protein antibodies in different populations of patients with systemic lupus erythematosus. *British Journal of Rheumatology*, **32**, 663–5.

Thompson, D., Juby, A., and Davis, P. (1993). The clinical significance of autoantibody profiles in patients with systemic lupus erythematosus. *Lupus*, **2**, 15–19.

Tishler, M., Golbrut, B., Shoenfeld, Y., and Yaron, M. (1994). Anti-Ro(SSA) antibodies in patients with rheumatoid arthritis—a possible marker for gold induced side effects. *Journal of Rheumatology*, **21**, 1040–2.

Totoritis, M.C., Tan, E.M., McNally, E.M., and Rubin, R.L. (1988). Association of antibody to histone complex H2A–H2B with symptomatic procainamide-induced lupus. *New England Journal of Medicine*, **318**, 1431–6.

Van Venrooij, W.J. and Sillekens, P.T.G. (1989). Small nuclear RNA associated proteins: autoantigens in connective tissue diseases. *Clinical and Experimental Rheumatology*, **7**, 635–45.

Van Zeben, D. *et al*. (1994). Early agalactosylation of IgG is associated with a more progressive disease course in patients with rheumatoid arthritis: results of a follow-up study. *British Journal of Rheumatology*, **33**, 36–43.

Wallace, D.J., Lin, H.C., Shen, G.Q., and Peter, J.B. (1994). Antibodies to histone (H2A–H2B)–DNA complexes in the absence of antibodies to double-stranded DNA or to (H2A–H2B) complexes are more sensitive and specific for scleroderma-related disorders than for lupus. *Arthritis and Rheumatism*, **37**, 1795–7.

Wiik, A. (1995). Anti-neutrophil cytoplasmic antibodies in Wegener's granulomatosis. Some clinical and pathogenetic aspects. *Seminars in Clinical Immunology*, **9**, 5–13.

Woolf, A.D. *et al*. (1991). Predictors of the longterm outcome of early synovitis: 15-year follow-up study. *Journal of Rheumatology*, **30**, 251–4.

Young, A. *et al*. (1991). Agalactosyl IgG—an aid to differential diagnosis in early synovitis. *Arthritis and Rheumatism*, **34**, 1425–9.

Zouali, M. and Stollar, B.D. (1986). A rapid ELISA for measurement of antibody to nucleic acid antigens using uv-treated polystyrene microplates. *Journal of Immunological Methods*, **90**, 105–10.

4.6 Joint fluid

J. Lawrence Houk

Formation, composition, and function of joint fluid

Joint fluid acts as a lubricant for the lining of the joint cavity. It is particularly effective for decreasing the coefficient of resistance during low-impact joint loading and motion, while its adhesiveness enhances joint stability and tracking. Joint fluid also is a vehicle providing nutrients to and removing metabolic products from the articular cartilage (Simkin 1991).

The fluid is composed of molecules filtered from plasma plus a glycoprotein, hyaluronate, produced by synovial cells. The composition and concentrations of synovial components are determined primarily by their molecular weights. Small molecules such as glucose, amino acids, uric acid, bilirubin, and several enzymes are filtered freely through the endothelium and synovial tissues. Consequently the concentrations of small molecules normally reflect those seen in plasma. The influx of larger molecules such as fibrinogen requires increased vascular permeability and is impeded normally by filtration through the synovial tissue meshwork of hyaluronate proteins. Accordingly large molecules found in serum are not normally present in joint fluid. Molecules leave the joint fluid via lymphatics. This process is usually unaffected by joint disease. Inflammation can change the normal distribution of joint fluid molecules. The inflammatory response can increase vascular permeability, allowing larger molecules such as fibrinogen to enter. These molecules account for the clotting of synovial fluid samples in abnormal fluid. At the same time, inflammation decreases vascular perfusion, thus reducing the number of small molecules that are available for diffusion into the joint fluid.

Indications for joint fluid examination

Why should joint fluid be examined? Even when performed by an experienced physician, the procedure is time-consuming. A busy outpatient practice does not allow for many delays, and consequently, joint fluid examination is often omitted from the evaluation of the patient with arthritis.

None the less, the results of a carefully conducted joint fluid examination usually have important diagnostic value. In fact, it may be the only source of a conclusive diagnosis. For example, a diagnosis of gout or calcium pyrophosphate crystal deposition disease cannot be certain unless the appropriate intracellular crystals are observed. Information gained from analysing joint fluid often provides more important data than any combination of blood tests or radiological procedures. In one report, treatment plans were changed in 53 per cent of 180 patients after synovial fluid analysis (Eisenberg *et al*. 1984).

The coexistence of two or more arthropathies is common, and at times only joint fluid testing can distinguish the diseases. Rheumatoid arthritis and osteoarthritis frequently occur simultaneously. Although conventional radiography can demonstrate the osteoarthritis, joint fluid evaluation may be necessary to reveal concomitant rheumatoid inflammation and suggest the correct decisions regarding therapy. Similarly, the risk of septic arthritis is increased in joints affected by rheumatoid arthritis and other arthropathies. A culture of joint fluid is the most sensitive and specific test for the recognition of a septic arthritis. Any articular site that is markedly more inflamed than other areas must undergo arthrocentesis to rule out infection. Gout, particularly the polyarticular form, can involve the hands, and tophi can occur in the region of the distal interphalangeal joints commonly involved in osteoarthritis (see Chapter 5.16). Microscopic examination of joint fluid for tophaceous material may be necessary to reveal this association. Likewise, gout and calcium pyrophosphate crystal deposition disease can coexist (Dieppe *et al*. 1988).

When should joint fluid be obtained? Generally joint fluid should be aspirated whenever a diagnosis has not been established or if there is a new articular event. Objective findings in a previously uninvolved area or increased swelling, pain, or erythema in an already affected site are indications for oint fluid aspiration. There is no absolute contraindication to an arthrocentesis unless the needle insertion site is in a potentially infected area. Joint fluid analysis should be considered a mandatory procedure in any individual with an unexplained articular effusion (Cohen *et al*. 1975).

When the examination of joint fluid is believed to be of diagnostic or therapeutic value it should be obtained without delay. In inflammatory disorders such as rheumatoid arthritis, simple rest can decrease inflammation; the joint effusion may resolve, and consequently, an opportunity for a diagnosis is lost. Monosodium urate and hydroxyapatite crystals can decrease over a period of time, occasionally in as little as 6 h (Kerolus *et al*. 1989). Thus, a delay in synovial fluid analysis may eliminate the only source of a diagnosis. Septic joints have to be diagnosed immediately. Postponement of fluid aspiration can result in delay of treatment, possibly resulting in joint destruction. On the other hand, antibiotics started before fluid aspiration can prohibit growth of the infecting organism in culture; the diagnosis may never be established with certainty, and the patient may be subjected to inappropriate treatment.

The arthrocentesis itself can be therapeutic. Joint capsule distension can produce pain, loss of motion, or both, particularly in the knee. Removal of fluid can rapidly reduce discomfort and allow restoration of movement. However, excess fluid should not be aspirated repeatedly just because it is present. The additional medical expense and even the small risk of arthrocentesis must be justified by the possibility of correcting the diagnosis or improving the patient's condition.

Sites accessible to obtaining joint fluid

Potentially, any diarthroidal joint or bursal space can be entered with a needle. However some areas are easier than others and some sites usually are not aspirated in a medical clinic setting. The knee is generally the simplest joint to enter but aspiration of the elbow, ankle, shoulder, and wrist as well as the olecranon or prepatellar bursas also should be a clinic procedure. A hip joint arthrocentesis is usually performed with fluoroscopic guidance. Aspiration of fluid from smaller joints is possible but frequently difficult. Pharmacological substances can be instilled into interphalangeal joint spaces with small (e.g. 30-gauge) needles but these joints are usually too small to enter with needles that are large enough for aspiration of viscous fluid. This is a particular problem when podagra is suspected. The bunion joint can be extremely tender to the slightest pressure, let alone the insertion of a hypodermic needle, even with topical anaesthesia.

Arthrocentesis should not be performed unless the precise site of inflammation is clear. Bursas or joints can be seeded with pathogenic microorganisms if the needle passes through an infected area such as an adjacent cellulitis. If there is any doubt as to the site and source of inflammation, the injection should be done through an adjacent uninflamed area. If this is not possible, then other methods such as magnetic resonance imaging should be employed to determine whether the joint or bursa

is involved.

Techniques of arthrocentesis

Aspiration

Equipment required for arthrocentesis is minimal. A 5 ml sterile disposable syringe will usually hold an adequate amount of the local anaesthetic. A 3 ml syringe is suitable for steroid injections. A 10 or 20 ml syringe may be necessary for the collection of large joint effusions. Small gauge needles, preferably a 30-gauge, are required for injecting the skin with the anaesthetic and entering small joints including the interphalangeal, metacarpal, and tarsal joints. Larger needles are used for aspirating fluid from larger joints. A 1&DFSF;1 2' 21-gauge needle will usually suffice for aspiration from the knee, ankle, shoulder, and wrist. An 18-gauge needle may be necessary if the fluid is thick. A 3&DFSF;1 2' 21-gauge spinal needle can be used to enter the knee joints of obese patients. A haemostat clamp is helpful for supporting the needle during syringe changes and also for removing the needle in the unlikely event that it breaks off during the procedure. Disposable gloves are not only advisable, but in the United States they are mandated by the Occupational Safety and Health Act.

An arthrocentesis should not be omitted because of the absence of absolute sterile conditions. Although some authors state that total sterile technique with sterile gloves and drapes is compulsory, other authorities believe that clean procedures, with sterile syringes and needles, are adequate (Cohen *et al*. 1975; Hasselbacher 1985; Gattar and Schumacher 1991a).

Descriptions of the entry sites for specific joints have been detailed elsewhere (Krey 1992b) (see Chapter 6.6). The entry site can be marked by pressure from the tip of the haemostat, the end of the needle cover, or the point of a retracted pen. The area is first cleaned with alcohol; an antiseptic solution such as povadine (Betadine) is then applied, and alcohol is wiped on the entry site. The skin can be anaesthetized with chloroethane (ethyl chloride) or fluroimethane spray. Alternatively, after the patient has been asked about a history of allergic reactions to the anaesthetic agent, the area can be anaesthetized by intradermal injection through a 30-gauge needle creating a weal with an agent such as 1 per cent lidocaine. If the injection is subdermal, it may be difficult to identify the site when inserting the larger centesis needle. To enter the joint, Gattar and Schumacher recommend that a 'quick decisive thrust through the skin' produces 'the least discomfort' (Gattar and Schumacher 1991a). I prefer an alternative, though time-consuming, approach that allows more cautious entry without causing pain (Fig. 1). When the skin has been successfully anaesthetized with the 30-gauge needle, the subcutaneous area is then infused with more of the anaesthetic agent. At this stage it is essential to suspend the procedure until the anaesthetic agent has taken effect. There is no point in using an anaesthetic drug unless time is allowed for it to work. The next target is the joint capsule. Careful probing can usually reveal this area without piercing the periosteum or articular cartilage. Furthermore, if anaesthesia is adequate and the needle is advanced slowly, accidental striking of the periosteum or articular cartilage will not cause significant pain. Painful needle injections can and should be avoided. Not only is the pain an unnecessary experience for the individual, but frequently the patient will never allow this procedure to be performed again.

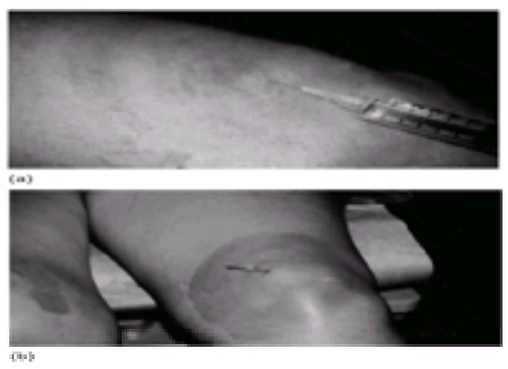


Fig. 1 (a) Anaesthetize skin with a 30-gauge needle. (b) Allow time for anaesthetic to take effect.

Aspiration of fluid from the joint cavity can be difficult if the fluid is loculated or very viscid. If fluid cannot be aspirated and infection is suspected, the joint can be irrigated with sterile saline and reaspirated. The needle may come out of the joint space when the distended joint capsule contracts during aspiration if it is not fully inserted. Leave a small amount of the anaesthetic in the joint (1 to 2 ml for large joints) as this may reduce the risk of infection because of the bacteriostatic properties of the solution.

Joint fluid collection and transportation

Synovial fluid for cell counts, differential, and Gram stain is put in a collection container with an anticoagulant, preferably with sodium heparin. Fluid for crystal examination is also stored this way. If a tube without anticoagulant is used, inflamed fluid may clot, making cell counts and crystal detection difficult. Lithium heparin or calcium oxalate anticoagulant tubes should not be used because they can form crystals which may cause confusion during microscopy. Aerobic and anaerobic cultures should be ordered routinely. Fungal and mycobacterium cultures should be requested only when there is clinical evidence of these diseases. Fluid for culture can be collected in the syringe, a Vacutainer tube, or blood culture bottles; however, if infection by *Neisseria gonorrhoeae* is a possibility, special care must be taken (Borenstein 1991). The joint fluid from suspected gonorrhoeal arthritis should be immediately plated onto chocolate agar and incubated in 5 to 10 per cent carbon dioxide. When an open laboratory is available, a technician should be summoned to put the fluid in culture as soon as it is aspirated. When the fluid must be transported to a laboratory, warm anaerobic transfer can be ensured by sending the fluid in the syringe.

Laboratories require that the syringe be covered with a sterile plastic cap to prevent accidental exposure to blood products. If there is only a small amount of fluid in a single collection container, make sure that the bacteriology laboratory does not centrifuge the specimen as, subsequently, an accurate cell count cannot be obtained.

Joint fluid tests

All specimens undergo a macroscopic examination for determination of fluid volume, clarity, colour, and viscosity. These evaluations are easy to do and provide essential diagnostic information. Clarity and colour should be observed in a clear glass container such as a Vacutainer tube. Plastic syringes are translucent and interfere with the accuracy of the determinations. Print can be read through transparent fluid; if the print appears as shades or blurs, the fluid is considered translucent. Sepsis can make the fluid opaque. Normal fluid is colourless or straw coloured. Viscosity is determined by the string test: joint fluid is made to drip from the syringe, and normal fluid will stretch out for 5 to 8 cm. Water-like drops indicate abnormally decreased viscosity.

All fluid should be examined for the presence of crystals and leucocytes. Gram stain and both aerobic and anaerobic cultures are indicated whenever infection is suspected. Infections can occasionally be associated with low white blood-cell counts. Thus, cultures should not be omitted because the fluid appears clear.

Although many other tests can be performed on joint fluid, few of them are of diagnostic value. Most authors agree regarding the relative value of joint fluid tests. Krey and Lazaro (1992) state that cultures, crystal examination, and leucocyte count with differential are the most useful tests. Shmerling *et al*. (1990) came to the same conclusion in a prospective study of joint fluid from 100 consecutive patients undergoing arthrocentesis. Lactate dehydrogenase levels were a sensitive marker of inflammation but had no greater diagnostic value than did leucocyte counts; protein measurements were likely to present 'misleading information'. The authors concluded that such tests should be omitted from routine diagnostic evaluation of synovial fluid. However, certain additional tests are helpful in specific situations. Glucose levels below 20 mg per 100 ml suggest infection but low levels can also occur in rheumatoid arthritis (Owen 1978). Accurate determination requires obtaining fluid after an 8-h fast and comparing it to serum glucose levels. The fluid should be placed in a fluoride tube to prevent further glucose metabolism. Rheumatoid factor can be present in the joint fluid and absent in the serum, but Baker (1991) states that this finding 'is not disease specific and does not require further evaluation'. Rodnan likewise noted that this finding is seen in other types of arthritis, including ankylosing spondylitis. However, Rodnan *et al*. (1963) further stated that 'demonstration of a positive agglutination reaction of synovial fluid may assist in the early diagnosis of rheumatoid arthritis when the serum has a negative reaction and the disease is confined to one or two joints'. Gas–liquid chromatography may detect volatile and non-volatile short-chain fatty acids from bacterial sources including gonococci. Thus, it may support the diagnosis of septic arthritis if antibiotics were started before cultures were obtained. Nucleic acid probes can detect DNA or RNA from microbes such as *Escherichia coli, Campylobacter jejuni*, or hepatitis B, but have not had extensive evaluation in joint fluid. Measurement of lipids,

urates, complement, circulating immune complexes, immunoglobulins, and antinuclear antibodies is of no value (Baker 1991).

One should decide the priorities of the tests to be done before aspirating the fluid. This forethought may prove to be crucial if only a small amount of joint fluid (less than 0.5 ml) is obtained. A delay in order to make these decisions or to find the correct tubes, culture transport vehicle, or syringe caps, can result in clotting or possible contamination of the fluid. If a microscope and trained personnel are available, small amounts of fluid should be analysed on site as follows: A drop of the fluid is placed on a clean slide, overlaid with a coverslip, and immediately examined for crystals and abnormal cells. The number of white blood cells is estimated. The coverslip and slide are than separated. A Gram stain can be performed on the slide smear and a Wright stain is done on the coverslip smear (Krey 1992a). The remaining fluid is then sent to the laboratory for culture.

If a microscope and other relevant supplies are not available, the fluid can be placed in a sterile vacuum container with anticoagulant and sent for white cell count with differential, culture, and crystal examination. Instructions accompanying the specimen should specify which test has the highest priority, should there be insufficient fluid for all three procedures. As mentioned previously, it must be ensured that the laboratory does not centrifuge any specimen destined for complete blood count and crystal analysis.

Reliability of test procedures

Unless joint fluid is examined promptly and reliably, the results can be worthless or even misleading. The macroscopic examination should be done immediately by the person performing the procedure. Ideally, the leucocyte count with a differential and examination for crystals should also be done on site. However, most outpatient medical facilities do not have the equipment or experienced staff needed for these tests. In this situation the synovial fluid must be sent to a hospital or commercial laboratory for further evaluation.

Unfortunately, studies indicate that many laboratories produce inaccurate results. Hasselbacher (1987) surveyed 39 hospitals in New Hampshire and Vermont regarding their experience and skill in joint fluid analysis. Their median frequency of joint fluid evaluation was only 1.5 times per month. The laboratories of 26 hospitals examined a reference synovial fluid, and there was a wide disparity in leucocyte counts and cell differentials. The results of crystal determinations were also frequently erroneous, because the crystals were not recognized or were identified inaccurately. Schumacher (1986) reported similar results from a study in which aliquots of 30 joint fluids were examined in four different laboratories. In 4 of the 30 specimens studied, the reported variations in the leucocyte count were large enough to change the fluid classification from an inflammatory to non-inflammatory type. The accuracy of the joint fluid analysis depends also on how soon testing is performed after the fluid has been aspirated. The leucocyte count starts to decline as soon as 1 h after aspiration; by 6 h, the drop is often large enough to result in misclassification. Refrigerating the fluid may slow the dissolution of monosodium urate crystals and reduce the formation of artefactual crystals (Kerolus *et al*. 1989).

Techniques for leucocyte and differential counting and for crystal identification

The risk that a hospital or commercial laboratory will produce inaccurate results, has led to a renewed interest in performing synovial fluid analysis on site. The following overview of fluid preparation methods includes references to more detailed descriptions.

Leucocytes are counted in the same way as peripheral white blood cells, except that normal saline, rather than acetic acid solution, is used for the diluent. Acetic acid at 2 mg per 100 ml will cause the fluid to clot. If red blood cells are present, they can be lysed by use of hypotonic saline (one part saline mixed with three parts distilled water) as the diluent. For the white cell count, aspirate the synovial fluid into a white blood-cell pipette to the 0.5 mark, dilute it with the saline to the 11 mark, and shake the pipette for 1 min by hand. Discard the first three drops from the pipette, then fill the clean haemocytometer. Count the leucocytes in all four large corner squares and multiply by 50 for the total cells per cubic millimetre (Krey 1992a). However, if the fluid appears clear on macroscopic examination, it can be placed directly without diluting onto the haemocytometer for counting.

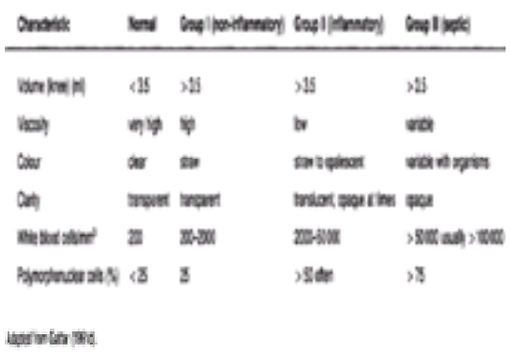
To prepare a specimen for a differential count, place a drop of sodium-heparinized joint fluid on a coverslip, overlay with a second coverslip and wait for a few seconds to allow the fluid to spread. Then separate the coverslips and allow them to dry. The smears are stained with Wright's stain for 3 min, and the stain is then drained off. Two drops of Giemsa stain are added, and the smear is flooded with distilled water which sits for 5 min. Following this, the stain is washed off with distilled water and passed for 1 s through a 95 per cent ethanol solution. The coverslip is washed again, dried, and mounted on a slide (Krey 1992).

A wet preparation can be made by placing one drop of fluid on a clean slide and applying a clean cover slip. If the amount of fluid in the syringe is not adequate, the needle can be flushed with either air or 95 per cent alcohol. The wet preparation is used to look for crystals and abnormal cells. A polarizing microscope with a red compensator is the best device for identifying crystals. Unfortunately, this apparatus is too expensive to be considered cost effective in most outpatient clinical settings. As an alternative, a standard microscope can usually be fitted with two polarizing filters, which will permit crystals on a wet smear to be visualized, and sometimes even identified.

Interpretation of examination results and differential diagnoses

Changes in joint fluid volume may indicate alterations in an inflammatory disorder. For example, a reduced amount of fluid on repeat aspiration of a septic joint may indicate an improvement. However, the aspirated fluid volumes can be misleading indicators, since complete evacuation of joint fluid is often difficult, particularly when it is very viscid or if loculation has occurred.

Decreased clarity of the synovia (i.e. turbidity) may indicate inflammation. Transparent fluid is characteristic of normal and group 1 non-inflammatory disease (Table 1). Translucent is characteristic of group II inflammatory (and occasionally, septic) fluids. Opaque fluid occurs in group III and is frequently septic.



Characteristic	Normal	Group I (non-inflammatory)	Group II (inflammatory)	Group III (septic)
Volume (knee) (ml)	<3.5	>3.5	>3.5	>3.5
Viscosity	very high	high	low	variable
Colour	clear	straw	straw to opalescent	variable with organisms
Clarity	transparent	transparent	translucent, opaque at times	opaque
White blood cells/mm³	200	200–2000	2000–50 000	>50 000, usually >100 000
Polymorphonuclear cells (%)	<25	25	>50 often	>75

Adapted from Gardner 1984 [5].

Table 1 Joint fluid characteristics

Although decreased clarity is usually due to increased leucocytes, other substances which can reduce clarity are crystals, lipids, fibrin, amyloid, and rice bodies (glistening white fragments produced by fibrin and collagen from degenerated synovial villi). Therefore, decreased clarity is suggestive of inflammation but is not absolutely diagnostic. Furthermore, normal clarity can occur during early inflammation that includes infection (Table 1). Normal joint fluid is colourless, or is straw coloured due to the presence of bilirubin. Inflammatory fluids are yellow. Septic fluids can be yellow, brown, or green. Cholesterol in joint fluid can cause a golden discoloration. A white or yellow pasty-looking fluid can result from urate or apatite crystals. Fluid speckled with dark particles indicates ochronosis (Hunter *et al*. 1974). Black or grey debris can be metal or plastic particles from total joint replacement. Lipids floating on top of the fluid may indicate an osteochondromal fracture (Lawrence and Seife 1971). The lipid layer should be investigated for the presence of bone marrow spicules, which is a further clue to existence of a fracture (Reginato *et al*. 1985). Increased viscosity occurs in hypothyroidism, acromegaly, and with mucinous cysts in osteoarthritis. Decreased viscosity is due usually to inflammation but oedema and bursal fluid also have low viscosities.

The mucin clot test is no longer considered to be a routine procedure, but it may be useful in distinguishing joint fluid from the anaesthetic or oedema fluid. The test is performed by putting a drop of joint fluid in acetic acid solution (2 mg/100 ml). Normal joint fluid clots, but when the hyaluronate molecules in the mucin have been

depolymerized by inflammation, the precipitate will be loose or will break up on agitation.

Joint fluid that clots when left in a syringe or non-anticoagulated tube indicates inflammation. Clotting occurs when fibrinogen and other clotting factors (prothrombin and factors V and VII) are present. These molecules are absent in normal fluid, occurring only in inflammation. The size of the clot is thought to be proportional to the extent of the inflammatory process (Cohen *et al*. 1975).

A microscopic examination may visualize abnormal cells. A finding of so-called Reiter's cells, i.e. macrophages that have phagocytosed a neutrophil, is not specific for Reiter's syndrome. Ragocytes are leucocytes, usually polymorphonuclear leucocytes, with large intracytoplasmic granules which can be visualized by light microscopy (Cherian and Schumacher 1983). These intracytoplasmic inclusions result from the phagocytosis of complement and immunoglobulins which, presumably, came from immune complexes. In 1965, Hollander and co-workers suggested that although cells with inclusion bodies are seen in several diseases, those containing complement and immunoglobulins are a feature of rheumatoid arthritis; accordingly, these authors referred to them as 'R.A. cells' (Hollander *et al*. 1965). Davis *et al*. (1988) noted that joints in patients with rheumatoid arthritis which contain a persistently elevated percentage of ragocytes are at greater risk for poor outcomes such as synovectomies and total joint replacement. Interestingly, ragocytosis might be seen in only a single joint, and, therefore, the results could not always be extrapolated to the entire disease process. Consequently the ragocyte status of a joint might influence treatment at least locally, or explain the lack of response of a single site when other joints seem well controlled with systemic therapy.

The joint-fluid leucocyte count allows a preliminary differential diagnosis. An increase of leucocytes, generally defined as a count exceeding 2000 cells/mm ^{3}, indicates inflammation. Normal joint fluid has fewer than 200 cells, and non-inflammatory fluids (group I) of arthritic joints have a leucocyte count between 200 and 2000/mm ^{3}. Conditions associated with inflammation (as defined by elevated white-cell counts) are listed in Table 2, but exceptions exist. The synovial white blood-cell count in systemic lupus erythematosus and other autoimmune rheumatic diseases can be in the non-inflammatory range. When rheumatoid arthritis is in remission or is controlled with anti-inflammatory medications, counts can be below 2000/mm ^{3} (Gattar and Schumacher 1991b). In one study, 27 per cent of 82 synovial fluid examinations in persons with gouty arthritis yielded results in a non-inflammatory range. Fluid from calcium pyrophosphate crystal deposition disease also can be non-inflammatory. The degree of leucocytosis seems to be proportional to the quantity of crystals (Cohen *et al*. 1975). Counts higher than 100 000 are usually seen in septic arthritis, but can also occur in rheumatoid arthritis, juvenile chronic arthritis, gout, calcium pyrophosphate disease, and Reiter's syndrome. A differential white-cell count can be very helpful in establishing a differential diagnosis. Krey and Bailen (1979) observed that when the leucocyte count was higher than 25 000 in a septic joint or in cases of gout, the average polymorphonuclear leucocyte count was at least 86 per cent. When the white cell count in these cases was higher than 50 000, the average polymorphonuclear leucocyte count exceeded 90 per cent. Although polymorphonuclear cells increased in individuals with rheumatoid arthritis who had more than 25 000 synovial white cells per cubic millimetre, the averages never exceeded 90 per cent, even when counts were greater than 50 000. Accordingly, polymorphonuclear counts higher than 90 per cent suggest gout or sepsis.

Group I — non-inflammatory	Group II — inflammatory
Osteoarthritis	Rheumatoid arthritis
Trauma	Crystal-induced synovitis
Internal derangement	Psoriatic arthritis
Osteochrondritis dissecans	Reactive arthritis
Osteochondromatosis	Whipple's disease
Charcot's arthropathy	Autoimmune rheumatic diseases
Subsiding inflammation	Polymyalgia rheumatica
Villonodular synovitis	Polychondritis
Hypertrophic osteoarthropathy	Sarcoidosis
Myxoedema	Behçet's syndrome
Acromegaly	Ankylosing spondylitis
Haemochromatosis	Juvenile rheumatoid arthritis
Gaucher's disease	Rheumatic fever
Ochronosis	Lyme disease
Paget's disease	Leukaemia
Sickle cell disease	
Milwaukee shoulder	

Adapted from Gardner (1981).

Table 2 Partial differential diagnosis of non-inflammatory and inflammatory arthropathies by joint fluid classification

Joint fluid eosinophilia was described in three individuals with Lyme disease by Kay *et al*. (1988). Other diseases associated with joint fluid eosinophilia are listed in Table 3.

- Rheumatic diseases:
  - Rheumatoid arthritis
  - Psoriatic arthritis
- Hypereosinophilic syndroma
- Infectious arthritides:
  - Tuberculosis arthritis
  - Lyme disease
- Allergic disease with arthritis:
  - Angio-oedema
  - Atopic history
  - Dermatographism
  - Urticaria
- Parasitic arthritides
- Haemorrhagic joint effusions:
  - Adenocarcinoma metastatic to synovium
  - Bilateral protrusio acetabuli following pelvic irradiation
- Posterthrography
- Postpneumarthrography
- Idiopathic

From Kay *et al*. (1988).

Table 3 Diseases associated with joint fluid eosinophilia

Blood in the joint fluid can be the result of a traumatic arthrocentesis or the conditions listed in Table 4. If the blood in the fluid is non-homogeneous or clumped it is more likely the result of trauma than of disease.

- Trauma with or without fracture
- Charcot's arthropathy
- Haemorrhagic diathesis:
  - Anticoagulant therapy
  - von Willebrand's disease
  - Haemophilia
  - Scurvy
  - Thrombocytopenia
- Haemangioma
- Tumour:
  - Pigmented villonodular synovitis
  - Synoviona
- Ruptured aneurysm
- Arteriovenous fistula
- Myeloproliferative disease with thrombocytosis
- Joint prosthesis

Table 4 Causes of haemorrhagic joint fluid

A variety of crystals may be identified by compensated polarized light microscopy. Monosodium urate crystals are usually needle shaped and visible under light microscopy, but sometimes they are small and difficult to visualize. Intracellular monosodium urate crystals confirm the diagnosis of gout. However, the presence of these crystals in an inflamed joint should not preclude obtaining cultures. Calcium pyrophosphate dihydrate crystals may be either rod-like or rhomboid. Their presence in phagocytic cells indicates calcium pyrophosphate deposition disease, of which there are several forms (McCarty 1975). When the laboratory reports the

presence of other crystals, the physician must determine whether or not these substances are of clinical significance. Although individual basic calcium phosphate (calcium hydroxyapatite) crystals can only be seen by electron microscopy, they form clumps that look like shiny coins under polarized light microscopy. They have been associated with periarticular problems including tendinitis and bursitis. Basic calcium phosphate crystals are also the pathological agent in 'Milwaukee shoulder' disease (Halverson *et al*. 1990). Other crystals associated with arthropathies include brushite (calcium hydrogen phosphate dihydrate), octacalcium phosphate, and calcium oxalate (Gattar and Schumacher 1991c).

The results of a Gram stain cannot be used to exclude the presence of a joint infection. Gram stains detect 75 per cent of staphylococcal infections, 50 per cent of most Gram-negative infections, and only 25 per cent of gonococcal infections (Goldenberg and Reed 1985).

Joint fluid findings in specific diseases

Osteoarthritis

Osteoarthritis is associated with a group I non-inflammatory fluid. It is usually clear and straw coloured with normal viscosity. The cell count is less than 2000 and the type is predominantly mononuclear. However, slight decreases in viscosity and a mild leucocytosis with cell counts up to 8000 can occur, with an even number of neutrophils to lymphocytes. The fluid volume varies from 1 to 50 ml.

Fragments of bone and cartilage may exist in the joint fluid specimen. Calcium pyrophosphate, basic calcium phosphate (calcium hydroxyapatite), and cholesterol crystals also may be present. Gibilisco *et al*. (1985) described knee fluid examined from 100 consecutive persons with osteoarthritis. In 60 per cent calcium pyrophosphate and/or calcium hydroxyapatite crystals were found. The individuals with crystals had more severe osteoarthritis on radiography and also tended to be older than persons without crystals. There was also an increased frequency of previous corticosteroid injections in the knees of the persons with crystals present. Patients with a higher percentage of crystals within cells tended to have more severe osteoarthritis. However, there was no difference in the leucocyte count between the fluids with and without crystals, but fluids with cell counts in excess of 2000 were excluded from the study. In some cases the radiological changes may precede the identification of synovial fluid crystals. In this situation, the crystals may be the result of the degenerative process rather than the cause. Possibly, in disease patterns not typical for primary osteoarthritis, calcium hydroxyapatite or calcium pyrophosphate crystals might be part of the pathological process (see below— calcium pyrophosphate crystal deposition disease).

Rheumatoid arthritis

The joint fluid in rheumatoid arthritis is usually inflammatory, and white cell counts can exceed 100 000/mm ^{3}. However, fluid occasionally may be in the non-inflammatory range, particularly if the disease has been suppressed with medications (Thomas and Carroll 1993). Neutrophils are usually predominant, but patients treated with non-steroidal anti-inflammatory drugs can demonstrate a preponderance of lymphocytes (Bahremand and Schumacher 1991). Fluid examined within the first 6 weeks of disease onset may also show more lymphocytes than polymorphonuclear cells, particularly when the synovial fluid white cell count is below 10 000 (Gattar and Richmond 1975). Extremely high white cell counts and a marked increase in polymorphonuclear cells, particularly when the knee is involved, suggest bacterial infection. Kortekangas *et al*. (1992) noted the median range of leucocytosis in infected knees from eight individuals with rheumatoid arthritis to be 118 600/mm^{3}, with a median polymorphonuclear leucocyte count of 93 per cent. As mentioned previously, a positive rheumatoid factor test of synovial fluid may help the early diagnosis of rheumatoid arthritis when the serum has a negative reaction and the disease affects only one or two joints (Rodnan *et al*. 1963). Glucose levels can be low, but this finding also occurs in septic arthritis.

Rheumatoid arthritis and crystal arthropathies rarely can coexist. Talbot *et al*. (1978) reported on 87 patients who were suspected of having both gout and rheumatoid arthritis.

Juvenile chronic arthritis

Synovial fluid is usually in the group II range with an average number of 11 400, but counts greater than 100 000 can occur. Also, there may not be a correlation between the fluid leucocyte count and the apparent extent of joint inflammation as group I results have been reported in inflamed joints. The main cell types are neutrophils but mononuclear cells can predominate. Glucose levels can be decreased (Cassidy and Petty 1995).

Punzi *et al*. (1992) suggested that higher polymorphonuclear counts in the synovial fluid may predict a polyarticular course (in contrast to a pauciarticular form) in persons with oligoarticular juvenile chronic arthritis.

Reiter's syndrome

Reiter's syndrome results in group II inflammatory fluid, but with white cell counts usually below 50 000 (Nordstrom *et al*. 1985). In 1967 Pekin and coworkers described an unusual macrophage in individuals with Reiter's syndrome, characterized by one or more intact, intracellular neutrophils (Pekin *et al*. 1967). This Pekin–Zvaifler, or Reiter's, cell was found subsequently in other arthropathies and is no longer considered diagnostic for this disorder.

Bacterial infection

Bacterial joint infections induce a marked inflammatory response in the joint fluid characterized by leucocyte counts higher than 50 000 cells/mm ^{3}, group III fluid; counts as high as 284 000 have been recorded. Lower counts can occur, however. Krey and Bailen (1979) examined 50 joint fluid specimens from individuals with culture-proven infections. Fifteen fluids had leucocyte counts below 50 000 cells/mm ^{3}. These individuals usually had mild, uncomplicated infections or serious associated disorders; some were on antibiotics. Regardless of the white cell count, the percentage of polymorphonuclear leucocytes is almost always over 90 per cent. As previously mentioned, negative Gram stains do not exclude a bacterial infection.

Viral arthropathies

The joint fluid white-cell counts in viral arthropathies can vary greatly from group I to group III ranges. The range in hepatitis B infection can vary from 145 to 90 000 cells/mm^{3} and in rubella from 1900 to 60 000/mm^{3} (Ytterberg 1993). Mononuclear cells sometimes predominate in the joint fluid, particularly in rubella and in parvovirus; this finding is rare in bacterial infections. Accordingly a viral infection should be considered whenever there is monocytosis in an inflammatory fluid.

Gout

The joint fluid in gout attacks is inflammatory. Leucocyte counts can exceed 50 000 with a predominance of polymorphonuclear cells. Viscosity is decreased. Crystals are needle or rod shaped in appearance but can be blunt. They are negatively birefringent, appearing yellow when aligned parallel to the axis of slow vibration of the red compensator on a polarizing microscope.

Crystals are present within polymorphonuclear leucocytes during acute attacks and visualization is required usually to confirm the diagnosis. However, the sensitivity of crystal analysis during an acute gout attack is 85 per cent (Wallace *et al*. 1977). Therefore, the failure to visualize crystals can not exclude the diagnosis of acute gout.

Monosodium urate crystals can be present extracellularly between episodes. Pascual (1991) analysed 74 joint fluid specimens from asymptomatic knees of 55 persons with gout. Monosodium urate crystals were still present in 36 of 37 persons who had previous attacks of gout in the sampled knee. However, only 8 fluid samples from 37 knees that had never been inflamed still showed crystals. Also, although the leucocytes counts were in a non-inflammatory range, they tended to be slightly higher in knees which had been the site of prior gout attacks.

In contrast, Rouault *et al*. (1982) did not detect a significant difference in the frequency of crystals between involved and non-involved joints in individuals with gout. They aspirated fluid from an asymptomatic first metatarsal phalangeal joint of 23 persons with gout. Crystals were present in 83 per cent of persons who had experienced previous podagra attacks in that joint compared with 67 per cent who had not.

Monosodium urate crystals can also be present in joint fluid in individuals with hyperuricaemia who have not experienced gout attacks, but this is uncommon (Rouault *et al*. 1982).

Gout attacks can occur in Heberden's and Bouchard's nodes (nodal osteoarthritis); a condition that is missed frequently (Simkin *et al*. 1983). Patients with coexisting disease are usually older than persons with gout. Many individuals are on diuretics. Lally *et al*. (1989) reported that 25 of 149 patients with gout developed episodes in distal interphalangeal or proximal interphalangeal joints affected with osteoarthritis. Thirteen of these individuals were women and the average age of onset was 71 years. In 11 persons involvement of these phalanageal joints was the intial or only manifestation of gouty arthritis. Seventy-two per cent were taking diuretics and 60 per cent had renal insufficiency. Most of the patients had radiographic changes indicative of gout in addition to the osteoarthritic changes, including soft tissue densities (tophi), large intra-articular erosions, non-marginal cortical erosions, and periarticular osteolysis. Crystal arthritis can coexist with septic arthritis. Baer *et al*. (1986) described 4 cases and studied 22 other patients from a literature review. Ten patients had coexisting gout, 13 had calcium pyrophosphate crystal inflammation, and in 3 persons both crystal diseases were present in the septic joint. The knee was the most commonly affected site. The authors stated that the combination of both diseases was often not recognized until late in the disease. They conclude that 'the detection of infection in a joint should not deter the physician from looking for a coexisting crystalline arthritis'.

Calcium pyrophosphate crystal deposition disease

McCarty *et al*. (1976) recognized six types of calcium pyrophosphate arthropathies. In a review on crystal deposition disease in the elderly, Doherty and Dieppe (1986) reduced this to two clinical forms—acute synovitis and chronic pyrophosphate arthropathy. The acute form resembles gout and was labelled the pseudogout syndrome by McCarty *et al*. (1962). The joint most often involved is the knee but attacks also involve the wrist, shoulder, and ankles. Episodes can include more than one joint but are usually not polyarticular. The synovial fluid is inflammatory (group II) with leucocyte counts in the range of 20 000 and the neutrophil count may exceed 90 per cent. The fluid may be bloodstained. Intracellular calcium pyrophosphate crystals confirm the diagnosis.

Chronic calcium pyrophosphate arthropathy occurs most often in elderly women and tends to involve primarily the knees but the wrists, shoulders, elbows, or ankles can also be involved. Individuals complain of pain and stiffness. Acute episodes superimposed on the chronic discomfort may be experienced. Physical examination may reveal bony joint enlargement with decreased range of motion. Patellofemoral crepitation may be present. Severe joint destruction can occur which may suggest a Charcot joint. Inflammatory findings with synovitis could resemble rheumatoid arthritis. The presence of inflammatory features, particularly synovitis in the knees, helps to distinguish this condition from osteoarthritis. Other features which differ between the two disorders include the sites of involvement. Osteoarthritis usually does not involve the wrists, elbows, shoulders, and ankles. Blood-stained fluid containing some neutrophils occurs more often in chronic calcium pyrophosphate arthritis. In addition, radiographic findings can be different in these two disorders (Doherty and Dieppe 1986). The presence of calcium pyrophosphate crystals supports the diagnosis. However, as mentioned above (osteoarthritis), these crystals as well as calcium hydroxyapatite can be seen in the fluid of joints afflicted by otherwise typical osteoarthitis. In this setting the crystals may be the result rather than the cause of the arthropathy.

Calcium pyrophosphate crystals may be linear or rhomboid. They have blunt ends in contrast to monosodium urate crystals that have sharp needle-like ends. Calcium pyrophosphate crystals are often more difficult to see than monosodium urate and can be confused with cholesterol, xanthine, or steroid crystals. Calcium oxalate arthropathy occurring in persons on haemodialysis can be clinically similar to calcium pyrophospate deposition diseases and the crystals may be confused on compensated polarized light microscopy (Reginato *et al*. 1986).

Although the presence of calcium pyrophosphate crystals can be suggested by ordinary light microscopy, a polarizing lens with a first order red compensator are usually necessary for conformation. Calcium pyrophosphate crystals are blue (weakly positively birefringent) when parallel to the axis of the compensator.

Whereas monosodium urate crystals can dissolve, particularly when left at room temperature, calcium pyrophosphate crystals remain stable in synovial fluid and can still be easily visualized up to 4 weeks after aspiration (McKnight and Agudelo 1991).

Chapter References

Baer, P.A., Tenenbaum, J., Fam, A.G., and Little H. (1986). Coexistent septic and crystal arthritis. Report of four cases and literature review. *Journal of Rheumatology*, **13**, 604–7.

Bahremand, M. and Schumacher, H.R. (1991). Effect of medication on synovial fluid leukocyte differentials in patients with rheumatoid arthritis. *Arthritis and Rheumatism*, **34**, 1173–6.

Baker, D.G. (1991). Chemistry, serology, and immunology. In *A practical handbook of joint fluid analysis*, (2nd edn) (ed. R.A. Gattar and H.R. Schumacher), pp. 70–8. Lea and Febiger, Philadelphia.

Borenstein, D.G. (1991). Stains, cultures, and other tests for infection. In *A practical handbook of joint fluid analysis*, (2nd edn). (ed. R.A. Gattar, and H.R. Schumacher), pp. 63–9. Lea & Febiger, Philadelphia.

Cassidy, J.T. and Petty, R.E. (1995). Juvenile rheumatoid arthritis. In *Textbook of pediatric rheumatology*, (3rd edn) (ed. J.T. Cassidy and R.E. Petty), pp. 133–223. W.B. Saunders, Philadelphia.

Cherian, P.V. and Schumacher, H.R. (1983). Immunoelectronmicroscopic characterization of intracellular inclusions in synovial fluid cells of patients with rheumatoid arthritis. *Ultrastructural Pathology*, **5**, 15–27.

Cohen, A.S., Brandt, K.D., and Krey, P.R. (1975). Synovial fluid. In *Laboratory diagnostic procedures in the rheumatic diseases*, (2nd edn) (ed. A.S. Cohen), pp. 1–62. Little, Brown and Co., Boston.

Davis, M.J., Denton, J., Freemont, A.J., and Holt, P.J.L. (1988). Comparison of serial synovial fluid cytology in rheumatoid arthritis: delineation of subgroups with prognositic implications. *Annals of the Rheumatic Diseases*, **47**, 559–62.

Dieppe, P., Campion, G., and Doherty, M. (1988). Mixed crystal deposition. *Rheumatic Disease Clinics of North America*, **14**, 415–26.

Doherty, M. and Deippe, P. (1986). Crystal deposition disease in the elderly. *Clinics in the Rheumatic Diseases*, **12**, 97–116.

Eisenberg, J.M., Schumacher, H.R., Davidson, P.K., and Kaufmann, L. (1984). Usefulness of synovial fluid analysis in the evaluation of joint effusions. *Archives of Internal Medicine*, **144**, 715–19.

Gattar, R.A. and Richmond, J.D. (1975). Predominance of synovial fluid lymphocytes in early rheumatoid arthritis. *Journal of Rheumatology*, **2**, 340–5.

Gattar, R.A. and Schumacher, H.R. (1991*a*). Joint aspiration: indications and technique. In *A practical handbook of joint fluid analysis*, (2nd edn) (ed. R.A. Gattar and H.R. Schumacher), pp. 14–23. Lea and Febiger, Philadelphia.

Gattar, R.A. and Schumacher, H.R. (1991*b*). Clinical significance of joint fluid findings. In *A practical handbook of joint fluid analysis*, (2nd edn) (ed. R.A. Gattar and H.R. Schumacher), pp. 85–96. Lea and Febiger, Philadelphia.

Gattar, R.A. and Schumacher, H.R. (1991*c*). Microscopic findings under compensated polarized light and phase light microscope. In *A practical handbook of joint fluid analysis*, (2nd edn) (ed. R.A. Gattar and H.R. Schumacher), pp. 45–58. Lea and Febiger, Philadelphia.

Gibilisco, P.A., Schumacher Jr., H.R., Hollander, J.L., and Soper, K.A. (1985). Synovial fluid crystals in osteoarthritis. *Arthritis and Rheumatism*, **28**, 511–15.

Goldenberg, D.L. and Reed, J.I. (1985). Bacterial arthritis. *New England Journal of Medicine*, **312**, 764–71.

Halverson, P.B., Carrera, G.F., and McCarty D.J.(1990). Milwaukee shoulder syndrome: fifteen additional cases and a description of contributing factors. *Archives of Internal Medicine*, **150**, 677–82.

Hasselbacher, P. (1985). Synovial fluid analysis. In *Rheumatoid arthritis; etiology, diagnosis, management*, (ed. P.D. Utsinger, N.J. Zvaifler, and G.E. Ehrlich), pp. 193–208. J.B. Lippincott, Philadelphia.

Hasselbacher, P. (1987). Variation in synovial fluid analysis by hospital laboratories. *Arthritis and Rheumatism*, **30**, 637–42.

Hollander, J.L., McCarty, D.J., and Rawson, A.J. (1965). The 'R.A. cell,' 'ragocyte,' or 'inclusion body cell'. *Bulletin of the Rheumatic Diseases*, **16**, 382–3.

Hunter, T., Gordon, D.A., and Ogryzlo, M.A. (1974). The ground pepper sign of synovial fluid: a new diagnostic feature of ochronosis. *Journal of Rheumatology*, **1**, 45–53.

Kay, J., Eichenfield, A.H., Balu, H.A., Doughty, R.A., and Schumacher Jr., H.R. (1988). Synovial fluid eosinophilia in lyme disease. *Arthritis and Rheumatism*, **31**, 1384–9.

Kerolus, G., Clayburne, G., and Schumacher, H.R. (1989). Is it mandatory to examine synovial fluids promptly after arthrocentesis. *Arthritis and Rheumatism*, **32**, 271–8.

Kortekangas, P., Aro, H.T., Tuominen, J., and Toivanen, A. (1992). Synovial fluid leukocytosis in bacterial arthritis vs. reactive arthritis and rheumatoid arthritis in the adult knee. *Scandinavian Journal of Rheumatology*, **21**, 283–8.

Krey, P.R. (1992*a*). Practical guide to the analysis of synovial fluid. In *Analysis of synovial fluid*, (ed. P.R. Krey and D.M. Lazaro), pp. 74–86. Ciba-Geigy Corp., Summit, NJ.

Krey, P.R. (1992*b*). Arthrocentesis. In *Analysis of synovial fluid*, (ed. P.R. Krey, and D.M. Lazaro), pp. 74–86. Ciba-Geigy Corp., Summit, NJ.

Krey, P.R. and Bailen, D.A. (1979). Synovial fluid leukocytosis: a study of extremes. *American Journal of Medicine*, **67**, 436–42.

Krey, P.R. and Lazaro, D.M. (1992). Specific tests. In *Analysis of synovial fluid*, (ed. P.R. Krey and D.M. Lazaro), pp. 28–52. Ciba-Geigy Corp., Summit, NJ.

Lally, E.V., Zimmermann, B., Ho G., and Kaplan, S.R. (1989). Urate-mediated inflammation in nodal osteoarthritis: clinical and roentgenographic correlations. *Arthritis and Rheumatism*, **32**, 86–90.

Lawrence, C. and Seife, B. (1971). Bone marrow in joint fluid: a clue to fracture. *Annals of Internal Medicine*, **74**, 740–2.

McCarty, D.J. (1975). Diagnostic mimicry in arthritis: patterns of joint involvement associated with calcium pyrophosphate dihydrate crystal deposits. *Bulletin of the Rheumatic Diseases*, **25**,804–9.

McCarty, D.J. (1976). Calcium pyrophosphate dihydrate crystal deposition disease—1975. *Arthritis and Rheumatism*, **19**, 275–86.

McCarty, D.J., Kohn, N.N., and Faires, J.S. (1962). The significance of calcium phosphate crystals in the synovial fluid of arthritis patients: the 'pseudogout syndrome' I. Clinical aspects. *Annals of Internal Medicine*, **56**, 711–37.

McKnight, K.M. and Agudelo, C. (1991). Comment on the article by Kerolus *et al. Arthritis and Rheumatism*, **34**, 118.

Nordstrom, D., Konttinen, Y.T., Bergroth, V., and Leirisalo-Repo, M. (1985). Synovial fluid cells in Reiter's syndrome. *Annals of the Rheumatic Diseases*, **44**, 852–6.

Owen Jr., D.S. (1978). Synovial fluid glucose. *Journal of the American Medical Association*, **239**, 193.

Pascual, E. (1991). Persistence of monosodium urate crystals and low-grade inflammation in the synovial fluid of patients with untreated gout. *Arthritis and Rheumatism*, **34**, 141–5.

Pekin, T.J., Malinin, T.I., and Zvaifler, N.J. (1967). Unusual synovial fluid findings in Reiter's syndrome. *Annals of Internal Medicine*, **66**, 677–84.

Punzi, L., Ramonda, R., Glorioso, S., and Schiavon, F. (1992). Predictive value of synovial fluid analysis in juvenile chronic arthritis. *Annals of the Rheumatic Diseases*, **51**, 522–4.

Reginato, A.J., Feldman E., and Rabinowitz J.I. (1985). Traumatic chylous knee effusion. *Annals of the Rheumatic Diseases*, **44**, 793–7.

Reginato, A.J., Seoane, J.L.F., Alvarez, C.B. *et al*. (1986). Arthropathy and cutaneous calcinosis in hemodialysis oxalosis. *Arthritis and Rheumatism*, **29**, 1387–96.

Rodnan, G.P., Eisenbeis Jr., C.H., and Creighton, A.S. (1963). The occurrence of rheumatoid factor in synovial fluid. *American Journal of Medicine*, **35**, 182–8.

Rouault, T., Caldwell, D.S., and Holmes, E.W. (1982). Aspiration of the asymptomatic metatarsalphalangeal joint in gout patients with hyperuricemic controls. *Arthritis and Rheumatism*, **25**, 209–12.

Schumacher, H.R. (1986). Reproducibility of synovial fluid analysis. *Arthritis and Rheumatism*, **29**, 770–4.

Shmerling, R.H., Delbanco, T.L., Tosteson, A.N., and Trentham, D.E. (1990). Synovial fluid tests, what should be ordered. *Journal of the American Medical Association*, **264**, 1009–14.

Simkin, P.A. (1991). Physiology and pathophysiology. In *A practical handbook of joint fluid analysis*, (2nd edn) (ed. R.A. Gattar and H.R. Schumacher), pp. 8–23. Lea and Febiger, Philadelphia.

Simkin, P.A., Campbell, P.M., and Larson, E.B. (1983). Gout in Heberden's nodes. *Arthritis and Rheumatism*, **26**, 94–7.

Talbot, J.H., Altman, R.D., and Yu, T. (1978). Gouty arthritis masquerading as rheumatoid arthritis or vice versa. *Seminars in Arthritis and Rheumatism*, **8**, 77–113.

Thomas, R. and Carroll, G.J. (1993). Reduction of leukocyte and interleukin-1B concentrations in the synovial fluid of rheumatoid arthritis patients treated with methotrexate. *Arthritis and Rheumatism*, **36**, 1244–52.

Wallace, S.L., Robinson, H., Masi, A.T., Decker, J.L., McCarty, D.J., and Yu, T. (1977). Preliminary criteria for the classification of the acute arthritis of primary gout. *Arthritis and Rheumatism*, **20**, 895–900.

Ytterberg, S.R. (1993). Viral arthritis. In *Arthritis and allied conditions: a textbook of rheumatology*, (12th edn). (ed. D.J. McCarty and W.J. Koopman), pp. 2047–65. Lea & Febiger, Philadelphia.

4.7 Genetic marker analysis

Rafal Ploski and Øystein Førre

Introduction

Many rheumatic diseases have a genetic component. In some of them the genetic component is very strong and is accountable directly for all the manifestations. In such cases there is usually a mutation that alters the function of one important gene, and these diseases are known as 'single gene diseases'. The pattern of inheritance of single gene diseases follows mendelian laws and the gene may be dominant or recessive. Osteogenesis imperfecta, Marfan syndrome, Ehlers–Danlos syndrome, and some forms of hereditary osteoarthritis are examples of single gene diseases that may be encountered by a rheumatologist.

The most prevalent rheumatic diseases, like rheumatoid arthritis, spondylarthropathies, systemic lupus erythematosus, and juvenile chronic arthritis, do not follow mendelian laws, although genetic factors play a role in all of them. The pattern of inheritance of these diseases is best explained by postulating that a number of different genes with relatively weak effects play a role in the pathogenesis in combination with environmental factors. These disorders will be referred to in this chapter as 'diseases with a complex inheritance'.

A polymorphic genetic marker is a trait which reflects variations in the DNA sequence between individuals. Phenotypic features and (glyco)protein, carbohydrate, or lipid molecules can all be used as genetic markers if they are determined by an allele of a single polymorphic gene. Many polymorphic genetic markers can be found by direct analysis of the DNA sequence. This is because all phenotypic variation is reflected at the level of coding or regulatory DNA sequences, and the analysis of DNA allows an exploitation of the polymorphism in the abundant non-coding regions of the genome.

The analysis of polymorphic genetic markers is essential for delineating the genetic component of any disease. For example, the recently developed approach of reverse genetics, which enables the responsible gene to be identified even when its function is not known, depends on an extensive analysis of a large number of polymorphic genetic markers. Once the genetic basis of a disease is known, genetic marker analysis can be useful clinically for the purposes of diagnosis and/or prediction of complications. Although genetic marker analysis is most rewarding in studies of single gene disorders, the example of HLA-B27 typing used in the diagnosis of ankylosing spondylitis illustrates that it may also be helpful in the investigation of diseases with a complex inheritance.

General approaches to identification of medically relevant genetic markers

A number of approaches exist aimed at finding genetic marker(s) located in the vicinity of the disease locus. Such marker(s) are needed usually for the eventual precise identification of the gene causing the disease. Further, in certain situations, they can be used in diagnosis even if the disease-gene is not known.

Genetic linkage, the classical approach

Genetic linkage is the associated inheritance of two or more alleles (of different loci) caused by their location close on the same chromosome. When any two loci are on different chromosomes their alleles segregate independently into the gametes and thus are transmitted independently to the progeny (Fig. 1(a)). However, if the loci are encoded on one chromosome their alleles are transmitted together unless crossing over occurs between them (Fig. 1(b)). The probability of crossing over between any two points on the chromosome generally increases in proportion to the distance between them. Thus, if loci are relatively close on a chromosome, crossing over between them will be rare. As a result, in related members of a family, the alleles of the considered loci will often occur together and linkage will be found.

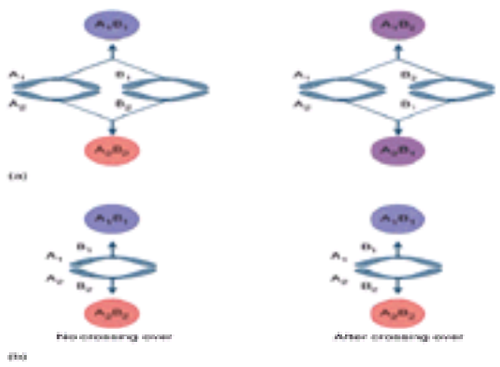


Fig. 1 Genetic linkage. The example used here is that of an individual who is doubly heterozygous for alleles of two loci A and B. (a) If A and B are on different chromosomes the A1, A2 and B1, B2 alleles segregate independently in the gametes and all their combination are equally frequent among the progeny. (b) If A and B are on the same chromosome the alleles are transmitted in parental combinations unless crossing over occurs. Linkage will be found between the A and B loci if they are so close that crossing over between them occurs at a frequency of less than 50 per cent.

The strength of the linkage is measured by the recombination fraction which corresponds to the percentage of gametes carrying combinations of alleles which were not present on the parental chromosomes. The recombination fraction is often expressed in centiMorgans. One centiMorgan, or 1 cM, corresponds to a recombination fraction of 1 per cent. In general, the stronger the linkage, the smaller the recombination fraction and the closer the locations of the two markers on the chromosome. Very roughly, a recombination fraction of 1 per cent can be expected if two loci are 1 Mb (1 million base pairs) apart.

A classical linkage study performed in medical genetics is based on an analysis of the cosegregation of a given marker with the disease phenotype in a family or families with many affected members. First, all the family members, both healthy and affected, are typed for a set of markers that are, as far as possible, evenly

distributed throughout the genome. Fewer markers can be typed for if the approximate location of the disease locus is known. The next step is to assess whether any of the markers cosegregate with the disease to a significant extent in the families concerned. A significant cosegregation indicates that the marker is located in the neighbourhood of the disease locus.

The statistical significance of any cosegregation observed is estimated by calculating the lod (logarithm of odds) score. The lod score is the logarithm of the ratio between the probability that the cosegregation is due to linkage and the probability that it has occurred by chance. A lod score of greater than 3 (corresponding to a ratio of probabilities of greater than 1000:1 in favour of linkage) is taken as evidence of linkage, while a lod score of less than –2 is usually sufficient to reject a hypothesis of linkage. A lod score of less than –2 means that chance alone is more than 100 times more likely to explain the data than a linkage model.

It is usual to calculate lod scores by assuming different recombination fractions between the marker and disease allele. The value which gives the highest score is accepted as being the most likely one, and can be used as a rough estimate of the distance between the marker and the disease locus.

Advanced methods of linkage analysis allow a large number of markers to be ordered in relation to the disease locus (multipoint analysis). It is also possible to study models that take into account incomplete penetrance of the predisposing gene and to test for genetic heterogeneity or for the presence of interactions between genes.

The power of the classical linkage approach lies in its sensitivity. Using a single informative marker, a relatively large region of chromosome (100 cM in theory, approximately 50 cM in practice) can be screened for the presence of a disease locus, provided a sufficient number of families are available.

The method's limitation lies in the fact that the results, including acceptance of the linkage together with the estimated recombination fraction, or rejection of the linkage hypothesis, are valid only if the exact pattern of inheritance of the disease is known. Thus, linkage analysis is a powerful tool in studies of classical mendelian traits, but is difficult to apply to diseases with a complex inheritance. Another disadvantage of this method as regards studies of diseases with a complex inheritance is that it requires relatively large pedigrees with many affected members.

Allele-sharing methods

Allele-sharing methods are a group of linkage-related methods which circumvent the limitations of the classical linkage approach because they do not require any presumptions about the way the disease is inherited and can be efficiently applied to very simple pedigrees. These methods are based on the idea that if a polymorphism predisposes for a disease, the chromosomal region where it is located will be shared by affected members of the family more often than expected.

The most commonly used method relies on the analysis of pairs of affected siblings. The first step is to identify an adequate number of families with at least two affected sibs (usually more than 50 such families are necessary). The sibs and the parents are then typed for a set of polymorphic markers and the number of alleles at each marker locus shared by each pair is determined. The sharing is defined as identity by descent (**IBD**), i.e. identity due to the inheritance of the same parental chromosome fragment and not just the presence of the same allele in two relatives (identity by state). Sharing in the form of IBD of no, one, or two alleles between two sibs should occur in 25, 50, and 25 per cent of cases, respectively. Whenever an increased frequency of pairs sharing one or two alleles is found, this indicates that the marker is linked to the gene predisposing for disease (Fig. 2).

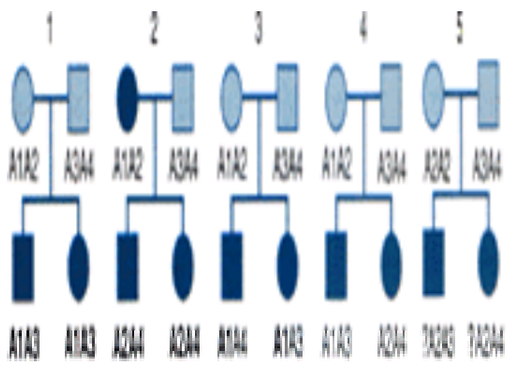


Fig. 2 Strategy of search for disease susceptibility genes based on affected sib pairs. In the example five pairs of affected sibs and their parents were typed for a marker A with four alleles (A1–A4). The alleles shared between the sibs due to identity by descent (IBD) are bold-faced. Two pairs (1 and 2) share two alleles, one pair (3) shares one, and one pair (4) shares no alleles as defined by IBD. One pair (5) is not informative because it cannot be resolved whether the sibs share one or no alleles as defined by IBD. The sharing in the form of IBD of no, one, or two alleles between two sibs should occur in 25, 50, and 25 per cent of cases, respectively. Whenever a significantly increased frequency of affected sib pairs sharing one or two alleles is found, this indicates that the marker is linked to the gene predisposing for disease. The squares and circles represent males and females, respectively. The black symbols denote affected individuals. Note that it is not important whether the parents are healthy or affected.

Allele-sharing methods have become increasingly popular in studies of diseases with a complex pattern of inheritance. Although in general these methods are less sensitive than classical linkage analysis performed in an optimal setting, this is more than compensated for by the fact that they are not dependent on information about the exact inheritance pattern of the disease or on a large and complex pedigree.

Association studies

This approach is known in epidemiology as a case–control study, and is based on determining the frequency of a polymorphic genetic marker in a group of unrelated patients and a group of healthy controls. When the marker occurs significantly more frequently among the patients than among the controls, it is said to be associated with the disease.

The statistical significance of an association is usually evaluated by the c^2 test or Fisher's exact test, while the strength of an association is often expressed in terms of an odds ratio (OR), sometimes also called Woolf's relative risk (Woolf 1956). Provided the disease is relatively rare, the odds ratio gives an approximate indication of how much more frequently the disease occurs among individuals carrying a given marker than in the population as a whole.

Frequent problems in association studies are unsuspected ethnic differences between patients and controls. Such differences can easily lead to false results and spurious associations because there is often considerable genetic variation between ethnic groups. As a solution to the problem, the patients' parents have been proposed as controls, since they are obviously well matched as regards ethnic background. A number of methods are available, all of which, in one way or another, take advantage of the pool of parental alleles that have not been transmitted to the patients (Falk and Rubinstein 1987; Julier *et al*. 1991).

An important factor to consider in designing and interpreting an association study is linkage disequilibrium (not to be confused with linkage). The statement that the alleles A1 and B1 (from different loci) are in linkage disequilibrium means that A1 and B1 are found together among a group of unrelated individuals more often than would be expected from their frequencies in the population. Linkage disequilibrium is especially strong in recently founded populations in which the haplotypes (i.e. sets of alleles encoded closely together on one chromosome) carried by the few founding individuals have not had time to 'mix' in the process of recombination. Similarly, linkage disequilibrium between A1 and B1 will also be found if the mutation which gave rise to the allele B1 has occurred relatively recently in an individual who carried the allele A1 on the same chromosome.

The occurrence of linkage disequilibrium has both advantages and disadvantages. The main disadvantage is that linkage disequilibrium complicates the interpretation of association studies. In general an association alone can never be accepted as a proof of significance of the given polymorphism for the disease pathogenesis because it may always be secondary to linkage disequilibrium with another gene in the region.

The advantage of linkage disequilibrium is that it provides an opportunity to detect an association even if the polymorphism being studied is not directly involved in the pathogenesis of the disease but is located somewhere in the vicinity of the disease locus. In favourable circumstances, and when a large number of markers are

studied, the presence of linkage disequilibrium can form the basis of an efficient strategy for precisely localizing the disease gene, known as linkage disequilibrium mapping (Hastbacka *et al*. 1992).

Linkage disequilibrium can be useful in genetic mapping provided it is not absolute, i.e. provided that certain alleles do not always occur together. The presence of a sufficient number of recombinant haplotypes is particularly important in the analysis of diseases with a complex pattern of inheritance in which the presence of single recombinants either among patients or controls is not very informative due to the low penetrance of the predisposing genes.

A very strong, practically absolute linkage disequilibrium occurs in the HLA complex between certain alleles encoding HLA-DR and -DQ molecules. As a result in many diseases which have long been known to be associated with certain alleles of these loci (for example juvenile chronic arthritis) it is still not clear which of the alleles is the one with primary association.

Linkage versus association

Linkage (found by the classical approach or by allele-sharing methods) and association both indicate that the gene predisposing for the disease is located in the neighbourhood of the studied marker. Linkage methods are more sensitive than those based on association and it often happens that a linkage but not an association is found (although the reverse is also possible) (Lander and Schork 1994). In general, a marker showing association with disease is likely to be located closer to the disease locus than a marker showing only linkage.

Markers showing association are potentially useful as a diagnostic test in sporadic individuals (for example B27 in diagnosis of ankylosing spondylitis). In contrast, markers showing linkage can be used to predict disease only if the proband comes from a family with many affected members (for example RFLP markers used in the prenatal diagnosis of osteogenesis imperfecta).

Detailed discussion of linkage and association methods can be found elsewhere (White and Lalouel 1987 ; Lander and Schork 1994).

Techniques used in genetic typing

Basic techniques of DNA analysis

Denaturation and hybridization

DNA is composed of four nucleotides which contain four different bases (guanine, **G**; cytosine, **C**; adenine, **A**; and thymine, **T**). In its native conformation a DNA molecule is a duplex consisting of two complementary strands kept together by specific non-covalent bonds formed between G and C as well as between A and T.

Denaturation or melting of DNA is the separation of the two strands of DNA molecule. Denaturation can be achieved by subjecting the DNA to high temperatures (95 to 100°C), alkaline pH, or certain chemicals such as urea or formamide. Generally, DNA rich in GC pairs melts at a higher temperature (or higher concentration of a denaturant) than DNA with a low GC content because the G:C bond is stronger than the A:T bond.

When the denaturing agent is removed, the complementary strands reform into duplexes and this process is called renaturation or hybridization. Under proper conditions, hybridization is very specific—a single strand of DNA will find its complementary molecule even in a highly complex mixture of other DNA fragments. Many molecular typing techniques rely in some way on hybridization of a specific single-stranded DNA fragment (called a probe) to its complementary sequence in the analysed sample.

Often hybridization assays are facilitated by blotting, i.e. immobilization of a sample of DNA on a solid foundation. Nylon or nitrocellulose membranes with high-affinity binding to DNA are most commonly used for this. The simplest method of blotting is the dot blot procedure, which consists of placing a droplet of DNA solution on to the surface of a membrane and leaving it to dry.

Gel electrophoresis

In its most commonly used form, gel electrophoresis enables nucleic acids to be separated according to their length. After going through the electrophoresis the molecules are grouped according to size in bands in the gel. A band can be visualized by staining with ethidium bromide (the easiest way) or by silver staining.

The gels are made from agarose or polyacrylamide. Agarose gels are simpler to prepare, but offer lower resolution than polyacrylamide gels, which can separate molecules that only differ in length by one nucleotide. Polyacrylamide gels, particularly those used for DNA sequencing, usually have a denaturing agent (urea) added, which prevents the formation of secondary structures and ensures that the DNA fragments migrate strictly according to their lengths.

Labelling

Labelling is the modification of macromolecules in a way that allows them to be easily detected. Many methods used in molecular genetics involve the labelling of nucleic acids. For example most hybridization assays are possible because the probe is labelled and can be detected with great sensitivity.

Labelling is achieved by attaching certain compounds to the end of the nucleic acid or by introducing modified nucleotides into the DNA molecule. These techniques are called end labelling and internal labelling respectively.

The classical labelling techniques rely on the use of radioactive isotopes of phosphorus or sulphur, followed by detection of the signal by autoradiography (i.e. exposure of a nylon membrane or the gel containing the labelled molecules on the photosensitive film).

Non-radioactive methods, particularly indirect immunoenzymatic methods, are also widely used (Guesdon 1992). These methods consist of labelling the nucleic acid with a small compound which does not affect the structure of the molecule and is relatively stable. Digoxigenin, biotin, acetylaminofluorene, and dinitrophenol are the most commonly used. The next step is to add an antibody specific for the label coupled with an enzyme like alkaline phosphatase or horseradish peroxidase. Finally, the labelled molecules are detected by adding a substrate from which the enzyme generates coloured or fluorescent products, or light (chemiluminescence).

Recently a group of direct fluorescent labels (usually referred to by their abbreviations: JOE, FAM, TAMRA, RO) have been developed. A useful property of these compounds is that they fluoresce at different wavelengths and can thus be differentiated when several of them are present.

Polymerase chain reaction (PCR)

PCR is an important method that enables millions of copies of a specific nucleotide sequence to be synthesized rapidly from a sample containing a minute amount of the relevant DNA (Mullis and Faloona 1987; Erlich *et al*. 1991). Before the invention of PCR, laborious methods of molecular cloning based on replication of DNA after it had been introduced into microorganisms had to be used. The principle of PCR is illustrated in Fig. 3.

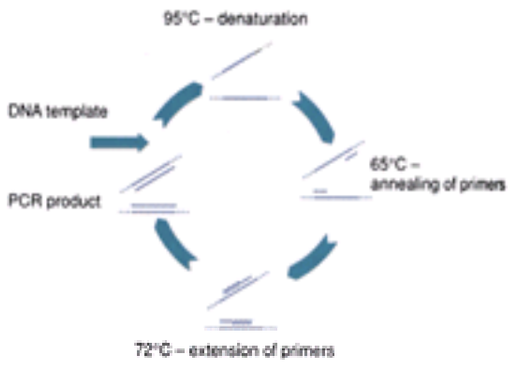


Fig. 3 The principle of the polymerase chain reaction (PCR). The PCR mixture contains sample DNA, two synthetic oligonucleotides or primers which anneal at the flanking regions of the sequence of interest, DNA polymerase (usually the thermostable Taq polymerase) and deoxyribonucleotides (i.e. substrates for the synthesis of DNA). The temperature of the reaction mixture is changed cyclically to permit denaturation of the target DNA, and annealing and extension of the primers. The PCR product synthesized in each cycle is used as a template in succeeding cycles and, therefore, the reaction is exponential. The temperatures shown are often modified to optimize the reaction.

PCR can also be used to modify the terminal regions of a DNA molecule. This is achieved by introducing the desired changes into the primers that are incorporated in the synthesized DNA. Although such primers are not fully matched with the target, they usually function well, especially if the mismatches are few and located far from the 3' end, where the polymerization begins. Examples of this type of modification are the introduction of a restriction site or the addition of a GC-rich sequence.

Despite its power, PCR has a number of limitations. (1) The synthesis is controlled by two oligonucleotide primers and, therefore, the sequences of the flanking regions of the amplified target need to be known. (2) PCR is very efficient when applied to sequences shorter than 2 kb, but longer sequences are usually difficult or impossible to amplify. (3) Owing to the relatively low fidelity of Taq DNA polymerase (the enzyme commonly used for PCR), artefactual mutations can be found in PCR products. A given mutation is present only in a small fraction of the synthesized molecules and, therefore, it generally does not pose a problem when the whole PCR product is analysed. However, when DNA produced by PCR is cloned (which allows single DNA molecules to be analysed) a large number of clones need to be analysed before any new sequence variant can be accepted as genuine. (4) The high sensitivity of PCR makes it prone to give false positive results when the sample is contaminated, especially when the contamination is caused by the products of previous reactions which are enriched enormously in the target sequence.

DNA for genetic analysis by PCR is isolated usually from peripheral blood. In our experience and that of Holodniy *et al*. (1991), it is important not to use heparin as an anticoagulant. Heparin interferes with the PCR, so that the samples have to be treated with heparinase before they can be used.

Principles of the main techniques of genetic marker analysis

Restriction fragment length polymorphism (RFLP)

This is used for the detection of variation in the distribution of recognition sites for restriction endonucleases in total genomic DNA with the aid of blotting and hybridization techniques.

Restriction endonucleases are enzymes which cleave DNA at certain sites when they recognize a specific sequence, usually four to eight nucleotides long. Owing to the large number of restriction enzymes, a single nucleotide substitution can often create or destroy the recognition site of one of them. The length of a DNA sequence between two restriction sites may also vary between individuals. In either case when genomic DNA from different individuals is digested by a given restriction enzyme, this gives rise to fragments of different lengths (Kan and Dozy 1978).

This restriction fragment length polymorphism, or RFLP, can then be analysed by the method described by Southern (1975). The method involves gel electrophoresis of digestion products followed by denaturation, blotting, and hybridization with a labelled probe complementary to the sequence in the neighbourhood of the polymorphic site (Fig. 4). Usually relatively long fragments of DNA such as cDNA (i.e. DNA obtained by reverse transcription of messenger RNA) or fragments of genomic DNA are used as probes, but synthetic oligonucleotides are also an alternative (Conner *et al*. 1983).

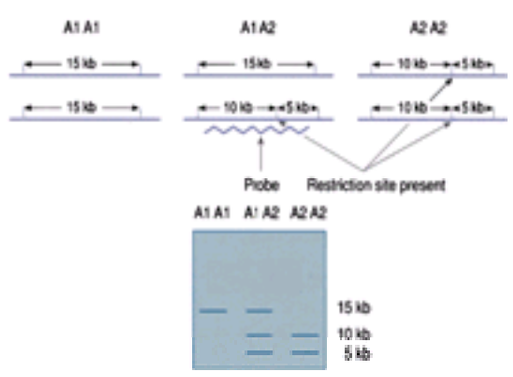


Fig. 4 The principle of analysis of restriction fragment length polymorphism (RFLP). The technique consists of digestion of genomic DNA with an appropriate restriction enzyme, gel electrophoresis of the resulting DNA fragments, denaturation, and blotting on to a membrane, followed by hybridization with a labelled probe annealing to the sequence around the polymorphic restriction site and autoradiography. The figure shows an example of band patterns corresponding to each of the three genotypes (designated A1A1, A1A2, and A2A2) generally found when there is one polymorphic restriction site and a probe derived from the surrounding region (as marked) is used. The A1 A1 and A2 A2 are homozygotes for the absence and the presence of the restriction site, respectively, whereas the A1 A2 is a heterozygote carrying the restriction site (the A2 allele) only on one chromosome.

The polymorphisms detected in the form of RFLP are often located in the non-coding DNA sequences such as introns or the flanking regions of a gene, and their exact positions are frequently unknown. This is because the majority of RFLP were discovered by random screening of different restriction enzymes until polymorphism detectable with a given probe was found. Such an approach is biased towards the detection of the polymorphism of the relatively poorly characterized non-coding regions of the genome because they are more prevalent and have a higher rate of mutation than the DNA sequences encoding proteins.

Before the discovery of PCR and the PCR-based techniques, the analysis of RFLP was the most important source of DNA-based polymorphic genetic markers.

PCR-based typing techniques

All the techniques from this group were made possible or were much simplified by the discovery of PCR. Unless otherwise stated, the first step in each method consists of PCR amplification of the appropriate region of DNA.

PCR-SSO (**SSO** stands for sequence-specific oligonucleotide) typing relies on the observation that under optimal conditions the stability of binding between a short DNA probe (typically 15 to 25 nucleotides) and the complementary target can be affected by a single nucleotide mismatch (Wallace *et al*. 1979; Saiki *et al*. 1986).

To perform the assay the PCR product is denatured, blotted on to a membrane, and hybridized with a labelled SSO probe. Usually the blotting is performed by the dot blot or related slot blot procedure. In the next step the membrane is washed in appropriately stringent conditions. If the blotted DNA does not contain sequences fully

complementary to the probe, the probe is removed, whereas the presence of a membrane-bound probe resistant to washing indicates that the DNA sample contains the relevant sequence (positive reaction) (Fig. 5(a)(i)).

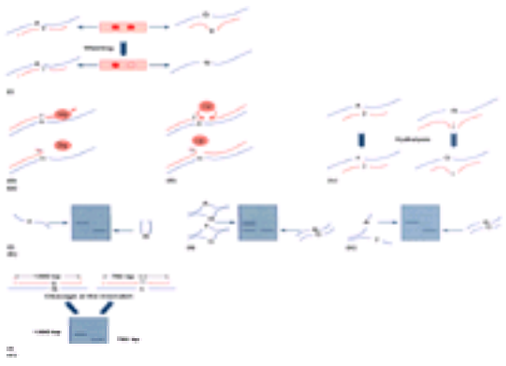


Fig. 5 Some of the commonest techniques used in the analysis of polymorphic genetic markers. The examples show analysis of DNA from two homozygous individuals carrying alleles differing by the substitution of G for A (or of G:C for A:T in double-stranded DNA) at the marker locus. The detection of the allele carrying A at the polymorphic position is shown as a positive reaction, while the G variant represents the negative control. The blue and red lines represent target DNA and probes respectively. The asterisk represents a label. (a) Typing techniques. (i) PCR-SSO typing. A labelled sequence-specific oligonucleotide (SSO) probe hybridizes with the blotted samples of denatured PCR product. During washing the probe bound to the mismatched sequences is removed. The red dots represent positive signals detected on the membrane. The detection method depends on the way the probe was labelled. (ii) Sequence-specific PCR. Under suitable conditions Taq polymerase will only extend a primer whose 3' end is fully complementary to the template DNA. (iii) Oligonucleotide ligation assay (OLA). The ligase (Lig) covalently joins two oligonucleotide probes only if their adjoining ends are correctly paired with the target DNA. (iv) Hybridization protection assay (HPA). An acridinium ester label is protected from hydrolysis only if the probe forms a fully matched hybrid with the target sequence. (b) Screening techniques based on detection of conformational changes in DNA molecules. (i) In single-strand conformation polymorphism analysis (SSCP), the single-stranded DNA is analysed by non-denaturing polyacrylamide gel electrophoresis. In the example the allele containing A migrates more slowly because of its different, 'more extended' conformation. (ii) In heteroduplex analysis, hybrids between the wild type and the mutated variant are created and their migration rate in non-denaturing polyacrylamide gel electrophoresis is compared with the migration rate of fully matched hybrids. The hybrids with mismatches (heteroduplexes) migrate more slowly because of their irregular conformation. (iii) In denaturing gradient gel electrophoresis (DGGE), the samples are analysed in a polyacrylamide gel containing an increasing concentration of denaturant. In the example shown the allele containing A starts to melt at a lower concentration of denaturant and thus migrates more slowly than the wild-type variant. (c) Techniques based on cleavage of mismatched nucleotides. A probe labelled at one end is added to the samples and allowed to hybridize. Next, the samples are treated so that cleavage occurs at the mismatched nucleotide and the products of the reaction are resolved by denaturing polyacrylamide gel electrophoresis. The presence of a band shorter than the length of the probe indicates the presence of a mismatch. The length of the band corresponds to the position of the mismatch relative to the beginning of the probe.

The PCR-SSO method allows quick analysis of a large number (up to a thousand) samples for the presence of particular polymorphism. If only a small number of samples are being analysed but many different polymorphisms need to be tested for, the reverse dot-blot PCR-SSO technique is useful (Saiki *et al*. 1989). In this assay a panel of SSO probes are immobilized on the membrane and hybridized with a suitably labelled PCR product. A modification of this approach is the microtitre plate oligotyping assay, which adapts the reverse dot-blot PCR-SSO protocol to the widely used techniques for protein analysis (enzyme-linked immunosorbent assay—ELISA), making the typing easy to automate (Cros *et al*. 1992).

Amplification refractory mutation system (**ARMS**) (Newton *et al*. 1989) and allele-specific PCR (**ASPCR**) (Wu *et al*. 1989) are similar although independently described assays based on the observation that under the appropriate conditions, Taq polymerase does not efficiently extend a primer with a mismatch at the 3' end (Fig. 5(a)(ii)). In this approach PCR is performed with the DNA sample using primers designed to be specific for sequences found only in a given allelic variant. The presence of a PCR product indicates that the analysed DNA contains the relevant sequence. A large number of polymorphisms can be analysed in a single PCR reaction if differently labelled primers are used (Gibbs *et al*. 1989). Although they are dependent on PCR, ARMS and ASPCR do not require prior PCR amplification.

Oligonucleotide ligation assay (**OLA**) (Landegren *et al*. 1988) is based on the particular features of a DNA-modifying enzyme called DNA ligase. In order to perform the assay, two oligonucleotide probes are designed: one probe is complementary to the target sequence downstream from the polymorphic position, while the other complements the polymorphic nucleotide and the immediately adjacent region upstream. The probes are mixed with the denatured sample of analysed DNA and allowed to hybridize with complementary sequences. In the next step the DNA ligase is added. The ligase will covalently link the two probes only if the nucleotides at their adjacent ends are correctly paired with the target DNA. Thus, the presence of the ligation product indicates the presence of a particular polymorphism in the analysed DNA (Fig. 5(a)(iii)).

By repeated cycles of denaturation, annealing, and ligation, the concentration of the ligated product can be increased linearly, which facilitates detection. An exponential accumulation of the product can be achieved by adding a second pair of primers that are complementary to the first (Wu and Wallace 1989). Owing to the analogies with PCR, this version of the technique has been called the ligase chain reaction (**LCR**). LCR is so sensitive that it can be applied directly to the sample of genomic DNA without the need for prior PCR amplification. The assay can be facilitated further by the introduction of thermostable DNA ligase (Barany 1991).

A simple typing assay can be designed if the polymorphism present in the PCR product creates a recognition site for a restriction enzyme (**PCR-RFLP** assay). Although the original protocol was relatively complex (Embury *et al*. 1987), the assay can usually be performed by digestion of the PCR product with the appropriate enzyme followed by agarose gel electrophoresis and visualization of the products by ethidium-bromide staining. The selection of the enzyme is facilitated by computer programs determining the distribution of all known restriction sites in a given sequence.

PCR-RFLP can sometimes be adapted to the analysis of polymorphisms which do not create a restriction site. This is done by introducing a mutation into the PCR product (via a partially mismatched primer) which creates the necessary restriction site together with a particular polymorphic sequence.

Hybridization protection assay (**HPA**) (Fig. 5(a)(iv)) (Arnold *et al*. 1989) is based on the use of oligonucleotide probes fluorescently labelled by acridinium ester. Under suitable conditions the acridinium-ester label is susceptible to hydrolysis when the probe is single-stranded or forms a hybrid with certain mismatches. However, when the probe hybridizes to the fully complementary sequence, this protects the label. To perform the assay, the labelled probe is mixed with the analysed DNA and allowed to hybridize with the complementary sequence. Next, a hydrolysing agent is added and the fluorescence is measured. A significant degree of fluorescence indicates that the sequence complementary to the probe was present in the analysed sample. This technique is very fast and easy to automate.

In general, the main feature of the methods discussed above is that the sequence of the genetic variants to be distinguished has to be known in advance. Thus they are useful in genetic typing but cannot be applied to screening for unknown polymorphisms.

All the above techniques can detect reliably even single nucleotide differences, and the choice of a particular method largely depends on the experience of the experimenter and the equipment available. PCR-SSO is the most widely used of the assays in this group, especially in the field of DNA-based HLA typing. Allele-specific PCR and PCR-RFLP are also popular because of their simplicity and speed.

Screening techniques

The methods discussed here are mainly used for the detection of novel sequence variations. All of them are facilitated by first amplifying the DNA by means of PCR.

The single-strand conformation polymorphism technique (**SSCP**) is based on the electrophoretic analysis of relatively short, single-stranded DNA molecules in a non-denaturing polyacrylamide gel (Fig. 5(b)(i)) (Orita *et al*. 1989). In such a gel, single-stranded DNA molecules adopt a three-dimensional conformation which depends on their nucleotide sequence. As the conformation can significantly affect the migration rate during electrophoresis, the polymorphic variants of the analysed sequence often appear as distinct bands. The method is quite sensitive and under optimal conditions almost 100 per cent of single nucleotide substitutions can be

differentiated (Sheffield *et al*. 1993).

The technique of heteroduplex analysis (Fig. 5(b)(ii)) exploits the conformational changes of double-stranded DNA molecules caused by the presence of mismatched nucleotides. The heteroduplexes (i.e. double-stranded DNA molecules containing mismatch(es)) can be detected because they often migrate more slowly than homoduplexes during electrophoresis in non-denaturing polyacrylamide gel.

Heteroduplexes can be created by adding the wild-type (i.e. 'normal') DNA to the sample expected to contain substitutions, followed by heating and cooling of the resulting mixture. The heating causes all the DNA molecules to be denatured, while the cooling causes renaturation. The heteroduplexes are formed because renaturation is possible not only between the fully matched strands but also between strands with some mismatches. The addition of exogenous DNA is not required if the analysed sample comes from a heterozygous individual and therefore contains both the normal and mutated variants.

Denaturing gradient gel electrophoresis (**DGGE**) (Fig. 5(b)(iii)) (Fischer and Lerman 1979) is based on the behaviour of double-stranded DNA molecules during their migration through a gel with an increasing gradient of chemical denaturant (urea and/or formamide). When a DNA molecule reaches sufficiently high concentrations of the denaturant, it starts to melt and slows down further migration through the gel. As a result of the complex interactions between nucleotides across the DNA helix, the exact concentration of denaturant required to initiate melting depends not only on the length and total base composition but also on the nucleotide sequence of the molecule. Even a single base-pair substitution may change the melting point of the whole DNA molecule sufficiently to allow separation from the original variant by DGGE.

The resolution of DGGE can be increased by attaching to the end of the analysed sequence a domain rich in guanine–cytosine pairs (a 'GC clamp') (Sheffield *et al*. 1989). This is usually done by incorporating the GC-rich domain into one of the PCR primers. The GC clamp acts by preventing complete denaturation of the analysed molecule during electrophoresis, which should be avoided because the retardation effect is lost after total separation of the strands.

DGGE is particularly powerful when combined with heteroduplex analysis. This is due to the fact that the presence of mismatch(es) destabilizes the DNA molecule and very often causes it to start melting in a lower concentration of denaturant than the fully matched duplex.

One general disadvantage of the above screening techniques is the decrease in sensitivity when long DNA fragments (more than 0.5 kb) are analysed. Further, even under optimal conditions none of the techniques are so sensitive that they are able to detect all cases of polymorphism. SSCP is the most widely used of these methods because of its technical simplicity and its relatively high sensitivity.

Screening techniques which help to localize polymorphisms

The techniques in this group enable the detection of unknown polymorphisms, and provide information about their location in the analysed DNA molecule.

In the approach based on selective cleavage of mismatched nucleotides (Fig. 5(c)), a labelled probe is added to the analysed sample and allowed to hybridize. The probes used can be relatively long (up to 1 kb). If the relevant sequence in the sample is not fully complementary to the probe, heteroduplexes are formed. In the next step the mismatched nucleotides are selectively cleaved by enzymatic or chemical methods, resulting in degradation of the probe into fragments that can be detected by gel electrophoresis. The occurrence of cleavage indicates the presence of polymorphism, and an analysis of the length of the cleavage products provides clues to the location of the polymorphism relative to the length of the probe.

Two cleaving agents can be used: ribonuclease A (Myers *et al*. 1985) or chemical cleavage with piperidine after treatment with hydroxylamine and osmium tetroxide (**HOT**) (Cotton *et al*. 1988). The ribonuclease A assay is simple, but has relatively low sensitivity and requires the use of an RNA probe. The chemical cleavage method has higher sensitivity, but is relatively laborious and necessitates the use of toxic substances. Both methods can be applied to the PCR-amplified DNA, although the use of total RNA isolated from suitable cells or total genomic DNA is also possible.

Another approach relies on the treatment of heteroduplexes generated as described in the previous section with a chemical (carbodiimide) which selectively modifies mismatched nucleotides (Ganguly and Prockop 1990). Heteroduplexes modified in this way are subsequently denatured and used as templates for the synthesis of complementary strands of DNA. The conditions are such that the synthesis terminates at the nucleotide modified by carbodiimide. Next, the extension products are analysed by gel electrophoresis. The presence of a fragment shorter than the analysed DNA molecule indicates that a mismatch was present whose position can be estimated from the length of the extension products.

The general advantage of the above group of techniques is that relatively long stretches of DNA (up to 1 kb) can be screened. However, because the assays are relatively laborious and do not have absolute sensitivity they are seldom used.

DNA sequencing

DNA sequencing is the only method which allows the definitive characterization of all the polymorphisms in a given DNA fragment. In general, sequencing methods are based on the generation of four groups of labelled DNA fragments whose lengths are related to the positions of each of the four nucleotides in the sequenced strand of DNA. The relative lengths of these four groups of fragments are subsequently determined by polyacrylamide gel electrophoresis (Maxam and Gilbert 1977; Sanger *et al*. 1977).

Recently DNA sequencing was considerably speeded up by the use of fluorescent labels and automated detection of the products of the sequencing reaction during the gel run. The application of four different fluorescent dyes allows the products of all four sequencing reactions to be run in a single lane. This increases the number of sequences that can be analysed by one gel and also improves the accuracy of the resolution by excluding artefactual lane-to-lane variation in electrophoresis. The detection and analysis process is automated and computer-based and, therefore, it is possible to get a printout of the sequence just after electrophoresis.

Before DNA can be sequenced it must be pure and present in relatively large amounts. This used to be achieved by laborious molecular cloning, but the discovery of PCR made it possible to omit this step, since the PCR product is generally pure enough to be sequenced directly, i.e. without prior cloning (direct PCR sequencing) (Rao 1994). In addition to its simplicity, this approach has the advantage of being insensitive to PCR artefacts.

Although DNA sequencing used to be a relatively laborious and difficult technique, the recent advances have made it a highly practical method, with many applications in genetic screening and typing.

Analysis of repetitive DNA sequences

Tandemly arranged, repetitive DNA sequences are an important class of genetic markers. Such sequences are very abundant, constituting about 10 per cent of the human genome. They frequently display length polymorphism caused by variations in the number of basic repeating units (Table 1).

	Satellites	Minisatellites (also called VNTR)	Microsatellites (also called STR)
Repeat unit	Depending on the family: a sequence of 5–6.25 bp, 68 bp, (Sau3A repeat) or 171 bp (alphoid repeat)	Usually a shared core consensus sequence similar to bacterial sequences known to promote recombination: GGTGGGCA-GAGACG	1–6 bp repeat Most often CA or CT or even a single A
Total length	Up to 5 Mb	Typically 0.5–30 Kb	Typically 100–250 bp
Distribution	Heterochromatin, mainly centromeres	Telomeric regions	Evenly throughout the genome
Polymorphism	+	+++	—
Significance	Basic biological research	Forensic medicine, linkage studies	Linkage studies, high-resolution genome mapping When expanded may cause disease

VNTR, variable number of tandem repeats.

Table 1 Characteristics of the main classes of the tandemly arranged, repetitive DNA sequences in the human genome (Cooper and Krawczak 1993)

The first repetitive sequences to be discovered were located mainly in the centromere regions of chromosomes and were very long. They were called 'satellite DNA' because they could be separated from the bulk of DNA by centrifugation.

Subsequently, another group of shorter, tandem, repetitive sequences characterized by very high polymorphism were found and christened 'minisatellites' (Table 1). Due to their high polymorphism, minisatellites are very useful in forensic medicine. In medical genetics, however, their application is limited because they occur mainly in the telomere regions of chromosomes (i.e. at the ends of the chromosomes).

Polymorphic minisatellites can be detected in the form of RFLP by Southern blotting. In this case the variation in the length of the DNA fragments generated by digestion with restriction enzyme is caused by differences in the distance between two restriction sites rather than by the presence or absence of one of the sites. Sometimes it is also possible to amplify the region containing the minisatellite by PCR and to assign alleles on the basis of the length of the PCR product (Jeffreys *et al*. 1988). However, this simple approach is limited by the low efficiency of PCR when applied to longer sequences. In particular, when a locus containing both short and relatively long alleles is analysed, it may happen that in a heterozygous individual only the shorter allele is amplified, falsely suggesting homozygosity.

The most recently discovered class of abundant repetitive DNA is represented by microsatellites or short tandem repeats (**STR**). The microsatellites consist of a number of copies of a very short sequence (1 to 5 bp long) and they are often polymorphic. In contrast to the previously discussed polymorphisms, the microsatellites are relatively evenly distributed throughout the genome. At a rough estimate, a microsatellite may be present every 20 kb in the genome, so that every gene may have one or more STR in its neighbourhood (Table 1).

The functional relevance of the microsatellite polymorphism found among healthy individuals is not clear. However, an excessive extension of the length of trinucleotide STR located within certain genes can seriously alter the function of these genes and cause disease. Fragile X syndrome, myotonic dystrophy, Huntington's disease, Kennedy's disease, and spinocerebellar ataxia are examples of diseases caused by such a mechanism.

A polymorphic STR can be analysed by PCR amplification of the block of repeat units in the presence of radioactively labelled nucleotides (internal labelling), followed by denaturing polyacrylamide gel electrophoresis and autoradiography (Weber and May 1989). However, several more efficient methods have recently been developed. A particularly powerful approach is based on PCR amplification of microsatellites using fluorescently labelled primers followed by electrophoresis with automated detection of the PCR products as they migrate through the gel (Ziegle *et al*. 1992). Both the principle and the equipment are the same as those used for automated DNA sequencing. The advantage of this method is that it is automated and enables three microsatellites with overlapping lengths to be run together in a single lane of the gel together with an internal molecular-weight standard.

Due to their abundance, even distribution throughout the genome, and polymorphism, the microsatellites are the single most important group of genetic markers known today. Their analysis has been crucial to the construction of the high-resolution map of the human genome (Weissenbach *et al*. 1992) and to the recent search for susceptibility genes to type I diabetes (Davies *et al*. 1994) and psoriasis (Tomfohrde *et al*. 1994). STR analysis is also likely to play a central role in any mapping of the genes that predispose to common rheumatic diseases.

Analysis of polymorphism in HLA genes and molecules

HLA molecules and the genes of the HLA complex

The HLA molecules are specialized receptors which bind fragments of antigen (peptides) and present them to the T lymphocytes. The interaction between the peptide–HLA complex and the T-cell receptor is the single most important step in specific immune responses. Depending on the circumstances, it can lead to the generation of cytotoxic T cells, the stimulation of antibody production by B cells, the recruitment of non-specific phagocytic cells in the immune response (delayed hypersensitivity response), or, in certain situations, the development of allergy. Apart from their role in the immune response, HLA molecules influence the selection of the T-cell repertoire in the thymus.

HLA molecules are subdivided into class I and class II molecules. The class I molecules interact with CD8 T lymphocytes and present, for the most part, peptides from the proteins located in the cytoplasm of the cell. The class II molecules interact with CD4 T lymphocytes and mainly present peptides derived from external antigens which have been internalized by the cell. Although there are certain differences between the HLA class I and class II molecules in the subunit composition (Fig. 6), crystallographic studies have shown that their three-dimensional structure is similar. Among their most notable similarities is that both types of molecules have characteristic peptide-binding grooves containing discrete sites (specificity pockets) that interact with side chains of the bound peptide (Brown *et al*. 1993; Guo *et al*. 1993).

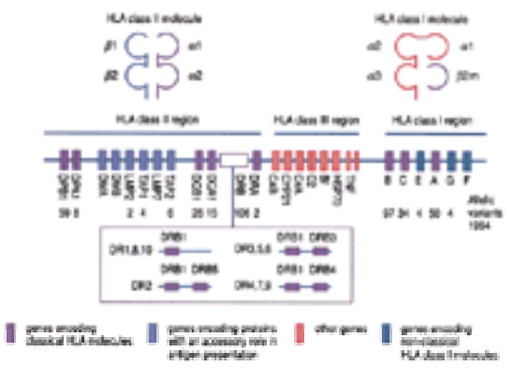


Fig. 6 HLA molecules (above) and schematic diagram of the HLA complex (below). HLA class I molecules consist of transmembrane glycoprotein (the heavy chain) non-covalently associated with a smaller non-polymorphic protein (b_2-microglobulin (b_2-m)). The extracellular portion of the heavy chain has three domains with an immunoglobulin-like structure (a_1–a_3 domains). There are three series of classical HLA class I molecules (HLA A, HLA B, and HLA C). The non-classical class I molecules, with a similar structure but with unclear functions, are called HLA E, HLA F, and HLA G. HLA class II molecules are heterodimers of two transmembrane units (a and b chain). The extracellular portions of each chain have two immunoglobulin-like domains (a_1, a_2 and b_1, b_2, respectively for the a and b chains). There are three different series of HLA class II molecules (HLA DR, HLA DQ, and HLA DP). In the class I molecules the peptide binding site (the groove) is formed by the a_1 and a_2 domains and in the class II molecules it is created by a_1 and b_1 domains. Only the loci known to be expressed have been included in the map of the HLA complex (i.e. not pseudogenes). The numbers of the allelic variants are according to Bodmer *et al*. (1994). The structure of the DR region differs according to the haplotype and in the majority of cases two functional DR molecules are encoded.

The genes encoding HLA molecules form a cluster called the HLA complex, which is located on chromosome 6 (Fig. 6). The only exception is the gene for b_2-microglobulin, which is found on chromosome 15. The HLA complex is subdivided into three regions, termed HLA classes I to III. The HLA class I region contains loci for both the classical (i.e. HLA A, B, and C) and the non-classical (i.e. HLA E, F, and G) class I antigens. The class II region harbours genes encoding the a and b chains of the DR, DQ, and DP molecules and the recently discovered genes for proteins involved in the processing of antigen. These include the *LMP2*, *LMP7*, *TAP1*, and *TAP2* genes as well as the *DMA* and *DMB* genes.

The products of both the *LMP* and *TAP* genes are important in the class I pathway of antigen presentation. The products of *LMP* genes form a part of a multisubunit proteinase complex (proteasome), which is involved in the degradation of cytoplasmic proteins into peptides. The *TAP* genes encode a transporter which delivers

peptides generated in this way into the lumen of the endoplasmic reticulum, where they bind to HLA class I molecules. The products of the *DMA* and *DMB* loci are necessary for the functioning of the class II presentation pathway, but their mechanism of action is not clear.

The HLA class III region encodes a number of proteins with immunological functions not directly involved in antigen presentation, such as complement components (C4, C2, Bf) and tumour necrosis factor (TNF). The genes for heat-shock protein 70 (HSP70) and an enzyme that participates in the synthesis of steroid hormones (CYP21) are also located in this region (Fig. 6).

In addition to the above-mentioned loci, the HLA complex harbours a large number of pseudogenes and expressed genes with unknown functions. All the available information on the genetic structure of the HLA complex has been compiled in a publicly accessible computerized database (Newell *et al*. 1994).

The HLA molecules and their genes are highly polymorphic. This polymorphism is reflected at the functional level by the differences in the repertoire of peptides that can be bound by various molecules. The polymorphism of HLA probably evolved because of the advantages it provides for the population. For example, the existence of a large number of different HLA molecules may decrease the chance that a single pathogen will cause a lethal disease in all members of a population. In keeping with the functional significance of HLA polymorphism, the majority of variable residues of both class I and class II molecules are clustered in domains which form the peptide-binding groove, especially at the sites determining structure and distribution of the specificity pockets (Guo *et al*. 1993).

Susceptibility to many diseases, including a number of common rheumatic disorders, is associated with certain HLA variants. It is not fully clear whether these associations reflect a direct role of the HLA molecules in the disease pathogenesis or whether they are secondary to linkage disequilibrium with a neighbouring gene. However, the central role of the HLA molecules in the immune response and the functional significance of their polymorphism suggests that the majority of disease associations are with the HLA molecules themselves.

HLA typing

The development of fast and accurate methods of HLA typing has always posed a challenge in the field of genetic marker analysis. The classical protocols for HLA typing rely on serological methods and are still in use, especially in typing for HLA class I, including the typing for HLA B27. The most widespread technique is the microcytotoxicity assay, which is based on assessing the viability of cells, usually leucocytes, which are incubated with a panel of antisera (and/or monoclonal antibodies) and complement. Cell death indicates the presence of the HLA type that the antiserum or monoclonal antibody was specific for. The microcytotoxicity assay was improved greatly by the application of antibody-coated immunomagnetic beads for purifying the cells used in the typing (Vartdal *et al*. 1986).

Unfortunately, serological typing, although simple to use, often fails to detect small differences between allelic variants and is particularly inefficient in the analysis of HLA class II polymorphism.

HLA class II typing with better resolution was made possible by a method based on the observation that T cells react strongly to all foreign (allogenic) but not self HLA class II molecules. In one version of this approach, lymphocytes from a typed individual are incubated with a panel of irradiated homozygous stimulator cells whose HLA class II types are known. Whenever lack of proliferation is observed it indicates that the individual carries the HLA type of certain given stimulator cells. The method is rather laborious and is seldom used today.

Isoelectrofocusing of isolated HLA molecules is another relatively sophisticated technique for HLA typing. Isoelectrofocusing is a form of electrophoresis that enables proteins to be separated on the basis of their net charge, and it is particularly powerful when combined with traditional electrophoresis in an approach known as two-dimensional gel electrophoresis. While this purely biochemical approach offers good resolution, it is laborious and technically difficult.

The major advance in HLA typing occurred when methods based on the analysis of DNA were introduced. The first widely used approach relied on the RFLP analysis of genomic DNA using locus-specific cDNA probes. The typing was made possible because correlations could be empirically established between certain RFLP and HLA types that have been determined by other methods.

DNA-based HLA typing benefited enormously from the discovery of PCR. Today PCR-based methods have come to dominate the analysis of HLA class II and are starting to be applied to the analysis of class I polymorphism. The most popular approach in class II typing relies on PCR amplification of the second exon of the analysed locus (where all polymorphic residues are located) followed by blotting and hybridization with a panel of 10 to 20 oligonucleotide probes (the PCR-SSO technique). PCR-SSO typing has been also applied to HLA class I typing, including assays for B27 (Hill *et al*. 1991) and its subtypes (Dominguez *et al*. 1992).

A number of other PCR based methods have been adapted to allow HLA typing. These include reversed dot-blot PCR-SSO, sequence-specific PCR, PCR-RFLP, techniques based on single-strand conformation polymorphism, heteroduplex analysis, hybridization protection assay, and direct sequencing of PCR products (Bidwell 1994). The latter technique, especially in an automated format, may well represent the ultimate solution to high-resolution HLA typing.

Nomenclature of the HLA system

One consequence of the tortuous history of HLA typing methods is the rather complex nomenclature, which differs according to whether serological specificities, T-cell-defined specificities, or the HLA genes are being discussed (Bodmer *et al*. 1994). The nomenclature of serologically defined HLA molecules is based on the name of the particular series (i.e. A, B, C, DR, DQ, or DP) followed by a number. The T-cell-defined specificities (only HLA class II) are designated as 'Dw', for the DR and DQ molecules, and 'DPw', for the DP molecules, followed in each case by a number.

The most precise nomenclature is based on the sequences of the HLA genes. Each gene is designated by the name of the locus followed by an asterisk ('*') and a four-digit number. The name of the locus is related to the associated series of serological specificities, while the number denotes the allelic variant. The first two digits of the number often correspond to the associated serological specificity, so that for example DRB1*0101 corresponds to a subset of DR1. There are, however, three exceptions to this rule: DR2, corresponding to DRB1*15 or *16; DR5, corresponding to DRB1*11 or *12; and DR6, corresponding to DRB1*13 or *14. The third and fourth digits denote the subtype. If the subtype is not known, only the first two digits are used: for example DRB1*01 denotes any subtype of DR1. In the case of alleles identical at the amino-acid level but differing at the nucleotide sequence, a fifth digit is added (for example DRB1*08031).

The three systems of nomenclature applied to the genes and/or molecules associated with serological specificity 'DR4' are shown in Table 2.

Serological specificity	T-cell defined specificity	DRB1 allele[a]
DR4	Dw4	DRB1*0401
DR4	Dw10	DRB1*0402
DR4	Dw13	DRB1*0403 DRB1*0407
DR4	Dw14	DRB1*0404 DRB1*0408
DR4	Dw15	DRB1*0405
DR4	Dw'KT2'	DRB1*0406
DR4	?	DRB1*0409 DRB1*04010 DRB1*04011 DRB1*0412 DRB1*0413 DRB1*0414 DRB1*0415 DRB1*0416 DRB1*0417 DRB1*0418 DRB1*0419

[a]Although each DR4 molecule is heterodimer of the α and β chains encoded by alleles of DRA and DRB1 loci, usually it is sufficient to specify only the DRB1 allele because of the very limited polymorphism of DRA.

Table 2 The three systems of nomenclature for the factors of the HLA system applied to the genes and molecules associated with serological specificity 'DR4' (Bodmer *et al*. 1994)

Genetic marker analysis in rheumatology

Prenatal diagnosis and genetic counselling in osteogenesis imperfecta

General features

Osteogenesis imperfecta is caused by defects in the structure or amount of collagen type I, which results in brittle bones. In the majority of cases the disease is inherited as an autosomal dominant trait. The molecular defect in osteogenesis imperfecta can be located at either of two unlinked loci which encode different chains of collagen type I (i.e. the COL1A1 and COL1A2 loci encoding the a_1 and a_2 chains of collagen type I, respectively).

Among the main features of osteogenesis imperfecta are the relatively high frequency of new (i.e. not inherited) mutations, the occurrence of mosaicism, and the fact that different mutations are usually found in different families. Mosaicism is a condition in which the mutated gene is only present in a fraction of the cells, and it results from new mutations occurring at an early embryonal stage of development. Mosaicism poses a problem in genetic counselling, since the probability that the child of a mosaic individual will inherit the disease may vary from 0 to 50 per cent, depending on the fraction of germ cells carrying the mutated allele.

Prenatal diagnosis and genetic counselling

In some cases the prenatal diagnosis of osteogenesis imperfecta can be performed by ultrasonography, which enables skeletal abnormalities to be demonstrated directly (Thompson 1993). A more sensitive method of diagnosis is the biochemical analysis of fetal cells obtained by chorionic villus biopsy or amniocentesis. The biochemical approach is based on the observation that the presence of defective chain(s) in the type I collagen leads to an excessive glycosylation of the mature molecules, which can be detected by gel electrophoresis (Cohn and Byers 1990). The method is rather time-consuming, however, since the fetal cells need to be expanded *in vitro* for approximately 1 month. It is also limited to chorionic villus biopsies because the cells obtained by amniocentesis produce a collagen that is different from the type I collagen necessary for the analysis.

An alternative approach is based on genetic analysis to determine whether the fetus carries the mutated gene. It is usually not known which mutation or even which locus (*COL1A1* or *COL1A2*) causes the disease in a particular family because of heterogeneity of mutations found in osteogenesis imperfecta. However, a number of markers defined by RFLP are known to be located in or very close to the COL1A1 and COL1A2 genes, making it possible to perform linkage analysis. In favourable circumstances, especially if a sufficient number of affected family members are available, linkage analysis may show a significant cosegregation (as measured by the lod score) between one of the RFLP and the disease. In such cases the RFLP provides a tag which enables the mutated gene in this family to be traced and a diagnosis to be made by typing the fetal DNA.

It may also be possible to find a marker linked with the disease gene on the basis of exclusion. The RFLP markers are so close to the *COL1A1* or *COL1A2* locus that the chance of recombination is negligible, and because osteogenesis imperfecta is known to be caused always by a mutation at one of these loci, it follows that one of the markers must be linked with the disease in every family. Thus, if other markers can be excluded, the one that remains has to be linked, even if it is not associated with a significant lod score. Such an approach may be successful even in relatively small pedigrees because a marker can be excluded on the basis of a single recombination event (Sykes 1993).

Although the linkage methods can be very powerful, their usage is limited to relatively extensive families with a history of disease whose members are available for collection of DNA samples. Furthermore, even when the family is large, the analysis may still be impossible if some of the key individuals are homozygous for the markers used. Another disadvantage of the method is that it cannot be used to estimate the chance of recurrence of disease among the children of a mosaic individual.

The best strategy in situations when linkage analysis fails is to focus on the direct identification of the mutation causing the disease in a given family. This has been achieved successfully with the use of methods based on cleavage of mismatched nucleotides in heteroduplexes as well as by PCR-based methods such as PCR-SSCP. Given the fast development of the DNA techniques, in the future, the direct identification of mutations is likely to play an increasing role in the genetic counselling in osteogenesis imperfecta.

Ankylosing spondylitis and other spondylarthropathies

Ankylosing spondylitis and other spondylarthropathies are associated with HLA B27 (Brewerton *et al*. 1973; Schlosstein *et al*. 1973; Benjamin and Parham 1990). In Caucasians, the B27 molecule is found typically in approximately 95 per cent of patients with ankylosing spondylitis compared with approximately 10 per cent of the whole population. The prevalence of B27 is similarly high among patients with juvenile ankylosing spondylitis, and is lower but still higher than normal among patients with reactive arthritis (60 to 80 per cent), arthritis associated with inflammatory bowel disease (50 per cent), and psoriatic arthritis (20 per cent). Typing for B27 has an established place in the clinical diagnosis of spondylarthropathies, and the recently proposed set of diagnostic criteria for spondylarthropathies includes B27 positivity (Amor *et al*. 1991).

Typing for B27 can be useful in a clinical situation provided the test results are interpreted correctly. It is important to realize that despite the strong association between B27 and ankylosing spondylitis, the majority (approximately 90 per cent) of B27-positive individuals do not have the disease. Thus, screening of the general population for the presence of B27 would not be helpful in identifying cases of ankylosing spondylitis.

However, as with any diagnostic test, the situation is different if there is some evidence that the patient has the disease before the test is performed (Khan and Khan 1982). For example, if on the basis of the clinical presentation the pre-test probability of ankylosing spondylitis is 50 per cent and the patient also turns out to be B27-positive, the post-test probability rises to 95 per cent, whereas in the case of a negative result it would be only 10 per cent. A reader familiar with the issue will realize that the above is generally true of any test with a sensitivity of 95 per cent and a specificity of 90 per cent (in the case of B27 typing, the sensitivity of the test equals the frequency of B27 among the patients, whereas the specificity is equal to the percentage of controls who are negative for B27).

The association between spondylarthropathies and B27 probably means that the B27 molecule is directly involved in the pathogenesis of the disease. This has been suspected for a long time because the increase in B27 has been found among patients with ankylosing spondylitis irrespective of their ethnic background, and also because all the subtypes of B27 that have been studied appear to be associated with the disease (Ivanyi 1992). Recently more direct evidence implicating the B27 molecule was obtained by showing that rats transgenic for human B27 develop a disease similar to ankylosing spondylitis (Hammer *et al*. 1990).

Whereas the presence of B27 is clearly important for the pathogenesis of spondylarthropathies, less is known about the role of other genetic factors. In two studies a relatively weak but distinct association with Bw60 was found among patients with ankylosing spondylitis in addition to the association with B27 (Robinson *et al*. 1989; Rubin *et al*. 1994). Recently it was suggested that homozygosity for a variant of the *LMP2* gene (*LMP2b*) could predispose to the development of acute anterior uveitis in the course of ankylosing spondylitis (Maksymowych *et al*. 1994). Our own studies have shown that both Bw60 and LMP2b homozygosity as well as HLA DRB1*08 and DPB1*0301 were all increased among patients with juvenile but not adult ankylosing spondylitis. Although the interpretation was complicated by linkage disequilibrium, there may indeed exist a non-B27, HLA-associated genetic component in certain forms of ankylosing spondylitis (Ploski *et al*. 1994c).

Rheumatoid arthritis

Susceptibility to rheumatoid arthritis is probably determined by the presence of a particular stretch of amino acids (the shared epitope) found in positions 67 to 74 of certain subtypes of DR4 (DRB1*0401, *0404, *0405, and *0408), DR1 (DRB1*0101), and DR6 (DRB1*1402) (Stastny 1978; Gregersen *et al*. 1987; Hansen and Nelson 1990). The shared epitope is found in up to 90 per cent of Caucasian patients with rheumatoid arthritis and in about 50 per cent of healthy controls.

All the DR molecules which carry the shared epitope show association with rheumatoid arthritis, although some associations are only apparent in certain ethnic groups. DRB1*0401, DRB1*0404, and DRB1*0408 are associated with the disease in Caucasians, DRB1*0101 is found to predispose to rheumatoid arthritis among Israeli Jews (and to some extent among Caucasians), and DRB1*0405 and DRB1*1402 are found in association with rheumatoid arthritis in Oriental populations and Yakima Indians, respectively. The different associations found among various ethnic groups are probably caused by differences in the prevalence of the various alleles in the general population. For example, DRB1*0405 and DRB1*1402 are relatively frequent among Orientals and Yakima Indians respectively, but occur only rarely among Caucasians.

Although the shared epitope determines susceptibility to rheumatoid arthritis, other regions of the HLA molecules may determine the severity of the disease. Patients who have severe disease that is positive for rheumatoid factor relatively often carry DR4, whereas those with mild, seronegative, rheumatoid arthritis tend to be DR1 positive (Dobloug *et al*. 1979; Winchester 1994; Ploski *et al*. 1994b). A particularly high risk of severe disease with extra-articular complications is associated with the carriage of two *DRB1* alleles encoding DR4, especially two copies of *DRB1*0401*, or *DRB1*0401* plus *DRB1*0404* (Weyand *et al*. 1992; Wordsworth *et al*. 1992).

The importance of the double dose of DR4 in determining the severity of rheumatoid arthritis is supported by studies of polyarticular juvenile chronic arthritis that is positive for rheumatoid factor. This subset of juvenile chronic arthritis probably represents relatively severe rheumatoid arthritis with an early onset and, interestingly, it shows an association with the carriage of two DR4-encoding alleles, like severe rheumatoid arthritis in adults (Nepom *et al*. 1984; Ploski *et al*. 1993b).

Typing for the shared epitope cannot help in the diagnosis of rheumatoid arthritis because of the relatively high prevalence of this sequence among healthy controls. However, typing aimed at identifying patients carrying a double dose of DR4 may have some potential in predicting the severity of the disease (Gough *et al*. 1994).

Early-onset, pauciarticular, juvenile chronic arthritis

Early-onset, pauciarticular, juvenile chronic arthritis, the most prevalent of the subsets of juvenile chronic arthritis, is a disease with relatively abundant HLA associations. DRB1*08 (Stastny and Fink 1979), DRB1*1301, DRB1*11, DRB1*12 (Stastny and Fink 1979; Glass *et al*. 1980; Reekers *et al*. 1983; Fernandez-Vina *et al*. 1994), and DPB1*0201 (Odum *et al*. 1986; Begovich *et al*. 1989) are all found in a larger proportion of patients than healthy controls. Among the patients, DPB1*0201 is frequently found together with DRB1*11, DRB1*12, DRB1*1301, or DRB1*03, suggesting that positive interaction predisposing to the disease occurs between these alleles (Van Kerckhove *et al*. 1990; Ploski *et al*. 1994a). The DRB1*04 and DRB1*07 alleles may have a protective effect against the development of this disease.

Early-onset, pauciarticular, juvenile chronic arthritis is often complicated by the occurrence of chronic iridocyclitis. The presence of DR5, in particular of the DRB*1104 subtype, may be a risk factor for the development of the eye disease, while DR1 shows a negative association with this complication (Giannini *et al*. 1991; Melin-Aldana *et al*. 1992; Ploski *et al*. 1993b). A polymorphism located in the promoter region of the interleukin 1a gene (chromosome 2, non-HLA-linked) may also predispose to the eye disease (McDowell *et al*. 1994). Interleukin 1a and the related interleukin 1b are proinflammatory cytokines, which play a role in the induction of fever, the acute-phase response, and the local inflammatory fibrosis and tissue destruction.

Another complication occurring in this disease is the development of relatively severe polyarticular joint involvement. Susceptibility to polyarticular disease in early-onset, pauciarticular, juvenile chronic arthritis may be increased by DQA1*0101 (or DRB1*0101, which is in strong linkage disequilibrium with it) (Van Kerckhove *et al*. 1991; Ploski *et al*. 1993b).

An interaction has been suggested between HLA and the genes encoding the antigen receptor of T cells in early-onset, pauciarticular, juvenile chronic arthritis. In two studies the frequency of the T-cell receptor *TCRVB6S1*2* allele (the Vb6.1b allele according to the previous nomenclature, encoded on chromosome 7) has been found to be increased among patients with pauciarticular, juvenile chronic arthritis who were positive for *DQA1*0101* (Maksymowych *et al*. 1992; Charmley *et al*. 1994). Interestingly, the polymorphism of *TCRVB6S1* has a distinct significance: the *TCRVB6S1*2* allele, in contrast to *TCRVB6S1*1*, is non-functional because it lacks one of the cysteine residues necessary for the correct folding of the protein. However, because in some cohorts of patients the association with *TCRVB6S1*2* could not be found, more studies are needed before definite conclusions can be drawn (Ploski *et al*. 1993a).

Unfortunately, although the immunogenetic associations in early-onset, pauciarticular, juvenile chronic arthritis are quite abundant, at present none of them can be regarded as sufficiently strong or well documented to be used as a test either in the diagnosis or in the prediction of the course of the disease.

Chapter References

Amor, B. *et al*. (1991). Evaluation des criteres des spondylarthropathies d'Amor et de l'European Spondylarthropathy Study Group (ESSG). Une etude transversale de 2228 patients. *Annales de Medicine Interne*, **142**, 85–9.

Arnold, L.J., Jr., Hammond, P.W., Wiese, W.A., and Nelson, N.C. (1989). Assay formats involving acridinium-ester-labeled DNA probes. *Clinical Chemistry*, **35**, 1588–94.

Barany, F. (1991). Genetic disease detection and DNA amplification using cloned thermostable ligase. *Proceedings of the National Academy of Sciences* (*USA*), **88**, 189–93.

Begovich, A.B., Bugawan, T.L., Nepom, B.S., Klitz, W., Nepom, G.T., and Erlich, H.A. (1989). A specific HLA-DP beta allele is associated with pauciarticular juvenile rheumatoid arthritis but not adult rheumatoid arthritis. *Proceedings of the National Academy of Sciences* (*USA*), **86**, 9489–93.

Benjamin, R. and Parham, P. (1990). Guilt by association: HLA-B27 and ankylosing spondylitis. (Review.) *Immunology Today*, **11**, 137–42.

Bidwell, J. (1994). Advances in DNA-based HLA-typing methods. *Immunology Today*, **15**, 303–7.

Bodmer, J.G. *et al*. (1994). Nomenclature for factors of the HLA system. *Tissue Antigens*, **11**, 1–18.

Brewerton, D.A., Hart, F.D., Nicholls, A., Caffrey, M., James, D.C., and Sturrock, R.D. (1973). Ankylosing spondylitis and HLA-27. *Lancet*, **i**, 904–7.

Brown, J.H. *et al*. (1993). Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. *Nature*, **364**, 33–9.

Charmley, P., Nepom, B.S., and Concannon, P. (1994). HLA and T cell receptor b-chain DNA polymorphism identify a distinct subset of patients with pauciarticular-onset juvenile rheumatoid arthritis. *Arthritis and Rheumatism*, **37**, 695–701.

Cohn, D.H. and Byers, P.H. (1990). Clinical screening for collagen defects in connective tissue diseases. (Review.) *Clinical Perinatology*, **17**, 793–809.

Conner, B.J., Reyes, A.A., Morin, C., Itakura, K., Teplitz, R.L., and Wallace, R.B. (1983). Detection of sickle cell beta S-globin allele by hybridization with synthetic oligonucleotides. *Proceedings of the National Academy of Sciences* (*USA*), **80**, 278–82.

Cooper, D.N. and Krawczak, M. (1993). *Human gene mutation*. BIOS Scientific Publishers, Oxford.

Cotton, R.G., Rodrigues, N.R., and Campbell, R.D. (1988). Reactivity of cytosine and thymine in single-base-pair mismatches with hydroxylamine and osmium tetroxide and its application to the study of mutations. *Proceedings of the National Academy of Sciences* (*USA*), **85**, 4397–401.

Cros, P., Allibert, P., Mandrand, B., Tiercy, J.M., and Mach, B. (1992). Oligonucleotide genotyping of HLA polymorphism on microtitre plates. *Lancet*, **340**, 870–3.

Davies, J.L. *et al*. (1994). A genome-wide search for human type 1 diabetes susceptibility genes. *Nature*, **371**, 130–6.

Dobloug, J.H., Forre, O., and Thorsby, E. (1979). HLA-DRw4 and rheumatoid arthritis. *Lancet*, **i**, 548–9.

Dominguez, O., Coto, E., Martinez-Naves, E., Choo, S.Y., and Lopez-Larrea, C. (1992). Molecular typing of HLA-B27 alleles. *Immunogenetics*, **36**, 277–82.

Embury, S.H. *et al*. (1987). Rapid prenatal diagnosis of sickle cell anemia by a new method of DNA analysis. *New England Journal of Medicine*, **316**, 656–61.

Erlich, H.A., Gelfand, D., and Sninsky, J.J. (1991). Recent advances in polymerase chain reaction. *Science*, **252**, 1643–51.

Falk, C.T. and Rubinstein, P. (1987). Haplotype relative risks: an easy reliable way to construct a proper control sample for risk calculations. *Annals of Human Genetics*, **51**, 227–33.

Fernandez-Vina, M., Fink, C., and Stastny, P. (1994). HLA associations in juvenile arthritis. (Review.) *Clinical and Experimental Rheumatology*, **12**, 205–14.

Fischer, S.G. and Lerman, L.S. (1979). Length-independent separation of DNA restriction fragments in two-dimensional gel electrophoresis. *Cell*, **16**, 191–200.

Ganguly, A. and Prockop, D.J. (1990). Detection of single-base mutations by reaction of DNA heteroduplexes with a water-soluble carbodiimide followed by primer extension: application to products from the polymerase chain reaction. *Nucleic Acids Research*, **18**, 3933–9.

Giannini, E.H. *et al*. (1991). Longitudinal analysis of HLA associated risks for iridocyclitis in juvenile rheumatoid arthritis. *Journal of Rheumatology*, **18**, 1394–7.

Gibbs, R.A., Nguyen, P.N., and Caskey, C.T. (1989). Detection of single DNA base differences by competitive oligonucleotide priming. *Nucleic Acids Research*, **17**, 2437–48.

Glass, D. *et al*. (1980). Early-onset pauciarticular juvenile rheumatoid arthritis associated with human leukocyte antigen-DRw5, iritis, and antinuclear antibody. *Journal of Clinical Investigation*, **66**, 426–9.

Gough, A. *et al*. (1994). Genetic typing of patients with inflammatory arthritis at presentation can be used to predict outcome. *Arthritis and Rheumatism*, **37**, 1166–70.

Gregersen, P.K., Silver, J., and Winchester, R.J. (1987). The shared epitope hypothesis. An approach to understanding the molecular genetics of susceptibility to rheumatoid arthritis. (Review.) *Arthritis and Rheumatism*, **30**, 1205–13.

Guesdon, J.L. (1992). Immunoenzymatic techniques applied to the specific detection of nucleic acids. A review. *Journal of Immunological Methods*, **150**, 33–49.

Guo, H.C. *et al*. (1993). Comparison of the P2 specificity pocket in three human histocompatibility antigens: HLA-A*6801, HLA-A*0201, and HLA-B*2705. *Proceedings of the National Academy of Sciences* (*USA*), **90**, 8053–7.

Hammer, R.E., Maika, S.D., Richardson, J.A., Tang, J.P., and Taurog, J.D. (1990). Spontaneous inflammatory disease in transgenic rats expressing HLA-B27 and human beta 2m: an animal model of HLA-B27-associated human disorders. *Cell*, **63**, 1099–112.

Hansen, J.A. and Nelson, J.L. (1990). Autoimmune diseases and HLA. *Critical Reviews in Immunology*, **10**, 307–28.

Hastbacka, J., de la Chapelle, A., Kaitila, I., Sistonen, P., Weaver, A., and Lander, E. (1992). Linkage disequilibrium mapping in isolated founder populations: diastrophic dysplasia in Finland. [Published erratum appears in *Nature Genetics* (1992), 2 (4), 343]. *Nature Genetics*, **2**, 204–11.

Hill, A.V., Allsopp, C.E., Kwiatkowski, D., Anstey, N.M., Greenwood, B.M., and McMichael, A.J. (1991). HLA class I typing by PCR: HLA-B27 and an African B27 subtype. *Lancet*, **337**, 640–2.

Holodniy, M., Kim, S., Katzenstein, D., Konrad, M., Groves, E., and Merigan, T.C. (1991). Inhibition of human immunodeficiency virus gene amplification by heparin. *Journal of Clinical Microbiology*, **29**, 676–9.

Ivanyi, P. (1992). Immunogenetic basis of the seronegative spondyloarthropathies. (Review.) *Current Opinion in Rheumatology*, **4**, 484–93.

Jeffreys, A.J., Wilson, V., Neumann, R., and Keyte, J. (1988). Amplification of human minisatellites by the polymerase chain reaction: towards DNA fingerprinting of single cells. *Nucleic Acids Research*, **16**, 10953–71.

Julier, C. *et al*. (1991). Insulin-IGF2 region on chromosome 11p encodes a gene implicated in HLA-DR4-dependent diabetes susceptibility. *Nature*, **354**, 155–9.

Kan, Y.W. and Dozy, A.M. (1978). Polymorphism of DNA sequence adjacent to human beta-globin structural gene: relationship to sickle mutation. *Proceedings of the National Academy of Sciences* (*USA*), **75**, 5631–5.

Khan, M.A. and Khan, M.K. (1982). Diagnostic value of HLA-B27 testing in ankylosing spondylitis and Reiter's syndrome. *Annals of Internal Medicine*, **96**, 70–6.

Landegren, U., Kaiser, R., Sanders, J., and Hood, L. (1988). A ligase-mediated gene detection technique. *Science*, **241**, 1077–80.

Lander, E.S. and Schork, N.J. (1994). Genetic dissesction of complex traits. (Review.) *Science*, **265**, 2037–48.

Maksymowych, W.P. *et al*. (1992). Polymorphism in a T-cell receptor variable gene is associated with susceptibility to a juvenile rheumatoid arthritis subset. *Immunogenetics*, **35**, 257–62.

Maksymowych, W.P. *et al*. (1994). Polymorphism in an HLA-linked proteosome gene influences phenotypic expression of disease in HLA-B27 individuals. *Journal of Rheumatology*, **21**, 665–9.

Maxam, A.M. and Gilbert, W.A. (1977). A new method of sequencing DNA. *Proceedings of the National Academy of Sciences* (*USA*), **74**, 560–4.

McDowell, T.L., Symons, J.A., Ploski, R., Forre, O., and Duff, G.W. (1994). A genetic association between juvenile rheumatoid arthritis and a novel interleukin 1 alpha polymorphism. *Arthritis and Rheumatism*, **38**, 221–8.

Melin-Aldana, H. *et al*. (1992). Human leukocyte antigen-DRB1*1104 in the chronic iridocyclitis of pauciarticular juvenile rheumatoid arthritis. *Journal of Pediatrics*, **121**, 56–60.

Mullis, K.B. and Faloona, F.A. (1987). Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods in Enzymology*, **155**, 335–50.

Myers, R.M., Larin, Z., and Maniatis, T. (1985). Detection of single base substitutions by ribonuclease cleavage at mismatches in RNA:DNA duplexes. *Science*, **230**, 1242–6.

Nepom, B.S., Nepom, G.T., Mickelson, E., Schaller, J.G., Antonelli, P., and Hansen, J.A. (1984). Specific HLA-DR4-associated histocompatibility molecules characterize patients with seropositive juvenile rheumatoid arthritis. *Journal of Clinical Investigation*, **74**, 287–91.

Newell, W.R., Trowsdale, J., and Beck, S. (1994). MHCDB–database of the human MHC. *Immunogenetics*, **40**, 109–15.

Newton, C.R. *et al*. (1989). Analysis of any point mutation in DNA. The amplification refractory mutation system (ARMS). *Nucleic Acids Research*, **17**, 2503–16.

Odum, N. *et al*. (1986). Increased frequency of HLA-DPw2 in pauciarticular onset juvenile chronic arthritis. *Tissue Antigens*, **28**, 245–50.

Orita, M., Iwahana, H., Kanazawa, H., Hayashi, K., and Sekiya, T. (1989). Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. *Proceedings of the National Academy of Sciences* (*USA*), **86**, 2766–70.

Ploski, R., Hansen, T., and Forre, O. (1993*a*). Lack of association with T-cell receptor TCRBV6S1*2 allele in HLA-DQA1*0101-positive Norwegian juvenile chronic arthritis patients. *Immunogenetics*, **38**, 444–5.

Ploski, R. *et al*. (1993*b*). HLA class II alleles and heterogeneity of juvenile rheumatoid arthritis. DRB1*0101 may define a novel subset of the disease. *Arthritis and Rheumatism*, **36**, 465–72.

Ploski, R., Flato, B., Vinje, O., Maksymowych, W., Førre, Ø., and Thorsby, E. (1994 *a*). Association to HLA-DRB1*08, HLA-DPB1*0301 and homozygosity for an HLA-linked proteasome gene in juvenile ankylosing spondylitis. *Human Immunology*, **44**, 88–96.

Ploski, R., Mellbye, O.J., Ronningen, K.S., Førre, Ø., and Thorsby, E. (1994 *b*). Seronegative and weakly seropositive rheumatoid arthritis differ from clearly seropositive rheumatoid arthritis in HLA class II associations. *Journal of Rheumatology*, **21**, 1397–402.

Ploski, R., Vinje, O., Maksymowych, W., Thorsby, E., and Førre, Ø. (1994 *c*). Immunogenetic difference between juvenile and adult ankylosing spondylitis (JAS and AS). *Scandinavian Journal of Rheumatology*, Suppl. **98**.

Rao, V.B. (1994). Direct sequencing of polymerase chain reaction-amplified DNA. (Review.) *Analytical Biochemistry*, **216**, 1–14.

Reekers, P., Schretlen, E.D.A.M., and van de Putte, L.B.A. (1983). Increase of HLA-'DRw6' in patients with juvenile chronic arthritis. *Tissue Antigens*, **22**, 283–8.

Robinson, W.P. *et al*. (1989). HLA-Bw60 increases susceptibility to ankylosing spondylitis in HLA-B27+ patients. *Arthritis and Rheumatism*, **32**, 1135–41.

Rubin, L.A. *et al*. (1994). Investigating the genetic basis for ankylosing spondylitis. Linkage studies with the major histocompatibility complex region. *Arthritis and Rheumatism*, **37**, 1212–20.

Saiki, R.K., Bugawan, T.L., Horn, G.T., Mullis, K.B., and Erlich, H.A. (1986). Analysis of enzymatically amplified beta-globin and HLA-DQ alpha DNA with allele-specific oligonucleotide probes. *Nature*, **324**, 163–6.

Saiki, R.K., Walsh, P.S., Levenson, C.H., and Erlich, H.A. (1989). Genetic analysis of amplified DNA with immobilized sequence-specific oligonucleotide probes. *Proceedings of the National Academy of Sciences* (*USA*), **86**, 6230–4.

Sanger, F., Nicklen, S., and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences* (*USA*), **74**, 5463–7.

Schlosstein, L., Terasaki, P.I., Bluestone, R., and Pearson, C.M. (1973). High association of an HLA antigen, W27, with ankylosing spondylitis. *New England Journal of Medicine*, **288**, 704–6.

Sheffield, V.C., Cox, D.R., Lerman, L.S., and Myers, R.M. (1989). Attachment of a 40-base-pair G + C-rich sequence (GC-clamp) to genomic DNA fragments by the polymerase chain reaction results in improved detection of single-base changes. *Proceedings of the National Academy of Sciences* (*USA*), **86**, 232–6.

Sheffield, V.C., Beck, J.S., Kwitek, A.E., Sandstrom, D.W., and Stone, E.M. (1993). The sensitivity of single-strand conformation polymorphism analysis for the detection of single base substitutions. *Genomics*, **16**, 325–32.

Southern, E.M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology*, **98**, 503–17.

Stastny, P. (1978). Association of the B-cell alloantigen DRw4 with rheumatoid arthritis. *New England Journal of Medicine*, **298**, 869–71.

Stastny, P. and Fink, C.W. (1979). Different HLA-D associations in adult and juvenile rheumatoid arthritis. *Journal of Clinical Investigation*, **63**, 124–30.

Sykes, B. (1993). Linkage analysis in dominantly inherited osteogenesis imperfecta. (Review.) *American Journal of Medical Genetics*, **45**, 212–6.

Thompson, E.M. (1993). Non-invasive prenatal diagnosis of osteogenesis imperfecta. (Review.) *American Journal of Medical Genetics*, **45**, 201–6.

Tomfohrde, J. *et al*. (1994). Gene for familial psoriasis susceptibility mapped to the distal end of human chromosome 17q. *Science*, **264**, 1141–5.

Van Kerckhove, C. *et al*. (1990). HLA-DP/DR interaction in children with juvenile rheumatoid arthritis. *Immunogenetics*, **32**, 364–8.

Van Kerckhove, C. *et al*. (1991). HLA-DQA1*0101 haplotypes and disease outcome in early onset pauciarticular juvenile rheumatoid arthritis. *Journal of Rheumatology*, **18**, 874–9.

Vartdal, F. *et al*. (1986). HLA class I and II typing using cells positively selected from blood by immunomagnetic isolation–a fast and reliable technique. *Tissue Antigens*, **28**, 301–12.

Wallace, R.B., Shaffer, J., Murphy, R.F., Bonner, J., Hirose, T., and Itakura, K. (1979). Hybridization of synthetic oligodeoxyribonucleotides to phi chi 174 DNA: the effect of single base pair mismatch. *Nucleic Acids Research*, **6**, 3543–57.

Weber, J.L. and May, P.E. (1989). Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *American Journal of Human Genetics*, **44**, 388–96.

Weissenbach, J. *et al*. (1992). A second-generation linkage map of the human genome. *Nature*, **359**, 794–801.

Weyand, C.M., Xie, C., and Goronzy, J.J. (1992). Homozygosity for the HLA-DRB1 allele selects for extraarticular manifestations in rheumatoid arthritis. *Journal of Clinical Investigation*, **89**, 2033–9.

White, R. and Lalouel, J. (1987). Investigation of genetic linkage in human families. (Review.) *Advances in Human Genetics*, **16**, 121–228.

Winchester, R. (1994). The molecular basis of susceptibility to rheumatoid arthritis. *Advances in Immunology*, **56**, 289–466.

Woolf, B. (1956). On estimating the relation between blood group antigens and disease. *Annals of Human Genetics*, **20**, 309–11.

Wordsworth, P. *et al*. (1992). HLA heterozygosity contributes to susceptibility to rheumatoid arthritis. *American Journal of Human Genetics*, **51**, 585–91.

Wu, D.Y. and Wallace, R.B. (1989). The ligation amplification reaction (LAR)–amplification of specific DNA sequences using sequential rounds of template-dependent ligation. *Genomics*, **4**, 560–9.

Wu, D.Y., Ugozzoli, L., Pal, B.K., and Wallace, R.B. (1989). Allele-specific enzymatic amplification of beta-globin genomic DNA for diagnosis of sickle cell anemia. *Proceedings of the National Academy of Sciences* (*USA*), **86**, 2757–60.

Ziegle, J.S. *et al*. (1992). Application of automated DNA sizing technology for genotyping microsatellite loci. *Genomics*, **14**, 1026–31.

4.8 Immune complex detection

Bryan D. Williams

Introduction

Immune complexes are formed when antibody interacts with the corresponding antigen(s); the antigen(s) may be tissue fixed or free in serum and other body fluids. This is a dynamic process and their clearance, which depends on many variables, normally plays a physiological role which is beneficial to the host and leads to the elimination of antigen. However, under certain circumstances immune complexes may persist in the circulation and subsequently localize in target tissue where they may elicit an inflammatory response. The basement membranes of blood vessels, glomeruli, the choroid plexus, and other organs are common locations for immune complex deposition.

The clinical setting in which circulating immune complexes deposit at multiple tissue sites and produce lesions is referred to as an immune complex disease. These diseases share certain common pathogenic mechanisms although the underlying causes may vary due to the different origins and nature of the antigens present in the immune complexes. The immune complex diseases have common pathological mechanisms and their clinical manifestations include glomerulonephritis, vasculitis, arthritis, skin eruptions, pleuritis, and pericarditis. Such multiple organ involvement is frequent in disorders that result from deposition of immune complexes from the circulation. Diseases resulting from local immune complex formation are associated usually with involvement of a single organ such as the kidney or thyroid gland. Considerable variations in the expression of disease features are seen between patients and in a given patient during the course of the disease.

Ideally to incriminate circulating immune complexes directly in the pathogenesis of human diseases certain criteria should be fulfilled. These would include the following.

1. Identification of the antigen and its binding to a specific antibody within the immune complex.
2. Components of the circulating immune complex should be identified within involved tissues which show histological features of vasculitis or glomerulonephritis.
3. If the antibody is of the appropriate isotype, evidence of hypercatabolism of complement should occur.
4. Circulating immune complex levels should correlate with disease features.

Few diseases in humans fulfil all of these criteria. However, it is widely believed that immune complexes are involved directly in the pathogenesis of several of the rheumatic diseases and contribute to the development and persistence of inflammation in blood vessels and other tissues.

Pathogenicity of immune complexes

Evidence of the pathogenicity of immune complexes has been well reviewed (Van Es *et al*. 1984; Salvido and Andres 1986; Virella 1990; Virella 1993). Characteristics of the immune complexes which predispose to their tissue deposition and pathogenicity are summarized in Table 1. Variation in the exposure time and the nature and persistence of the antigen significantly changes the biological properties of the immune complex formed. DNA of high molecular weight is removed very quickly from the circulation by the liver and as a consequence immune complexes containing such DNA are also eliminated rapidly from the circulation. Glycoprotein antigens containing exposed galactose residues may hasten the removal of small lattice immune complexes by interacting with galactose receptors on hepatocytes, and cationic (positively charged) antigens can interact directly with negatively charged tissue thereby contributing to local immune complex formation.

Size
Valency
Combining ratios
Affinity
Immunoglobulin class
Physicochemical nature of the antigen

Table 1 Characteristics of immune complexes which are relevant to pathogenicity

Antibodies are the other essential constituent of immune complexes and they also influence the pathogenicity of antigen–antibody complexes. They achieve this by variations in isotype as well as antibody avidity. Isotype influences the glomerular deposition of DNA–anti-DNA complexes in NZB/NZW mice, a strain which develops

a lupus-like syndrome. The switch from IgM to IgG isotype occurs much earlier in the female than in male mice and coincides with the onset of clinical disease. IgG anti-DNA, rather than IgM anti-DNA, is associated with glomerular deposits and the isotype switch, therefore, favours the formation and glomerular deposition of pathogenic immune complexes. The production of antibodies of low avidity, another characteristic of NZB/NBW mice, also interferes with the effective elimination of antigen and predisposes to the formation of circulating immune complexes which subsequently deposit in tissues. Other factors which influence the formation and removal of immune complexes are summarized in Table 2. The complement system also plays an important physiological role in preventing the tissue deposition of immune complexes in normal individuals. The ability of the complement system to prevent immune complexes from precipitating from plasma and the role of the CR1 receptor on the erythrocyte membrane are two crucial components of this process (see Chapter 3.1). Only when defects occur in this system will immune complex deposition occur and an immune complex disease be produced (Inman and Day 1981).

Event	Variables
Exposure to antigen	Single event or sustained exposure Nature of the antigen Route of exposure Quantity
Antibody production	Previous exposure to antigen Genetically determined host immune response Affinity and isotype Quantity
Immune complex formation	Circulating or in situ
Fate of immune complexes	Modulation of lattice size by complement activation Inherited or acquired complement deficiencies Inhibition complement activation by drugs
Clearance	Capacity of the reticuloendothelial system Interaction with complement or Fc receptors
Deposition in tissue	Binding to erythrocyte CR1 receptors Blood flow and vessel structure Changes in permeability Affinity of the antigen for tissue components Prior tissue damage Pharmacological agents

Table 2 Dynamics of immune complex disease

Considerable research has been directed towards the development of specific, sensitive, and reliable techniques for the demonstration of immune complexes because of their pathogenic role in certain diseases. A large number of methods are applicable to the detection of complexes involving unknown antigens, but these are essentially based on either the physical or biological properties of the immune complexes. Historically the development of assays for soluble immune complexes started with physicochemical rather than immunological techniques, although currently nearly all the assays used in clinical practice are based on the latter methods. Techniques currently in use can be conveniently divided into two main groups, those that are antigen specific and a much larger, although more readily applicable group, the antigen non-specific tests. Not all of these assays, however, are suitable for routine clinical use. The ideal test should be both specific for the immune complex and sensitive for a range of immune complex characteristics, be simple to perform, use stable and easily stored reagents, require small quantities of serum, and allow the quantity of the immune complex to be estimated accurately and reproducibly.

Antigen-specific methods

The provocative antigen in a pathogenic immune complex is rarely known but when it is, either the antigen or its corresponding antibody can be detected in the circulation. Virus–antibody complexes have been visualized by electron microscopy in patients with hepatitis associated with the B antigen and in those patients who have developed polyarteritis nodosa; hepatitis B virus particles have been shown to be part of the circulating immune complex (Almeida and Waterson 1969). The presence of virus–antibody complexes in the blood of animals chronically infected with the lymphocytic choriomeningitis (**LCM**) virus had been inferred following a decrease in the infectivity of serum after the virus–antibody complex was precipitated by an antiglobulin antibody.

In the sera of patients with systemic lupus erythematosus, free DNA has been detected during phases of disease activity, often alternating in sequential longitudinal studies with the appearance of anti-DNA antibody (Koffler *et al*. 1974). Others have claimed that DNA–anti-DNA complexes are present in the serum of patients with active systemic lupus erythematosus since the concentration of anti-DNA antibody increases following digestion of the complex by DNAse (Harbeck *et al*. 1973). In other human diseases specific antibody has been identified by immunofluorescence in the granular immune complex deposits which accumulate in the renal glomerulus (Williams 1981). Methods of isolating and analysing the constituents present in immune complexes have been reviewed by Jones and Orlans (1981).

Antigen non-specific methods

Since the antigens in human sera in the majority of the immune complex diseases are unknown, the presence of immune complexes are identified most frequently on the basis of different physical, chemical, and biological properties of antigen-complexed IgG compared with free IgG. Most immune complex assays, therefore, employ molecular or cellular reagents that react preferentially with antibody when it forms part of an immune complex and less readily with free monomeric antibody. A problem associated with all of these methods is the possibility that non-specifically aggregated immunoglobulin may be detected and interpreted as reflecting the presence of immune complexes *in vivo*. Over 40 assays have been developed to detect immune complexes and these have varying degrees of specificity, sensitivity, and reproducibility; factors which, therefore, preclude the detection of all species of immune complexes with a single assay (Theofilopoulos and Dixon 1979).

Methods based on the physical characteristics of immune complexes

These techniques, which are now largely of historical interest only, are based on the increased molecular size or altered solubility of the immune complex compared with free antigen or antibodies. The methods used, summarized in Table 3, include ultracentrifugation, gel filtration, sucrose density gradient centrifugation, polyethylene glycol precipitation, and cryoprecipitation.

Table 3 Methods for detecting immune complexes

Ultracentrifugation

The use of analytical and preparative ultracentifugation has been discussed (Stanworth and Johns 1977). The important contribution of this technique to the demonstration of circulating immune complexes containing rheumatoid factor and to the understanding of the handling of radiolabelled antigen in animals in undeniable. However, its application is largely as a research tool and because of its low sensitivity it has little clinical use.

Gel filtration

Gel filtration allows macromolecular material of defined size and, therefore, distinguishable from free IgG and IgM to be separated and analysed for the characteristics

of an immune complex. A sensitive detection system for complement component C3, bound to serum fractions larger than IgG and free C3 has been used to identify immune complexes in the sera of patients with rheumatoid arthritis and other diseases. Although very sensitive this method is not readily adaptable for extensive clinical use (Williams and Slaney 1977).

Polyethylene glycol precipitation

Polyethylene glycol (**PEG**), an uncharged, water soluble, linear polymer, has unusual properties that are useful in detecting immune complexes. Twenty per cent PEG precipitates most monomeric IgG and many other proteins from serum but when its concentration is reduced to less than 3 per cent the precipitation of monomeric IgG decreases significantly without preventing the precipitation of immune complexes. Precipitation of the latter depends on the steric exclusion of the immune complex from the domain of the polymer. This technique has been used to evaluate the presence of immune complexes in the sera of patients with a variety of autoimmune rheumatic disorders on the assumption that IgG complexed with antigen precipitates more readily than monomeric IgG. Two main problems associated with this technique are its lack of specificity and the results obtained are influenced by the abnormally high concentrations of IgG present in the sera of patients with autoimmune rheumatic diseases. Furthermore the addition of PEG to serum also induces the aggregation of immune complexes (Creighton *et al*. 1973).

Cryoprecipitation

Cryoprecipitation refers to the reversible precipitation of proteins or protein complexes at reduced temperatures. These cryoprecipitates (cryoglobulins) frequently contain IgG and they have been classified into three types: type I composed of monoclonal IgG; type II composed of mixed cryoglobulins with a monoclonal component possessing antibody activity against polyclonal IgG; and type III containing mixed polyclonal cryoglobulins in which none of the components are monoclonal. Type III cryoglobulins occur in the sera of patients with a variety of autoimmune rheumatic disorders and their measurement is often useful in patient evaluation. Since cryoglobulins will precipitate from serum at different temperatures, samples should be either transferred to the laboratory within minutes or if this is not possible transferred in a thermos flask containing water at 37°C. After clotting at 37°C, serum is then stored at 4°C for up to 7 days before isolating the cryoprecipitate by centrifugation. After repeated washing the protein concentration is determined and expressed per millilitre of serum. The largest quantities of cryoglobulin are found with type 1 or type 11. Cryoglobulins containing IgM rheumatoid factor and IgG rheumatoid factor are present in sera from patients with rheumatoid arthritis, particularly in those with extra-articular disease (Schreiber and Maini 1984). In patients with severe rheumatoid arthritis, cryoglobulinaemia is common and frequently associated with nodules, digital vasculitis, leg ulcers, lung disease, and scleritis. There is a strong positive correlation between the presence of extra-articular disease and cryoglobulinaemia, and therefore the detection of cryoglobulin is useful in identifying a group of patients with a high morbidity and mortality. Recently type II cryoglobulinaemia has been strongly associated with concomitant hepatitis C virus infection and a high rate of false-negative serological tests. Hepatitis C virions and hepatitis C antibody are concentrated in the cryoprecipitate, most commonly in association with a rheumatoid factor of a specific idiotype (Agnello *et al*. 1992).

Identification of immune complexes by interaction with biological recognition units

Assays for the detection of immune complex can be divided into three major categories and the principles underlying each of these are shown in Table 4. In the solid phase assays the recognition unit is coated onto a plastic surface, this then interacts with the sample and the interaction is demonstrated finally by the binding of radiolabelled anti-immunoglobulins. In fluid phase assays the recognition unit is incubated with the sample and the bound and free units are separated by differential precipitation. In the cellular assays, cells bearing the appropriate receptors are incubated with the sample, the complexes attach to the cell surface receptors and are quantified by the addition of radiolabelled anti-immunoglobulins (reviewed by Kilpatrick and Weston 1985).

	Recognition	Interaction	Quantification
Solid phase binding assays	Coating plate with receptor molecule	Serum immune complex with coated receptor	Develop with labelled anti-immunoglobulin; express as amount of antibody bound to plate
Fluid phase binding assays	Serum immune complex with labelled receptor molecules	Separation of bound from free receptor using a precipitation step	Percentage of precipitated receptor
Cellular binding assays	Cells bearing the appropriate receptor	Serum immune complex interacts with cell receptor	Develop with labelled anti-immunoglobulin; express as amount of antibody bound to the cell

Table 4 Principles underlying immune complex assays

Most immune complex assays employ biological reagents and use molecular or cellular probes which react with antibody in the immune complex but not with free monomeric antibody. The reaction of immune complex with complement components afford several methods for immune complex detection. Complement-fixing immune complexes may be detected either by indirect measurement of complement consumption (e.g. anticomplementary activity) or by direct evaluation of the binding of immune complex to C1q, the first component of complement (adapted to produce to C1q-binding fluid phase or solid phase assays). Immune complexes that have reacted with the complement system have molecules of C3 deposited on their surface, such complexes can be identified in those assays which utilize conglutinin (e.g. conglutinin-binding assay) or antibodies to fixed complement components (anti-C3 assays). All these methods have in common a dependence on the reactivity of the immune complex with complement, although many circulating complexes fulfil this requirement those containing only IgG4, IgA, or IgE would not be detected.

A second group of assays that discriminate aggregated from monomeric immunoglobulins is based on the interaction of immune complexes with low avidity anti-immunoglobulins antibodies. Most of these methods are sensitive only for immune complexes containing IgG antibodies and all are influenced to varying degrees by the concentration of monomeric IgG. A third group utilizes the interaction of immune complexes with receptors present on cell membrane. The use of living cells makes standardization difficult, however, and the results may be influenced by antibodies present in the patient's serum which either bind directly to the cell membrane or interfere with cell metabolism. One lymphoblastoid cell line, known as the Raji cell line, is particularly suitable for detecting immune complexes because of the high density of complement receptors on its surface and the lack of surface immunoglobulin, coupled with small numbers of low affinity receptors for IgG Fc.

Methods commonly used for the detection of immune complexes in clinical samples

The five assays outlined in Fig. 1 have been found to be the most acceptable for clinical use since they are sensitive for a variety of immune complexes, relatively simple to perform, reproducible, and give a quantitative estimation of the concentration of immune complexes in serum (WHO/IUIS Working Group 1981).

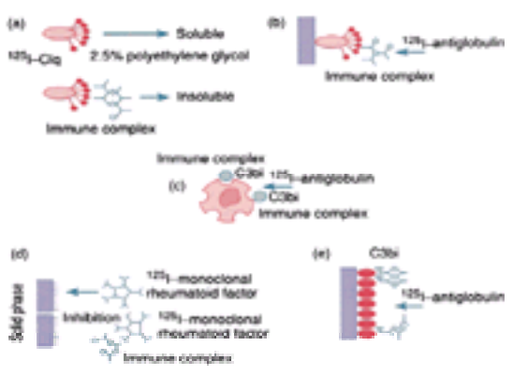


Fig. 1 Five methods found most acceptable by the WHO/IUIS Working Group (1981). (a) C1q binding assay. (b) C1q solid phase assay. (c) Immune-complex Raji cell

assay. (d) Monoclonal rheumatoid factor inhibition assay. (e) Solid phase conglutinin binding assay.

C1q molecules bind weakly to monomeric IgG1, IgG2, IgG3, and IgM, but this binding is much enhanced when these protein complex with antigen. However, C1q assays fail to detect immune complexes made with non-complement activating antibodies (IgG4) and also fail to detect complexes made with antibodies that activate complement preferentially by the alternative pathway. Whichever assays are used the results must be interpreted with a knowledge that various polyanions, DNA, endotoxin, heparin, and viruses may bind directly to C1q thereby reducing its specificity as a probe for immune complex detection.

C1q binding assay

This assay depends on the differential precipitation by PEG of free radiolabelled C1q from the C1q bound to the immune complex. In a modified assay, pathological sera are treated with **EDTA** (ethylene diamine tetra-acetic (ethanoic) acid) in order to prevent the incorporation of radiolabelled C1q into the endogenous C1qrs complex. The percentage of radioactivity precipitated corresponds to the C1q binding activity of the sample and provides a measure of the concentration of immune complexes present in the sample. PEG also favours the binding of C1q to immune complexes and once bound prevents their dissociation. This assay has advantages, therefore, in detecting low avidity immune complexes. The radiolabelled C1q is only present in small quantities but in the presence of EDTA there is competition between endogenous C1q and the labelled C1q for binding to the immune complex. This method has advantages over other C1q tests in that no positive results arise from the presence of DNA or endotoxin in the sample, presumably because complexes of these substances with C1q are not precipitable by the concentration of PEG employed. Free DNA, however, strongly inhibits binding of C1q to aggregated immunoglobulin. In addition there is no assurance that labelled C1q binds all complement-fixing immune complexes if, for example, the IgG sites are fully occupied with endogenous C1q. This assay is also liable to interference by binding of C1q to the anticoagulant heparin (Zubler *et al*. 1976a).

C1q solid phase assay

This test is based on the interaction of immune complexes or aggregates with C1q bound to polystyrene tubes. In these assays pathological sera are incubated in tubes internally coated with C1q; thereafter the amount of bound IgG is assessed with radiolabelled or enzyme-linked anti-IgG antisera. These tests are very sensitive and have the distinct advantage in that they are IgG specific. Since the developing antiserum is one directed against IgG, false-positive results from DNA, endotoxin, or other interfering substances are avoided (Hay *et al*. 1976).

Raji cell binding assay

This assay is based on the binding of complement-activating immune complexes to complement receptors on an lymphoblastoid cell line (Raji cells). These cells have large number of receptors for C3b but lack surface immunoglobin and have low numbers of low affinity receptors for the Fc portion of IgG. Immune complexes bound to the cells are quantified using a labelled anti-immunoglobulin antibody, the concentration of immune complexes being calculated by extrapolation from a standard curve prepared by adding increasing amounts of aggregated immunoglobin to normal human serum which provides a source of complement. This assay is sensitive but very high immune complex levels are recorded spuriously when antilymphocyte antibodies, found frequently in sera from patients with systemic lupus erythematosus, bind to the cell membrane.

Rheumatoid factor inhibition assay

This assay is based on the inhibition by immune complexes of the binding of radiolabelled rheumatoid factor to plastic wells which have been coated with immunoglobulin. The degree of inhibition achieved with the test sample is compared with that obtained by dilutions of a standard preparation of aggregated IgG. This assay is very sensitive, can detect IgG complexes as small as 8S, and is independent of the complement fixing properties of the immune complex. Disadvantages associated with this technique are the fact that it is influenced by high concentrations of monomeric IgG, it can only detect those immune complexes which contain IgG, and the presence of rheumatoid factor in the test sample may produce a false-positive result (Luthra *et al*. 1975).

Conglutinin binding assay

The binding of conglutinin to immune complexes is complement and calcium dependent and is specific for the inactivated form of C3, C3bi. Conversion of immune complex bound C3b to C3bi occurs very rapidly *in vivo* and a subsequent degradation of the C3bi molecule to C3c and C3d leads to loss of binding of the immune complex with conglutinin. The assay is performed by placing the test sera in plastic plates which have been coated with bovine conglutinin and determining the amount of bound IgG with a radiolabelled anti-IgG antibody. Aggregated human immunoglobulin incubated with normal human serum is used as a standard. The advantages of this assay are its specificity, the relatively simple technology involved, and the stability of the conglutinin at 4°C. Disadvantages are its preferential reactivity with large immune complexes and, given the specificity of the conglutinin, reactivity with immune complexes is lost following the degradation of the C3bi *in vivo* (Casali *et al*. 1977).

The sensitivity, the causes of false-positive and false-negative results, and the advantages, disadvantages, and main areas of clinical use of these assays are summarized in Table 5 and Table 6.

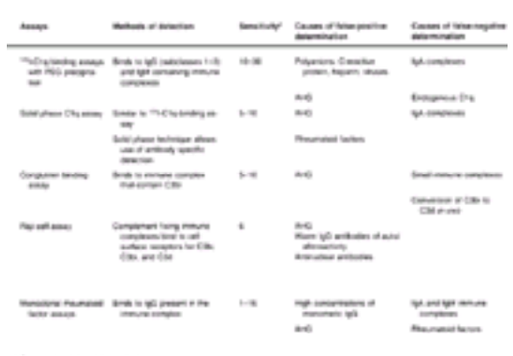


Table 5 Commonly used assays for detecting immune complexes

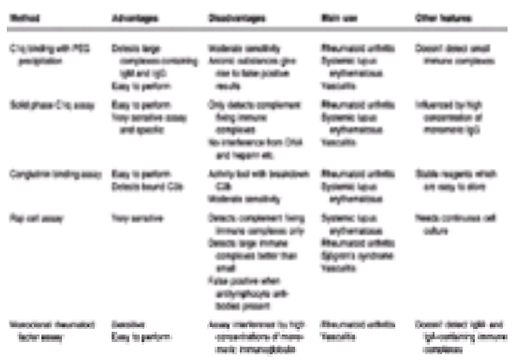


Table 6 Commonly used assays for detecting immune complexes

Problems associated with the use of antigen non-specific methods for detecting immune complexes

Specificity

There is no reagent which is capable of detecting all types of immune complexes. C1q reacts mainly with the complement-fixing IgG and IgM complexes whereas rheumatoid factor reacts only with IgG complexes, both complement fixing and non-complement fixing. Conglutinin and the Raji cells react only with complement-fixing immune complexes. In some instances the reactivity of the reagent may be transient, immune complex bound C3b is cleaved to produce C3bi thereby conferring reactivity with conglutinin, further degredation, however, results in the cleavage to C3c and C3d with subsequent loss of conglutinin binding activity. In general reagents that react with complement-fixing immune complexes will react more effectively with the large immune complexes that are more efficient activators of the complement system. Rheumatoid factor will react with immune complex as small as 8S.

Interfering substances

None of these methods differentiate antigen–antibody complexes from non-specifically aggregated immunoglobulin generated during sample collection or storage. A general problem with the use of biological reagents is that they have unknown or uncharacterized reactivities that give rise to false-positive results. C1q reacts with polyanions such as DNA, endotoxin, and heparin whilst some monoclonal rheumatoid factors used in the monoclonal rheumatoid factor binding assays cross-react with DNA. These assays are also subject to interference by endogenous rheumatoid factors that are present frequently in the sera of patients with autoimmune rheumatic disorders. The Raji cell assay reacts with both antilymphocyte antibodies and with certain antinuclear antibodies, the latter reaction being due to the presence of small quantities of DNA on the cell membrane. Assays which employ polyclonal rheumatoid factors are also prone to interference by the rheumatoid factors frequently present in the patients' sera.

A number of substances present in normal sera can also interfere with the immune complex assays. The greater reactivity of both C1q and rheumatoid factor reagents with aggregated IgG compared with monomeric IgG has been shown to be relative when quantitative assays are used. High concentrations of monomeric IgG will, therefore, interfere to varying degrees with several assays, and in EDTA-treated sera endogenous C1q interferes with the radiolabelled C1q in the C1q binding assay.

Standardization

Another general problem associated with these assays is that they are only semiquantitative; in addition to quantity, the size range of the immune complex will determine the extent of its reactivity. Therefore, expressing the results of a reference curve gives only relative results which differ with each assay and standard curve. The use of heat-aggregated IgG as a standard which is then used to quantify immune complex in pathological sera introduces additional variability into comparative studies of different assays. Since heat-aggregated immunoglobulin contains a heterogeneous population of aggregates, some of these will be detected preferentially by some assays. The variation in the shape of the standard curves obtained with a single heat-aggregated preparation renders assessment of the concentration of immune complexes in pathological sera somewhat imprecise. Additional practical problems involve the availability and stability of the biological reagents used, in particular the Raji cell assay requires the maintenance of living cell lines.

Manipulation of samples

In order to minimize the possibility of introducing artefacts, serum should be allowed to clot for 2 h without cooling in order to prevent the cryoprecipitation of some immune complexs. Samples ideally should be kept frozen, if possible at –70°C, and repeated freezing and thawing should be avoided. Use of heparinized plasma samples should be avoided.

Removing rheumatoid factor reactivity by immunoabsorbance by preincubation of the sample with IgG is to be avoided since it may partially deplete immune complexes or change the immune complex structure. For all tests performed on living cells toxic preservatives should be avoided.

Finally, because all antigen non-specific methods for detecting immune complexes are indirect, controlled experiments are necessary whenever immune complexes are demonstrated in new clinical disease. These should include the characterization of the size of the immune complex; fractionation in physical or chemical conditions known to dissociate immune complexes, thereby decreasing their size and biological activity; and confirmation that the assay becomes negative after removal of immunoglobulins from the test sample or after treatment known to alter the biological properties of immune complexes (e.g. reduction and alkylation).

Immune complexes in the rheumatic diseases

For the majority of the diseases listed in Table 7, many do not have the typical clinical manifestations which are associated with the tissue deposition of immune complexes. Most studies have simply demonstrated a positive result in one or more of the antigen non-specific immune complex assays. There have been, however, more detailed studies on systemic lupus erythematosus, rheumatoid arthritis, and the rheumatic syndromes associated with hepatitis B antigen. In addition some notable findings have emerged from more limited studies in patients with Lyme arthritis and vasculitis. Several comprehensive reviews of this subject are available (Williams 1981; Van Es *et al.* 1984; Virella 1990).

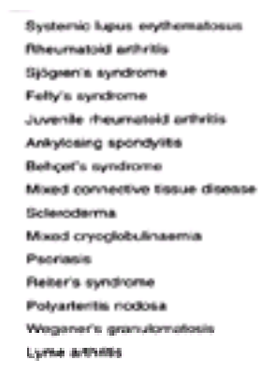


Table 7 Rheumatic diseases in which circulating immune complexes have been detected

Rheumatoid arthritis

Rheumatoid arthritis is a systemic disease characterized by a symmetrical erosive polyarthritis associated with the presence of IgM and IgG rheumatoid factor in the serum. The main antigenic site for activity of rheumatoid factor is the Fc portion of IgG, these molecules being able to form immune complexes in serum and synovial fluid either by self association or by interacting with monomeric IgG. These immune complexes are involved in the pathogenesis of the synovial lesions in rheumatoid arthritis. Synovial fluid from patients with rheumatoid arthritis frequently contains immune complexes composed of IgG–IgM rheumatoid factor and IgG–IgG rheumatoid factor which reacts strongly with C1q and IgM in the rheumatoid factor binding assay. The high molecular-weight complexes consist mainly of self-associating IgG rheumatoid factor, the intermediate-size immune complexes of monomeric IgG and IgG rheumatoid factor, while the 22S complexes demonstrated in serum consist of IgM rheumatoid factor bound to IgG.

Cryoglobulins containing either IgM or IgG rheumatoid factor are present in serum from patients with rheumatoid arthritis, particularly those with extra-articular disease. Levels are elevated in patients with severe rheumatoid arthritis and the frequency of cryoglobulinaemia exceeds 70 per cent in patients with extra-articular disease features which include nodules, digital vasculitis, rash, neuropathy, lung disease, and scleritis. Cryoglobulins are absent in patients with joint disease alone.

Cryoglobulinaemia is therefore a useful correlate of extra-articular disease and serves to define a group of patients in whom the morbidity and mortality are higher.

A variety of immune complex assays have been evaluated in a multicentre study, and of these six appear particularly effective in detecting the immune complexes present in sera from patients with rheumatic disease. These differences are highlighted in Table 8, which shows considerable variations in the degree of positivity recorded with the various assays in rheumatoid sera (Lambert *et al*. 1978). The fluid phase C1q binding assay and the Raji cell assay detect immune complexes much more frequently than the solid phase C1q, the monoclonal rheumatoid factor binding, and the conglutin binding assays. The lack of concordance between the three assays may be explained on the basis that the reactivity of rheumatoid sera is dependent on the function of both aggregate size and concentration. IgG aggregates of a defined size behave differently upon dilution in the three assays whereas IgG aggregates of similar size behave differently upon dilution in any one particular assay. The C1q binding, the monoclonal rheumatoid factor, and the Raji cell assays exhibit a high prevalence of positive results in both the synovial fluid and serum obtained from patients with rheumatoid arthritis. All studies which involve a comparison between serum and synovial fluid levels show higher concentrations of immune complexes in synovial fluid (Zubler *et al*. 1976b).

	[125]I-C1q	C1qSP	KGBSP	Raji	MMRI
Systemic Lupus erythematosus, percentage abnormal	58	58	32	85	5
Rheumatoid arthritis, percentage abnormal	82	25	33	73	30

[125]I-C1q, C1q binding in polyethylene glycol; C1qSP, C1q solid phase; KGBSP, conglutinin binding solid phase; Raji, Raji cell assay; MMRI, inhibition of fluid phase monoclonal rheumatoid factor binding to IgG aggregates.
Data obtained from WHO collaborative study (WHO/IUIS Working Group 1981).

Table 8 Results of different immune complex tests in systemic lupus erythematosus and rheumatoid arthritis

The immune complex assay using monoclonal rheumatoid factor has shown a positive correlation between the presence of immune complexes and the anatomical stage and functional class of the disease (Luthra *et al*. 1975); C1q binding activity however correlates best with extra-articular disease features rather than the clinical stage of the disease (McDougal *et al*. 1982; McDougal and McDuffie 1991). Attempts to evaluate serum C1q binding activity with disease activity have produced variable results, positive correlations with disease activity being exhibited in only some studies.

Detecting immune complexes in the serum may however be helpful in the following situations.

1. In early arthritis, circulating immune complexes may be detected in serum several months before a definite diagnosis of rheumatoid arthritis can be made (Jones *et al*. 1981).
2. Circulating immune complexes are found in 70 per cent of sera from patients with seronegative rheumatoid arthritis and, therefore, their presence may help to distinguish seronegative rheumatoid arthritis from other seronegative arthropathies, e.g. Reiter's disease.

Carefully conducted prospective clinical studies show that indices which reflect the activity of the arthritis correlate with the fluid phase C1q binding assay (Nydegger *et al*. 1977). However, the correlation coefficient obtained was weaker than that from measuring the erythrocyte sedimentation rate. Some studies in which specimens were obtained serially from individual patients showed a correlation between C1q positivity and either the number of swollen joints or a composite joint index (Halla *et al*. 1979); many studies, however, emphasize the weak correlation existing between various immune complex assays and the systemic features of rheumatoid arthritis. Some therapies such as corticosteroids, slow acting antirheumatic agents (Forster and McConkey 1986), and cryofiltration (Krakauer *et al*. 1983) produce improvement in the clinical state and a parallel decline in the materials reacting in the various immune complex assays.

Systemic lupus erythematosus

The nature of the circulating immune complexes in the sera of patients with systemic lupus erythematosus has not been well defined, although occasionally immune complexes have been isolated and shown to contain both DNA and anti-DNA antibodies. Increased levels of circulating immune complexes generally correspond with active disease, high levels of antidouble-stranded DNA antibody, and reduced complement levels (Morrow *et al*. 1982). In a comparative study the highest frequency of immune complexes in the sera from patients with systemic lupus erythematosus was found using the Raji cell binding assay (91 per cent), the fluid phase C1q binding test (78 per cent), and platelet aggregation (74 per cent) (Lambert *et al*. 1978).

Clinical studies addressing the relationship between immune complexes and disease activity in systemic lupus erythematosus have produced conflicting results (Cano *et al*. 1977; Levinski *et al*. 1977; Inman *et al*. 1980; Lloyd and Schur 1981). In systemic lupus erythematosus no particular clinical disease manifestation has been found to correlate with the presence of antigen non-specific immune complexes detected using the solid phase C1q binding assay. Patients with lupus nephritis had a higher level of reactivity with the solid phase C1q binding assay if their renal disease was of the diffuse proliferative type rather than membraneous nephritis (Wener *et al*. 1987).

In systemic lupus erythematosus much uncertainty remains whether immune complex assays reflect changes in the activity of the disease. Even for a single assay, the fluid phase C1q binding assay, positive (Inman *et al*. 1980) and negative (Hamburger *et al*. 1982) correlations have been observed between disease activity and immune complex levels. The monoclonal rheumatoid factor binding assay shows a positive correlation with activity of extrarenal disease features of systemic lupus erythematosus, but no correlation could be shown between immune complex levels and renal function as defined by changes in the serum creatinine (Lloyd and Schur 1981). The role of immune complexes in inducing renal injury has been reviewed by Gauthier and Abrass (1992).

In a well-conducted prospective study (Abrass *et al*. 1980) using the solid phase C1q binding assay, immune complex levels were correlated with disease activity and the presence of nephritis and arthritis. In this study, changes in solid phase C1q binding predicted a change in disease activity leading to admittance to hospital or a change in medication. Changes in fluid phase C1q binding activity, anti-DNA antibody, and C3 levels were not predictive, however. Attempts at grouping features of the disease into cycles of activity and correlating either renal or extrarenal disease with the fluid phase C1q binding were unsuccessful. These authors concluded that the determination of C4 and CH50 levels were superior to the C1q binding activity in reflecting the evolving clinical course in patients with systemic lupus erythematosus. High levels of immune complexes, detected using the Raji cell assay, predicted the later development of lupus nephritis (Boyd *et al*. 1983).

Much disagreement remains, therefore, between immune complex levels and disease activity in systemic lupus erythematosus (see Inman 1982 for a review). Likely explanations include the use of different methods of immune complex detection, the retrospective nature of the studies, and the differing criteria adopted for patient selection. The introduction of bias into the study whereby patients with milder disease are excluded is an important factor. At present it seems premature to base therapeutic decisions on measurements of immune complex levels alone.

Other autoimmune rheumatic diseases

Sjögren's syndrome, mixed connective tissue disease, and vasculitis

Serum from patients with Sjögren's syndrome frequently contain immune complexes and these are best demonstrated using the C1q binding and the Raji cell assay (Lawley *et al*. 1979; Fishbach *et al*. 1980). The immune complex levels do not correlate with disease activity or the development of lymphoid infiltrates in organs (Lawley *et al*. 1979).

In patients with mixed connective tissue disease, circulating immune complexes have been found using the C1q binding, the Raji cell, and the monoclonal rheumatoid factor inhibition assays; approximately 81 to 90 per cent of the sera examined being positive with two assays and 65 per cent positive with one assay (Halla *et al*.

1979).

In patients with vasculitis the most useful test is the monoclonal rheumatoid factor binding assay. In patients with cutaneous vasculitis, the presence of fluid phase C1q binding activity was found to a far greater extent in the group with necrotizing vasculitis. In polyarteritis nodosa, the first report highlighted that up to 30 per cent of patients had circulating immune complexes containing hepatitis B antigen both in the circulation and tissue (Eye *et al*. 1977). Subsequent studies have shown a much lower incidence of hepatitis B antigen and many of the tests for immune complexes have been negative. Correlations between changing immune complex assay levels and the activity of the vasculitis are few in number, but in patients with necrotizing vasculitis the C1q binding and monoclonal rheumatoid factor assays appeared to predict changes in disease activity (Mackel *et al*. 1979). Changes in the above assays also parallel disease activity in patients with polyarteritis nodosa and Wegener's granulomatosis; effective therapy substantially diminished the frequency of positivity in the Raji cell assay but not the fluid phase C1q binding assay (Ronco *et al*. 1983). Other factors, however, may also contribute to the pathogenesis of vasculitis (reviewed by Smiley and Moore 1989).

Scleroderma

Approximately half of the patients with scleroderma have circulating immune complexes; patients with progressive systemic sclerosis have elevated serum immune complex levels when detected by the Raji cell assay. A positive assay result was more often associated with diffuse scleroderma, tendon friction rub, and positive antinuclear antibodies. Individual with immune complexes detected by the C1q binding test had a higher incidence of pulmonary involvement and were more likely to have rheumatoid factor present in their serum (Seiboid *et al*. 1982).

Juvenile rheumatoid arthritis

In this group, circulating immune complexes have been detected particularly in those patients whose disease was marked by systemic onset or polyarticular disease (Rossen *et al*. 1977).

Lyme arthritis

This disorder is characterized by a distinct skin legion, myalgia, and monoarticular or oligoarticular arthritis which can be transient, recurrent, or persistent. Almost all patients with Lyme arthritis have abnormally elevated C1q binding assays, but in those patients with skin lesions alone the positivity disappears as the rash resolves. A persistent positive result in the C1q binding assay is associated with neurological or cardiac complications, and persistent synovitis is associated with a high titre of C1q binding reactant in the synovial fluid. This association of persistent positivity in the C1q assay with neurological or cardiac complications is clinically useful information that is not provided by any other laboratory test (Hardin *et al*. 1979).

Clinical usefulness

The majority of the assays are antigen specific and, therefore, their diagnostic value is limited since the presence of immune complexes is not specific to any particular disease.

A single measurement in a given patient does not provide a good indicator of the clinical activity of the disease in patients with either systemic lupus erythematosus or rheumatoid arthritis. However, in suspected but not definite rheumatoid arthritis, the absence of immune complexes may be of diagnostic value in ruling out this disease, provided that the test used is of high sensitivity. Despite much published work on correlations between immune complex levels and disease activity, considerable uncertainty remains of the clinical usefulness of these assays. Although correlations are found in patients with rheumatoid arthritis between clinical activity and serum measurements of immune complex levels, these correlations are often weak and are dependent on the assay used. Patients with more severe extra-articular disease do however have higher levels, often with cryoglobulinaemia, in their sera. Resolution of the disease activity following the introduction of therapy is usually associated with a progressive decrease in immune complex levels. The finding of a raised level of immune complexes in a given patient, however, does not have any predictive value as to the type of disease that patient will subsequently develop. In patients with systemic lupus erythematosus there does continue to be a role for some of the immune complex assays, such as the solid phase C1q binding assay in conjunction with measurements of complement and anti-DNA antibody levels in assessing and predicting disease activity.

There is no ideal assay to measure immune complex levels in sera and in most clinical situations the correlations between the different techniques are poor. It would be advantageous therefore to use several assays, preferably those that depend on different principles for the identification of complexes in the circulation. The failure to demonstrate significant advantages of these techniques over simple measures of disease activity, such as the erythrocyte sedimentation rate, has lead to a limitation of their use in clinical laboratory practice. If used, the results obtained must be interpreted cautiously because of the number of interfering substances that are present in the sera or may be generated *in vitro*.

Chapter References

Abrass, C.K., Nies, K.M., Louie, J.D., Border, W.A., and Glassock, R.J. (1980). Correlation and predictive accuracy of circulating immune complexes with disease activity in patients with systemic lupus erythematosus. *Arthritis and Rheumatism*, **23**, 273–82.

Agnello, V., Chung, R.T., and Kaplan, L.M. (1992). A role for hepatitis C virus infection in type II cryoglobulinemia. *New England Journal of Medicine*, **327**, 1490–5.

Almeida, J. and Waterson, A.P. (1969). Immune complexes in hepatitis. *Lancet*, **ii**, 983–6.

Boyd, R.D., Birchmore, D.A., Kaiser, D.L., Young, A.C., and Davis, J.S.I. (1983). Acute effects of steroids on immune complex profile of patients with systemic lupus erythematosus. *Arthritis and Rheumatism*, **26**, 637–44.

Cano, P.O., Jerry, L.M., Sladowski, J.P., and Osterland, C.K. (1977). Circulating immune complexes in systemic lupus erythematosus. *Clinical and Experimental Immunology*, **29**, 197–204.

Casali, P., Bossus, A., Carpentier, N.A, Lambert, P.H., and Miescher, P.A. (1977). Solid phase enzyme immunoassay or radioimmunoassay for the detection of immune complexes based on their recognition by conglutinin. Conglutinin binding test. *Clinical and Experimental Immunology*, **29**, 342–54.

Creighton, W.D., Lambert, P.H., and Miescher, P.A. (1973). Detection of antibodies and soluble antigen–antibody complexes by precipitation with polyethylene glycol. *Journal of Immunology*, **111**, 1219–27.

Eye, K., Becker, M., Theofilopoulos, A., Montsopoulous, H., Feldman, J.L., and Talal, N. (1977). Immune complexes in hepatitis B antigen-associated periarteritis nodosa. *American Journal of Medicine*, **62**, 783–91.

Fischbach, M. *et al*. (1980). Immune complexes in Sjögren's syndrome. *Arthritis and Rheumatism*, **23**, 791–5.

Forster, P.J. and McConkey, B. (1986). The effect of antirheumatic drugs on circulating immune complexes in rheumatoid arthritis. *Quarterly Journal of Medicine*, **58**, 29–42.

Gauthier, V.J. and Abrass, C.K. (1992). Circulating immune complexes in renal injury. *Seminars in Nephrology*, **12**, 379–94.

Halla, J.T., Volanakis, J.E., and Schroenloher, R.E. (1979 *a*). Immune complexes in rheumatoid arthritis sera and synovial fluids. *Arthritis and Rheumatism*, **22**, 440–8.

Halla, J.T., Volanakis, J.E., and Schroenloher, R.E. (1979 *b*). Circulating immune complexes in mixed connective tissue disease. *Arthritis and Rheumatism*, **22**, 484–9.

Hamburger, M.I., Lawley, T.J., Kimberly, R.P., Plotz, P.H., and Frank, M.M. (1982). A serial study of splenic reticuloendothelial system Fc receptor functional activity in systemic lupus erythematosus. *Arthritis and Rheumatism*, **25**, 48–54.

Harbeck, R.J., Bardona, E.J., Kohler, P.F., and Carr, R.I. (1973). DNA anti-DNA complexes: their detection in systemic lupus erythematosus sera. *Journal of Clinical Investigation*, **52**, 789–95.

Hardin, J.A., Steere, A., and Malawista, S. (1979). Immune complexes and the evolution of Lyme arthritis: dissemination and localization of abnormal C1q binding activity. *New England Journal of Medicine*, **301**, 1358–63.

Hay, F.C., Nineham, L.J., and Roitt, I.M. (1976). Routine assay for the detection of immune complexes of known immunoglobulin class using solid phase C1q. *Clinical and Experimental Immunology*,

24, 396–400.

Inman, R.D. (1982). Immune complexes in SLE. *Clinics in the Rheumatic Diseases*, **8**, 49–62.

Inman, R.D. and Day, N.K. (1981). Immunologic and clinical aspects of immune complex disease. *American Journal of Medicine*, **70**, 1097–106.

Inman, R.D. *et al*. (1980). The C1q binding assay in systemic lupus erythematosus: discordance with disease activity. *Arthritis and Rheumatism*, **23**, 1282–6.

Jones, V.E. and Orlans, E. (1981). Isolation of immune complexes and characterisation of their constituent antigens and antibodies in human diseases: a review. *Journal of Immunological Methods*, **44**, 249–70.

Jones, V.E., Jacoby, R.K., Wallington, T., and Holt, P. (1981). Immune complexes in early arthritis. I. Detection of immune complexes before rheumatoid arthritis is definite. *Clinical and Experimental Immunology*, **44**, 512–21.

Kilpatrick, D.C. and Weston, J. (1985). Immune complex assays and their limitations. *Medical Laboratory Sciences*, **42**, 178–85.

Koffler, D., Agnello, V., and Kunkel, H.G. (1974). Polynucleotide immune complexes in serum and glomeruli of patients with systemic lupus erythematosus. *American Journal of Pathology*, **74**, 109–24.

Krakauer, R.S., *et al*. (1983). Therapeutic trial of cryofiltration in patients with rheumatoid arthritis. *American Journal of Medicine*, **74**, 951–5.

Lambert, P.H. *et al*. (1978). A collaborative study for the evaluation of eighteen methods for detecting immune complexes in serum. *Journal of Laboratory and Clinical Immunology*, **1**, 1–22.

Lawley, T.J., Moutsopoulos, H.M., Katz, S.I., Theofilopoulos, A.N., Chused, T.M., and Frank, M.M. (1979). Demonstration of circulating immune complexes in Sjögren's syndrome. *Journal of Immunology*, **123**, 1382–7.

Levinsky, R.J., Cameron, J.S., and Soothill, J.F. (1977). Serum immune complexes and disease activity in lupus nephritis. *Lancet*, **i**, 564–7.

Lloyd, W. and Schur, P.H. (1981). Immune complexes, complement and anti-DNA in exacerbations of systemic lupus erythematosus (SLE). *Medicine*, **60**, 208–17.

Luthra, H.S., MacDuffie, F.C., Hunder, G.G., and Samayoa, E.A. (1975). Immune complexes in sera and synovial fluids with rheumatoid arthritis. Radioimmunoassay with monoclonal rheumatoid factor. *Journal of Clinical Investigation*, **56**, 458–66.

McDougal, J.S. and McDuffie, F.C. (1985). Immune complexes in man: detection and clinical significance. *Advances in Clinical Chemistry*, **25**, 1–60.

McDougal, J.S. *et al*. (1982). Comparison of five assays for immune complexes in the rheumatic diseases: an assessment of their validity for rheumatoid arthritis. *Arthritis and Rheumatism*, **25**, 1156–66.

Mackel, S.E., Tappeiner, G., Brumfield, H., and Jordan, R.E. (1979). Circulating immune complexes in cutaneous vasculitis: detection with C1q and monoclonal rheumatoid factor. *Journal of Clinical Investigation*, **64**, 1652–60.

Morrow, W.J.W., Isenberg, D.A., Todd-Pokropek, A., Parry, H.F., and Snaith, M.L. (1982). Useful laboratory measurements in the management of systemic lupus erythematosus. *Quarterly Journal of Medicine*, **51**, 125–38.

Nydegger, U.E. *et al*. (1977). Circulating complement breakdown products in patients with rheumatoid arthritis: correlation between plasma C3d, circulating immune complexes, and clinical activity. *Journal of Clinical Investigation*, **59**, 862–8.

Ronco, P. *et al*. (1983). Immunopathological studies of polyarteritis nodosa and Wegener's granulomatosis: a report of 43 patients with 51 renal biopsies. *Quarterly Journal of Medicine*, **52**, 212–23.

Rossen, R.D., Brewer, E.J., Person, D.A., Templeton, J.W., and Lipsky, M.D. (1977). Circulating immune complexes and antinuclear antibodies in juvenile rheumatoid arthritis. *Arthritis and Rheumatism*, **20**, 1485–90.

Salvidio, G. and Andres, G. (1986). Immune deposits and immune complex disease. *Clinical and Experimental Rheumatology*, **4**, 281–8.

Schreiber, I. and Maini, R.N. (1984). Circulating immune complexes in connective tissue diseases. *Netherlands Journal of Medicine*, **27**, 327–39.

Seiboid, J.R., Medsger, T.A., Winkelstein, A., Kelly, R.H., and Rodnan, G.P. (1982). Immune complexes in progressive systemic sclerosis (scleroderma). *Arthritis and Rheumatism*, **25**, 1167–73.

Smiley, J.D. and Moore, S.E. (1989). Immune complex vasculitis: role of complement and IgG-Fc receptor functions. *American Journal of Medical Sciences*, **298**, 267–77.

Stanworth, D.R. and Johns, P. (1977). Ultracentrifugal methods of detecting soluble immune complexes. *Annals of the Rheumatic Diseases*, **36** (Suppl. 1), 12–16.

Theofilopoulos, A.N. and Dixon, F.J. (1979). The biology and detection of immune complexes. *Advances in Immunology*, **28**, 89–220.

Van Es, L.A., Daha, M.R., Valentijn, R.M., and Kauffmann, R.H. (1984). Pathogenic significance of circulating immune complexes. *Netherlands Journal of Medicine*, **27**, 350–8.

Virella, G. (1990). Immune complex disease. *Immunology Series*, **50**, 395–414.

Virella, G. (1993). Immune complex disease. *Immunology Series*, **58**, 379–96.

Wener, M.H., Mannik, M., Schwartz, M.M., and Lewis, E.J. (1987). Relationship between renal pathology and the size of circulating immune complexes in patients with systemic lupus erythematosus. *Medicine*, **66**, 85–97.

Williams, B.D. and Slaney, J.M. (1977). Macromolecular C_3—an indirect marker of circulating immune complexes. *Annals of the Rheumatic Diseases*, **36** (Suppl. 1), 37–9.

Williams, R.C. (1981). Immune complexes in human diseases. *Annual Review of Medicine*, **32**, 13–28.

WHO/IUIS Working Group (1981). Use and abuse of laboratory tests in clinical immunology: critical consideration of eight widely used diagnostic procedures. *Bulletin of the World Health Organization*, **59**, 717; and *Clinical and Experimental Immunology*, **46**, 662.

Zubler, R.H., Lange, G., Lambert, P.H., and Miescher, P.A. (1976 *a*). Detection of immune complexes in unheated sera by a modified ^{125}I-C1q binding test. Effect of heating on the binding of C1q by immune complexes and application of the test to systemic lupus erythematosus. *Journal of Immunology*, **116**, 232–5.

Zubler, R. *et al*. (1976*b*). Circulating and intra-articular immune complexes in patients with rheumatoid arthritis. Correlation of ^{125}I-C1q binding activity with clinical and biological features of the disease. *Journal of Clinical Investigation*, **57**, 1308–19.

4.9.1 Imaging in adults

Peter Renton

Introduction

It is the aim of this chapter to provide a general overview of techniques currently available for the imaging of the musculoskeletal system. Changes in individual diseases will be described in the appropriate sections.

Radiology has undergone a major technological revolution in this generation and new methods of imaging are now available (Table 1). These are generally expensive and time-consuming. Investigations have to be tailored to the individual needs of the patient. Radiologists and clinicians have to work together to ensure that the needs of all are best matched.

Plain films
Linear tomography
Arthrography
Computed tomography (CT) scanning
Radionuclide scanning, including DEXA (bone densitometry)
Ultrasonography
Magnetic resonance (MR) imaging

Table 1 Radiological investigative techniques

Plain radiography

Conventional radiology is available generally and is usually the first imaging technique used. Changes in film and X-ray technology have occurred in recent years, often to the detriment of image quality. Faster films, rapid automatic processing, the use of rare earth screens, and ultimately the move away from analogue to digital imaging have not necessarily meant an improvement in image quality, but a lower radiation dose.

Plain film images of bones and joints are usually obtained in two planes. For many joints, e.g. the knee, this will mean at least an anteroposterior and lateral radiograph; for some joints, however, an oblique view is necessary, e.g. at the hip, as the lateral view of the pelvis superimposes the two hip joints.

Plain film imaging of soft tissues

Soft tissue swelling is an unfailing indicator of underlying musculoskeletal disease. It should always be searched for, especially in the target areas. It follows that images must be correctly exposed for both bone and soft tissue. Digital imaging allows operator control of these factors.

Soft tissue swelling is seen as an increase in width over involved bones or joints, i.e. over the ulnar styloid (Fig. 1) or over the 5th or 1st metatarsal head. The thickened soft tissues are also radiologically denser and, therefore, the involved area looks 'greyer' and the underlying bone detail is less well seen, as it is 'filtered' by the soft tissue thickening. Comparison can be made with an unaffected or normal joint.

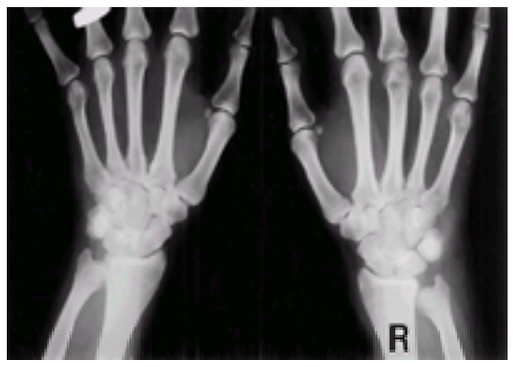


Fig. 1 Rheumatoid arthritis. Soft tissue swelling over the right ulnar styloid.

If a swollen joint is not at a margin but more centrally situated, e.g. the 2nd to 4th metatarsophalangeal joint, then a general increase in density is seen together with alteration of local contours. Filling in of web spaces in between fingers, which become convex rather than concave, or similar soft tissue swelling over metatarsal heads projecting over basal phalanges may be seen.

Fat planes overlie joints and separate muscle bundles. They are seen on radiographs as fat is of lesser density and attenuates the beam less than muscle. Displacement of fat planes is an indicator of muscle thickening, oedema, or haemorrhage or, conversely, muscle atrophy. Joint effusions also displace overlying fat lines (Fig. 2).

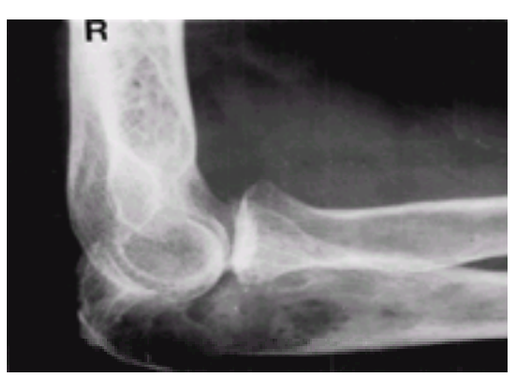


Fig. 2 Rheumatoid arthritis. Pathological fracture through a large cyst is associated with elevation of the anterior fat pad (sail sign).

Plain film changes with soft tissue thickening and fat plane displacement

Skull

These changes occur following trauma, over fractures, and at sites of foreign bodies. They may be found in the paranasal air sinuses with trauma, infection, or tumour.

Cervical spine

On a lateral view of the cervical spine, the retropharyngeal soft tissues are applied closely to the anterior aspect of the upper four cervical vertebral bodies. Below C4 the tracheal translucency is situated around 1 cm anterior to the cervical spine. Air is seen occasionally in the oesophagus, and an anterior cervical fat stripe may also be present. Soft tissue thickening following infection, trauma, or malignancy is demonstrated by displacement of the air shadows (Fig. 3).

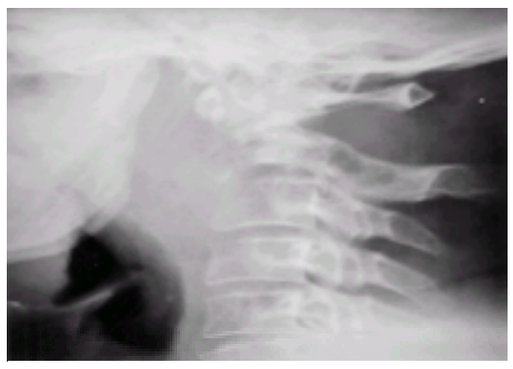


Fig. 3 Pathological collapse of the body of C2 with an associated large anterior soft tissue mass. The destructive lesion in this 27-year-old male was a chordoma.

Thoracic spine

In the thoracic region, soft tissues closely approximate to the vertebral bodies and are well demonstrated on a correctly exposed anteroposterior view. These are best seen on the left, are mainly pleural, and are separate from the descending aorta. These paraspinal lines are displaced by osteophytes and syndesmophytes but especially by bleeding in malignancy, trauma, infection, etc. Generally, the widest point of the soft tissue swelling lies at the site of maximal skeletal change (Fig. 4).

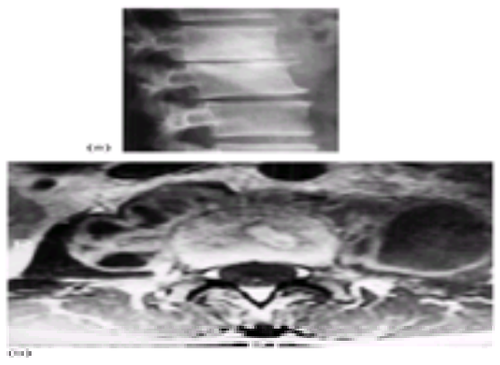


Fig. 4 Paraspinal soft tissue mass in spinal tuberculosis. (a) The radiograph demonstrates narrowing of an upper lumbar disc space with reversal of the curve. There is destruction of the adjacent end-plates with reactive new bone formation anteriorly, and a soft tissue mass is demonstrated anterior to the affected disc. (b) The axial MR images demonstrate not only the erosive destruction of the vertebral body but also psoas abscesses. The images are obtained after intravenous gadolinium and show both areas of enhancement and necrosis in the psoas muscles as well as anteriorly in the retroaortic tissues.

Lumbar spine

In the lumbar region the spine is not surrounded by air, but by soft tissues. Nonetheless, masses can still be assessed by inspection of the psoas and its overlying fat plane on an anteroposterior view, and by displacement of gut gas shadows on the lateral view. Similarly, sacral disease displaces the rectal gas shadow.

Shoulder

The axillary recess of the joint can be assessed on a plain radiograph and effusion diagnosed. Similarly, a large subdeltoid bursal effusion or, conversely, deltoid atrophy may also be assessed (Fig. 5).

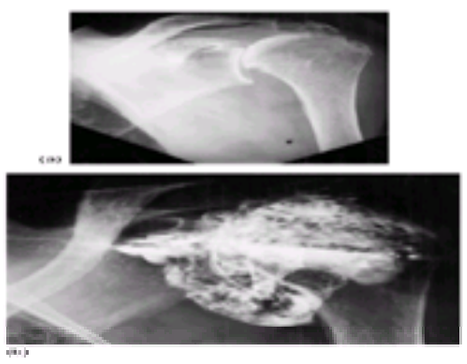


Fig. 5 Rheumatoid arthritis with a large effusion which can be seen on the plain film. (a) Upward subluxation of the humeral head is associated with marginal erosions, as well as erosion of the under surface of the acromion. The displaced fat plane beneath the glenoid and humeral head is seen on the subsequent arthrogram to represent the margin of the distended capsule. (b) A rotator cuff tear is demonstrated at arthrography with filling of the subacromial bursa. The distended joint capsule is demonstrated, especially inferiorly. Numerous loose bodies lie within the joint space. These may represent areas of synovial proliferation or osteochondral fragments but, in the main, represent fibrinous loose bodies.

Elbow

Anterior and posterior fat pads lie on the synovium and are usually just visible on the lateral radiograph. This is taken with the arm in the flexed position. Under these circumstances, the posterior fat pad is pressed inwards by the triceps tendon while the anterior capsule and fat pad are redundant. Effusions elevate the anterior fat pad more easily, therefore, than the posterior (Fig. 2). A 'sail' sign results, and is suggestive of a fracture given a history of trauma, especially if the posterior fat plane is elevated; this only happens with larger quantities of fluid.

Wrist

There are numerous bursae and tendons around the wrist. The tendons are surrounded by synovial sheaths. The wrist joint itself is a complex multicompartmental structure. Soft tissue swelling may occur at many sites (Fig. 1). These should be actively searched for in the arthritides, for instance, over radial and ulnar styloids. The lateral view should also be inspected. Pronator quadratus muscle lies on the palmar aspect of the distal radius and ulna, and a fat plane lies on it. Trauma to the distal radius or ulna elevates this fat pad, which assumes an abnormal convex palmar appearance (Fig. 6). Similarly, soft tissue swelling may be seen over interphalangeal joints, and especially at distal interphalangeal joints in cases of osteoarthritis, and over the dorsum of the carpus.

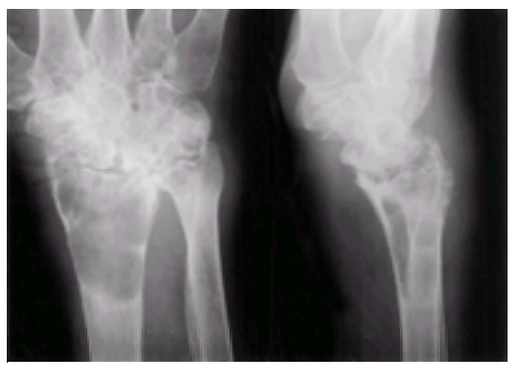


Fig. 6 Displacement of pronator quadratus. On the lateral view there is diffuse soft tissue swelling over the wrist in this patient with severe rheumatoid arthritis. The fat plane over pronator quadratus is displaced in a palmar direction over the large cyst in the distal radius.

Pelvis and hips

The acetabulum is difficult to visualize in a conventional anteroposterior film. For this reason, 45°-oblique views are needed to demonstrate the anterior and posterior pillars satisfactorily (Judet views). The obturator internus lies just inside the acetabulum and the fat line internal to it is seen as a thin stripe lying on, and parallel to, the iliopectineal line. Trauma to the acetabulum, malignant disease, or cortical bone destruction elevates the fat pad (Fig. 7).

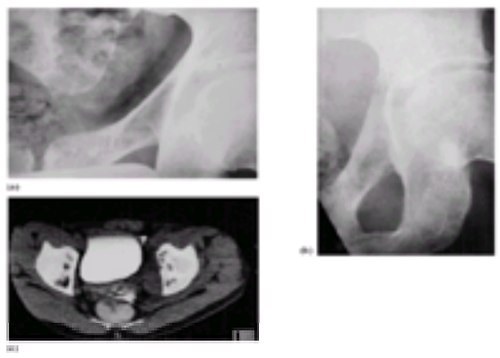


Fig. 7 (a) The initial radiograph shows early displacement of the fat pad over obturator internus. (b) Subsequent images of this region, including a CT scan (c), show progressive enlargement of this soft tissue mass with destruction of the overlying bone due to a reticulum cell sarcoma.

An effusion in the hip joint also displaces two commonly seen fat pads round the medial and lateral aspects of the femoral neck and, in the case of a monarthritis, comparison can be made with the normal side.

Knee

The suprapatellar pouch is seen in the undistended state arising almost vertically from the patellofemoral joint space. It is a thin (2 to 3 mm) parallel-walled or rectangular structure surrounded by fat. With progressive distension, it enlarges to a lentiform shape of soft tissue density surrounded by fat. On the anteroposterior view, fat planes lying on the synovium are projected on either side of the distal femur, concave outward and parallel to the distal femoral contour. With an effusion, these are displaced progressively laterally.

Ankle

Both malleoli lie directly subcutaneously. Bony trauma causes soft tissue swelling more proximally than damage to the collateral ligaments, which is associated with swelling distally.

Joint effusions are demonstrated on the lateral view. Anterior and posterior soft tissue masses project from the joint space and are seen as soft tissue densities, especially posteriorly where they project into the fatty triangle bounded by tendo-Achillis, calcaneus, and posterior tibia (Kager's triangle).

Behind this triangle lies the Achilles tendon, which inserts into the posterior aspect of the calcaneus, some 2 to 3 mm beneath its upper angle. Between tendon and bone lies a small, normally non-visualized bursa. The tendon is some 6 mm thick at the insertion, and sharply defined anteriorly and posteriorly by overlying fat. In the presence of Achilles tendinitis or posterior calcaneal erosive disease, the distal tendon thickens and its insertion, being oedematous, becomes poorly defined. The retrocalcaneal bursa enlarges, first fills the space known as the retrocalcaneal recess, then further encroaches into Kager's triangle (Fig. 8).

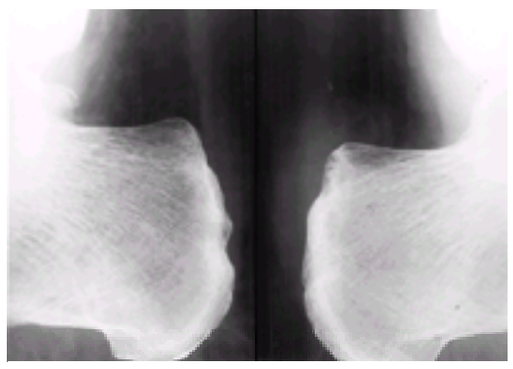


Fig. 8 Rheumatoid arthritis. Thickening of the retrocalcaneal bursa at the junction of tendo-Achillis and the calcaneus. A diffuse soft tissue mass protrudes into Kager's triangle.

Foot

Changes at the metatarsal heads mirror those seen in the hands. Early on in an erosive arthritis, therefore, particular attention should be paid to the soft tissues over the metatarsal heads.

Arthrography

The injection of contrast into joints has been in use for many years prior to the advent of image intensification. Knee arthrography was reported as early as 1904. Lindblom (1939) performed shoulder arthrography using Myodil by a technique of direct injection of the joint from above.

In the recent past the major use of arthrography has been in the knee, to demonstrate meniscal tears. In experienced hands, arthrography is as accurate as arthroscopy in the diagnosis of meniscal lesions (Table 2(a)) and probably more accurate in the diagnosis of lesions of the posterior horn of the medial meniscus. Arthroscopy is much more accurate, however, in assessing the status of the cruciate ligaments (Thijn 1982), even if intra-articular adrenaline (to diminish dilution of contrast by effusion) and linear tomography are used (Table 2(b)) (Ng *et al*. 1989).

	Arthrography (%)	Arthroscopy (%)
Medical meniscus	94	81
Lateral meniscus	90	94.5
Patellar chondropathy	55	99.5
Cruciate ligament lesions	69	97

After Thijn (1982).

Table 2a Comparison of accuracy rates

	Arthrography (%)	Arthroscopy (%)
Correct	69	97
Equivocal	1	3
False positive	12	0
False negative	18	0

After Thijn (1982)

Table 2b Accuracy of arthrography and arthroscopy for cruciate ligament rupture

In the shoulder, arthrography is mainly used to diagnose rotator cuff tears and restrictive capsulitis—tasks for which the examination is admirably suited, being cheap, quick, and requiring no particular expertise or expensive equipment (Fig. 5).

Both of these examinations, and possibly arthroscopy, are becoming extinct owing to the advent of magnetic resonance imaging (**MRI**). Indeed, our younger radiologists now never even see enough arthrograms to be trained adequately to perform them in the future.

Similarly, the use of arthrography of the temporomandibular joint to display the meniscus, and of the wrist to assess the triangular cartilage, is also in decline.

CT scanning

This technique uses X-rays to obtain axial images of the body in serial slices. The technique is therefore a tomographic one and because the attenuated beam, having passed through the patient, is detected and analysed by computer a much wider grey scale is available on the computer-based reconstructions than on X-ray film.

Axial tomography not only enables spatial relationships to be obtained in the sagittal and coronal planes simultaneously, but allows densities to be assessed visually and also directly from the computer, measured in Hounsfield units.

Enhancement with intravenous water-soluble iodine-based contrast media further opacifies vascular tissues, enhancing their intrinsic differences. This is especially the case with vascular tumours and infections while, conversely, avascular lesions do not enhance.

Using a radiographic technique, therefore, it is possible to differentiate between fat (which shows low attenuation), serous fluid, and blood, and between muscle and cortical and medullary bone.

Radionuclide scanning

Until the advent of MRI, isotope scanning was the most accurate method of demonstrating the presence of pathological change in bones and joints (Fig. 9). Even so, it seems to have been underused and plain films were, and still are, the basic diagnostic imaging tool. Technetium-99m phosphate images demonstrate increased blood flow to inflamed synovium in the early or vascular phase of the scan and, on delayed images obtained 3 h after injection, increased uptake is demonstrated at sites of increased bone turnover due to isotope deposition on hydroxyapatite crystals.

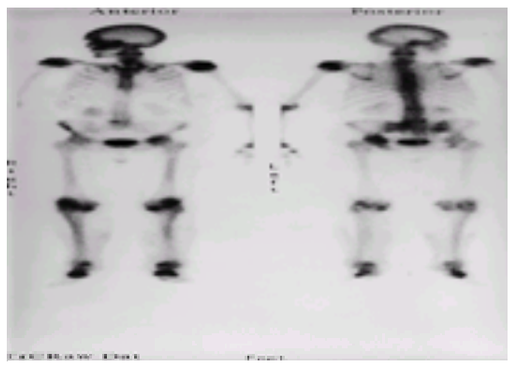


Fig. 9 Rheumatoid arthritis. Radioisotope bone scan showing increase in uptake in both shoulder joints, the left hip, both knees, and the hindfeet. (By courtesy of Dr A. Hilson, Royal Free Hospital.)

Radioisotope scanning is an invasive technique which, of course, depends on radioactivity—unlike MRI—and while there is good sensitivity, there is poor specificity for individual diseases. In one study, technetium-99m phosphate imaging was 11 per cent more sensitive than clinical examination and 72 per cent more sensitive than conventional radiographs. Similarly, early and persistently positive scans are a good prognosticator for subsequent erosive change, but chronic inactive erosions may often be negative at phosphate scanning (Rosenthall 1991). Spatial discrimination has been improved by use of **SPECT** (single photon emission computed tomography) imaging. The various arthropathies can only be distinguished by the distribution of the abnormal foci of increase in uptake.

Measurement of bone density

It is recognized that up to 50 per cent of bone mineral in a given area must be lost before being radiologically visible on plain films. In addition, subjectivity results in observer variation. Early methods for assessing bone density included photographic densitometry, an aluminium wedge being included on the film, or a cadaver metacarpal matched for age and sex. Images of the hand could also be used to measure cortical thickness in the metacarpal, and resorption of the cortex gives an indication of demineralization.

Unfortunately, there is poor correlation between bone mineralization in the axial and appendicular skeleton, as the hands and feet are subject to external influences—osteoarthritis at the wrist or body weight at the foot.

CT scanning can be used to assess bone density, especially in the spine, by comparing attenuation of single or dual energy CT beams with attenuation by phantoms which may be scanned simultaneously or separately. This technique does involve irradiation, especially superficially.

Photon absorptiometry

This uses a single source of photon energy through the radius or calcaneus, and absorption of the isotope by bone is compared with that in adjacent soft tissues.

Dual photon densitometry

This utilizes photons of two energies from a single isotope (gadolinium-153). The results are independent of thickness of local soft tissues (whereas single photon

absorptiometry requires a uniform thickness of soft tissue over the bony part being investigated). Radiation dose is relatively low and double photon absorptiometry is in particular accurate, with only 2 to 3 per cent error. Double photon absorptiometry enables both central and peripheral bone to be assessed.

Quantitative dual energy radiography (X-ray absorptiometry; DEXA)

This technique uses an X-ray tube rather than an isotope as the source, which results in greater resolution and speed of investigation (Sartoris and Resnick 1990). The X-ray source provides a beam alternating between two energies. The beam is finely collimated and also passes through a disc containing attenuating materials, used as a reference. A ratio of attenuation is given. This is now the technique of choice as it is precise, scanning time is short, and radiation dose low. Accurate measurements are taken most commonly from the lumbar spine and proximal femur and results are usually expressed in terms of the number of standard deviations above or below the mean of an age-related control population (Z-score) or of a young healthy adult population (T-score) (Fig. 10).

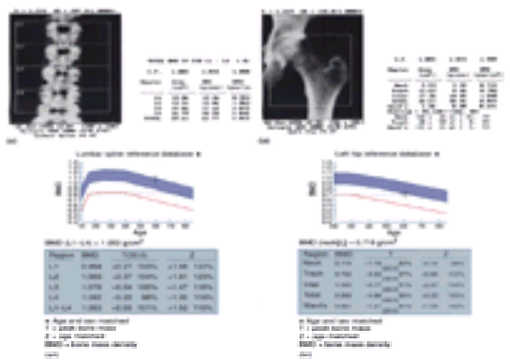


Fig. 10 (a) DEXA scan of lumbar spine. The scan shows an increase in density in the lumbar spine above the expected norm for the age, but this is probably because of the presence of osteophytes in degenerative disease; (b) further imaging of the neck of femur, however, shows the patient to lie exactly on the predicted value for bone density at her age.

Ultrasonography

Diagnostic ultrasound utilizes sound frequencies above those audible to the human ear, i.e. 20 000 cycles per second (Hertz). Frequencies in use clinically range from 2 to 10 MHz.

Ultrasound waves are generated by a piezoelectric crystal in the usually hand-held transducer, which both transmits the sound waves and receives back the resultant echoes. The returning signal impinges upon the crystal which converts the attenuated sound waves into electronic impulses, processed via a computer.

Images are displayed on a screen, film, or paper using a grey scale, or colour to show venous/arterial flow, or a combination of both. Real-time imaging allows dynamic continuous scanning of body parts. Linear images can be obtained in any plane of the body (Van Holsbeeck and Introcaso 1991).

Images obtained depend on the following:

1. The frequency of the ultrasound. The higher the frequency, the better is the resolution of the images, but the penetration of the body by sound waves is limited.
2. The nature of the tissues into which the sound waves must pass. Bone is totally refractory to sound waves (as is air) and so deep structures in joints are inaccessible. At the knee, therefore, the superficial ligaments and tendons are well imaged, as are the superficial parts of the menisci, but not the deep structures (Fig. 11).

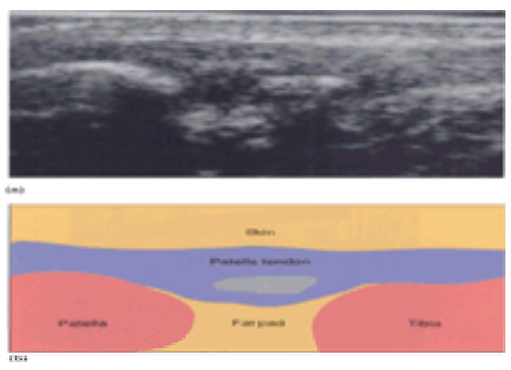


Fig. 11 Calcific tendinitis in a 27-year-old chauffeur with pain in his knee. Ultrasound of the patellar tendon shows a thickened proximal part with a hypoechoic area inferiorly. A large focus of calcium is seen. (By courtesy of Dr S. Burnett, St Mary's Hospital.)

3. The angle of incidence of the beam. Sound waves pass through matter and the speed of sound in a tissue varies with the nature of the tissue, materials of the greatest density transmitting sound at the highest velocity. Sound is also reflected at tissue interfaces, while the amount of reflected sound also depends on the angle of incidence of the beam, the least reflection occurs with a beam which is at right angles to the reflecting surface (Van Holsbeeck and Introcaso 1991). The returning pulses of sound characterize the tissues and tissue interfaces. A cyst is echo-free, calculi cast an acoustic shadow, and soft tissues show numerous interfaces, giving a complex but regular pattern if normal and a more irregular pattern in the presence of disease.

Ultrasound is thus used in:

1. joint disease to show effusions, intra-articular loose bodies or other structures, and the capsule and synovium;
2. surrounding ligaments and tendons;
3. muscles, tears, haematomas, and tumours;
4. vascular lesions—arterial and venous flow using Doppler.

It is not used in the diagnosis of bone disease.

The shoulder region is thus much more amenable to ultrasound investigation than the knee, as the rotator cuff is more accessible to the ultrasonic waves (Fig. 12).

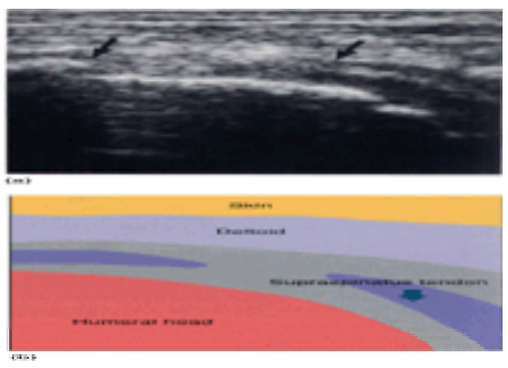


Fig. 12 Rotator cuff tear in a 56-year-old woman who fell off her horse. The free ends of the supraspinatus tendon (arrows) are seen deep to the deltoid. (By courtesy of Dr S. Burnett, St Mary's Hospital.)

Magnetic resonance imaging (MRI)

MRI does not involve ionizing radiation, but utilizes the different biophysical properties of the various constituents of the body tissues, especially the number of hydrogen nuclei or protons in a given tissue. MR images reflect the distribution of these nuclei. Grey scale changes on film reflect the number of protons in a given tissue, and reflect their behaviour in the externally applied magnetic fields.

These field strengths are generally between 0.5 and 1.5 T (Tesla). Such field strengths cause the hydrogen nuclei to precess or gyrate at frequencies in the range of radiowaves or radiofrequency (**RF**).

RF stimulation (i) excites protons and (ii) also brings them into phase with each other. Once the RF pulse is turned off, the system returns to normal. The protons return to a low energy state. A measure of the time taken for dephasing to occur is the T_2 (or transverse relaxation time) and the T_1 (or longitudinal relaxation time) is that taken for the protons to re-establish their equilibrium (Nixon 1987).

In most cases, the sequences used are simple T_1- and T_2-weightings. In T_1-weighted images fat is bright, fluid is grey, and bone is black. With T_2-weighted images, fat is less bright, fluid is very bright, and bone is black.

Fat suppression **STIR** (short tau inversion recovery) sequences cause the bright signal from fat to be suppressed completely. Muscle is therefore a more homogeneous grey (as the internal fat is black). Fat, and fatty marrow, is black. Fluid, however, remains very bright and so oedema, vascular tumours, fluid collections, and veins stand out clearly from the surrounding suppressed bone, muscle, and fat. Avascular and necrotic tissues are also black (i.e. of low signal) (Fig. 13).

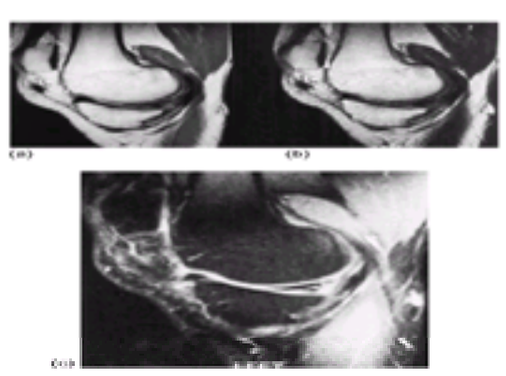


Fig. 13 Meniscal tears demonstrated with (a) T_1-, (b) T_2- and (c) STIR weightings. The STIR weighting has the added advantage of graphically demonstrating cystic change or oedema in the adjacent bone much more clearly than the other modes.

It should be noted that primary images can be obtained in any plane.

Fluid, fat, and medullary bone have the highest concentration of protons and, therefore, the signal from these structures is brightest, while those from cortical bone, tendon, ligaments, and menisci are lowest. Signal is also altered by disease, so that oedema, inflammation, or haemorrhage all cause local increase in signal. Cell death, however, decreases signal, as does calcification.

MRI is becoming the investigation of choice for the musculoskeletal system. It enables primary, not reformatted, images to be obtained in all planes, is non-invasive, and does not use ionizing radiation. Inherent differences in structure of the body tissues result in greater differences in grey-scale contrast than are obtainable by any other imaging technique.

Muscles, tendons, ligaments, fat, blood vessels, and marrow all have distinguishing features which can be manipulated further by using different weightings (Table 3). Cortical images are probably better obtained by using X-rays, however, either by conventional radiography or CT scanning.

Signal intensity	Tissue
Highest (white)	Fat Medullary bone
Grey	Articular cartilage
Intermediate intensity	Muscle
Low intensity (black)	Blood vessels* Ligaments, tendons Menisci Cortical bone

Reproduced from Benjuist, T. H. (1987). In *Magnetic resonance of the musculoskeletal system* (ed. T. H. Benjuist), by kind permission of the author and Raven Press.
*Normally flowing blood gives no signal. The appearance of vessels may vary with changes in flow and multiecho sequences.

Table 3 Signal intensity of musculoskeletal tissue

MR images the bone marrow well, especially fatty marrow. Fat generally gives a bright image both on T_1- and T_2-weighted sequences, but is suppressed, i.e. of low signal intensity, on STIR fat suppression sequences. Fluid, oedema, or inflammatory changes are bright on T_2-weighting and on fat suppression sequences, but approximate to muscle signal on T_1-weighted images. Changes in marrow due to oedema or inflammation are thus of reduced signal on T_1-weighting and increased signal on T_2-weighted and STIR sequences. Tumours, too, primary and secondary, alter the signal from marrow, and their extent can be assessed as well as other

features such as central necrosis or matrix mineralization, though this latter feature is also best assessed by X-ray (Table 4).

Disease process	T_1	T_2
Inflammation	Increased	Increased
Neoplasia	Increased	Increased
Fibrosis	–	Decreased
Fatty infiltration	Decreased	–
Interstitial haemorrhage	Increased	Increased

Reproduced from Ehman, R. L. (1987). In *Magnetic resonance of the musculoskeletal system* (ed. T.H. Berquist), by kind permission of the author and Raven Press.

Table 4 General trend of relaxation time changes for some musculoskeletal disease processes

MR scanning is not inevitably successful and may be stressful to the patient. Older scanners were somewhat restricting, the patient being closely confined in a dark and noisy environment, leading to a patient rejection rate of around 2 per cent and also to degradation of images due to patient movement.

Sedation is rarely necessary in adolescents and adults. Intravenous sedation needs to be performed in an environment where the patient can be monitored. Non-ferromagnetic anaesthetic and resuscitation equipment is not inevitably available. Modern scanners are more open, or totally so, and are thus more likely to lead to a successful scan.

Before scanning, the presence of ferromagnetic and electric implants should be excluded. Metallic implants cause artefacts due to signal void or, on occasion, marked hyperintensity. Ferromagnetic implants cause a greater artefact than non-ferromagnetic, and are seen as an irregular area of low signal in and around the implant. Non-ferromagnetic implants may be associated with signal void, which does not significantly impair image quality. Larger field strength magnets also cause a larger artefact. A low signal artefact is thus seen to be due to both the presence of the metal itself and the surrounding effect on the magnetic field.

Pacemakers may be affected by the magnetic field and their presence creates image degradation. The pacemaker may move and pacemaker function may be interfered with. Current may be induced in leads sufficient to induce fibrillation or burns. Prosthetic heart valves are not generally a contraindication to MR scanning.

Ferromagnetic surgical clips can move in a magnetic field and their presence in a vulnerable situation, e.g. at an aneurysm, is a contraindication to MR scanning. Similarly, cochlear implants may be ferromagnetic. Some cochlear implants are held in place by magnets; others may be electronically activated. MRI should not be performed on these patients. Copper IUCDs are not ferromagnetic and are safely included in a scan.

Skin and body temperature are not increased by MRI but conductive materials may be heated and currents induced in leads resulting in burns.

Patients should also be screened for the presence of metallic foreign bodies, such as shrapnel, pellets, or bullets. It is often not possible to say whether pellets, bullets, or shrapnel are ferromagnetic. Movement may cause damage to adjacent vital structures. Blindness may result if intraocular metal fragments are displaced by the magnetic field. Metal workers should have radiographs of the orbit performed prior to MRI (Berquist 1991).

Imaging in rheumatoid arthritis

In patients with rheumatoid arthritis it is recognized that disease is often first seen in the feet rather than in the hands (Fletcher and Rowley 1952; Thould and Simon 1966; Brook and Corbett 1977). Erosions are first described at the metatarsophalangeal joints, and most commonly at the 5th metatarsophalangeal joint. The metatarsal head is eroded before the adjacent phalangeal base. The attention of the radiologist should therefore be drawn to this region first in a patient suspected of having rheumatoid arthritis.

Resnick *et al*. (1977) have drawn our attention to the concept of 'target areas'. Understanding of where change is likely to occur should draw our attention to those areas specifically in a proactive manner. The clinician has the benefit of having the patient before him, and so will know the sites of pain, swelling, and deformity. Erosions are a relatively late feature in the arthritides.

Radiological changes at joints include:

1. soft tissue swelling, local or general;
2. alteration in bony density;
3. joint narrowing and alignment changes;
4. periostitis;
5. surface irregularity, cortical loss, and erosions;
6. muscle wasting;
7. calcification or ossification in surrounding soft tissues.

The plain film diagnosis of erosive changes

Erosions have been studied extensively by Buckland-Wright (1983) and coworkers using macroradiography. This technique is probably impractical for use in conventional imaging departments. Images are obtained using a very small target (less than 15 m diameter). The object is placed close to the X-ray tube and the object-to-film distance is increased. According to Dacre and Buckland-Wright (1992), the advantages of using such a fine X-ray tube focus are (i) large magnification (up to × 10 or 20) and (ii) high spatial resolution, the smallest objects recordable being between 25 and 50 µm. According to these authors, changes are seen as early as 2 weeks after onset of symptoms (Fig. 14).

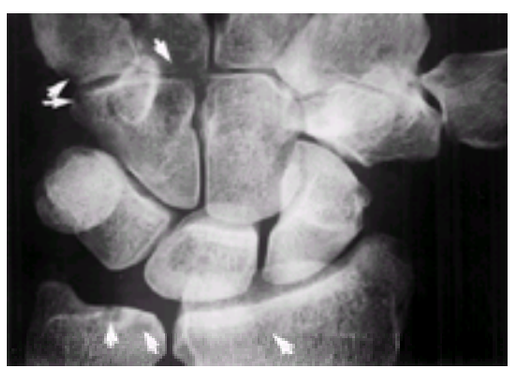


Fig. 14 Rheumatoid arthritis. Erosions demonstrated at macroradiography. (By courtesy of Dr J. C. Buckland-Wright, Guy's Hospital, London.)

Martel *et al*. (1980) described 'bare areas' of bone between articular cartilage and synovium, and pointed out that pannus logically would erode bone at this site which was free of cartilage. Erosions therefore commence as small areas of juxta-articular radiolucency adjacent to the margin of the articular cartilage. These lucencies are due to trabecular loss and cortical thinning. Small defects (0.1 to 0.3 mm diameter) begin to be seen in the cortex at sites of ingrowth of pannus. With progression of these changes, visible classical erosions form.

Migration of pannus below the articular cortex results in subchondral erosions and, with weakening of local bone, the articular cortex collapses (Fig. 15). Alternatively, direct invasion by pannus from above destroys articular cartilage and also causes subchondral invasion.

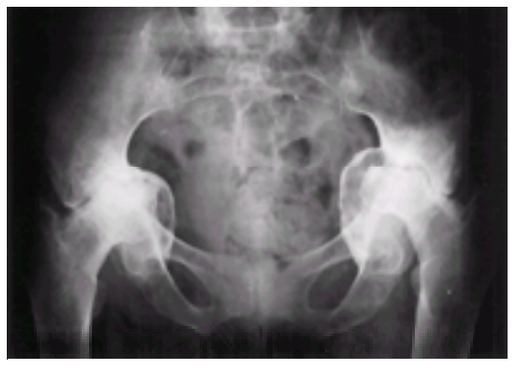


Fig. 15 Rheumatoid arthritis. Collapse of the femoral head associated with protrusio acetabuli.

Pressure erosions

Where two normally non-congruous bony surfaces approximate, and one moves on the other, well-corticated pressure erosions develop. With large rotator cuff tears, upward subluxation of the humeral head results in it eroding the undersurface of the acromion. Subsequently, the inferior glenoid lip erodes the adjacent humeral neck (Fig. 16).

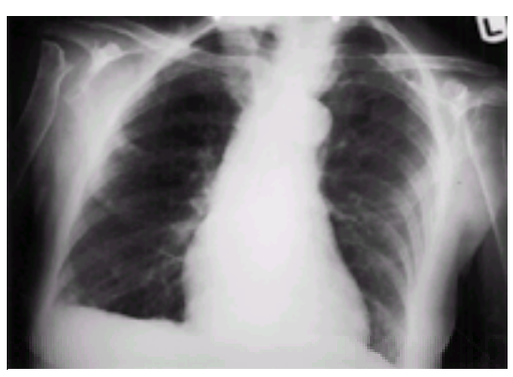


Fig. 16 Rheumatoid arthritis. Upward subluxation of the resorbed humeral head has eroded the inferior aspect of the acromion and clavicle. The neck of the humerus is eroded by the glenoid, and the superior aspects of the 3rd and 4th ribs are also eroded superiorly.

In patients with wasting of the thoracic cage musculature, the scapula erodes the upper surface of the 2nd, 3rd, and 4th ribs (Fig. 16). This occurs with chronic rheumatoid arthritis, but also in old age and after polio. Proximal phalangeal subluxation in scleroderma and after Jaccoud's arthritis similarly allows the phalangeal base to erode the metacarpal neck (Fig. 17). In the foot, impinging malleoli cause defects on the adjacent talus.

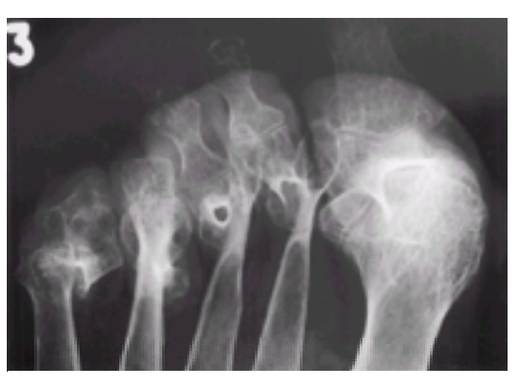


Fig. 17 Rheumatoid arthritis. Phalangeal subluxation causes pressure erosion of the metatarsal necks.

Sites of erosions in hands and wrists (Table 5)

	Wrist	Hand
Very early rheumatoid arthritis	Medial side hamate and base of 5th metacarpal joint Junction between hamate, capitate, base of 3rd and 4th metacarpal joints Medial side triquetral Lateral side head of ulna (inferior radioulnar joint) Styloid process of radius	2nd and 3rd metacarpophalangeal joints 3rd and 4th proximal interphalangeal joints
Early rheumatoid arthritis	Styloid process of ulna Medial surface of radius (inferior radioulnar joint) Scapholunate articular surfaces Lateral surface of scaphoid Lateral surface of base of 1st metacarpal joint Lateral surface of trapezium Lateral surface of capitate Subchondral erosion/cyst in epiphysis of radius	Remaining metacarpophalangeal and interphalangeal joints

From Blackburn-Wright (1984), reproduced by kind permission of the author and the British Medical Association

Table 5 Approximate order of the site of onset of erosions seen in the magnification radiographs of patients with 'very early' and 'early' rheumatoid arthritis

Buckland-Wright (1984) also demonstrated that erosions are symmetrical and bear no relation to hand dominance, though others do not agree. Erosions also vary in size, with the largest at the radiocarpal, medial carpometacarpal, and second and third metacarpophalangeal joints. It should be noted that in one study of patients with early rheumatoid arthritis, 23 per cent never developed erosions while, in those who did, in 48 per cent the erosions became static after 22 months, indicating how variable the process of erosive disease is (Brook and Corbett 1977). Moreover, radiographic progression of erosions can occur even when treatment successfully halts the acute clinical phase of the disease.

Much effort has been expended in assessing the status of rheumatoid disease at presentation and subsequently after treatment. Both conventional and special projections, e.g. of hands (Brewerton 1967), have been used. Films are often of different quality on different occasions, especially in multicentre trials. The author has participated in three such surveys; it is not an enjoyable experience. Films may be of such poor quality that an accurate staging is not possible. Observer variation even occurs when one radiologist is involved, let alone more than one, especially when viewing serial radiographs.

The importance of accurate imaging and reporting in assessing disease progress is self-evident, but using radiographic scoring methods is not only tedious but liable to intrinsic errors because of the factors discussed above. Different scoring methods can give contradictory results. The whole subject of scoring has been well researched (Kaye 1991). What is evident, however, is that MR is much more sensitive in demonstrating erosions and cysts than plain films.

Arthrography in rheumatoid arthritis

Arthrography in patients with rheumatoid disease demonstrates synovial proliferation and allows contrast to enter bone at sites of cortical defects and erosions. Larger subarticular cysts or geodes may fill with contrast. Debris is shown in the joint space as filling defects, often floating around freely. In the shoulder, a rotator cuff tear allows filling of the subacromial bursa (Fig. 5), while at the knee, adventitial cysts in communication with the joint may be shown at arthrography. The largest of these—the Baker's cyst—is an enlargement of the gastrocnemio-semimembranosus bursa which communicates with the posterior joint space. This cyst can be seen on a plain lateral radiograph as a fairly discrete, lobulated soft tissue shadow posterior to the knee. Loose bodies may lie in it, or if multiple densities are present inferoposterior to the knee, the presence of local synovial osteochondromatosis can be inferred (Fig. 18).

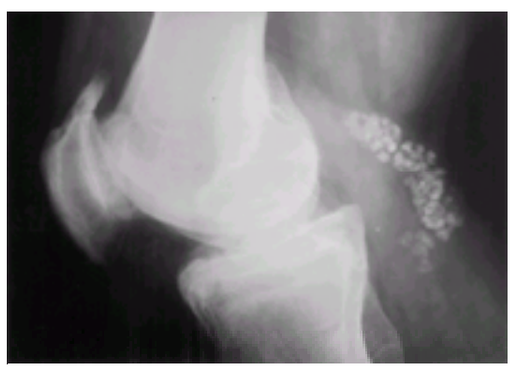


Fig. 18 Synovial osteochondromatosis. Numerous small ossicles are demonstrated posterior to the knee joint. There are so many that they cannot be held to be loose bodies; they must be in synovial osteochondromatosis. The distribution is that of a Baker's cyst.

Rupture of this cyst causes leak of synovial fluid into the calf, between skin and muscles or between soleus and gastrocnemius (Fig. 19). Rupture results in calf pain and swelling, mimicking a deep vein thrombosis. Should a venogram be performed as the first examination because of the need to exclude a deep vein thrombosis, the popliteal vein is seen to be displaced. Ultrasound is preferable as it confirms the presence of a cyst, the leak, and the absence of a deep vein thrombosis (Gompels and Darlington 1982). Arthrography followed by exercise fills the cyst, unless the defect in the cyst wall is filled with debris, but MRI of course demonstrates the cyst, the effusion, and the leak as well as the meniscal abnormality (Fig. 20).

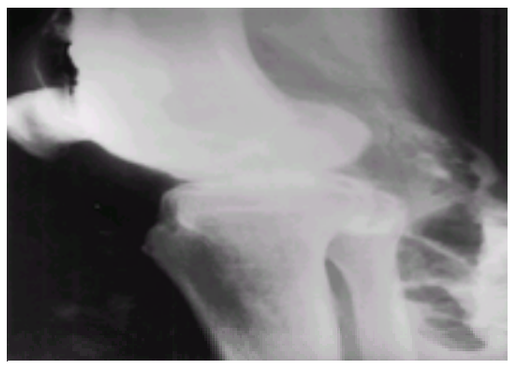


Fig. 19 Ruptured Baker's cyst allows air or contrast to pass into the posterior musculature of the calf.

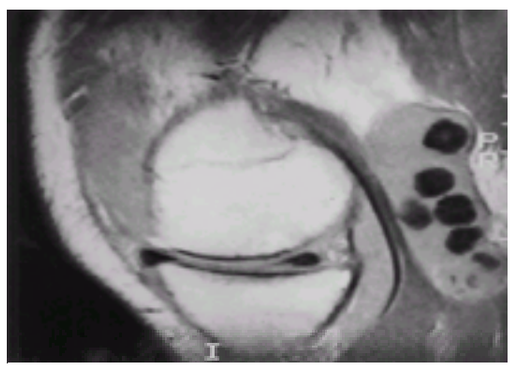


Fig. 20 Baker's cyst. MR scan of the knee demonstrating a large posterior cyst containing numerous loose bodies. This is associated with a tear of the posterior horn of the medial meniscus. Comparison can also be made with a plain film of a similar lesion (see Fig. 18).

Baker's cysts are associated with meniscal tears and other forms of internal derangement, presumably because of increased amounts of joint fluid and pressure in the joint. Rheumatoid and osteoarthritis are also associated with cysts and also with communicating cysts at the elbow and ankle. MRI shows a 5 per cent incidence of popliteal cysts in patients with internal derangement but a 13 per cent association with a ruptured anterior cruciate ligament (Fielding *et al*. 1991). Arthrography shows a much higher incidence of cysts, possibly because of the additive effect of distension by contrast and gas.

MRI in rheumatoid arthritis

As a result of the inherent contrast of MR images, the components of a joint, bone, and soft tissue are better defined than in any other modality. Synovial fluid,

synovium, cartilage, ligaments, tendon, and bone can all be identified, and pathology seen in them (Fig. 20).

Erosions are the pathognomonic change in rheumatoid disease and are seen at MRI as marginal low-signal defects in bone, while pannus shows on conventional images as an intermediate or soft tissue signal mass adjacent to the erosions. Effusions in joints or tendon sheaths are bright (brighter than pannus) on T_2-weighted views (Fig. 20).

It seems well accepted that erosions are seen earlier and more often on MR scans than on plain films. This is so for both large and small joints. The superiority of MR imaging in demonstrating erosions, subchondral cysts, effusions, tendinous and bone change, extra-articular collections of fluid, and pannus is so marked that MRI is, where available, the method of choice for detecting early change in rheumatoid arthritis (Beltran *et al*. 1987; Gilkeson *et al*. 1988; Poleksic *et al*. 1993).

The use of intravenous gadolinium with T_1-weighted sequences further distinguishes inflammatory from non-inflammatory fibrous pannus by demonstration of synovial enhancement where synovial proliferation is hypervascular. MRI is probably therefore the investigation of choice for the initial diagnosis of erosive disease and its progression (Fig. 21) but, because of its cost, complexity, and relative non-availability, is not likely to be in general use.

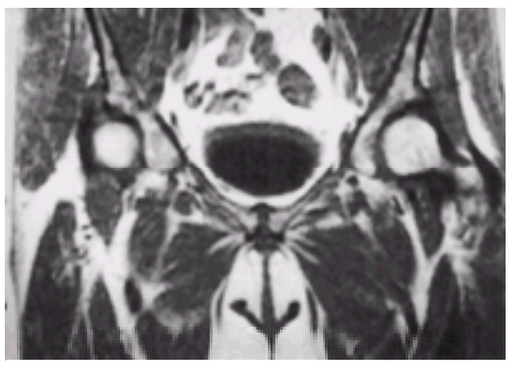


Fig. 21 Rheumatoid arthritis. Thickening of the capsule of the right hip joint is demonstrated with irregularity of the acetabular articular surface (T_1-weighting).

It has, however, been found that MR is better at demonstrating cartilage loss than any other imaging modality (Chan *et al*. 1991), so that a joint which appears normal on plain films is shown to have cartilage loss on MRI. This technique is also better than conventional radiography in demonstration of osteophytes, though CT does this equally well. MRI has the added benefit of demonstrating effusions and loose bodies without the necessity of joint injection, and demonstrates subarticular cysts and extra-articular cartilage change better than the other modalities (Fig. 20).

Changes at the C1–C2 interface: plain radiography, tomography, CT scanning, and MRI

Cervical spine involvement has been seen in up to 86 per cent of patients with rheumatoid arthritis and atlantoaxial subluxation occurs in up to 40 per cent. Ligamentous laxity and rupture may be associated with erosive change around the peg due to local pannus. Instability may involve anterior displacement of C1 or posterior displacement of the peg. A conventional lateral radiograph taken in flexion should not result in a space between arch and peg of more than 2.5 mm in an adult. Superior and lateral displacement of the peg also occur.

Erosion of the lateral masses at C1–C2 allows the odontoid peg to shift upward. The distance from the anteroinferior surface of C2 to McGregor's line should be more than 34 mm in men and 29 mm in women (McGregor's line is a line drawn from the back of the hard palate to the lowest part of the occiput).

Changes at C1–C2 can be clarified further by linear tomograms in the anteroposterior and lateral projections and, subsequently, by CT scanning, which also demonstrates local soft tissue change. A further development was the combination of CT with radiculography, i.e. after cerebrospinal fluid opacification. Such scans could be obtained in the flexed and extended states. This was, until the advent of MRI, the investigation of choice (Fig. 22).

Fig. 22 CT radiculography. (a) Normal relationship of peg, arch of atlas, theca, and cord in the axial mode, together with sagittal reconstruction. (b) In this patient with rheumatoid arthritis the odontoid peg is eroded (see lateral reconstruction) and no longer has an intimate relationship with the arch of atlas. There is a large soft-tissue mass interposed between the two. The cord is displaced posteriorly and the theca impressed upon. (By courtesy of Dr J. Steven, National Hospital, Queen Square.)

MRI of the craniocervical region

MRI is now the investigation of choice for imaging the craniocervical region. Pannus, cerebrospinal fluid, bone, and cord are all shown (Fig. 23). Scans can also be obtained in flexion and extension, but there is probably no advantage to this over a single MR study with plain lateral flexion and extension views and tomography to show erosions. In addition, the status of the cord can be assessed at MRI for oedema, gliosis, or atrophy (Aisen *et al*. 1987; Semble *et al*. 1988; Einig *et al*. 1990).

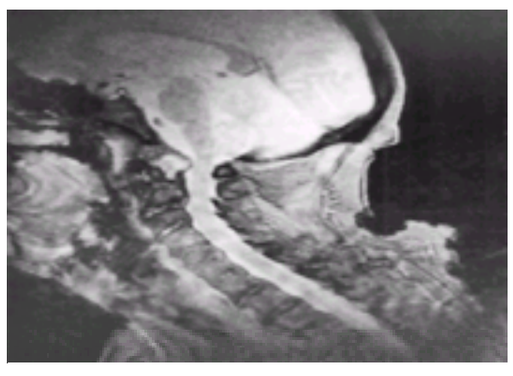


Fig. 23 Rheumatoid arthritis. MR sagittal T_2-weighted image shows erosion of the odontoid peg with pannus and compression of the cervical cord.

Imaging the spine

Anatomy of the lumbar spine

The lumbar spine is lordotic, convex anteriorly. In part this is because the lower lumbar vertebral bodies are greater in height anteriorly. At L5, the difference may be as much as 0.9 cm. At L1, however, the reverse is the case and the body of L1 may normally be wedged anteriorly. Similarly, the mid- and lower lumbar discs may be significantly thicker anteriorly (Farfan 1973).

Although it is often stated that the height of the L5/S1 disc is around 50 per cent of that at L4/5, in one small study 40 per cent of L5/S1 discs were thicker posteriorly that the L4/5 disc, and an even larger number showed the discs to be equal in height posteriorly.

With segmentation anomalies the intervening disc is rudimentary and diminished in height. Discs of smaller heights have diminished movement. The L5 disc may be high, i.e. at or just below the level of the iliac crest, or deep, i.e. much lower than the crest. Higher discs are said to be more vulnerable to degeneration. Disc degeneration in the lumbar spine has the highest incidence at the lumbosacral junction and the incidence decreases to be lowest at L1/2. If the lower lumbar disc is also sacralized, wholly or in part, i.e. associated with large L5 transverse processes, that disc will be small, 'protected', and less likely to be diseased.

The transverse processes at L5 may be the bulkiest and they may unilaterally or bilaterally articulate with, or be totally or partially fused to, the sacrum in some 15 to 30 per cent. Large transverse processes at L5 thus protect the disc below, but seem to be associated with increased levels of degeneration at the level above, which is thus high, 'free', and potentially unstable.

The sacrum also contributes to the lordosis as its upper surface is angled distoanteriorly. The L3/4 disc is aligned horizontally. Loss of the lordosis is shown by reduction of the angle that the lumbosacral disc makes with the plane of the L3/4 disc. Also, in a normal spine, the anterior surfaces of the vertebral bodies are always tangential to each other, while in a spastic painful or degenerate and immobile segment, adjacent anterior vertebral surfaces are in alignment.

The transverse processes are seen on a lateral view of the spine along a line drawn from the L1 transverse process to the anterior surface of the femoral head. On the anteroposterior film, the transverse processes differ in size and shape. The transverse process at L3 is often the most lateral and its inferior surface often the lowest to be inclined horizontally (Fig. 24). This is of benefit in trying to decide which lumbar vertebra is L3; of course the definitive way is to count down from C1.

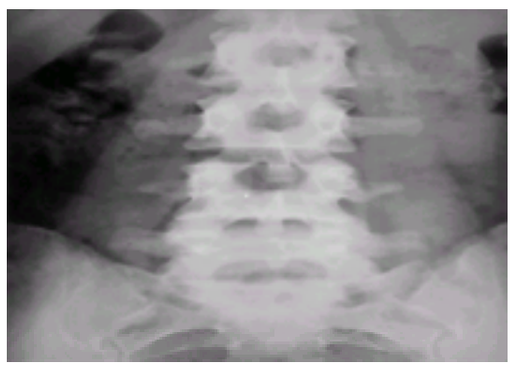


Fig. 24 Anteroposterior view of lumbar spine. The transverse processes of L3 are the lowest to be inclined horizontally.

The facet joints

The facet joints in the lumbar spine vary in their orientation. At L1, the facet joints have an almost vertical orientation but, passing inferiorly, change direction so that at L5/S1 the superior facet faces superiorly, medially, and posteriorly. Facets are generally symmetrically orientated. The angle of the facets to the midline is around 52° at the L5/S1 joint, and 10° less at L4/5. Facet asymmetry is also greatest at L5/S1. A high correlation has been reported between the side of increased facet rotation and disc protrusion with sciatica (Farfan 1973).

Facetal changes of osteoarthritis follow discal height loss. These changes are seen on plain films, especially with oblique views, and very well demonstrated at CT scanning, where bony hypertrophy and synovial thickening are seen to narrow the lateral recess and exit foramina (Fig. 25). Yellow ligament thickening is seen at axial imaging to indent the theca from behind, and posterolaterally.

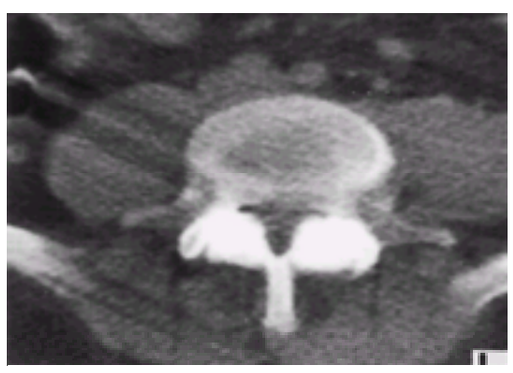


Fig. 25 A rather old scan, but a good example of facet hypertrophy causing lateral recess stenosis and a trefoil deformity.

Plain film demonstration of disc narrowing, facet slip and rotation, foraminal encroachment, and marginal new bone formation are all indicative of established discal degeneration.

Scoliosis associated with vertebral asymmetry is seen commonly on routine chest radiographs and occurs in up to 13 per cent of spines. It has been stated that these curves are due to primary vertebral asymmetry with a diminished body height on the concavity, as well as shorter pedicles and a flatter neural arch.

Pain sources in the spine

Pain may arise (i) in the motion segment and surrounding structures, comprising the discal nucleus and annulus, vertebral end-plates, and the two local facet joints, as well as surrounding ligaments and muscles; (ii) in the superficial structures around the spinous processes; and (iii) referred from spinal nerves and the sympathetic chain. Pain may arise in isolation but, because all three areas are interdependent, a complex pain pattern may result (O'Brien 1984).

The disc is the largest structure in the motion segment. The nucleus and inner annulus have no innervation but the outer annulus has a rich multilevel sensory nerve

supply. Similarly, the vertebral body has a rich sensory nerve supply.

The emerging nerve root occupies about 50 per cent of the exit foramen and so is compromised easily by changes in the local bone and soft tissue. The dura around the nerve is itself innervated anteriorly, but not posteriorly. The capsule of the facet joint (but not the synovium) has a three-level innervation. It is also likely that each anatomic level of the anterior and posterior longitudinal ligaments derive their innervation from the level above.

Patterns of spinal and referred pain are thus seen to be complicated because of the overlap of innervation (Fig. 26). Pain fibres are also present in the sympathetic chain which lies in close contact with the spine, and they too are affected by local disease, bony osteophytes, and lesions of the annulus (Fig. 27).

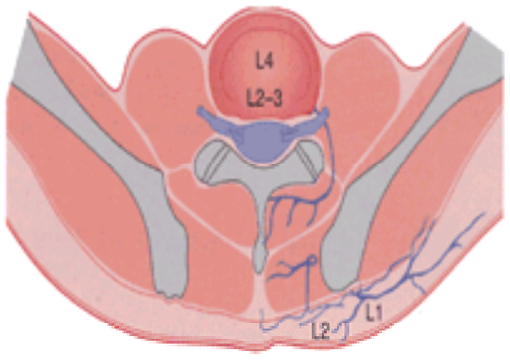


Fig. 26 Transverse section of the spine and associated structures at L4 to illustrate the different innervations of structures at the same anatomical level. (Reproduced from *Textbook of pain*, by courtesy of Mr J.P. O'Brien FRCS and Churchill Livingstone.)

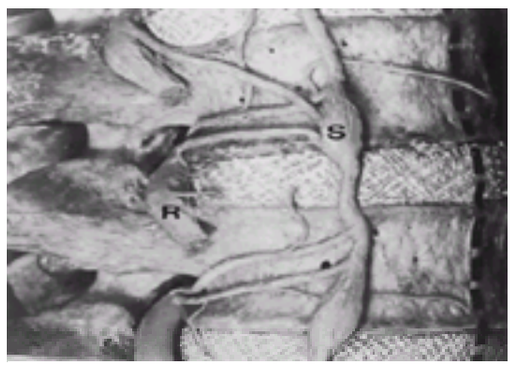


Fig. 27 Superficial view of lumbar spine demonstrating the nerve supply of the disc. Note the close proximity of the sympathetic chain to the disc and vertebral margin and also the communication between the nerve root and sympathetic chain just above the level of the disc. R, nerve root; S, sympathetic trunk. (Reproduced from *Textbook of pain*, by courtesy of Mr J.P. O'Brien FRCS and Churchill Livingstone.)

At least 50 per cent of the population suffers from back pain at some time. Abnormal anatomy predisposes to abnormal stresses on the disc. Discal degeneration is an inevitable fact of life; no spine remains unaffected. Discal arrowing causes alteration of facet alignment and subsequent degeneration. Exit foraminal encroachment by soft tissue and osteophytes compromise the emerging nerve roots.

Pathology in the disc

Many disc protrusions regress spontaneously with time and conservative therapy (Cowan *et al*. 1992). Diagnosis of degeneration based solely on MR or CT criteria often does not address the cause of pain and, indeed, inappropriate surgery may worsen the patient's condition.

Posterior annular circumferential fissures start to form in the lower lumbar spine as early as puberty (Fig. 28). Subsequently, the nucleus starts to dehydrate and radial tears occur, perhaps related to the orientation of facets.

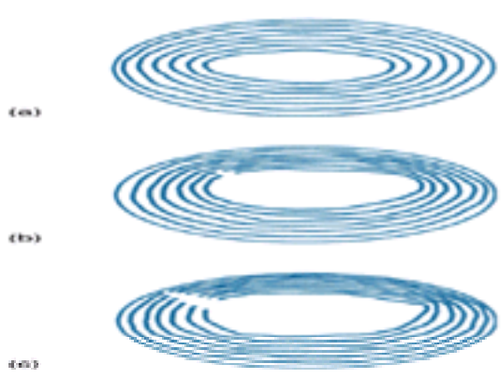


Fig. 28 Formation of radial fissure: (a) concentric laminae of equal number at all points around the circumference; (b) some of the inner lamellations become discontinuous and the clefts in the annulus become larger; (c) the annular clefts and the tears of the minor annular layers continue to form a radial fissure. (Reproduced from Farfan, H.F., *Mechanical disorders of the low back*, by courtesy of Lea and Febiger.)

Forces acting on the nucleus can cause both nuclear protrusion and annular fissuring, as well as end-plate fractures (Fig. 29 and Fig. 30); not to be confused with Schmorl's nodes, which are nuclear herniations through corticated defects in the end-plates which had transmitted blood vessels in infancy (Fig. 31).

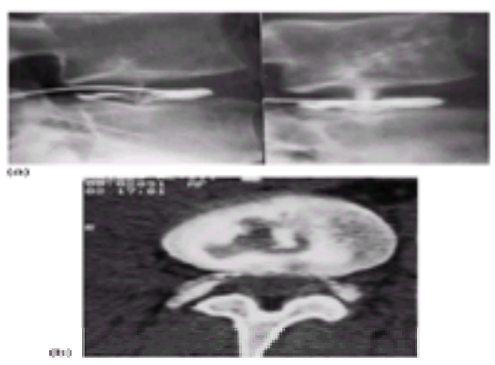


Fig. 29 End-plate fracture. (a) Discography. The initial radiograph shows a spurt of contrast going upwards from the nucleus into the end-plate of the overlying vertebral body. Subsequently, the defect is shown to be much larger, but patient movement occurred because of pain during injection. (b) CT scan showing contrast leaving the nucleus and entering the undersurface of the bone.

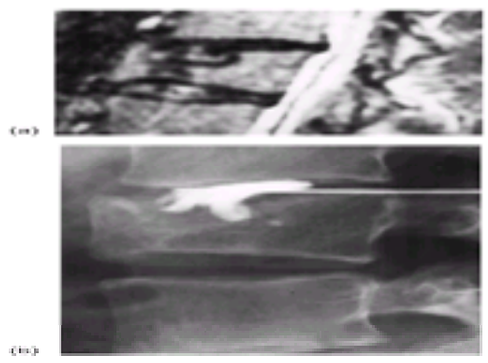


Fig. 30 End-plate fracture. (a) Defects are demonstrated on the upper surface of a lumbar vertebral body with depression of the cortex at MR scanning together with some loss of signal in the local disc. (b) At discography the defects are shown to fill with contrast injected into the nucleus.

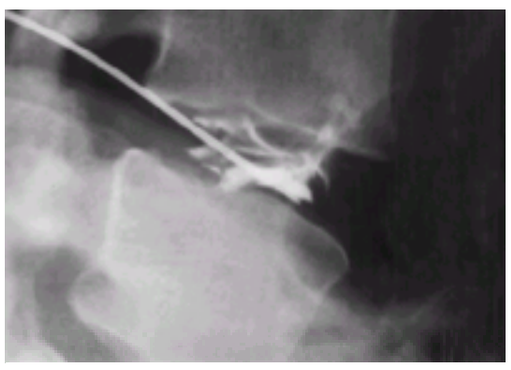


Fig. 31 Schmorl's node. At discography defects around the injected disc are demonstrated at the upper and lower end-plates.

Osteophyte formation

Osteophytes are seen at vertebral marginal edges and so-called traction spurs some 2 to 3 mm away from the vertebral margin. Osteophytes are intimately related to, and are usually indicative of, local annular tears (Fig. 32). Traction spurs may be related to change in the more superficial layers of the annulus and adjacent anterior longitudinal ligament.

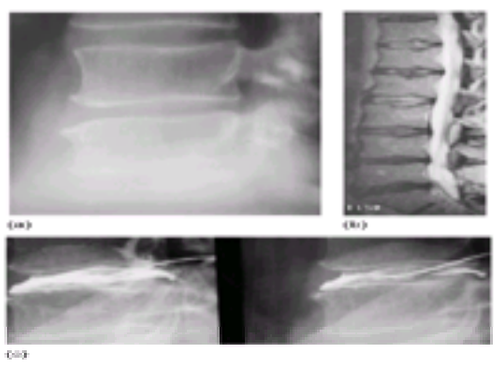


Fig. 32 Disc degeneration. (a) Plain tomography showing disc narrowing and anterior osteophytes. (b) MRI showing disc degeneration especially at L4/5. (c) Anterior and posterior annular tears at discography. Note vascular filling, presumably the result of superficial inflammation. Anteriorly, the contrast in the tear creeps over an osteophyte on the superior surface of the lower end-plate at that disc level.

Radiological investigation of low back pain

Griffiths (1991) has given us an algorithm for the radiological investigation of a patient with low back pain (Fig. 33). He recommends simple anteroposterior and lateral views as adequate for the first examination for the general practitioner or accident and emergency department, while the specialist may require further views. I habitually use lateral views in flexion to demonstrate loss of movement in both the cervical and lumbar regions.

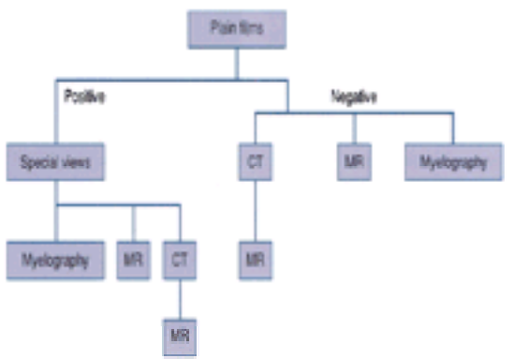


Fig. 33 Algorithm for radiological investigation. (Reproduced from *Imaging of the lumbar spine*, by courtesy of Dr H. Griffiths and Raven Press.)

Plain radiography

The plain film of the spine is rapidly obtained, cheap, and ubiquitously available. Much information can be obtained on the status of the motion segment. Is alignment normal and are disc heights preserved? Are there erosions, osteophytes, or traction spurs, indicating disc pathology? Rotational abnormalities can be assessed in both planes, together with an assessment of bony density.

End-plate density is lost with infection and increased with degeneration and instability. It is recognized, however, that up to 50 per cent of bone must be resorbed before foci of demineralization can be seen on the film. Marrow changes are thus better seen with isotope and MR scanning.

Myelography and radiculography

Intrathecal negative-contrast investigation of the spinal canal using air was first performed in 1921 and this was followed by the use of positive-contrast iodinated oil (Lipiodol/Myodil) in 1922. The bulging disc protrudes into the canal and causes a filling defect to be seen in the contrast. Oily media cause arachnoiditis and theoretically should have been removed after the study but, in practical terms, this was not always possible. Even though Myodil is very slowly absorbed naturally, remnants may be seen in the canal after many years. Early water-soluble media could only be used to examine lumbar nerve roots. If run higher up the body, fits and possible skeletal trauma resulted.

Using modern non-ionic media and fine non-bevelled needles, the incidence of meningism and other complications are lessened, but postlumbar-puncture headaches still occur, though these can be minimized if the patient is well hydrated by drip infusion, and given prophylactic cortisone in the drip. Only small volumes (10 ml of iohexol-240) need to be used to study lumbar roots. Unfortunately, only the areas reached by contrast can be investigated, i.e. the nerves are shown as far as the tip of the root sheaths, but even parts of these were not seen with oil-based media which were too viscid to enter the root sheath.

CT scanning

Axial scanning using CT added a new dimension to the conventional water-soluble contrast study. Delayed CT scanning allows total mixing of the contrast with cerebrospinal fluid. CT on its own demonstrates bone detail well and, in the axial mode, the discs, facets, lateral recesses, and pedicles are all clearly seen together with their relationships to theca and roots, even in the absence of intravenous contrast, so that CT alone has been used as a screening test, as a non-invasive study with no side-effects (other than those of a cumulative irradiation dose). If sufficient fine contiguous slices are obtained, reformatting allows visualization in any plane.

The examination is limited in extent by time, cost, and radiation exposure, so that a standard examination of the lumbar spine is usually only of the L3/S1 motion segments, and plain films are still needed to show higher levels, but most significant degenerative disease of the spine is below L3.

The combination of CT and radiculography gives better appreciation of the soft tissue interfaces, especially between disc, root, and theca, while at the same examination, the lateral recesses, facets, and yellow ligaments are better imaged than at radiculography. Some surgeons still request CT/radiculography, partly to show the recesses, as bone is better seen at CT, and also perhaps because of inherent conservatism (Fig. 34).

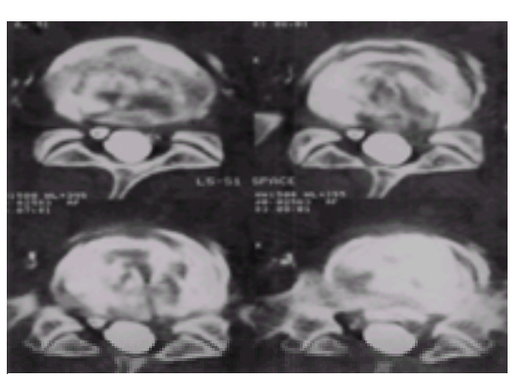


Fig. 34 CT discoradiculography. The degenerate disc is demonstrated with total loss of the normal nuclear/annular configuration. The bulging disc is demonstrated centrally and on the left side of the canal. It is associated with facet hypertrophy and lateral recess stenosis. The soft tissue of the disc is only partially opacified by contrast in the left lateral recess, but there is non-visualization of the S1 root on the left compared with its normal visualization on the right.

MRI of the spine

MR imaging of the spine has now largely replaced CT and radiculography. It is, as previously stated, non-invasive and allows multiplanar imaging. Metal implants in the spine sometimes contraindicate MRI but this technique can still be used to assess change in discs above, for instance, a fused level. In the postoperative spine, enhancement by gadolinium occurs with recurrent disc protrusions, but fibrotic masses do not enhance as they are relatively water-free.

Patients now undergo MR scanning as a primary investigation for back pain, and even plain films may be omitted. Apart from the fact that much can be learned from the plain film, the basic anatomy (see above) often cannot be inferred from a multisection scan—the wood cannot be seen for the trees. MRI shows changes as never before, and the number of positive diagnoses increases. There is apparently a resulting increase in the amount of spinal surgery performed the United States (Deyo 1994) but, as much recent work has shown, disc protrusion imaged at a particular level may not be the cause of the patient's pain as many such discs are asymptomatic (Jensen *et al.* 1994).

At MR scanning, bony changes are not as well demonstrated as at CT, though the low signal of cortical bone is shown. The disc height may be diminished, and nuclear signal decreased on T_2-weighted images. Disc protrusions, both posterior and anterior, are seen, especially against bright cerebrospinal fluid on T_2-weighted images (Fig. 35), and the displaced low-signal posterior longitudinal ligament on T_1-weighting.

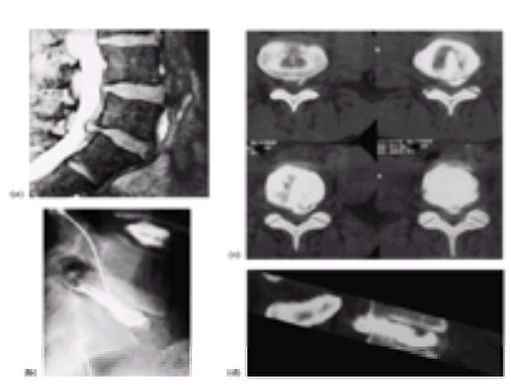


Fig. 35 Disc prolapse. (a) MR scan, T_2-weighted image. There is dorsal protrusion of a disc impinging upon the theca and narrowing the sagittal diameter of the canal. (b) The discogram demonstrates this lesion, showing a posterior annular tear with partial opacification of the bulge. A normal disc is demonstrated at the L4/5 level. (c) CT scanning following this study shows the degenerate disc with the central and left-sided protrusion. Sagittal reconstructions are also obtained (d).

Sagittal slices through the pedicles demonstrate the emerging nerve roots in exit foramina and show them against bright fat when normal, or impinged upon by soft

tissue or bony masses. Axial scans confirm these changes.

Discography

The role of discography in the diagnosis of spinal disc disease is controversial. Discography provides an image of disc anatomy and pathology. Discal injection also aims either to reproduce the patient's pain or to exclude that disc level as a pain source.

As far as disc anatomy is concerned, it might be stated with some justification that MRI demonstrates spinal anatomy and pathology more clearly, as (i) extradiscal structures are also displayed; (ii) the entire cervical, thoracic, or lumbar spine is imaged in one study; and (iii) the technique is non-invasive, not painful, and no risk is involved (these can include discitis or anaphylaxis at discography).

Prior to the advent of MRI, when plain films and myelography were standard techniques in the investigation of spinal pain, the limitations of myelography meant that discography had a perhaps greater diagnostic role. The relatively limited filling of root sheaths and opacification of roots meant that lateral disease could not be imaged. The theca is also not closely applied to the L5/S1 disc, but is separated from it by fat, so that the compressive effect of a protrusion at that level was not always seen.

MRI demonstrates disc degeneration and protrusion well, and multiple levels of often unsuspected abnormality may be seen. Discography can then indicate which particular disc or discs reproduce in part or whole the patient's pain. Thus, a pain-free level at discography, even in the presence of an MR abnormality, is an indication that this level need not necessarily be treated, while multiple levels of pain reproduction may indicate that more extensive surgery is indicated. In short, discography indicates which levels need or do not need intervention.

It has also become increasingly recognized that a painful disc can be associated with a normal MRI scan. Radial or circumferential tears of the disc or end-plate fractures may all be associated with pain in the presence of normal plain films, radiculograms, CT, and MR scans (it will be remembered that the outer fibres of the annulus are innervated) (Zucherman *et al*. 1988; Collins *et al*. 1990; Osti and Fraser 1992). MRI is said to be more sensitive than discography only in the diagnosis of disc herniation (Birney *et al*. 1992)—a finding that I am not wholly in agreement with.

A partial radial tear need not be associated with MR change and, as stated above, this tear antedates rather than follows degenerative change. It should also be recognized that interventional techniques, whether radiculography or discography, are not performed early on in the course of disease but in patients in whom delay, conservative treatment, and imaging have been of little benefit.

Normal and abnormal patterns at discography (Fig. 36)

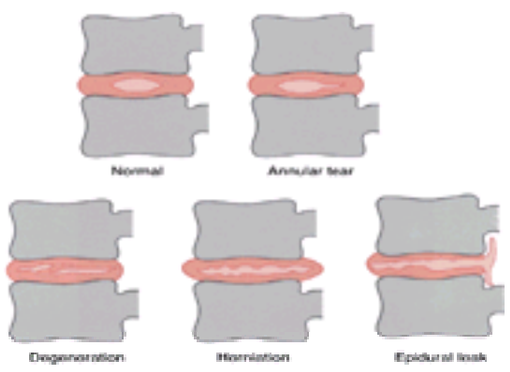


Fig. 36 Diagrammatic representation of the progression of discal degeneration, starting with an annular tear.

A plain film is an essential adjunct. Experience with discography leads to the realization that even so-called minor abnormalities on the plain film may be associated with discal pathology and pain. Osteophytes in particular, however small, can be associated with painful radial tears and do indicate underlying disc disease, as does loss of lordosis and minor malalignment on a lateral view, especially even minimal retrolisthesis of L5 on S1 (Fig. 37).

Fig. 37 An annular tear is associated with, and leads to, the osteophyte on the subjacent end-plate anteriorly. The osteophyte is the site of considerable reactive bone sclerosis. The disc is otherwise normal.

Obvious plain film changes are almost inevitably associated with discal pathology, which can be demonstrated at discography, but not inevitably with pain at injection.

MRI is said to be more sensitive than discography in the detection of herniation or prolapse of the disc because, of course, the posterior aspect of the disc is imaged directly. In practice, a torn annulus is associated with leak of contrast behind the disc, anterior to the displaced theca, displaying the bulge (Fig. 35), and very often vascular filling is demonstrated, presumably a result of chronic inflammatory changes around a degenerate disc (Fig. 29). Many studies are now in agreement that discography aids the diagnosis, plans treatment, and demonstrates pathology not shown by any other method. Discography may be combined with CT and even radiculography to give a total image of the diseased motion segment (Fig. 34) (Birney *et al*. 1990).

The spondylarthritides

New bone formation around the spine occurs in ankylosing spondylitis as well as in the other seronegative spondylarthritides. In ankylosing spondylitis, the changes may occur first at the thoracolumbar junction and spread in both directions. The new bone, or syndesmophyte, is usually vertically aligned in the annulus, gracile, and vertebromarginal (Fig. 38). This change follows the demonstration of vertebromarginal erosions which may be sclerotic and can give an impression of vertebral 'squaring' (Fig. 39). Non-marginal and floating syndesmophytes also occur in ankylosing spondylitis, but seem more common in the other seronegative spondylarthropathies (Fig. 40).

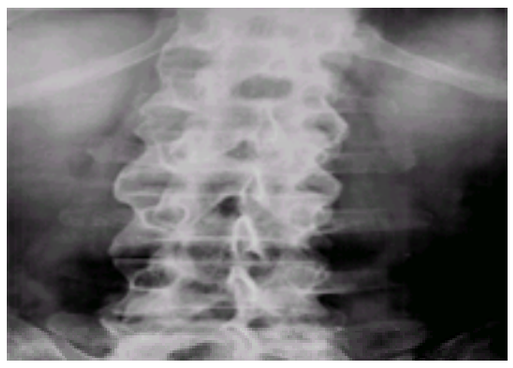


Fig. 38 Ankylosing spondylitis. Gracile, vertically orientated new bone runs between adjacent vertebral body margins.

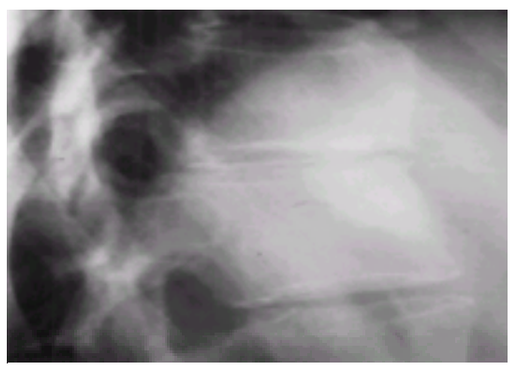


Fig. 39 Ankylosing spondylitis. Vertebral squaring results from margin erosion and also new bone laid down on the anterior concavity of the vertebral body.

Fig. 40 Psoriatic spondylitis. The presence of non-marginal and also 'floating' (non-attached) syndesmophytes is demonstrated in this seronegative spondylarthritis.

In the cervical spine, change occurs from C2 down, in continuity, and cervical spondylarthritis may be the presenting symptom in juvenile chronic arthritis and ankylosing spondylitis.

Ankylosing spondylitis

Paradiscal ossification in ankylosing spondylitis arises when the patient is relatively young, i.e. before discal degeneration. Disc heights are often then well-preserved, even in the elderly. On the lateral view of the spine in ankylosing spondylitis, the disc shows increased density, partly because of marginal new bone, but also because of discal nuclear calcification. This is seen as an added central density and occurs in other situations where there is bony fusion across discs (Fig. 41).

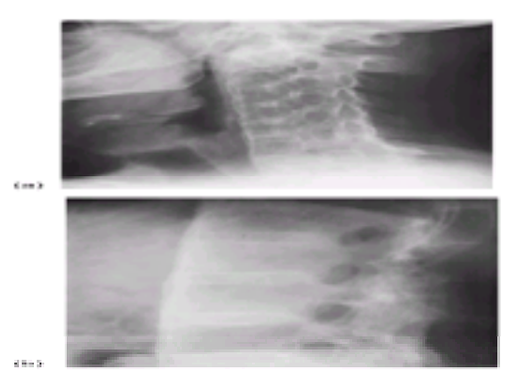


Fig. 41 Ankylosing spondylitis showing bony bridging with nuclear calcification.

Ankylosing hyperostosis (diffuse idiopathic skeletal hyperostosis, DISH; Forestier's disease)

In ankylosing hyperostosis, the new bone is accreted later on in life, is much more extensive, thicker, and florid, and is not necessarily seen in association with degeneration (Fig. 42).

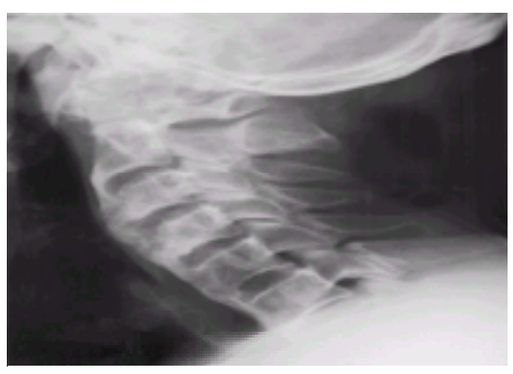


Fig. 42 Diffuse idiopathic skeletal hyperostosis. Preserved disc spaces are demonstrated in association with anterior ankylosis across many segments with a continuous bar of anterior new bone formation, which considerably increases the sagittal diameter of the vertebral body complex.

Paraspinal new bone can be seen on plain films and, as expected, on CT scanning. In ankylosing hyperostosis, it is especially well demonstrated in the thoracic spine on the right side on CT scanning, the left being spared presumably because pulsation in the descending aorta prevents its formation (Jones *et al*. 1988).

Ossification of the posterior longitudinal ligament (OPLL)

New bone formation is also seen in the distribution of the posterior longitudinal ligament of the spine as an isolated phenomenon (ossification of the posterior longitudinal ligament; OPLL), often in the Japanese race, but also in ankylosing hyperostosis. The bone here, as elsewhere, is well shown on CT scanning (Fig. 43) (Resnick *et al*. 1978; Tsuyama 1984).

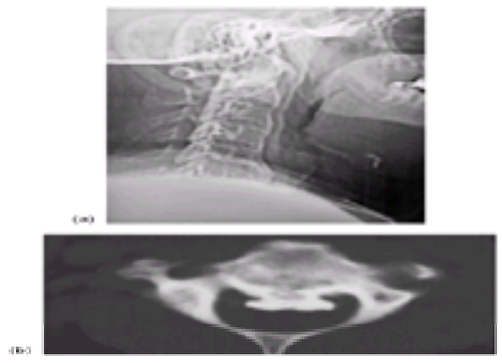


Fig. 43 Ossification of the posterior longitudinal ligament. To show the presence of new bone formation applied to the posterior aspect of the vertebral bodies of C2 and C3, considerably narrowing the sagittal diameter of the canal and indenting the theca.

Where bone formation is gross, increase in uptake of isotope is also seen in the spine in ankylosing hyperostosis and ankylosing spondylitis.

Imaging of the sacroiliac joints

Erosive change at the sacroiliac joints is the pathognomonic feature in ankylosing spondylitis. In children, the sacroiliac joints are poorly defined as the cortex is not readily visualized and, as a result, the joint seems wider than in the adult. Also, at isotope scanning, the images of the adolescent sacroiliac joints are more prominent—'hotter' and wider than in the adult. Fortunately, the sacroiliac joints are not normally affected early on in juvenile ankylosing spondylitis but, even so, the plain film diagnosis of erosive disease can be difficult to make in patients under 18 years of age.

In children and adults, infection of the sacroiliac joints is usually unilateral (Fig. 44). As with all infections, clinical onset of disease can antedate radiological change by up to 2 weeks. Gut gas, especially in sick children, also obscures the joint on plain radiographs.

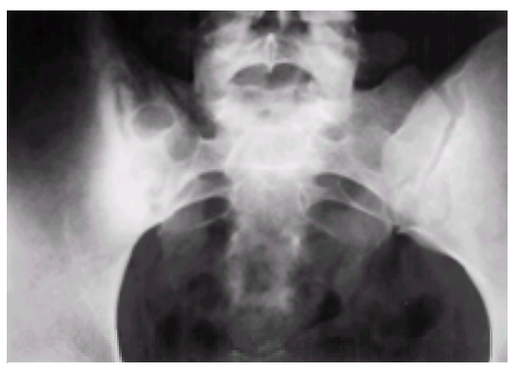


Fig. 44 Unilateral sacroiliiitis in tuberculosis.

Plain radiography

Conventional imaging is a simple anteroposterior radiograph centred on the sacrum (Fig. 45). This is probably the least effective means of imaging the joint as the beam diverges from, and the joint converges to, the midline. It is therefore more logical to image the sacroiliac joint prone, as the beam then passes down the joint space parallel to the articular surfaces. Oblique views and a 30° shoot-up anteroposterior image demonstrate the sacroiliac joint cortices much better than a single anteroposterior vertical beam.

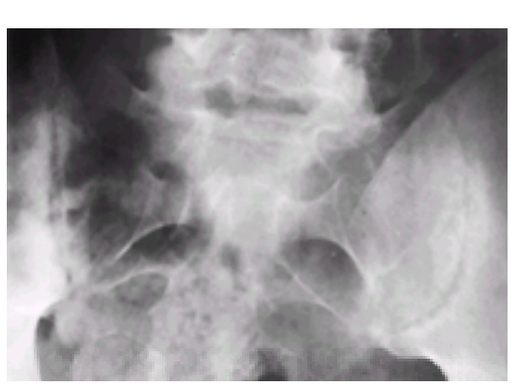


Fig. 45 Bilateral erosive sacroiliiitis in ankylosing spondylitis.

Linear tomography

Linear tomography is generally available and images are very satisfactory but it has been superseded by CT scanning (Fig. 46). CT scanning defines cortical bone

well; MRI is superior for assessment of soft tissues.

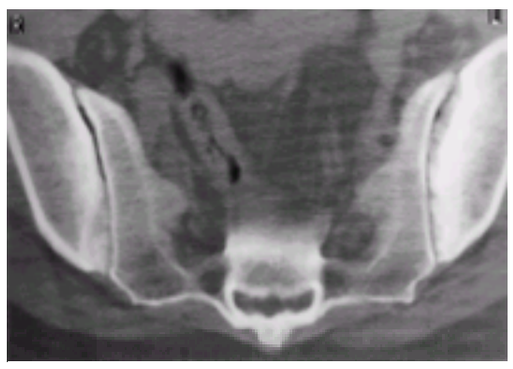


Fig. 46 Ankylosing spondylitis. Erosive sacroiliitis at CT scanning. There is narrowing of the sacroiliac joints bilaterally with erosive changes, initially confined to the lateral side of the joint. Reactive sclerosis is demonstrated in the underlying bone.

CT scanning

CT scans are useful in assessing the sacroiliac joints as the images are free of overlying gut shadows and the articular cortices can be assessed for erosions and fusion and subcortical cysts. Detail is generally better with CT scans than with plain films and, by altering windows, small erosions can be visualized (Fig. 46). Only a few images need be taken of what is quite a long structure as excess irradiation is to be avoided. Four to five 10 mm slices through the true joint at the mid- and lower thirds will generally enable the presence of erosions to be confirmed (Fig. 46).

Radionuclide bone scanning

Radioisotope scanning gives poor spatial resolution of sacroiliac joint disease but demonstrates pathology, albeit in a non-specific way by showing increase in uptake unilaterally or bilaterally. Unilateral sacroiliitis should be seen clearly because of asymmetric increase in uptake. Bilateral change can be more difficult to assess if uptake is symmetrically increased. On the delayed scan an increase in the ratio of counts between sacroiliac joints and sacral body of more than 1.4:1 was held to be significant (Fig. 47) (Ho *et al*. 1979), but it seems that there is too great a range of variation in the normal population for this test to be sufficiently accurate (Dequeker *et al*. 1978; Goldberg *et al*. 1978; Forrester 1990). Quantitative assesssment of increased blood flow to affected joints may be more sensitive. CT is probably better at defining bony change, even with unilateral disease.

Fig. 47 Ankylosing spondylitis. Radioisotope bone scan to show the sacroiliac joint/sacral uptake ratios. There is greater uptake in the right sacroiliac joint than on the left. The mean uptake on both sides is given, together with the mean uptake in the sacrum. (By courtesy of Dr A. Hilson, Royal Free Hospital.)

Fusion is a well-recognized sequel to endstage sacroiliitis—bilateral in ankylosing spondylitis but often unilateral in other seronegative spondylarthritides and following tuberculosis and other infections. Fusions at the sacroiliac joints are also seen in the elderly in whom no other evidence of ankylosing spondylitis exists, often in elderly females who might not be expected to have had sacroiliitis. It is also seen in patients with ankylosing hyperostosis, due to ligamentous and osteophytic ankylosis in the absence of erosions (Fig. 48) (Durback *et al*. 1988), and in patients with Paget's disease, thalassaemia, and X-linked hypophosphataemic osteomalacia.

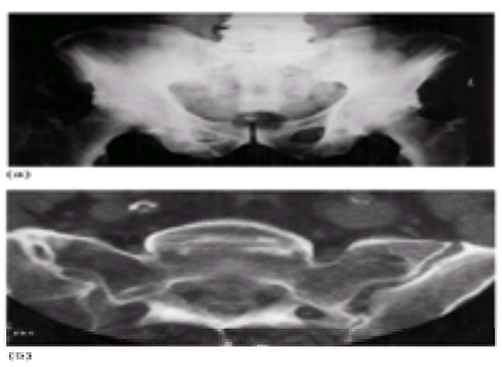


Fig. 48 Diffuse idiopathic skeletal hyperostosis. (a) There is new bone formation on the iliac crest and, in particular, superiorly at the sacroiliac joints, giving an impression of fusion across these. (b) CT scan to show new bone formation anteriorly across the sacroiliac joints in another patient with DISH. There are no erosions.

Imaging of joints

The shoulder

This topic has recently been reviewed (Gold *et al*. 1993). Ninety-five per cent of rotator cuff tears result from impingement.

Plain radiography

The presence of an impinging subacromial osteophyte and upward subluxation of the humeral head can be seen on the plain film. In the evaluation of impingement, an elevated humeral head on the plain film is compatible with tendinous narrowing or total rupture (Fig. 49).

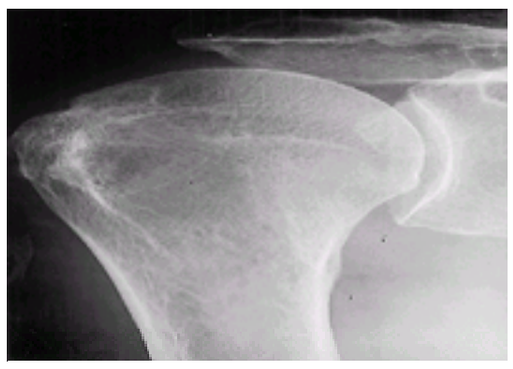


Fig. 49 Plain film prior to arthrography shows upward subluxation of the humeral head. Note the discontinuity of Mahoney's line (equivalent to Shenton's line at the hip). The undersurface of the acromion is irregular and there is a distal subacromial osteophyte.

Arthrography

The question of the presence or absence of a rotator cuff tear is answered easily at arthrography (Fig. 50 and Fig. 51), while restrictive capsulitis is shown at arthrography by demonstration of the contracted capsular outline and volumetric assessment of capacity (Fig. 52) (Resnik *et al*. 1984; Renton 1991a).

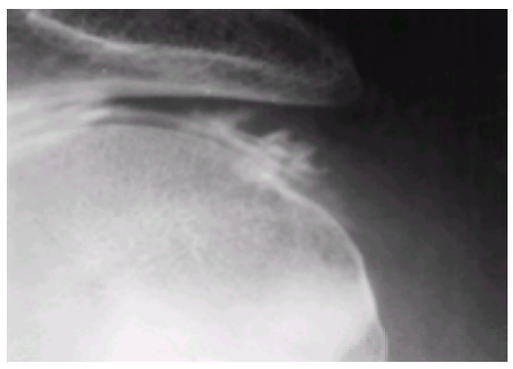


Fig. 50 Arthrogram of the shoulder showing partial tear of the rotator cuff. Irregularity of the inferior surface of the rotator cuff tendon is demonstrated. Contrast fills the tendon but does not reach the subacromial space.

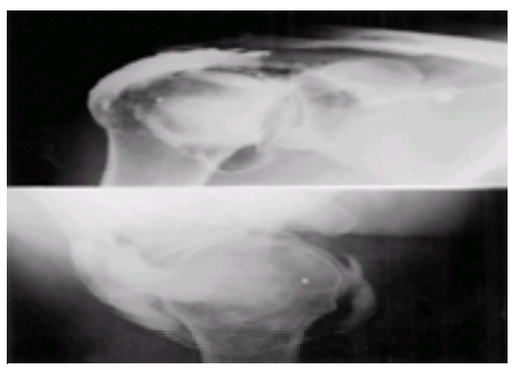


Fig. 51 Rotator cuff tear. Contrast medium now fills the subacromial bursa in continuity with the shoulder joint proper. On the axial view, the contrast medium in the subacromial bursa overlies that in the joint space proper.

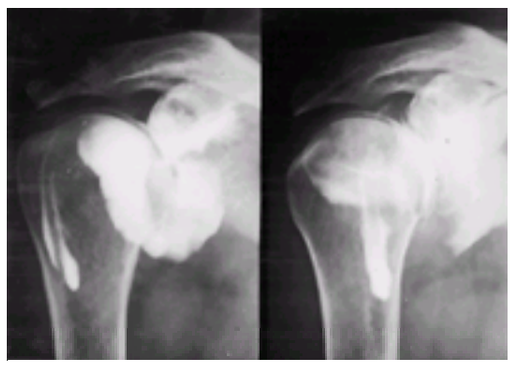


Fig. 52 Restrictive capsulitis. The capsule is reduced in size and is irregular. The long head of the biceps tendon sheath is filled as contrast medium seeks to escape from the tight shoulder joint capsule.

MR imaging

This is said to have 100 per cent sensitivity and 95 per cent specificity in the diagnosis of complete tears, while in the differentiation of a normal tendon from one affected by tendinitis with impingement, the sensitivity and specificity are 93 per cent and 87 per cent, respectively. For labral tears, these are also around 90 per cent (Iannotti *et al*. 1991). Changes of hypertrophy of bone and soft tissue at the acromioclavicular joint are seen to cause compression of the underlying supraspinatus, often associated with oedema (Fig. 53). The demonstration of a low-signal subacromial spur or osteophyte directs attention to the subjacent tendon, which may be thickened and oedematous, thinned, frayed, or frankly torn and distracted, allowing the humeral head to sublux upwards (Fig. 54).

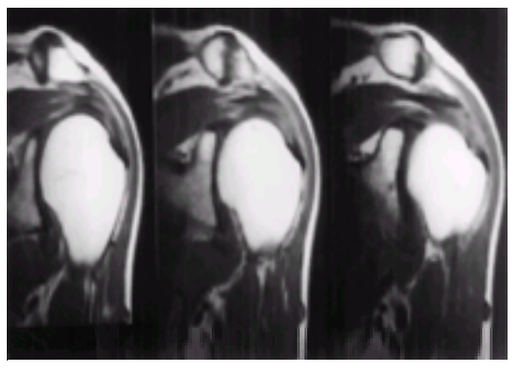


Fig. 53 MR scan of the shoulder. Hypertrophy of the acromioclavicular joint causes changes of oedema in the subjacent supraspinatus tendon. More distally there is a tear of the rotator cuff tendon shown as an area of increased signal replacing the normal low signal of the tendon.

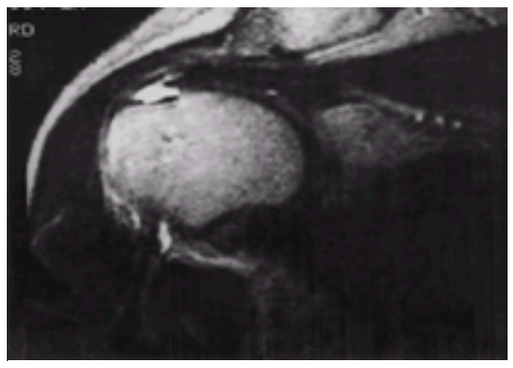


Fig. 54 MR scan of the shoulder. The tendon is seen to be retracted and fluid occupies the gap just proximal to the insertion into the greater tuberosity in a patient with a total rupture of the rotator cuff tendon.

Fluid is seen clearly on T_2-weighted or STIR sequences in the tendon, as well as in degenerative bone cysts at the greater tuberosity, and in the subacromial bursa. The deltoid may be atrophied.

Axial scans show the long head of the biceps and its relationship to the bicipital groove. Fluid may be present in the tendon sheath. The glenoid labrum is clearly seen in the axial plane (Fig. 55).

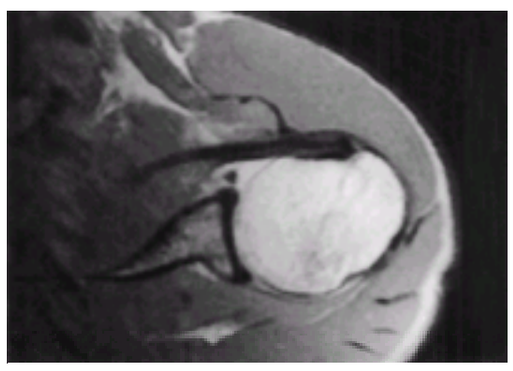


Fig. 55 Anterior labral tear. The detached portion is clearly demonstrated separately from the rest of the glenoid and the biceps tendon is showing entry into the bicipital groove.

CT arthrography/ultrasound

CT arthography does demonstrate these findings but is invasive and has been superseded by MRI, as has arthrography (see above), while ultrasound is operator dependent and the images may be difficult for others to interpret.

Necrotic changes or tendinitis in the supraspinatus and its tendon, on the other hand, are diagnosed at MRI but not at arthrography. The long head of the biceps tendon is seen at arthrography and, if CT is available concurrently, axial imaging by CT postarthrography will show the long head of the biceps, the glenoid labra (Fig. 56), and a possible Hill–Sachs or hatchet defect after recurrent dislocation (Fig. 57) (all, of course, are also seen at MRI) (Kieft *et al*. 1988).

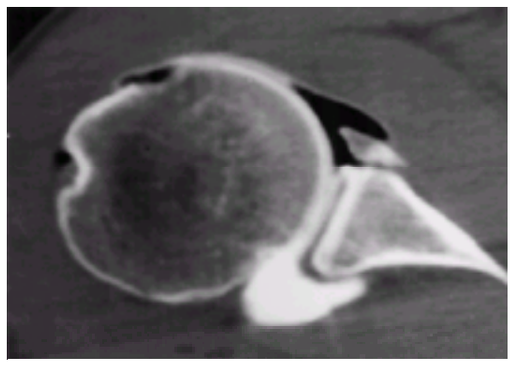


Fig. 56 CT arthrogram following trauma. The appearances are almost identical to those seen in Fig. 55. There is avulsion of the anterior lip of the labrum and the underlying corner of the glenoid.

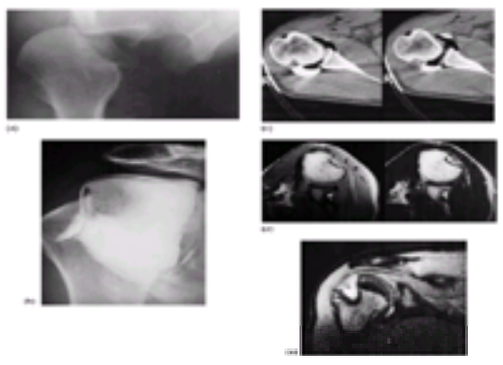


Fig. 57 Hill–Sachs or hatchet defect. Following anterior dislocation of the humeral head, especially if recurrent, a defect arises posteriorly following impingement upon the anterior lip of the glenoid. This may be demonstrated (a) on plain film, (b) at arthrography, (c) at CT, and (d,e) at MR scanning.

Arthrography is still in use at the shoulder as a quick, cheap screening test, though no doubt it too will be almost entirely superseded by MRI.

The elbow joint

Plain films have been discussed above.

Arthrography

Arthrography of the elbow has always been of limited use. It has been employed mainly for the diagnosis of loose bodies. Osteochondritis dissecans can also be shown. Unfortunately, the introduction of even the smallest of bubbles into a small joint makes the diagnosis of a loose body rather difficult.

Arthrography at the elbow can be single or double contrast. Using a single contrast technique, the contrast used must be diluted substantially as a dense contrast medium will hide a small lucent loose body, such as a fragment of cartilage avulsed into the joint. A double contrast technique, on the other hand, tends to cause bubbling, so that the diagnosis of a loose body is again made difficult. For this reason, a single contrast technique followed by conventional linear tomography has been utilized.

CT scanning

Opaque bodies in the joint are probably better demonstrated using CT. The cortex in particular of a detached loose body would be better shown by this technique. It may be that a single contrast technique using air alone might be more helpful or a double contrast technique using air and a little bit of contrast. In any case, with the advent of arthroscopy, it is likely that the use of arthrography, which never was very popular, will decline further.

MR imaging

There is no doubt that MR imaging of the elbow joint represents a substantial advance in the diagnosis, not merely of abnormalities of bone, articular cartilage, synovium, and joint space, but also in showing the related structures around the elbow joint. Fat suppression studies, especially of the elbow, are useful in showing changes of a local tendinitis at the medial and lateral condyles and adjacent flexor and extensor origins. As is generally the case, changes within the joint at the elbow are well shown at MR imaging.

The wrist

Plain films have been dicussed above.

Arthrography

Arthrography has been used at the wrist to demonstrate the integrity of the triangular cartilage. The radiocarpal joint is injected with contrast and a tear or defect in the triangular cartilage is then demonstrated by the passage of contrast medium into the distal radioulnar joint. Unfortunately, this tends to occur in around 15 per cent of normal individuals and tends to increase with age.

In addition, the integrity of the midcarpal joints is shown by failure of opacification of the mid- and distal carpal joints following radiocarpal injection. Opacification, therefore, of the distal joints of the carpus indicates that there is compartmental ligamentous damage.

CT scanning

CT scanning of the wrist has some use in the demonstration of rheumatoid change. In theory, it could also be used to demonstrate the integrity of the triangular cartilage, but only after the use of injected contrast.

MR imaging

MR imaging of the wrist is of course of great use in rheumatoid arthritis (see earlier section). In addition, however, MRI is used to demonstrate the integrity of the triangular cartilage and no doubt replaces arthrography where available.

On coronal images the cartilage should be seen extending from the ulnar styloid to the radius, on the radial aspect of the distal radioulnar joint. Failure to do so may indicate avulsion. Bright signal on T_2 images, and especially on fat suppression studies, demonstrates either internal degeneration of the meniscus or a complete tear.

The carpal tunnel syndrome is investigated using MR imaging. Bowing of the flexor retinaculum distally, oedema of the contents, and especially compression of the median nerve, which can be identified separately within the tunnel, are all indicators that the carpal tunnel syndrome is present. The flattened and compressed nerve becomes enlarged, both above and distal to the abnormal area.

The hip joint

Confirmation or exclusion of septic arthritis

The presence of fluid in the joint can be seen on a plain radiograph (Guerra *et al*. 1978). Ultrasound and MRI are more sensitive in diagnosing hip effusions and, using either fluoroscopy or ultrasound, the joint can be aspirated and contents analysed.

Arthrography

In both adults and children arthrography can be used after aspiration in the diagnosis of septic arthritis. Though contrast is said not to be bacteriostatic, joint fluid should be aspirated before contrast injection. Contrast demonstrates the state of the synovium, shows its contraction and irregularity, and also outlines areas of bone and cartilage destruction (Fig. 58).

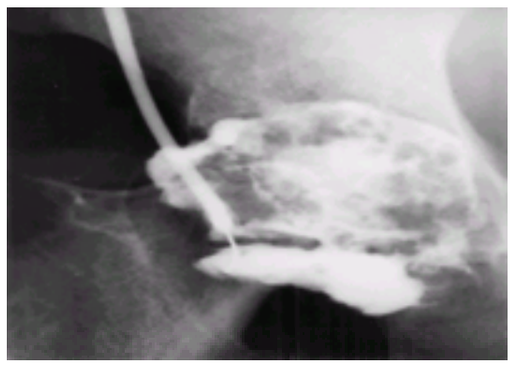


Fig. 58 Tuberculosis of hip. The presence of an irregular shrunken synovium is demonstrated together with erosion of the underlying bone at arthrography.

Arthrography is still very much in use in the assessment of prosthetic loosening and infection. Again, the joint is aspirated, the specimen sent for microbiological investigation, then contrast, usually around 20 ml, injected. Loosening is shown by contrast filling the lucent gap between bone and cement or prosthesis. The width of this gap, if present at all, should be less than 1 mm or less.

Tumours of synovium are infrequent. Benign lesions are pigmented villonodular synovitis (Goldman and Dicarlo 1988; Jelinek *et al*. 1989) and synovial osteochondromatosis. In pigmented villonodular synovitis, a distended capsule is associated with synovial proliferation and bone scalloping (Fig. 59). The joint space is preserved initially as cartilaginous destruction is a later phenomenon. Aspiration may give a serosanguinous fluid. Synovial chondromatosis is associated with multiple (four or more) pearl-like chondral tumours, either floating loose, or in the fronds of proliferating synovium (Fig. 18). When these pearls ossify, they may be seen on a plain film. Often the ossification is total, but a more irregular pattern may be found. Arthrography confirms the major radiological features—erosions, synovial proliferation, and loose bodies. Aspiration may give a clear fluid.

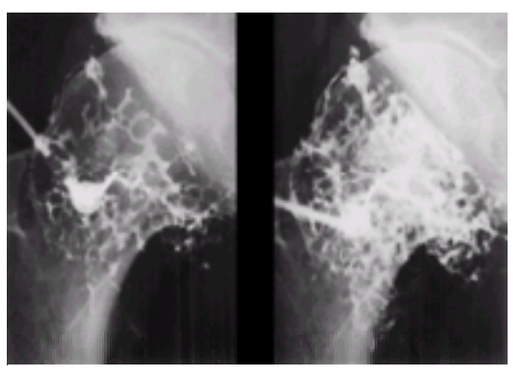


Fig. 59 Pigmented villonodular synovitis. At arthrography the frond-like proliferation of the synovium is demonstrated. This is associated with corticated erosion of the underlying bone.

With malignant synovioma, a patchily ossifying articular mass is associated with bone invasion.

The knee joint

It is easy to understand why MRI of the knee has superseded arthrography. This latter technique needs an experienced operator and preferably a fine-focus X-ray tube, which is not always available. MRI not only demonstrates superficial meniscal lesions, i.e. those to which contrast and air would have access, it also shows change within the meniscus not otherwise accessible to X-rays, as well as demonstrating the cruciate and collateral ligaments and the other extra-articular structures (see below).

Use of MRI

Accuracy rates are of the order of 95 per cent for the diagnosis of meniscal lesions (Mink *et al*. 1988), while the reported accuracy of arthroscopy is similar (Table 2) (Ireland *et al*. 1980; Thijn 1982). Fluid in the meniscus is shown as a band, or disc, of increased signal in the meniscus and may be graded accordingly (Fig. 60) (Mink and Deutsch 1990).

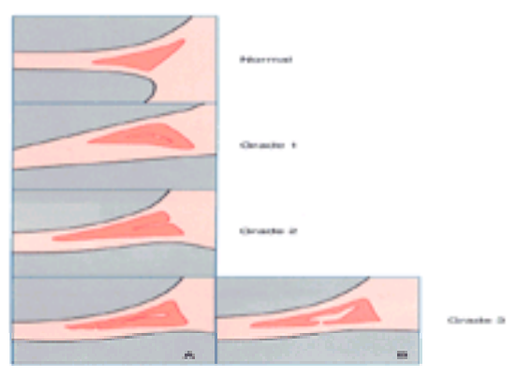


Fig. 60 Classification of meniscal change at MRI from normal to tear, according to Mink *et al*. (1993).

Areas of increased signal may be seen in adolescents and in older asymptomatic knees and are usually the result of myxoid degeneration (Stoller *et al*. 1987). Should the bright signal extend to the meniscal surface, a tear should be diagnosed and would be seen at arthroscopy. Equivocal or doubtful extension to the articular surface of the meniscus is responsible for false-positive diagnoses and, if in doubt, a tear should not be diagnosed (Kaplan *et al*. 1991).

Studies on American football players have shown progression from grade II to grade III lesions during the season, so that extension of intrameniscal degeneration to tear occurs (Reinig *et al*. 1991). Discoid lateral menisci and tears in them are also easily recognized (Fig. 61). Collateral ligament lesions in particular are well defined and disruption of the ligaments, local haemorrhage, and subsequent fibrosis may all be seen. The relationship between meniscal tears and degenerative changes in the underlying bone is clearly demonstrated (Fig. 62).

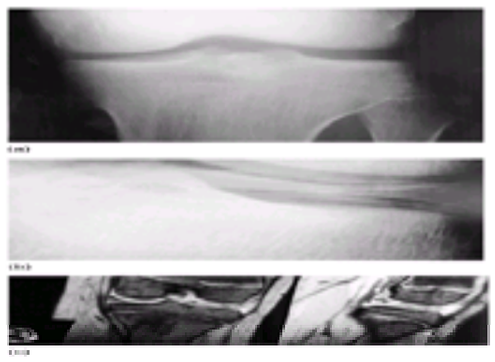


Fig. 61 Discoid meniscus. (a) The plain radiograph shows widening of the lateral compartment and often dishing of the lateral tibial plateau due to the mass effect of the enlarged mensicus. (b) On arthrography, instead of the normal triangular shape of the meniscus, its internal aspect is elongated and expanded. (c) The MR scan of a discoid meniscus shows the expansion of the meniscus, but also increased signal in it. This is extensive and presumably indicates widespread internal degeneration of the meniscus.

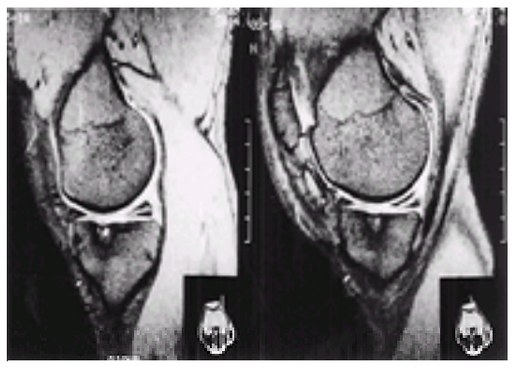


Fig. 62 MR scan of the knee showing a tear in the posterior horn of the medial meniscus and associated degenerative changes in the tibial plateau.

MRI is said not to be as accurate as arthroscopy in the diagnosis of stages I and II of chondromalacia patellae, but compares favourably for stages III and IV (Brown and Quinn 1993).

Arthroscopy was always more accurate than arthrography in the diagnosis of cruciate ligament tears (Table 2(b)). MRI demonstrates changes in the cruciate ligaments. The posterior ligament is better imaged as it is larger, while the anterior, besides being thinner, is also inclined at an angle to the midsagittal plane. It is thus necessary to image the anterior cruciate ligament with the limb in around 15° of external rotation.

Tears, in the acute phase, may be partial or rupture of a cruciate ligament may be total and are associated with fluid seen with T_2-weighted or STIR sequences. The tear may be seen, or total retraction may occur, so that the ligament is no longer visible along its normal course (Fig. 63).

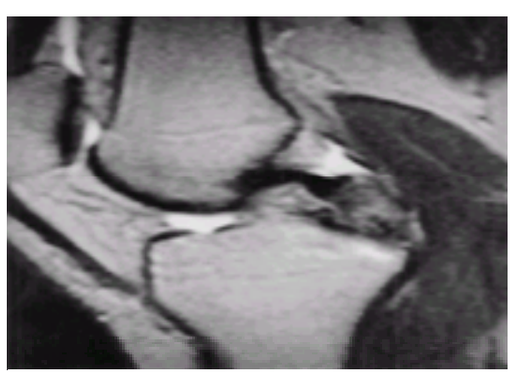


Fig. 63 MR scan of the knee. The posterior cruciate ligament is avulsed and has also taken the subadjacent cortical bone from the upper tibia with it. Note the oedema in the underlying tibial plateau.

When chronic, the torn ligaments fibrose and have an abnormal contour but need no longer be oedematous. The accuracy of MRI in the diagnosis of cruciate ligament lesions is up to 100 per cent in the literature (Lee *et al*. 1988; Mink *et al*. 1988).

The ankle

Arthrography at the ankle has been used to demonstrate tears of the collateral ligaments, even acutely after injury.

Tenosynovitis at the ankle tendons has been diagnosed by contrast injection of the tendon sheath, which is performed easily by direct puncture of the palpated tendon (Fig. 64). Tendon rupture, thickening, and synovial irregularity may all be seen and the images complemented by CT scanning but, again, this has been superseded by ultrasound and, especially, MRI (Fig. 65).

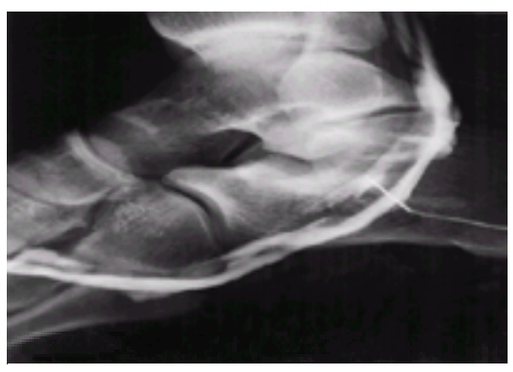


Fig. 64 Tenosynovitis shown at arthrography; peroneal tenosynogram. The irregularity of the synovial sheath is demonstrated in synovitis. The filling defect in the middle of the sheath is of course the tendon.

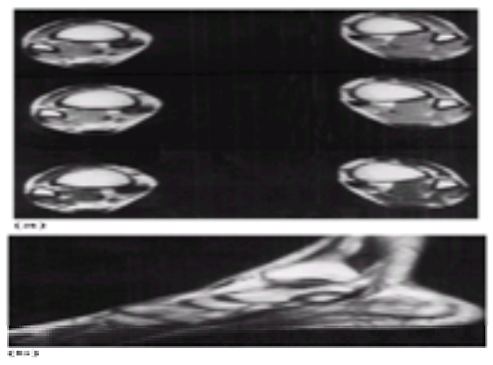


Fig. 65 Ruptured tibialis posterior tendon; sagittal and axial MR images. (a) The axial slices show both ankles. On the normal left side, the tendons behind the medial malleolus are well defined and compact. On the right, there is obvious thickening. This is due to retraction and the changes are easily identified then on the sagittal image (b).

MRI is good for evaluating structures at and around the ankle joint, including tendons and ligaments, as well as articular surfaces and the joint space.

Imaging of the hind foot

Pain in the heel is a recognized feature of seronegative and seropositive arthritis, described by Bywaters (1954). Erosions arise in varying sites around the heel, at target areas for the various arthritides, as described by Resnick *et al*. (1977). Erosions may be associated with a retrocalcaneal bursitis, seen on a soft-tissue lateral radiograph of the foot (Fig. 8).

Thickening of the distal Achilles tendon is also shown. Erosions occur above, at, and below its insertion on the posterior aspect of the calcaneus and will be associated with local increase in uptake on a radioisotope bone scan. Similarly, erosions form on the plantar surface of the calcaneus in the region of the origin of the plantar fascia in the seropositive and seronegative arthritides.

CT scanning demonstrates the bony erosions and can also show tendinous thickening, but the soft tissue changes are clearly demonstrated at MR imaging.

Resnick *et al*. (1977) demonstrated that 22 per cent of normal individuals have plantar spurs, which increase in incidence with age. Normal spurs are 4 mm or less in size and smooth in outline. Rheumatoid patients may have benign-looking spurs which may be painless, but fluffy irregular spurs occur in seronegative and seropositive patients, often associated with periostitis on the more distal inferior surface of the calcaneus, retrocalcaneal recess abnormalities, tendon thickening, and erosions. Benign spurs may become irregular with progression of disease and, similarly, irregular spurs may become smooth with remission of disease.

Spur formation is thus seen as a changing phenomenon in the long term. Larger, denser, and fluffier plantar spurs are strongly suggestive of Reiter's syndrome, in which posterior calcaneal spurs are rarely, if ever, found (Fig. 66) (Mason *et al*. 1959).

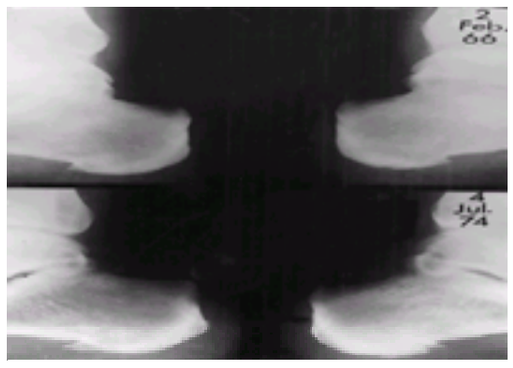


Fig. 66 Reiter's syndrome. A large, fluffy plantar periostitis is shown.

The tendo-Achillis and plantar fascia can be seen on a satisfactory conventional lateral radiograph, and the tendo-Achillis is also defined on CT (axial) images. The isotope scan is positive in the presence of erosions or enthesitis at musculotendinous insertions and may be positive in the early phase of the disease, before change is seen on CT or MR scans (when only the periosteum is affected).

MRI is of great value in defining soft tissue abnormality, showing thickening of tendons, local bursitis and oedema in tendon sheaths, plantar fascia, and calcaneus. These changes are seen on T_2-weighted images and especially fat suppression (STIR) sequences. Tears and tendinitis in the tendo-Achillis are especially well shown at MR scanning (Fig. 67).

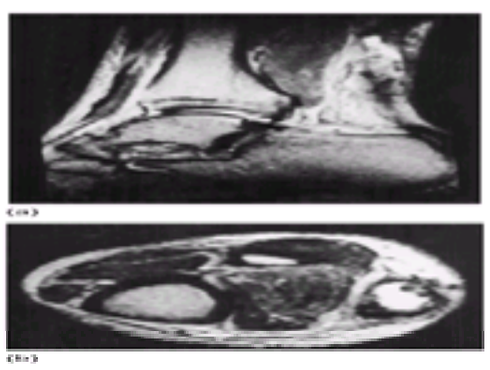


Fig. 67 MRI of the tendo-Achillis showing a degenerate and ruptured tendon. The capsule around the tendon seems intact but the space between the two frayed ends is filled with fluid.

Os trigonum

The os trigonum is seen in between 3 and 15 per cent of feet. It is almost certainly a tarsal accessory bone, rather than an old non-united fracture of the normal posterior process seen in 38 per cent of feet.

The ossicle is demonstrated clearly on a lateral view of the foot. It may be the cause of symptoms of pain and tenderness due to impingement in plantar flexion in

footballers or ballet dancers. The normal ossicle is smooth and well corticated, as indeed is the normal posterior tubercle, but with a normally lucent central medulla.

With chronic impingement, central sclerosis and cortical irregularity with pitting and local soft-tissue oedema result. These changes can all be seen on a plain lateral radiograph, as well as at tomography and CT scanning.

The radioisotope bone scan is of great value here as a screening test, an abnormal ossicle or tuberosity showing up as a strongly positive area (Fig. 68). Changes of osteonecrosis and soft tissue swelling are also seen at MR scanning (Renton 1991b).

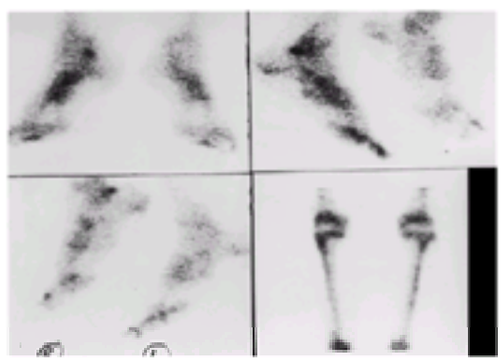


Fig. 68 Increase in uptake at the os trigonum.

The subtalar joint

There are three compartments of the subtalar joint—posterior, middle, and anterior. The posterior and middle lie parallel to each other, inclined at around 35 to 40° to the weight-bearing plane. Imaging on plain films of these compartments is performed with a lateral view in the weight-bearing plane. Weight-bearing films are more easily reproducible between examinations and between patients and, though static, demonstrate the foot in a position of function.

The axial view is taken with the X-ray beam projected along the line of the middle and posterior facets, i.e. at 35 to 40° to the weight-bearing plane. These views are easily performed. More complex views of the hindfoot are perhaps best avoided, especially as these joints are clearly demonstrated at CT scanning.

The peroneal spastic flat foot syndrome

Tarsal coalition exists in many forms. It has been shown to exist *in utero* and is the result of failure of segmentation of the cartilaginous primordia. The most common form is talocalcaneal synostosis, usually at the middle facet and well demonstrated on the axial view. Calcaneonavicular fusion is the next most commonly seen type, easily seen on an oblique view of the foot. These lesions become clinically manifest when the cartilaginous bridge between the ossific nuclei matures, usually into a continuous bar of bone (Fig. 69), though formes frustes are also found. The resulting fusion causes hindfoot rigidity and a flattening of the longitudinal arch, resulting in stretching pain and spasm in the peroneus longus tendon. The syndrome was seen in 2 per cent of Canadian army recruits (Harris and Beath 1948).

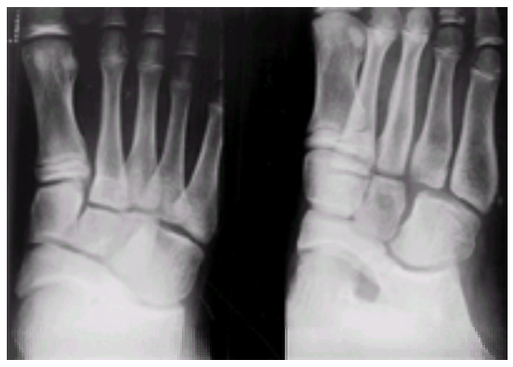


Fig. 69 Oblique radiograph of the foot demonstrates calcaneonavicular synostosis.

Plain radiographic changes are those of direct visualization of the fusion, and also indirect evidence of fusion—flattening of the longitudinal arch and talar beak formation.

CT scanning and, of course, MRI demonstrate the fusion directly. CT is superior in demonstrating cortical bone anatomy (Fig. 70). Isotope scanning demonstrates stress changes at the articular surfaces around the fusion.

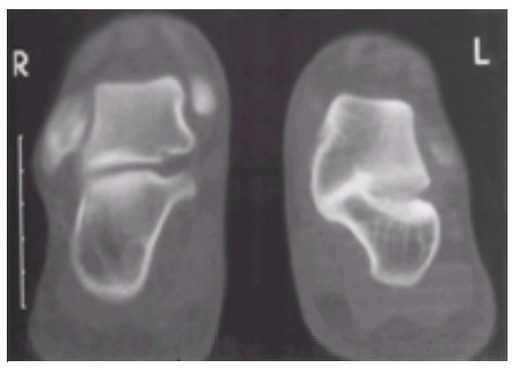


Fig. 70 Hindfoot fusion at the sustentaculotalar joint. Compare normal and abnormal sides. Best demonstrated at CT scanning.

Stress fracture in the foot

Stress fractures are caused by (i) regular, repeated submaximal stress on normal bone or (ii) normal stresses on abnormal, i.e. Pagetic, osteomalacic, or osteoporotic bone.

Stress fractures in the tarsus are usually calcaneal. As with stress fractures elsewhere, the history may be typical but no abnormality may be present on the initial radiograph. This is especially so in the elderly osteoporotic female or in those on long-term steroids in whom bone mass is much diminished and trabecular interruption cannot be seen. Under these circumstances, a radioisotope bone scan will be positive (Fig. 71), and a dense line representing bone compaction or

fracture healing will later show up on a lateral radiograph, tomograph or, especially, CT scan but, if the foot is rested, bony changes may never be seen on film.

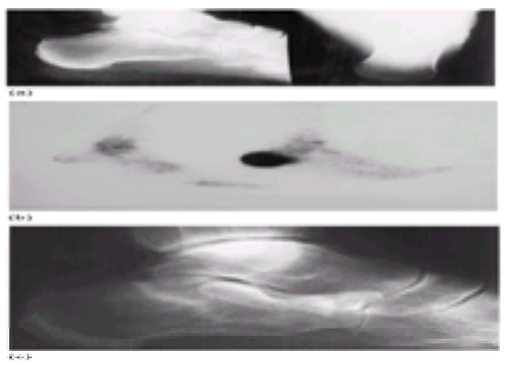


Fig. 71 Occult fracture of the calcaneum. (a) The initial radiograph shows no abnormality but a radioisotope bone scan (b) shows a gross increase in uptake in the calcaneum. (c) Another patient showing a stress or increment fracture in an osteoporotic calcaneum, seen as a serpiginous dense line extending to the upper cortex of the bone and reaching it at a right angle.

Stress lesions are well imaged at MR scanning, especially if the affected bone contains yellow marrow, i.e. fat. Stress lesions in bone show as ill-defined areas of decreased signal on T_1-weighted and increased signal on T_2-weighted and STIR fat suppression sequences (Fig. 72). These initial changes are the result of local oedema and are not as clearly shown in red marrow areas.

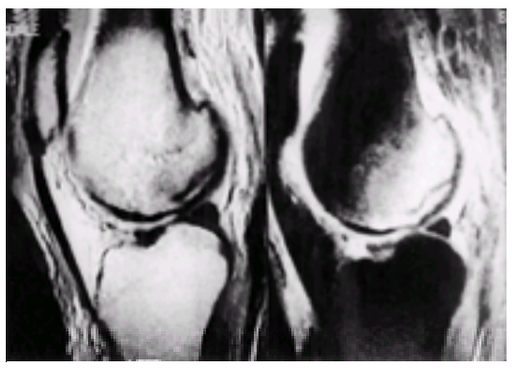


Fig. 72 Occult fracture of medial femoral condyle seen on a T_1-weighted image (a) as a serpiginous band of low signal extending to the posterior femoral cortex surrounded by some loss of the fatty bright signal. The bone oedema, however, is much better demonstrated on the STIR sequence (b).

Should a fracture line occur within the stress lesion (and this is not inevitable), it will be seen, as expected, as a serpiginous or straight, thin, low-signal band at right angles to the adjacent cortex, exactly as expected from a radiograph. It is likely that MRI is at least as sensitive as scintigraphy in the detection of early changes of trauma to bone (Yao and Lee 1988).

Osteoarthritis

It has been stated that MR is ideally the preferred imaging modality for the diagnosis and follow-up of rheumatoid arthritis and variants. Time, availability, and cost probably rule out its use in general clinical practice. The same can be said in the diagnosis of osteoarthritis which typically affects major joints. Again, plain films provide a cheap, quick, and simple method of imaging affected joints at onset and for follow-up. Changes in osteoarthritis include joint narrowing due to cartilage and meniscal destruction, marginal osteophytes, cyst formation and articular collapse, reactive sclerosis, and malalignment. Many of these changes relate to cortical bone and, therefore, they are well imaged on plain films and also at CT scanning (Fig. 73) or MRI (Fig. 74), which show the articular surfaces well. Axial imaging demonstrates the joint spaces, osteophytes, cysts, and changes in bony density especially well. Loose bodies are also well seen. The images may be enhanced by the use of single or double contrast arthrography.

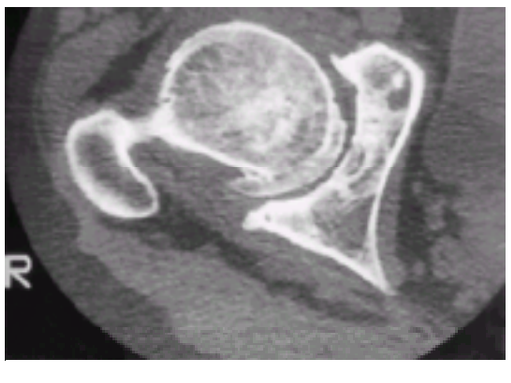


Fig. 73 Osteoarthritis demonstrated at CT scanning. New bone proliferation is seen around the articular surfaces and internally within the joint on the femoral head.

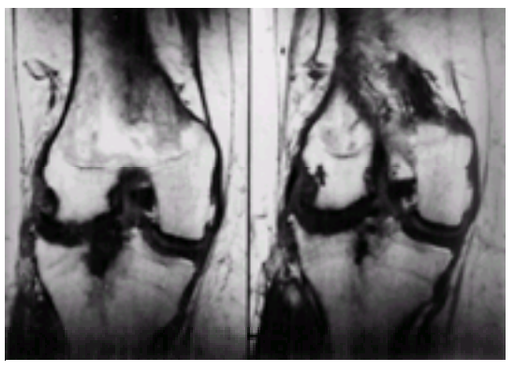


Fig. 74 Osteoarthritis demonstrated at MR scanning. Coronal images show irregular articular surfaces, marginal osteophytes, synovial proliferation and irregularity, or even loss of the adjacent meniscus.

Transient regional osteoporosis

Transient regional osteoporosis is an uncommon, painful condition first described in pregnant women, but more commonly seen in middle-aged men. It is associated with marked demineralization of bone around the hip seen on a plain film associated with preservation of the joint space, a feature which makes septic arthritis perhaps less likely in the presence of severe osteoporosis. The disease is self-limiting and clinical and radiographic changes revert to normal, usually within 6 months.

A radioisotope scan will show marked increase in uptake in the femoral head. The MR scan shows decreased marrow signal on T_1-weighted images, but generally increased signal on T_2-weighted studies—changes compatible with marked oedema (Bloem 1988)—and this may be seen, in the presence of a normal radiograph, early on in the disease.

Avascular necrosis

Deprivation of blood supply to cortex and medulla results in marrow cell death. No changes are seen on the plain film in the early phase, but this still has to be obtained to exclude other abnormalities and for a baseline study.

The early isotope scan demonstrates photopenia (Fig. 75) and subsequently, with repair, will show increased uptake around the abnormal devitalized area corresponding to healing at the zone of creeping substitution.

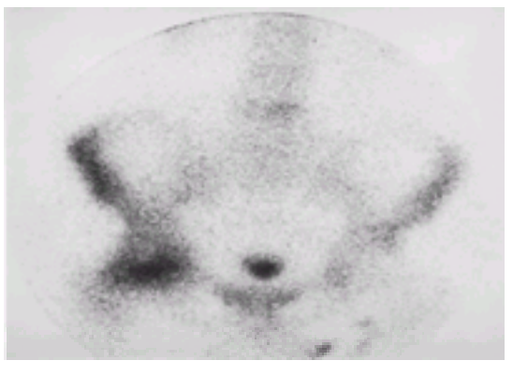


Fig. 75 Avascular necrosis. Radioisotope bone scan shows photopenia of the left femoral head.

Subsequent radiographs may demonstrate a subcortical crest of lucency which prognosticates structural failure and articular collapse (Fig. 76). The joint space remains intact until secondary osteoarthritis supervenes. The plain film may also subsequently show cyst formation and necrotic bone, which remains dense in the midst of demineralization and resorption.

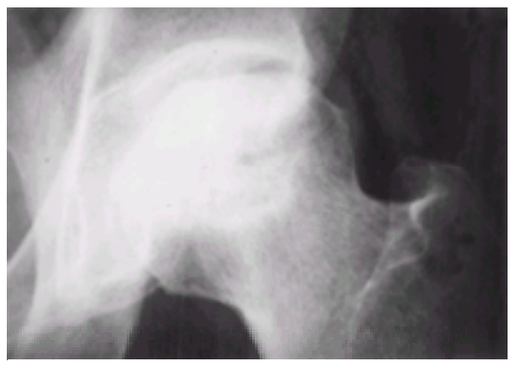


Fig. 76 Avascular necrosis. A sclerotic rim of creeping substitution is seen beneath an area of structural failure of the femoral head.

At MRI, changes in ischaemia precede plain film change, and probably isotope scan changes too (Brower and Kransdorf 1990). Linear or confluent areas of low signal in epiphyses replace bright marrow signal on T_1-weighted images. The linear bands may be subcortical or deep in the bone. The position of the avascular areas can be assessed on axial, sagittal, and coronal slices (Fig. 77).

Fig. 77 Avascular necrosis of both femoral heads seen at MR scanning. The areas of low signal are avascular.

On T_2-weighted images, the low signal band often has a parallel adjacent area of increase in signal representing oedema, or even neovascularity.

It should be noted that even a negative MR scan does not rule out avascular necrosis, as there is a temporal lag between the initial insult and subsequent cell death. Serial studies may be necessary (Mitchell *et al*. 1986; Beltran *et al*. 1988).

Chapter References

Aisen, A.M., Martel, W., Ellis, J.H., and McCune, W.J. (1987). Cervical spine involvement in rheumatoid arthritis. MR imaging. *Radiology*, **165**, 159–63.

Beltran, J. *et al*. (1987). Rheumatoid arthritis. MR imaging manifestations. *Radiology*, **165**, 153–7.

Beltran, J. *et al*. (1988). Femoral head avascular necrosis. MR imaging with clinical-pathologic and radionuclide correlation. *Radiology*, **166**, 215–20.

Berquist, T.H. and Morin, R.L. (1990). General technical considerations in magnetic resonance imaging. In *Magnetic resonance of the musculoskeletal system* (2nd edn) (ed. T.H. Berquist), pp. 53–73. Raven Press, New York.

Berquist, T.H. (1991). Magnetic resonance techniques in musculoskeletal disease. *Rheumatic Disease Clinics of North America*, **17**, 599–615.

Birney, T.J., White, J.J. Jr., Berens, D., and Kuhn, G. (1990). Lumbar discography followed by computed tomography. *Spine*, **15**, 690–707.

Birney, T.J., White J.J. Jr., Berens, D., and Kuhn G. (1992). Comparison of MRI and discography in the diagnosis of lumbar degenerative disc disease. *Journal of Spinal Disorders*, **5**, 417–23.

Bloem, J.L. (1988). Transient osteoporosis of the hip. MR imaging. *Radiology*, **167**, 753–5.

Brewerton, D.A. (1967). A tangential radiographic projection for demonstrating involvement of the metacarpal heads in rheumatoid arthritis. *British Journal of Radiolology*, **40**, 233–4.

Brook, A. and Corbett, M. (1977). Radiographic changes in early rheumatoid disease. *Annals of the Rheumatic Diseases*, **36**, 71–3.

Brower, A.C. and Kransdorf, M.J. (1990). Imaging of hip disorders. *Radiologic Clinics of North America*, **28**, 955–74.

Brown, T.R. and Quinn, S.F. (1993). Evaluation of chondromalacia of the patellofemoral compartment with axial magnetic resonance imaging. *Skeletal Radiology*, **22**, 325–8.

Buckland-Wright, J.C. (1983). Advances in the radiological assessment of rheumatoid arthritis. *British Journal of Rheumatology*, **22** (Suppl.), 34–43.

Buckland-Wright, J.C. (1984). Microfocal radiographic examination of erosions in the wrist and hand of patients with rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **43**, 160–71.

Bywaters, E.G.L. (1954). Heel lesions of rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **13**, 42–51.

Chan, W.P. *et al*. (1991). Osteoarthritis of the knee. Comparison of radiography, CT and MR imaging to assess extent and severity. *American Journal of Radiology*, **157**, 799–806.

Collins, C.D. *et al*. (1990). The role of discography in lumbar disc disease. Comparative study of magnetic resonance imaging and discography. *Clinical Radiology*, **42**, 252–7.

Cowan, N.C., Bush, K., Katz, D.E., and Gishen, P. (1992). The natural history of sciatica: a prospective radiological study. *Clinical Radiology*, **46**, 7–12.

Dacre, J.E. and Buckland-Wright, J.C. (1992). Radiological measures of outcome. *Ballière's Clinical Rheumatology*, **6**, 39–67.

Dequeker, J., Goddeeris, T., Walravens, M., and De Roo, M. (1978). Evaluation of sacroiliitis. Comparison of radiological and radionuclide techniques. *Radiology*, **128**, 687–9.

Deyo, R.A. (1994). Editorial: Magnetic resonance imaging of the lumbar spine. *New England Journal of Medicine*, **331**, 115–16.

Durback, M.A., Edelstein, G., and Schumacher, H.R. (1988). Abnormalities of the sacroiliac joints in diffuse idiopathic skeletal hyperostosis. Demonstration by computed tomography. *Journal of Rheumatology*, **15**, 1506–10.

Ehman, R.L. (1990). Interpretation of magnetic resonance images. In *Magnetic resonance of the musculoskeletal system* (2nd edn) (ed. T.H. Berquist), pp. 27–51. Raven Press, New York.

Einig, M., Higger, H.P., Meairs, S., Faust-Tinnefeldt, G., and Kapp, H. (1990). Magnetic resonance imaging of the craniocervical junction in rheumatoid arthritis: value, limitations, indications. *Skeletal Radiology*, **19**, 341–6.

Farfan, H.F. (1973). *Mechanical disorders of the low back*. Lea and Febiger, Philadelphia.

Fielding, J.R., Franklin, P.D., and Kustan, J. (1991). Popliteal cysts: a reassessment using magnetic resonance imaging. *Skeletal Radiology*, **20**, 433–5.

Fletcher, D.E. and Rowley, K.A. (1952). The radiological features of rheumatoid arthritis. *British Journal of Radiology*, **25**, 282–95.

Forrester, D.M. (1990). Imaging of sacro-iliac joints. *Radiologic Clinics of North America*, **28**, 1055–72.

Gilkeson, G. *et al*. (1988). Early detection of carpal erosions in patients with rheumatoid arthritis: a pilot study of magnetic resonance imaging. *Journal of Rheumatology*, **15**, 1361–6.

Gold, R.H., Seeger, L.L., and Yao, L. (1993). Imaging shoulder impingement. *Skeletal Radiology*, **27**, 555–61.

Goldberg, R.P., Genant, H.K., Shimshak, R., and Shames, D. (1978). Applications and limitations of quantitative sacroiliac joint scintigraphy. *Radiology*, **128**, 683–6.

Goldman, A.B. and Dicarlo, E.F. (1988). Pigmented villonodular synovitis. Diagnosis and differential diagnosis. *Radiologic Clinics of North America*, **26**, 1327–47.

Gompels, B.M. and Darlington, L.G. (1982). Evaluation of popliteal cysts and painful calves with ultrasonography. Comparison with arthrography. *Annals of the Rheumatic Diseases*, **41**, 355–9.

Griffiths, H.J. (1991). *Imaging of the lumbar spine*. Aspen, Gaitherburg, MA.

Guerra, J.R. *et al*. (1978). The adult hip: an anatomic study. *Radiology*, **128**, 11–20.

Harris, R.I. and Beath, T. (1948). Aetiology of peroneal spastic flat foot. *Journal of Bone and Joint Surgery*, **30B**. 624–34.

Ho, G. Jr., Sadovnikoff, N., Malhotra, C.M., and Claunch, B.C. (1979). Quantitative sacroiliac joint scintigraphy. A critical assessment. *Arthritis and Rheumatism*, **22**, 837–44.

Iannotti, J.P., Zlatkin, M.B., Esterhai, J.L., Kressel, H.Y., Dalinka, M.K., and Spindler, K.P. (1991). Magnetic resonance imaging of the shoulder. *Journal of Bone and Joint Surgery*, **73A**, 17–28.

Ireland, J., Trickey, E.L., and Stoker, D.J. (1980). Arthroscopy and arthrography of the knee. *Journal of Bone and Joint Surgery*, **62B**, 3–6.

Jelinek, J.S., Kransdorf, M.J., Utz, J.A., and Berrey, B.H. Jr. (1989). Imaging of pigmented villonodular synovitis with emphasis on magnetic resonance imaging. *American Journal of Roentgenology*, **152**, 337–42.

Jensen M.C., Brant-Zawadzki, M.N., Obuchowski, N., Modic, M.T., Malkasian, D., and Ross, J.S. (1994). Magnetic resonance imaging of the lumbar spine in people without back pain. *New England Journal of Medicine*, **331**, 69–73.

Jones, M.D., Pais, M.J., and Omiya, B. (1988). Bony overgrowths and abnormal calcifications about the spine. *Radiologic Clinics of North America*, **26**, 1213–34.

Kaplan, P.A., Nelson, N.L., Garvin, K.L., and Brown, D.E. (1991). MR of the knee. The significance of high signal in the meniscus that does not clearly extend to the surface. *American Journal of Roentgenology*, **156**, 333–6.

Kaye, J.J. (1991). Radiographic methods of assessment (scoring) of rheumatic disease. *Rheumatic Disease Clinics of North America*, **17**, 457–70.

Kieft, G.J., Bloem, J.L., Rozing, P.M., and Obermann, W.R. (1988). MR imaging of recurrent anterior dislocation of the shoulder. Comparison with CT arthrography. *American Journal of Roentgenology*, **150**, 1083–7.

Lee, J.K., Yao, L., Phelps, C.T., Wirth, C.R., Czajka, J., and Lozman, J. (1988). Anterior cruciate liagment tears. MR imaging compared with arthroscopy and clinical tests. *Radiology*, **166**, 861–4.

Lindblom, K. (1939). Arthrography and roentgenography in ruptures of tendons of the shoulder joint. *Acta Radiologica*, **20**, 548–62.

Martel, W., Stuck, K.J., Dworin, A.M., and Hylland, R.G. (1980). Erosive osteoarthritis and psoriatic arthritis. *American Journal of Roentgenology*, **134**, 125–35.

Mason, R.M.. Murray, R.S., Oates, J.K., and Young, A.C. (1959). A comparative radiological study of Reiter's disease, rheumatoid arthritis and ankylosing spondylitis. *Journal of Bone and Joint Surgery*, **41B**, 137–48.

Mink, J.H. and Deutsch A.L. (1990). The knee. In *MRI of the musculoskeletal system: a teaching file* (ed. J.H. Mink and A.L. Deutsch), pp. 251–387. Raven Press, New York.

Mink, J.H., Levy, T., and Crues, J.V. (1988). Tears of the anterior cruciate ligament and menisci of the knee. MR imaging evaluation. *Radiology*, **167**, 769–74.

Mink, J.H., Reicher, M.A., and Crues J.V. (1993). *MRI of the knee* (2nd edn), pp. 100–103. Raven Press, New York.

Mitchell, M.D., Kundel, H.L., Steinberg, M.E., Kressel, H.Y., Alavi, A., and Axel, L. (1986). Avascular necrosis of the hip. Comparison of MR, CT and scintigraphy. *American Journal of Roentgenology*, **147**, 67–71.

Ng, Y.Y., Khaw, K.T., Halpin, S., and Stoker, D.J. (1989). A comparative study to evaluate the role of intra-articular adrenaline in double-contrast arthrography. *Clinical Radiology*, **40**, 598–601.

Nixon, J.R. (1987). Basic principles and terminology. In *Magnetic resonance of the musculoskeletal system* (ed. T.H. Berquist), pp. 1–12. Raven Press, New York.

O'Brien, J.P. (1984). Mechanisms of spinal pain. In *Textbook of pain* (ed. P.D. Wall and R. Melzack), pp. 2–13. Churchill Livingstone, Edinburgh.

Osti, O.L. and Fraser, R.D. (1992). MRI and discography of annular tears and intervertebral disc degeneration. A prospective clinical comparison. *Journal of Bone and Joint Surgery*, **74B**, 431–5.

Poleksic, L., Zdravkovic, D., Jablanovic, D., Watt, I., and Bacic, G. (1993). Magnetic resonance imaging of bone destruction in rheumatoid arthritis. Comparison with radiography. *Skeletal Radiology*, **22**, 577–80.

Reinig, J., McDevitt, E.R., and Ove, P.N. (1991). Progression of meniscal degenerative changes in college football players. Evaluation with MR imaging. *Radiology*, **181**, 255–7.

Renton, P. (1991*a*). Arthrography. In *Surgical disorders of the shoulder* (ed. M. Watson), pp. 97–117. Churchill Livingstone, Edinburgh.

Renton, P. (1991*b*) Radiology of the foot. In *The foot and its disorders* (3rd edn) (ed. L. Klenerman), pp. 259–346. Blackwell Scientific, Oxford.

Resnick, D., Feingold, M.L., Curd, J., Niwayama, G., and Goergen, T.G. (1977). Calcaneal abnormalities in articular disorders. *Radiology*, **125**, 355–66.

Resnick, D., Guerra, J. Jr., Robinson, C.A., and Vint, V.C. (1978). Association of diffuse idiopathic skeletal hyperostosis and calcification and ossification of the posterior longitudinal ligament. *American Journal of Roentgenology*, **131**, 1049–53.

Resnik, C.S., Fronek, J., Frey, C., Gershuni, D., and Resnick, D. (1984). Intra-articular pressure determination during glenohumeral joint arthrography. Preliminary investigation. *Investigative Radiology*, **19**, 45–50.

Rosenthall, L. (1991). Nuclear medicine techniques in arthritis. *Rheumatic Disease Clinics of North America*, **17**, 585–97.

Sartoris, D.J. and Resnick, D. (1990). Current and innovative methods for noninvasive bone densitometry. *Radiologic Clinics of North America*, **28**, 257–8.

Semble, E.L., Elster, A.D., Loeser, R.F., Laster, D.W., Challa, V.R., and Pisko, E.J. (1988). Magnetic resonance imaging of the craniovertebral junction in rheumatoid arthritis. *Journal of Rheumatology*, **15**, 1367–75.

Stoller, D.W. *et al*. (1987). Meniscal tears. Pathologic correlation with MR imaging. *Radiology*, **163**, 731–5.

Thijn, C.J.P. (1982). Accuracy of double contrast arthrography and arthroscopy of the knee joint. *Skeletal Radiology*, **8**, 187–92.

Thould, A.K. and Simon, G. (1966). Assessment of radiological changes in the hands and feet in rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **25**, 220–8.

Tsuyama, N. (1984). Ossification of the posterior longitudinal ligament of the spine. *Clinical Orthopaedics and Related Research*, **184**, 71–83.

Van Holsbeeck, M. and Introcaso, J.H. (1991). *Musculoskeletal ultrasound*. Mosby Year Book, St Louis.

Yao, L. and Lee, J.K. (1988). Occult intraosseous fracture. Detection with MR imaging. *Radiology*, **167**, 749–51.

Zucherman, J. *et al*. (1988). Normal magnetic resonance imaging with abnormal discography. *Spine*, **13**, 1355–9.

Essential reading

Resnick, D. (1995). *Diagnosis of bone and joint disorders* (3rd edn). W.B. Saunders, Philadelphia.

Resnick, D. and Niwayama, G. (1989). *Bone and joint imaging* (2nd edn). W.B. Saunders, Philadelphia.

Weissman, B.N. (ed) (1991). Imaging of rheumatic diseases. *Rheumatic Disease Clinics of North America*, **17**, 457–816.

4.9.2 Imaging in children

G. H. Sebag and J. F. Quignodon

Introduction

The use of imaging in paediatric rheumatology reflects the old adage that a child is not a small adult. This chapter will illustrate the specificity of joint imaging in children. Successful imaging of children requires a thorough knowledge of the anatomical variants resulting from growth and development, knowledge of paediatric diseases and their manifestations, and finally the need for radiation protection (Mandell *et al*. 1990). Furthermore, in this era of advancing imaging technology, knowledge of the relative values of available imaging techniques is necessary to optimize the management of children with arthritis. The theoretical basis of imaging techniques, including conventional radiography, radionuclide scanning, ultrasonography, computed tomography (**CT**), and magnetic resonance imaging (**MRI**) are described in Chapter 4.9.1. Each is being used in paediatric radiology. Changes in individual diseases will be presented in the appropriate section. As in adult radiology, investigations have to be tailored to the individual needs of the child, and radiologists and paediatricians must collaborate to ensure that the needs of all are best matched.

Practical considerations

High quality, diagnostic examinations of sick children can be obtained consistently with care, time, and patience. The child's co-operation usually requires an explanation of exactly what is going to happen in language that can be understood. It is also crucial to reassure the parents that the procedure will go smoothly and that they will be informed of the results.

Expertise in immobilizing and imaging children is essential to perform the examination quickly and efficiently (Poznanski 1976). It is important that technologists are trained in the handling of neonates, small infants, and small children. A number of immobilization devices are available; the simplest ones are very efficient (velcro straps, wooden immobilization boards, bags). Immobilization need not be traumatic to the child. It will allow better positioning for radiography, fluoroscopy, CT, MR, and radionuclide scanning and make motion much less likely. This will result in less radiation to the child as smaller fields can be used and the chance of repeat examinations is reduced. In some cases (especially children under the age of 6 years) it is necessary to use sedation for imaging procedures such as arthrography, CT, and MRI. The choice of drug and route of administration must be made on an individual basis according to the type of examination, the child, the facility, and the experience of the radiologist and the support team.

Musculoskeletal radiography

The plain film is almost always the initial imaging technique used in the evaluation of the musculoskeletal system. Its predominance in diagnosing, for example, fractures, bone tumours, congenital dysplasias, metabolic and inflammatory disorders, is confirmed by the fact that it is the only technique needed for most children. It is the most widely available modality and can be obtained more quickly and easily than any other type of imaging (Fig. 1). Plain films play an important part in localizing and defining the nature of a lesion, especially when the clinical history and physical examination are uncertain or not contributory. This property is especially true in infants and young children who do not localize pain accurately or effectively communicate symptoms to the physician. For example, plain films are invaluable in the diagnosis of the aetiology of limping of recent onset for which a traumatic cause is not obvious (Blumhagen 1994) (Table 1).

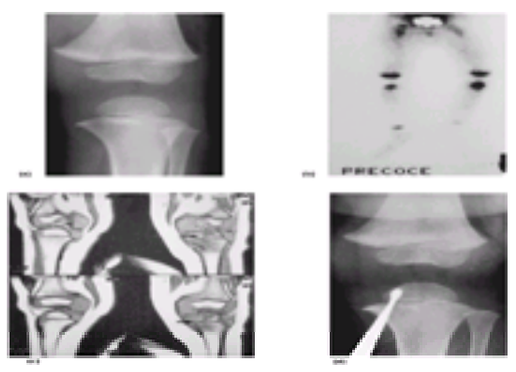


Fig. 1 Acute osteomyelitis in a 2-year-old boy. (a) Plain film: medial radiolucent area in the tibial metaphysis. (b) Bone scan: increased uptake in the upper left tibia. (c) T_1-weighted MR images confirm the previous findings and show the extension of the infection through the growth plate in the epiphysis (arrow). (d) Peroperative plain film confirms the MR findings.

Traumatic:
- fractures
- stress fractures
- non-accidental trauma
- foreign body

Inflammatory:
- transient synovitis
- osteomyelitis
- septic arthritis
- discitis
- soft tissue abscess

Congenital:
- congenital dislocation of the hip
- club foot
- spinal dysraphism

Avascular necrosis and related conditions:
- Legg–Calvé–Perthes' disease
- osteochondrosis
- osteochondritis dissecans

Slipped capital femoral epiphysis.

Neoplastic:

Table 1 Common causes of limping in children

As in adult radiology, conventional radiographs allow grouping of the various arthritides on the basis of the distribution and the pattern of joint space changes (Resnick 1988) (Fig. 2). However, radiological findings although very important are only one of the criteria by which arthritis is judged to be present and by which a specific diagnosis is established (Ansell 1990). Furthermore, assessment of disease progression (e.g. juvenile chronic arthritis) is made possible by repeating the plain films. In young children especially, radiographic changes (i.e. joint space narrowing indicative of cartilage loss) represent late and indirect signs of synovial disease.

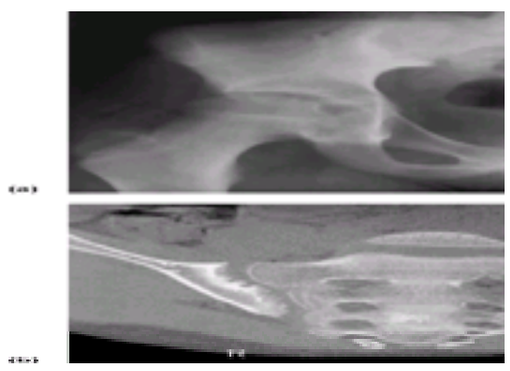


Fig. 2 Juvenile spondylarthropathy in an 11-year-old boy. (a) Plain film: lateral subluxation of the femoral head with lateral narrowing of the joint space. Sclerosis of the sacroiliac joint. (b) CT: sclerosis, irregularity, and erosion of the sacroiliac joint.

Technical considerations

Radiation protection must be a priority in children. Proper coning of the X-ray beam, use of high-speed intensifying screens, and the use of gonadal shielding all help reduce exposure. In the assessment of scoliosis, the frontal film is taken anteroposteriorly in order to reduce breast irradiation, and lateral view should be kept to the minimum in the follow-up. Other ways of reducing radiation to the patient include the use of digital radiography, if available. Electronic image manipulation permits viewing of segments of the image, adjustment of contrast, and measurement of curve or length. This reduces the need for repeat studies when the initial technique is unsatisfactory. Digital radiography has special applications for scoliosis and leg-length discrepancies, which require frequent follow-up examination (Kushner *et al*. 1986). Contrast manipulation allows the enhancement of soft tissue thickening and periarticular fat pad displacement. Evaluation of these structures are especially useful in infants and young children. At least two views at right angles, a true frontal and a lateral projection, should be obtained of the involved joint or extremity.

Correct positioning is important for accurate interpretation. A variety of radiographic guidelines (drawing lines, measurements of angle and distance) have been suggested to detect any distortion of the anatomical relationship. Use of guidelines is highly dependent upon proper positioning. For instance, tilting of the pelvis can result in false interpretation of hip position when evaluating an infant for congenital dislocation and/or dysplasia (Fig. 3).

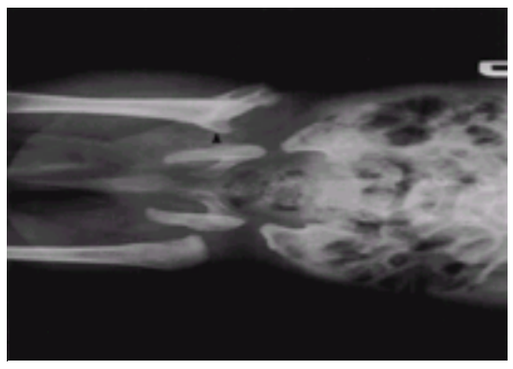


Fig. 3 Septic arthritis in a 4-month-old baby with dislocation of the left hip. Radiograph shows lateral displacement of the femoral head and new bone formation.

Recognition of even minimal pelvic and femoral rotation is necessary to avoid the false-positive interpretation of joint space widening, especially in a painful hip that is held typically in abduction and external rotation (Fig. 4). Unrotated pelvic views show symmetry of obturator foramina and of the medial acetabula as well as alignment of the pubic symphysis with the sacral midline. Joint space widening secondary to the accumulation of fluid in the hip joint has been found to be a more reliable sign in young children and infants. In older children, detecting fluid in the hip joint is best accomplished with ultrasound.

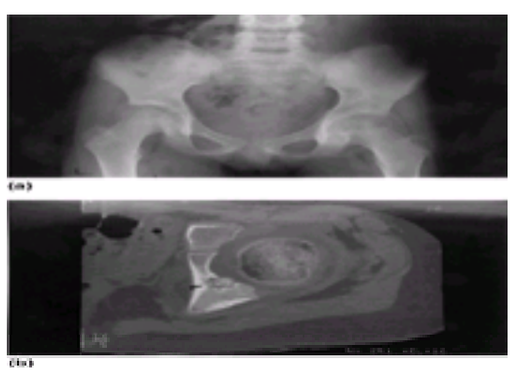


Fig. 4 Osteoid osteoma affecting the left hip in a 9-year-old boy. (a) Plain film: left hip effusion with severe osteoporosis. (b) CT scan: localizes the site of nidus in the acetabulum.

Frog-leg lateral views (in addition to anteroposterior neutral views) of the hip are mandatory when Legg–Calve–Perthes' disease or a slipped capital femoral epiphysis is suspected. The abnormal anterior subchondral surface of Legg–Calve–Perthes' disease and early posterior slippage are demonstrated better on these views.

In addition, it is useful to obtain comparative views of the normal side, especially in cases where abnormal findings are subtle. It is very useful in the assessment of the shape and size of the ossification centres (normal variants, advanced maturation). However, comparative views are not systematically required, although occasionally these, oblique, or other special views will facilitate a diagnosis. When choosing the right views, one must keep in mind that symptoms in children can be very confusing. For instance, children with discitis can experience symptoms suggesting lower extremity disease (limp, failure to bear weight, referred hip pain) and in

this case, a spinal view should be obtained.

Special views include :

- weight-bearing films, bending
- flexion views of the cervical spine (C1 to C2 dislocation)

These views are performed to determine whether a joint is stable or not. The C1 to dens distance is normally wider in children (as great as 4 to 5 mm) than in adults and flexion views are very useful to detect instability (Table 2).

Congenital hypoplasia of the dens and C1
Dislocation (anterior, rotatory)
Collagen vascular disease (juvenile chronic arthritis)
Trisomy 21
Storage diseases (Morquio disease)
Spondylo- and punctate epiphyseal dysplasia

Table 2 C1 to C2 instability

Growth abnormalities are an important issue in children with rheumatological disease. Epiphyseal overgrowth (caused by hyperaemia), epiphyseal deformities, accelerated ossification with premature fusion of growth plate, and compression fractures lead to malalignment and length discrepancy, especially in unilateral disease. Measurement is an important part in the evaluation of growth abnormalities. It includes estimation of valgus and varus angles, leg-length discrepancy, spinal curvature, femoral anteversion, and tibial torsion amongst others. Tables, charts, and graphs of standard measurements have been widely published and are referred to daily by practising orthopaedists and radiologists (Pettersson and Ringertz 1991; Ozonoff 1992). The use of these measurements correlated with the patient's skeletal age at each determination allows a prediction of future growth and effect and timing of surgery to be made (Pettersson and Ringertz 1991; Ozonoff 1992).

General growth abnormalities are influenced by the disease (type, onset, and duration), immobilization, and steroid therapy which lead to growth delay. Skeletal age assignment is the most common evaluation. After 1 year of age, the number, shape, and stage of development of the epiphyses in the left hand and wrist are compared with standards for various ages up to the time of epiphyseal fusion; the atlas of Greulich and Pyle (1959) is normally used. In addition, spinal growth in scoliosis can be estimated using Risser's method (Risser 1948) by assessing iliac crest development on the frontal view of the entire spine.

Osteoporosis, induced for example by immobilization and steroid therapy, can be assessed on plain films both in the appendicular and axial skeleton. Measurement of the cortical thickness of tubular bones is commonly performed (in metacarpals and tibia for instance).

Evaluation of musculoskeletal changes on plain films

Proper interpretation of plain films is pivotal in directing further diagnostic evaluation that can be both complex and expensive in children. Reading guidelines are therefore useful to enhance plain film interpretation skills. On a systematic basis, one should analyse:

- joint space (widening, narrowing, malalignment, ankyloses, calcification, gas) (Fig. 1 and Fig. 4)
- soft tissues (obliterated or displaced fat pads, calcification, ossification opacities, thickening and reticulation, tendon width changes) (Fig. 5 and Fig. 6)

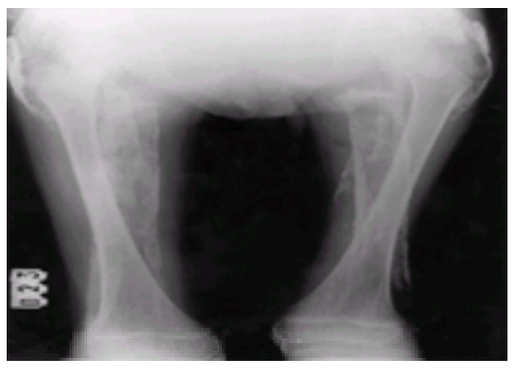


Fig. 5 Dermatomyositis in a 7-year-old girl. Plain film of extensive calcinosis in the subcutaneous tissue and in the muscles.

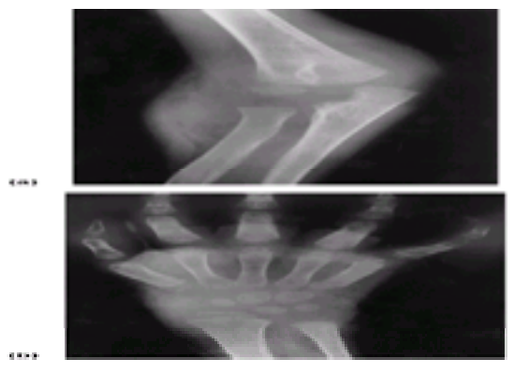


Fig. 6 Juvenile chronic arthritis in a 4-year-old boy at the early stage. Radiographs of the elbow (a) and wrist (b) show marked soft tissue swelling and joint effusion.

- bone density (increased or decreased, generalized or localized)
- epiphysis (shape, size, subchondral bone, defect, erosion)
- metaphysis (transverse bands, widening, cupping, destruction)
- growth plate
- diaphysis (cortical bone, periosteal reaction, tubulation)

The significance of these findings in children has been described widely (Thomas, P.S. *et al*. 1994).

By grouping the findings and using the complete range of the common causes of these findings, one can narrow the differential diagnosis (Swischuk *et al*. 1995)

(Table 2, Table 3, Table 4, Table 5, Table 6, Table 7, Table 8, Table 9, Table 10, Table 11, Table 12, Table 13, Table 14 and Table 15). This practical approach allows a proper interpretation of plain films, especially for clinicians or radiologists who practise paediatric rheumatology only occasionally.

Traumatic effusion and/or dislocation
Septic arthritis (hip, shoulder)
Transient synovitis (hip)
Legg–Calve–Perthes' disease
Developmental hip dysplasia
Juvenile chronic arthritis
Bleeding disorders
Joint laxity (neuromuscular) (Larsen syndrome, Ehlers–Danlos syndrome)
Collagen vascular disease
Pigmented villonodular synovitis
Synovial tumour
Synovial osteochondromatosis
Winchester–Grossman syndrome

Table 3 Joint space widening

Septic arthritis
Juvenile chronic arthritis
Haemophilic arthropathy
Degenerative arthritis (long-term complication of Legg–Calve–Perthes' disease, avascular necrosis, trauma)
Slipped capital femoral epiphysis (postoperative)
Pigmented villonodular synovitis

Table 4 Joint space narrowing

Juvenile chronic arthritis
Psoriatic arthritis
Haemophilic arthritis
Hypothyroidism
Rickets
Bone dysplasia:
achondroplasia
pseudoachondroplasia
dysplasia epiphysealis hemimelica
dysplasia epiphysealis multiplex
spondylepiphyseals dysplasia
mucopolysaccharidoses
enchondromatosis
Stickler syndrome

Table 5 Generalized epiphyseal abnormalities

Prematurity
Severe illness, trauma
Leukaemia, lymphoma
Metastasis (neuroblastoma)
Neonatal infections (toxoplasmosis, syphilis, cytomegalovirus, rubella)
Scurvy
Hypermagnesaemia
Cushing's syndrome
Hypervitaminosis D
Osteogenesis imperfecta
Idiopathic juvenile osteoporosis

Table 6 Transverse radiolucent metaphyseal band

Normal variants
Osteochondritis dissecans
Osteochondral avulsion
Juvenile chronic arthritis
Haemophilic arthritis
Septic arthritis (tuberculosis, fungus, pyogenic)
Epiphyseal osteomyelitis
Histiocytosis X
Tumours (synovial, chondroblastoma)

Table 7 Epiphyseal defect

Juvenile chronic arthritis

Neuromuscular disease

Larsen syndrome

Winchester–Grossman syndrome

Farber syndrome

Werner mesomelic syndrome

Stickler syndrome

Table 8 Generalized joint dislocation

Traumatic

Developmental hip dysplasia

Juvenile chronic arthritis

Neuromuscular disease

Congenital radial head dislocation

Madelung deformity

Genu recurvatum

Table 9 Single joint dislocation

Collagen disease (dermatomyositis)

Trauma

Infection

Hypervitaminosis D and A

Hyperparathyroidism

Tumoral calcinosis

Table 10 Periarticular calcifications

Traumatic avulsion

Osteochondritis dissecans

Idiopathic (in hips of infant after taps)

Synovial tumours

Synovial inflammation

Synovial chondromatosis

Oxalosis

Ochronosis

Table 11 Intra-articular calcifications

Trisomy 21

Marfan syndrome

Ehlers–Danlos syndrome

Morquio disease

Goltz syndrome

Stickler syndrome

Table 12 Joint hypermobility

Arthrogryposis

Juvenile chronic arthritis

Punctate epiphyseal dysplasia

Diastrophic dwarfism

Metatropic dwarfism

Storage disease

Table 13 Joint hypomobility

Rickets (privational, vitamin D resistant)
Osteogenesis imperfecta
Mucopolysaccharidosis
Hypophosphatasia
Renal osteodystrophy
Malabsorption disorders
Disuse (immobilization, neuromuscular)
Haematological disorders (malignant, haemolytic anaemia)
Endocrinological disorders (Cushing's disease)
Steroid therapy

Table 14 Diffuse decreased bone density

Fracture:
trauma
pathological
battered child
Infection:
osteomyelitis
neonatal infection (syphilis)
Infarction
Bleeding disorders
Gaucher's disease
Metabolic:
rickets
prematurity
scurvy
hypervitaminosis D, A
thyroid acropachy
Storage disease
Gaucher's disease
mucolipidosis
Tumours:
metastasis
leukaemia/lymphoma
primary malignant (sarcomas)
benign
histiocytosis X
soft tissue tumour
Vascular malformation (soft tissue)
Osteomyelography
Hypervitaminosis E (neonates)
Juvenile chronic arthritis

Table 15 Periosteal bone reaction

Finally, knowledge of the normal is a prerequisite for recognition of the abnormal. Familiarity with appearances at the various ages of development is learned by experience but considerable help may be gained from atlases both of normal variants and of normal development at different ages (Keats 1991; Swischuk *et al*. 1993). The following points should be borne in mind:

1. the bone-within-bone appearance: most commonly seen in the spine in newborns and premature infants;
2. periosteal new bone seen in normal neonates up to about 3 months of age, especially in the femoral diaphysis;
3. transverse sclerotic lines in the metaphysis (growth recovery lines);
4. irregular ossification centre seen especially in the medial aspect of the lower femoral epiphysis;
5. fibrous cortical defect occurring particularly around the knee and located in the superficial bony cortex;
6. gas in the joint: the 'vacuum' joint effect commonly seen in the shoulders and the hips of infants.

Role of plain films in the management of children with musculoskeletal disease

Diagnostic study

Plain radiographs provide crucial information in the early stages of diagnosing disease in children (Jacobs 1982; Mandell *et al*. 1990; Swischuk 1994; Thomas, P.S. *et al*. 1994) (Fig. 1, Fig. 2, Fig. 4, and Fig. 7). For instance, in cases with non-specific clinical presentation (such as limp, failure to bear weight, joint pain, back pain) the findings of plain film can be straightforward:

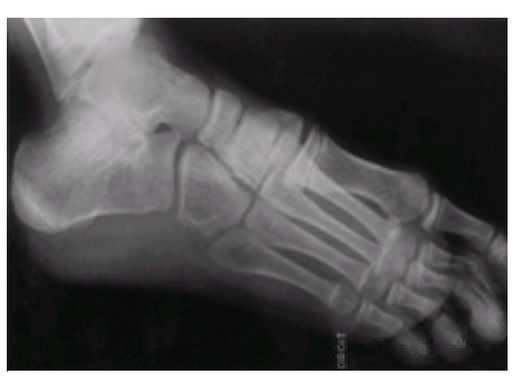


Fig. 7 Talocalcaneal coalition in a 10-year-old boy with foot pain.

1. stress and pathological fractures (Fig. 8, Fig. 9, and Fig. 10);

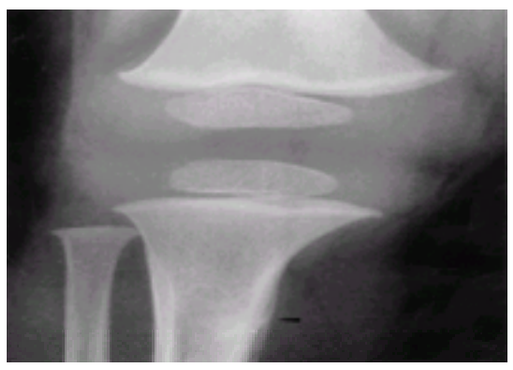


Fig. 8 Stress fracture in a 5-year-old boy. Area of increased sclerosis and periosteal bone deposition in the upper tibia.

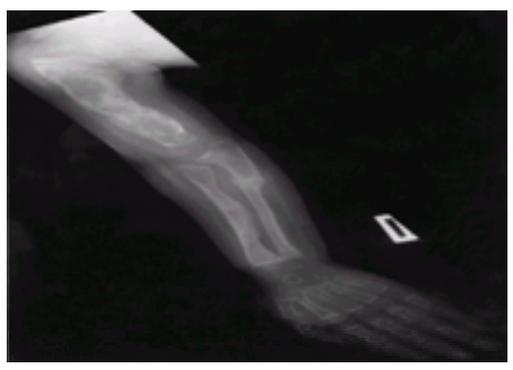


Fig. 9 Pathological fractures in a 2-year-old boy with osteogenesis imperfecta. Radiograph shows multiple fractures, diffuse decreased bone density, and transverse radiolucent metaphyseal bands.

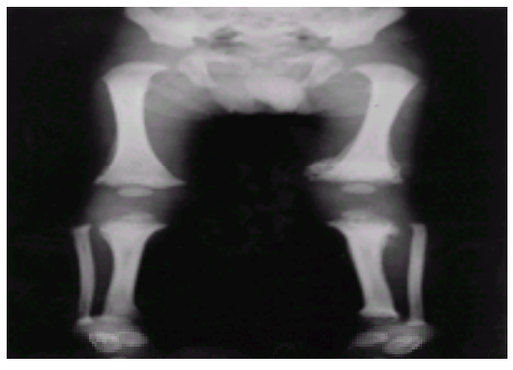


Fig. 10 Child abuse in a 4-month-old baby refusing to move his left limb. Radiograph shows bilateral metaphyseal corner fractures at different stages of healing.

2. transverse radiolucent metaphyseal band and osteolysis in metastatic neuroblastoma or leukemia (Fig. 11 and Fig. 12);

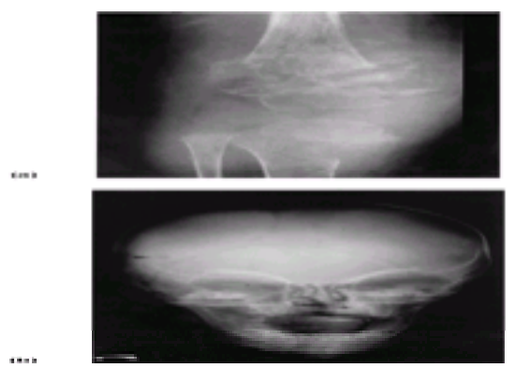


Fig. 11 Metastatic neuroblastoma in a 2-year-old boy with intense knee pain. (a) Radiograph of the knee. Extensive osteolysis and periosteal bone reaction. (b) Radiograph of the skull. Osteolytic area in the right parietal bone.

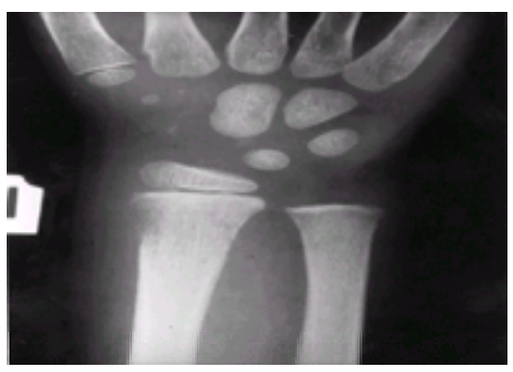


Fig. 12 Acute leukaemia in a 7-year-old-boy with wrist pain. Distal metaphyseal destruction of the cubitus on plain film.

3. joint effusion, subluxation of the femoral head with or without bone destruction in infant with septic arthritis of the hip;
4. posterior slippage of the femoral head indicative of slipped capital femoral epiphysis (Fig. 13);

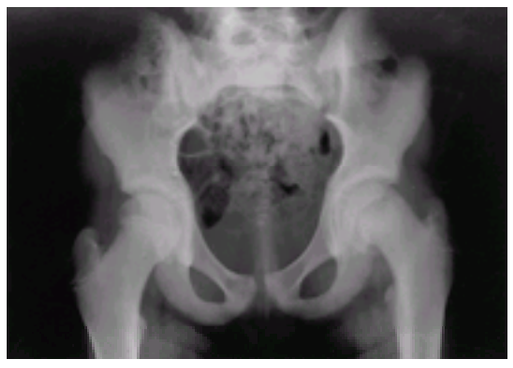


Fig. 13 Slipped capital femoral epiphysis in an 11-year-old boy.

5. disc narrowing with indistinct vertebral endplate indicative of spondylodiscitis;
6. small subchondral fragment within a defect in the medial femoral condyle indicative of osteochondritis dissecans of the knee;
7. small crescentic subchondral lucency of the femoral head on the frog-leg view indicative of Legg–Calve–Perthes' disease.

Furthermore, in the imaging evaluation of bone tumours or tumour-like disorders, plain radiographs are pivotal in directing further investigations. Plain films are very

often sufficient to separate lesions into two categories:

1. those that should be left alone, such as:
   - non-ossifying fibroma
   - fibrous dysplasia
   - chondroid lesions (Fig. 14 and Fig. 15)

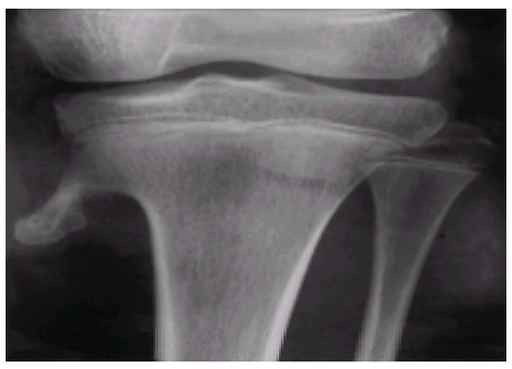


Fig. 14 Exostosis in a 13-year-old girl with a medial mass of the knee.

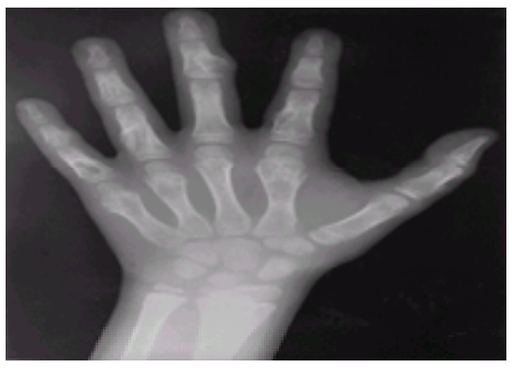


Fig. 15 Ollier disease in a 7-year-old boy with typical multiple enchondromas of the hand.

   - myositis ossificans
   - tumorous calcinosis
   - bone cysts
   - histiocytosis X (Fig. 16)

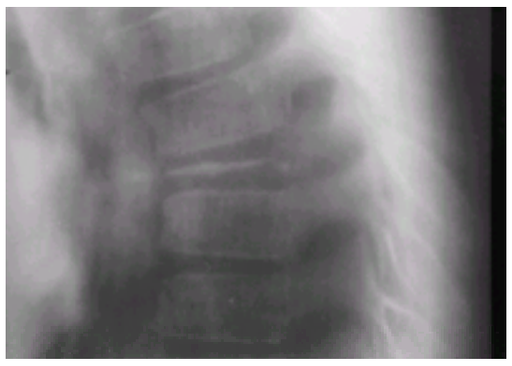


Fig. 16 Histiocytosis X in a 7-year-old boy with typical vertebra plana.

2. those that require intervention such as biopsy:
   - malignant lesions (sarcoma)
   - complicated lesions
   - undetermined lesions

It is important that developmental variations and benign lesions are not mistaken for malignancy. It is equally important that the diagnosis of malignancy is not delayed.

A skeletal survey is often required to establish if the lesion is solitary or polyostotic (histiocytosis X, fibrous dysplasia, malignancy) and to guide an eventual biopsy to the optimal site (Fig. 11, Fig. 12, and Fig. 16).

Chest films are often required in childhood rheumatic diseases. A frontal chest film is obtained most frequently. Lateral views are usually not much help but may be useful in the detection of pleural effusion and lymph node enlargement. Chest films are performed to rule out any pulmonary nodules, cavities or infiltrate, pleural effusion, mediastinal and/or hilar lymph node enlargement, and cardiomegaly (pericardial fluid, cardiac failure, myocarditis). Chest films are indicated especially in systemic juvenile chronic arthritis, systemic lupus erythematosus, scleroderma, dermatomyositis, Kawasaki disease, systemic vasculitis, sarcoidosis, rheumatic fever, the occurrence of a positive tuberculin test, or if otherwise clinically required.

Follow-up study

Conventional radiography is of major importance in evaluating disease progression and/or the effect of treatment in paediatric rheumatology in a reproducible and standardized manner (Jacobs 1982; Resnick 1988; Kaye 1990) (Fig. 2, Fig. 6, Fig. 17, and Fig. 18). In the literature, various methods of radiographic grading system have been proposed, especially in juvenile chronic arthritis and haemophilic arthritis (Pettersson *et al*. 1980; Pettersson and Rydholm 1984). Grading systems in adult rheumatology (such as the Steinbrocker method of evaluation of rheumatoid arthritis) have not be adapted for paediatric rheumatology (Steinbrocker *et al*. 1949).

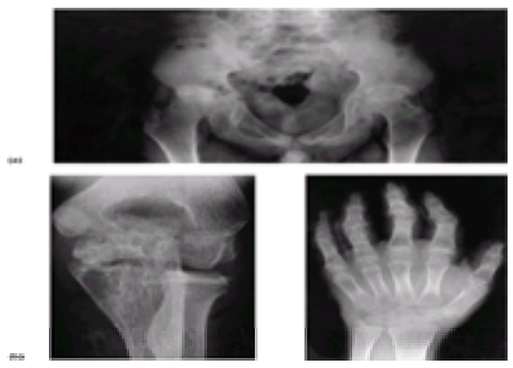


Fig. 17 Juvenile chronic arthritis in a 12-year-old girl with severe destructive abnormalities on the plain films of the hips (a), elbow (b), and the wrist and hand (c).

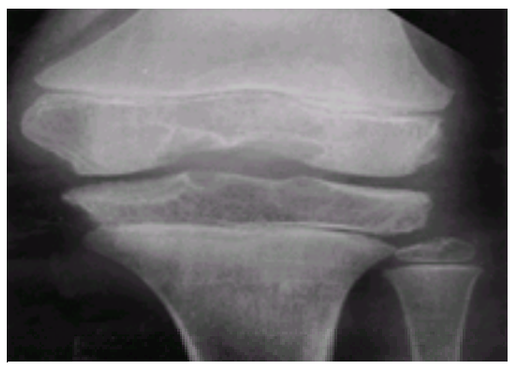


Fig. 18 Juvenile chronic arthritis in a 6-year-old girl with destructive abnormality on the plain films of the knee.

In children, less attention is paid to the loss of joint space. The reason is that estimating the joint space in children (especially young) is difficult because of :

- the thickness of the epiphyseal cartilage
- weight-bearing difficulties on the painful joint
- projectional error associated with flexion and/or valgus deformity
- joint space loss (often late and/or less prominent feature)

In children, scoring systems pay much attention to growth abnormality (Dale *et al*. 1994) (Table 16). In juvenile chronic arthritis, a radiographic classification system is often applied to the knee, which is the most commonly affected joint in childhood (Table 16 and Table 17).

Grade 0: normal condition

Grade I: slight abnormality
juxta-articular osteoporosis
± joint effusion

Grade II: growth abnormality
epiphyseal, patellar overgrowth

Grade III: destructive abnormality
overgrowth+marginal bony erosions
± malalignment

Grade IV: severe destructive abnormality
marginal bony erosions
epiphyseal deformations
malalignment

Grade V: mutilating abnormality
pronounced alignment changes and destruction
bony ankylosis

Table 16 Radiological grading system applied to the knee in juvenile chronic arthritis (Dale et al. 1994)

	Femur	Tibia	Patella
Osteoporosis	0–1	0–1	0–1
Enlargement of epiphysis or patella	0–1	0–1	0–1
Erosion	0–1	0–1	0–1
Subchondral cyst deformation	0–1	0–1	0–1
Deformed joint surfaces	0–1	0–1	0–1
Score	0–5	0–5	0–5
Total score		0–15	

Table 17 Radiological scoring of knee destruction in juvenile chronic arthritis (Petterson and Rydholm 1984)

Pettersson and Rydholm (1984) proposed radiological scoring of the knee joint destruction in juvenile chronic arthritis by dividing the joint into three separate compartments (i.e. femur, tibia, and patella). Each of the bones participating in the articulation is examined for signs of osteoporosis, enlargement, erosions, subchondral cysts, and deformity of the joint surface. The presence or the absence of these parameters in the joint is allotted 1 or 0 points, respectively (Table 17). This scoring method has a very significant correlation with the clinical status.

Finally, evaluation of complications of disease and treatment (especially therapy) is commonly performed by looking for: delayed bone age, decreased bone density, local growth abnormalities, compression fractures, avascular osteonecrosis, and instability of cervical spine.

Doppler ultrasonography

Doppler ultrasonography use has expanded greatly for the musculoskeletal system and has become a primary imaging tool for many aspects of paediatric rheumatology. In infants and young children, ultrasonography shows the cartilaginous forms of bones that cannot be visualized with plain radiography. Use of Doppler ultrasonography continues to increase as resolution of the equipment improves and anatomic structures become more easily identified with high frequency transducers (5 to 7 MHz). Ultrasonography is widely available, safe, non-ionizing and non-invasive, easily accepted by children, and can be performed at the bedside, in intensive care, or in a neonatal unit. Ultrasonography also allows dynamic stress examination of the hip joint.

Colour and pulsed Doppler facilitate identification of vascular anomalies. With the advent of power Doppler, visualization of femoral head vascularization is achievable consistently in newborns and shows promise in the detection of occlusion of feeding nutrient vessels (septic arthritis, congenital dislocation of the hip) (Fig. 19). Therefore, in some particular conditions, ultrasonography has become the primary imaging study performed, especially in the hip, replacing plain films. The early detection of hip dysplasia or dislocation allows prompt treatment and is the key to successful management. By treating this condition at a young age, most of the sequelae that occur when congenital dislocation of the hip goes unrecognized can be avoided.

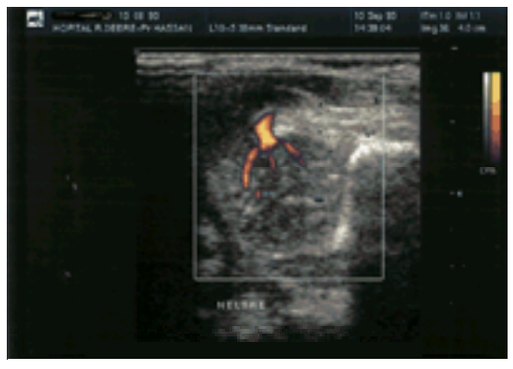


Fig. 19 Femoral head vascularization in a newborn. Power Doppler visualizes the branches of the circumflex arteries.

Clinical screening programmes have been instituted for all neonates. The use of imaging has been allied closely with programmes for detection and treatment. In the newborn period, plain films are unreliable since the key structures of the hip are composed of cartilage.

On ultrasonographic images, the cartilaginous components of the acetabulum and femoral head can be distinguished from other soft tissue structures. Real-time ultrasonography permits multiplanar evaluation which clearly defines femoral head dislocation in relation to the acetabulum (Keller *et al*. 1988). The ability to observe changes in hip position with movement is a further advantage of ultrasonography. The benefits of using ultrasonography in the diagnosis and management of congenital dislocation of the hip have been widely reported (Kalifa *et al*. 1991).

Variations in technique exist and, regardless of the method used, sufficient experience with both normal and abnormal hips should be gained. Initial assessment should be performed by 1 month of age (Keller *et al*. 1988; Kalifa *et al*. 1991). In most cases, referral for initial assessment is a questionable or abnormal physical examination or presence of an increased risk factor (first borns, breech position, oligohydramnios, family history, congenital torticollis, foot deformities). From a posterior approach, ultrasonography can be used to assess hip position after application of an abduction splint.

Ultrasonography is very useful in the detection of joint effusion especially in the hip, which is commonly evaluated in children. The common indications are the child with a limp who refuses to bear weight on the limb, and suspicion of septic arthritis in newborns or infants. Thus the crucial question is whether or not the hip is involved and specifically whether there is fluid in the hip joint. The examination of the hip is quick and within minutes anterior accumulation of fluid with joint capsule distension can be detected (Fig. 20). It is more important to rely on the configuration and shape than on absolute measurements. Furthermore, comparison with the opposite hip is always helpful unless bilateral effusions are present.

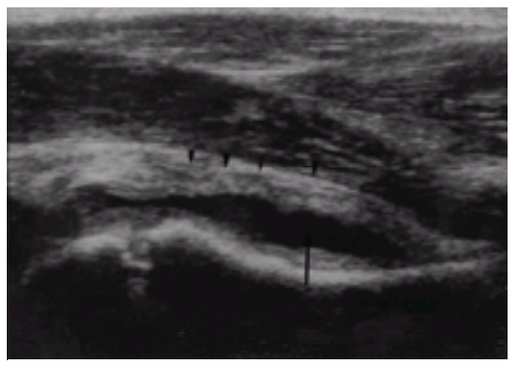


Fig. 20 Fluid in the hip joint. Ultrasonography shows anterior accumulation of fluid (arrow) with distension of the joint capsule and synovial thickening (arrowhead).

The use of ultrasonography is simple and more sensitive than plain films (reported sensitivity range: 88 to 100 per cent) (Marchal *et al*. 1987; Zieger *et al*. 1987). However, the nature of a joint effusion cannot be assessed accurately. We do not attempt to distinguish the different types of fluid and rely on the clinicians to determine which effusions should be tapped.

In the knee, ultrasonography can be used for evaluation of synovial hypertrophy, detection of synovial cysts, and exclusion of popliteal thrombosis (Fig. 21). In contrast, plain radiographs are not helpful in differentiating synovial effusions from synovial thickening. The distinction is easy with ultrasonography where synovial thickening appears as an irregular echogenic area surrounding the effusion (Fig. 20). Unfortunately, in chronic arthritis (such as juvenile chronic arthritis), ultrasonography does not permit reliable detection of joint loculation, which might make intra-articular injection of steroid difficult.

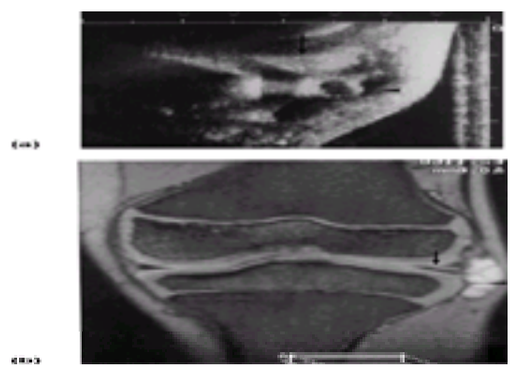


Fig. 21 Meniscal cyst in an 8-year-old boy with knee pain and medial mass. (a) Ultrasonography shows a cystic mass (arrowhead) adjacent to the medial meniscus (arrow). (b) Frontal T_2-weighted MR image shows a torn medial meniscus (arrow) and a meniscal cyst (arrowhead).

Cartilage erosions and thinning can be detected (Sureda *et al*. 1994). However, the whole surface of cartilage cannot be assessed. Ultrasonography does provide an objective method for observing disease course during therapy (Eich *et al*. 1994; Sureda *et al*. 1994).

Power Doppler also shows promise in evaluating the amount and the activity of pannus in juvenile chronic arthritis. Proliferative synovium, which is extremely

vascular, shows high power Doppler signal (Fig. 22).

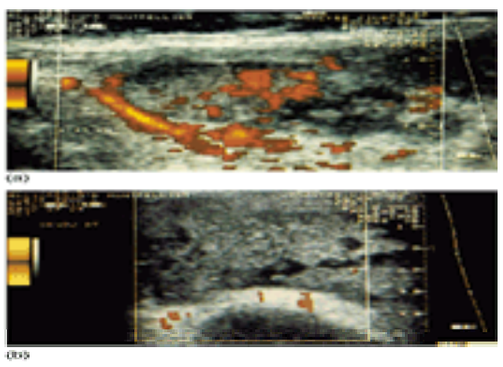


Fig. 22 Juvenile chronic arthritis affecting the knee in a 5-year-old girl (by courtesy of Dr A. Couture, Montpellier, France). (a) Power Doppler shows hypervascularization of the active pannus. (b) After intra-articular injection of steroid, there is a decrease of power Doppler signal within the pannus.

Finally, in a child presenting with a swollen red extremity, Doppler ultrasonography can be very useful by distinguishing:

- pyomyositis
- fluid collection (subperiosteal, abscess, bursitis)
- thrombophlebitis (Fig. 23)

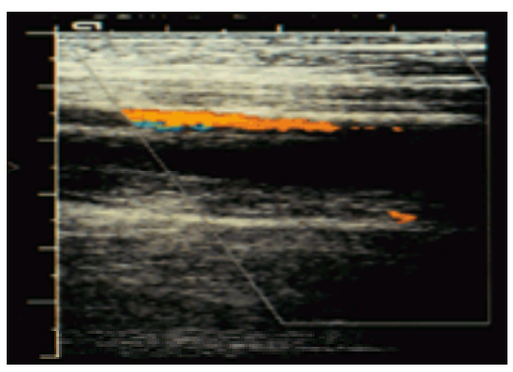


Fig. 23 Thrombophlebitis in a 15-year-old girl with Behçet's disease. Colour Doppler shows a thrombosis of the superficial femoral vein.

- non-opaque foreign body
- vascular malformation (Fig. 24)

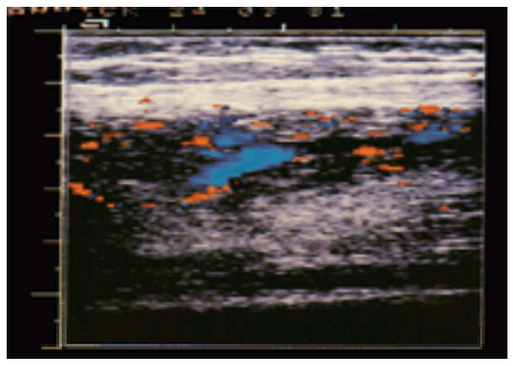


Fig. 24 Deep venous haemangioma of the calf in a 13-year-old girl with leg pain. Colour Doppler demonstrates a hypervascular lesion.

and guiding a needle for aspiration or catheter for drainage of fluid collection.

Ultrasonography can be employed easily in the initial evaluation and subsequent follow-up of soft tissue haemorrhage in trauma and bleeding disorders.

Abdominal ultrasonography can demonstrate peritoneal, pleural, and pericardial effusion, and hepatosplenomegaly in systemic juvenile chronic arthritis.

Echocardiography can demonstrate coronary aneurysms in Kawasaki disease or cardiac involvement in rheumatic fever.

Pulsed and colour Doppler can be useful in systemic vasculitis.

Arthrography and myelography

The use of arthrography and CT arthrotomography has decreased dramatically over the past few years in most paediatric practices because of the increasing availability of MRI and of arthroscopy. Arthrography requires general anaesthesia in young children and sedation in older children. The advantage of arthrography over CT and MRI is that it provides a dynamic study and depicts joints as they move. Hip arthrography is still an important test for assessing femoral head containment and assessing incongruity between the femoral head and the acetabulum (congenital hip dislocation, Legg–Calvé–Perthes' disease). It is an excellent tool in the operating room where it gives the surgeon an opportunity to visualize key relationships as the position of head is changed. It is a valuable aid for performing closed reduction in congenital dislocation of the hip and for guiding femoral or pelvic osteotomy in Legg–Calvé–Perthes' disease. Finally, arthrography is performed occasionally to aid detection of non-opaque loose bodies and the evaluation of articular surface irregularities.

Myelography has been widely replaced in children by CT and MRI, which are presently the first examinations to be performed. However, myelography is still very useful in the assessement of the cord in a few situations:

1. severe spinal deformities (spinal malformations, dysraphism, severe scoliosis) in which the interpretation of planar MR or CT images is very difficult;
2. surgical metallic spinal fixation in which the hardware creates CT and MR artefacts;
3. spinal instability (spondylolisthesis, spondylolysis) in which erected or dynamic films are crucial.

Angiography

In children, angiography is used only occasionally because of its invasive aspect and the low incidence of collagen vascular disease. In its place non-invasive,

vascular imaging modalities such as pulsed Doppler and colour Doppler ultrasonography, spiral CT angiography, and magnetic resonance angiography are now widely available. Multiple modalities can be performed including arteriography, phlebography, and direct percutaneous injection. Highly selective studies are now possible in infants using digital subtraction angiography, low osmolar contrast, and small catheters. Arteriography is indicated in the planning of resection of musculoskeletal tumours and can demonstrate neovascularity, displacement, encasement, or occlusion of major vessels.

Preoperative chemotherapy or embolization can be performed on children with sarcoma or an aneurysmal bone cyst. It also makes surgical procedure easier and less vascular. Therapeutic angiography can be particularly helpful in the management of synovial haemangioma. However, angiography cannot distinguish between malignant and benign disease. Cardiac angiography is indicated in Kawasaki disease to detect coronary aneurysms. Arteriography of the aorta and major arteries is useful in the evaluation of vasculitis, especially Takayasu's arteritis.

Balloon dilatation can be an effective treatment of arterial stenosis in Takayasu's arteritis. Phlebocavography is also very effective in the detection of thrombophlebitis and collateral flow in children at risk (Behçet's syndrome, coagulation disorders, lupus).

Radionuclide studies

Imaging the skeleton with bone-seeking radionuclide tracers offers the major advantage of high sensitivity in the detection of early pathological osseous changes in the entire skeleton (Van de Streck *et al*. 1994). In children, scintigraphy with bone-seeking and inflammatory agents, is the primary investigation to survey the skeleton for multifocal disease such as infection and malignancy (Fig. 1). Scintigraphy also provides critical information for the evaluation of musculoskeletal pain when plain radiographs are unrevealing. Disadvantages of bone scintigraphy are the lack of specificity and the poor spatial resolution. Correlating the results of a radionuclide study with information from other imaging techniques helps to overcome the lack of specificity inherent with this technique. The amount of radiation exposure during bone scintigraphy depends upon the radionuclide used and the dose administered. Children receive a portion of the adult dose calculated on the basis of body weight. The total body exposure from a technetium-99m bone study is comparable with that from a radiographic skeletal survey. Target organs such as the bladder and growth plates receive increased exposure. For bone scanning, the most commonly used and practical radiopharmaceutical in paediatrics is a phosphate compound labelled with technetium-99m.

The diagnostic accuracy of bone scintigraphy is closely related to the quality of the study. A technetium-99m bone scan is one of the most technically demanding nuclear imaging procedures. As with radiographs, meticulous attention should be given to positioning children for scintigraphic images. Positioning without rotation is essential. Homologous bones and joints must be shown with mirror-image positioning to allow careful comparison and detection of subtle asymmetry. Immobilization and/or sedation may be required. High-resolution techniques, using pinhole collimator views, are necessary to resolve small or subtle abnormalities that occur near the intense metabolic activity of the bone of the growing metaphyses of young children (Sullivan 1980). Long scanning times from 10 to 15 min per pinhole collimator view are required to obtain sufficient count density.

In general, images of the entire skeleton should be obtained even when symptoms are localized to a single site. This rule allows the recognition of unsuspected asymptomatic skeletal lesions. Pinhole views are used to resolve adequately the metaphyses, the epiphyses, and regions of the skeleton that are of specific clinical or radiological interest. Most false-negative diagnoses are secondary to metaphyseal lesions (osteomyelitis, corner fractures, and metastases) (Sullivan 1980). Familiarity with the scintigraphic appearance of the normal growth zone is necessary to detect subtle abnormalities in this region. The shape of the area of increased activity of the growth zone varies with age. In children of less than 18 months, the growth zone has an ovoid or elliptical form. After the age of 2 years, the growth zone is represented by a transverse linear band of increased activity that is always flat or slighty convex in the direction of the diaphysis. Moderate or marked convexity of this margin is abnormal. At all ages, the demarcation of the growth zone activity from the diaphysis is sharp.

Recent advances in nuclear radiology have improved the scintigraphic evaluation of musculoskeletal disease in children. A new generation of single-photon emission, computed tomography (**SPECT**) cameras can demonstrate subtle lesions that are undetectable by planar scintigraphy. State of the art SPECT systems feature multiple camera heads in a fixed ring design. More powerful post-processing has led to faster image reconstruction including sagittal, axial, and coronal planes and three-dimensional views. In paediatrics, SPECT has been particularly useful in the evaluation of the femoral heads for Legg–Calvé–Perthes' disease, in the evaluation of the lumbar spine for posterior element lesions (usually defects of the pars interarticularis), and finally in the evaluation of growth plate aberrations (varus and valgus deformity, cupped metaphysis, and post-traumatic arrest). In a study of 162 young patients presenting with back pain, SPECT demonstrated posterior element stress injury in 44 per cent, from which planar scintigraphy was abnormal in 20 per cent (Bellah *et al*. 1991).

Demonstration of occult fractures and sports injuries has increased in children with the use of bone scanning. Scintigraphic abnormalities, particularly of the tarsus and calcaneus, thought to reflect occult stress fracture, have been found in preschool children with unexplained lower extremity pain and normal radiographs (Englaro *et al*. 1992).

Bone scanning is useful in the evaluation of soft tissue and bone infection. Three phase bone scan is useful in the febrile infant or toddler when localization of infection can be difficult or in the older child with focal but non-specific complaints (Fig. 1). As in adults, scintigraphy help to distinguish a purely soft tissue infection from infection with bone involvement and to differentiate acute bone infarcts from acute osteomyelitis in children with sickle cell anaemia.

Leucocyte scans, using autologous cells labelled with indium-111 or technetium-99m detect osteomyelitis with high sensitivity (Lawson *et al*. 1994; Van de Streck *et al*. 1994). The use of leucocyte scans in paediatrics is more limited because of the moderately high radiation dose to the spleen. Leucocyte bone scans can be considered for children in whom the identification of a source of infection is critical and who can tolerate the removal of at least 20 ml of blood.

Other newer pharmaceuticals including polyclonal IgG, monoclonal antigranulocyte antibodies, and IgG fragments show promise as agents to detect inflammation (Oyen *et al*. 1992). A major advantage of these agents is the ease of preparation compared with that required for leucocyte labelling.

Finally, one disadvantage of scintigraphy is its low spatial resolution (Fig. 1). Therefore when a bone scans reveals a lesion that was not seen on plain films, the site of the lesion should be studied with a high resolution anatomical technique such as CT or MRI.

Computed tomography (CT)

The use of computed tomography has decreased dramatically over the past few years because of the superiority and increasing availability of MRI. However, CT is the better modality for imaging the mineralized portion of the bone, and is most effective when clinical evaluation, radiography, or bone scan have targeted a definite area.

The radiation dose received by a child during a CT examination is similar to that of multiple plain films and fluoroscopy. However, when multiple, overlapping thin slices are taken to obtain high resolution and/or three-dimensional studies, the volume of tissue irradiated is subject to more exposure than from a conventional radiograph.

Recent advances in CT software programs permit the data acquired from serial slices to be reformatted into any plane (even a curved plane) and to be reconstructed. The three-dimensional programs allow 'disarticulation' of joints and 'removal' or 'isolation' of any bone which may obscure the visualization of the underlying pathology or anatomy (Magid *et al*. 1988). Thus, three-dimensional CT provides a demonstration of complex anatomy, which can be of value in the management of complex malformations in children (Fig. 25).

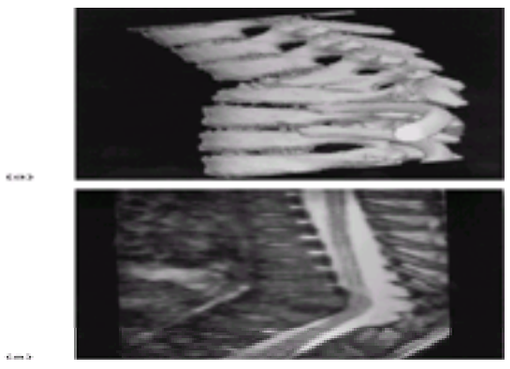


Fig. 25 Spondylodiscitis in a 12-year-old boy with dorsal kyphosis and vertebral block. (a) Three-dimensional CT shows the spinal deformity. (b) T_2-weighted MR image shows the cord compression.

The maximum preoperative information can be obtained by manipulating the data set in order to simulate the operative approach and the operation (osteotomy for instance). Three-dimensional CT can help in designing prostheses in children, notably in the hip (Swann 1993). This approach has proved to be valuable in a number of disorders such as neoplasms, malformation, or trauma affecting the pelvis, hip, feet, and spine.

Spiral CT, with the associated diminution of scanning times, facilitates high resolution and three-dimensional studies in children (Fig. 25).

CT techniques and software modifications have been developed so that quantitative measures of bone mineral can be performed. Evaluation of the lumbar spine for osteopenia is the usual procedure and normal values for children are available (Pettersson and Rydholm 1984; Gilsanz *et al.* 1988; Thomas, K.A. *et al*. 1991). CT offers an alternative to plain films for determining :

- femoral and tibial torsion (Hernandez *et al*. 1981)
- leg-length (using the scout view)

With careful positioning, almost any bone end can be imaged in a direct coronal or sagittal projection. In tarsal coalition, CT has proved to be the modality of choice. CT is the preferred method of evaluation in suspected cases of talocalcaneal coalition. The direct sagittal section is particularly valuable in demonstrating the anterior talocalcaneal articulation or coalition. For depiction of a calcaneonavicular coalition with CT, sections of the feet are obtained with the plantar surfaces perpendicular to the table.

The plane of the limb can be imaged directly with the gantry plane slightly off the true coronal or sagittal plane so as not to cut through the whole length of the bone (to avoid beam hardening artefacts). These views are particularly useful in the evaluation of partial closure of the growth plate. In these children, CT provides an excellent demonstration of the site and size of any bony bridges, but MR is the modality of choice for the evaluation of injured growth plate before the development of any bony bridge.

In sacroiliac arthritis, the orientation of sacroiliac joints and the ability to tilt the gantry to coincide with the plane of the joint make CT an excellent tool. The ability to demonstrate the joint completely with CT gives greater confidence in making a diagnosis of sacroiliac disease (Fig. 2).

CT remains the most satisfactory method for demonstrating sequestra in cases of chronic osteomyelitis (Hernandez 1985). CT is invaluable for definition of the osseous anatomy of lesions arising in bones that are difficult to image in two standard orthogonal planes, for example pelvis or vertebra, and for guiding a biopsy of such lesions when necessary.

CT has been an important adjunct to conventional radiography for the evaluation of the spine, however this role has now been challenged by MRI. CT offers the advantage of being more widely available especially in emergencies (trauma, spondylodiscitis). It is also the most sensitive modality for detecting calcification and ossification; this is invaluable for characterizing :

- calcifications in a chondroid matrix
- phlebolithis in soft tissue venous malformations or synovial haemangioma
- ossification in early myositis ossificans

CT is very effective in localizing precisely the site of the nidus of an osteoid osteoma (Fig. 4). Once the likely site of the osteoid osteoma has been determined with plain radiographs and/or bone scan, a high resolution CT study is performed in that limited area, and can be used to guide a biopsy. Attention should be directed to areas of high incidence of osteoid osteomas, such as the proximal femur where a reactive hip arthritis may be the presenting symptom (Fig. 4). Refinement of the technique has allowed percutaneous ablation of the nidus of osteoid osteoma.

Magnetic resonance imaging (MRI)

MR imaging has revolutionized the diagnostic evaluation of paediatric musculoskeletal disorders, becoming the modality of choice in many indications. The paediatric musculoskeletal system is ideally suited for investigation by MRI since it is the only technique which can clearly differentiate the individual components of the normal joints from one another (articular cartilage and fibrocartilage, growth plate, synovial membrane, joint effusion, ligaments). MRI also allows identification of bone marrow, muscles, tendons, fascia, nerves, and vessels with high contrast between these structures.

However one must not overlook the issue of clinical efficacy and cost-effectiveness. When deciding whether to use MRI, one should consider the financial cost, the time required, the need for sedation, and the operator dependency in comparison with good physical examination and conventional radiography. Therefore, physical examination must be performed and plain films obtained prior to any decision to use MRI. Thus cost-effectiveness is maximized and the diagnostic information of the MR study optimized by choosing the best protocol for a specific indication (Rubin and Kneeland 1994).

Technical considerations

Without the use of ionizing radiation, MR can safely produce high contrast, high resolution images of the paediatric musculoskeletal system in virtually any plane (Lawson *et al*. 1994). The success of MRI depends heavily upon multiple technical factors. The setting of the sequence needs to be adjusted for the individual examination. The magnetic field strengths and radiofrequencies currently employed with MRI produce no biological hazards. Immobilization is essential because of the intrinsic motion sensitivity of MRI. Unfortunately, the noise produced by the gradient coils of MR scanners disturbs young infants and children, albeit to a variable degree.

Some form of sedation may be required in young children but general anaesthesia is exceptional. MR examinations generally require more scan time (on the average 30 min) than do other imaging modalities such as CT and ultrasonography. Therefore, careful monitoring of children is essential (including heart rate, oxygen saturation). Newer magnet configuration designs such as small bore magnets for MRI of the extremities are being developed currently. They should be cheaper to build and run. 'Open' magnets are better tolerated by children and offer the potential for unrestricted motion studies and weight-bearing position (Lawson *et al*. 1994).

In children, the choice of the appropriate surface coil is crucial and is dependent on both the body part to be examined and the clinical question to be answered (Fig. 26). The smaller size of children may allow designated coils for adults to be used for a different body part of the child. In some instances, a coil designed to image an adult's head may be effective for imaging simultaneously both hips of an infant or a small child (or an 'orbit' coil for imaging a single joint). In plane spatial resolution of structures smaller than 1 mm with high signal and contrast-to-noise ratio is routinely achievable with this modality in children.

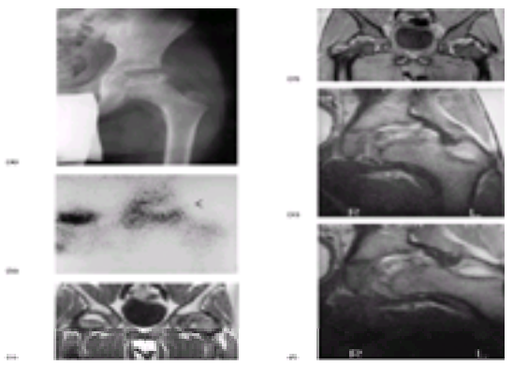


Fig. 26 Legg–Calvé–Perthes' disease involving the left hip in an 8-year-old boy. (a) Radiograph shows a typical fragmented appearance of the femoral head with lateral loss of containment. (b) Bone scan shows a decreased uptake in the necrotic bone. (c) T_1-weighted MR image shows the extent of necrosis (arrow). (d) Gradient-echo image shows the cartilaginous containment of the femoral head and the extent of the osteolytic area. (e–f) T_2-weighted MR images, without and with abduction, guide the femoral or pelvic osteotomy.

Pulse sequences

The appropriate choice of pulse sequences is critical to high-quality diagnostic MRI of the musculoskeletal system (Georgy and Hessenlink 1994; Rubin and Kneeland 1994). The ideal combination of parameters depends on the anatomy being examined and the suspected abnormality in addition to the available hardware, time constraints, and local preferences. Each sequence has relative advantages and disavantages. Spin-echo sequences are the mainstay of musculoskeletal MRI and are available on all MR systems producing either T_1, spin-density or T_2-weighted sequences (Fig. 27). Intravenous injection of paramagnetic contrast agent (gadolinium) is performed on T_1-weighted sequence (Fig. 27). The efficacy of spin-echo has been firmly established in various muskuloskeletal applications in the paediatric literature and any new sequence should be compared with spin-echo before its routine clinical use. The main disadvantage of spin-echo sequence is its relatively long acquisition time, especially for T_2-weighted sequences (7 to 14 min). More rapid imaging sequence is desirable in children in many cases.

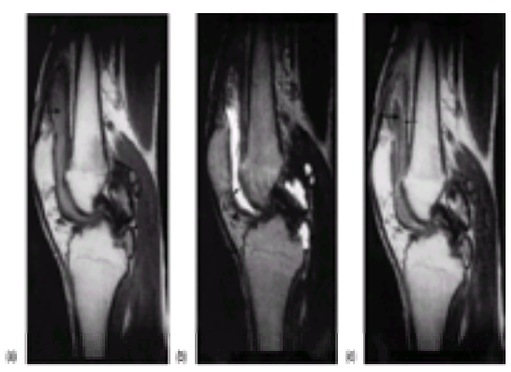


Fig. 27 Haemophilic arthropathy affecting the knee in a 14-year-old boy. (a) T_1-weighted MR image shows joint effusion (arrow) and synovial haemosiderin deposition (arrowhead). (b) T_2-weighted MR image shows articular surfaces with an arthrographic-like appearance (arrow). Haemosiderin deposition has a decreased signal (arrowhead). (c) Enhanced T_1-weighted MR image shows enhancement in the proliferative synovium (arrow) and allows a reliable distinction between active synovium and joint effusion (long arrow).

Gradient-echo and fast spin-echo allow faster imaging (Fig. 25). The use of gradient-echo with two-dimensional acquisition mode allows dynamic gadolinium-enhanced examinations studying musculoskeletal tissue vascularization and perfusion. This technique is maximized using post-processing software with pre- and post-contrast image subtraction (Fig. 28). The use of gradient-echo with a three-dimensional acquisition mode allows the highest spatial resolution (150 µm with a slice thickness of 0.7 mm) (Majundar *et al*. 1994). The acquisition of isotropic voxels allows reformatting in any plane and three-dimensional reconstruction.

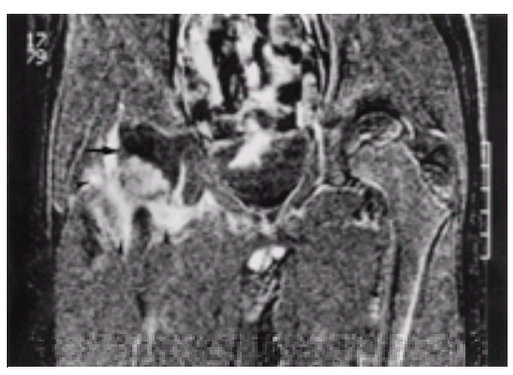


Fig. 28 Legg–Calvé–Perthes' disease mimicking a right hip arthritis in an 8-year-old boy. Enhanced substraction MR image shows an avascular right femoral head (arrow) with enhancing reactional synovitis (arrowhead).

The gradient-echo contrast can be advantageous for studying the trabecular bone structure and detecting haemorrhage (Sebag and Moore 1990) (Fig. 26). The use of the gradient-echo sequence is particularly useful in detecting osteolytic areas but can be problematic when identifying bone marrow oedema (Fig. 26). With certain gradient-echo parameters, epiphyseal, physeal, and articular cartilage appears brighter and better defined than it does on spin-echo images (Fig. 26). This contrast allows a very good demonstration of cartilage abnormalities in children. Fast spin-echo produces in a shorter time images with contrast similar to spin-echo. Fast spin-echo is especially advantageous for obtaining rapid, high resolution, heavily T_2-weighted images with a high signal-to-noise ratio. This is very useful for spinal studies with a myelographic effect (Georgy and Hessenlink 1994) (Fig. 25).

The transfer magnetization technique allows the contrast between joint fluid and articular cartilage to be increased, with an arthrogram-like effect (Wolf and Balaban 1994).

The fat-suppression technique by decreasing or eliminating the signal from fat is also increasingly used in musculoskeletal MR. It allows a good study of cartilage structures (which appear bright) and cartilaginous abnormalities, a better detection of gadolinium enhancement, and a better differentiation of subacute haemorrhage from fat (Peterfy *et al*. 1994).

Finally the short-TI inversion recovery (**STIR**) fat-suppressed sequence is a very sensitive screening method for detecting musculoskeletal disorders and especially marrow diseases in older children and teenagers (Fig. 29).

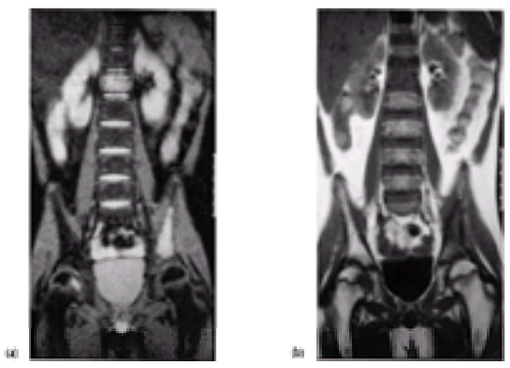


Fig. 29 Acute leukaemia relapsing in a 9-year-old boy with normal plain films. (a) Frontal T_1-weighted MR image. (b) Frontal STIR image. Infiltration of the T12 vertebral body, of the right femoral head, and of left iliac wing.

Normal and abnormal MR musculoskeletal appearance in children

It is important when assessing a joint that the observer is familiar with the appearance of a normal MR in the growing child. In the near-term fetus, the epiphyses are entirely cartilaginous. As ossified epiphyses develop with increasing age, there is gradual thinning of the epiphyseal cartilage. Therefore, the ossification centres have a bright signal representing the fatty marrow content. Growth plates and epiphyseal cartilage have an intermediate signal intensity on both T_1 and T_2 spin-echo sequences (Fig. 26 and Fig. 27). They have a bright signal on T_2* gradient-echo sequences and on fat-suppressed T_1 sequences (Fig. 26). Furthermore in children, the epiphyseal cartilage that surrounds the ossification centre, has a non-uniform signal pattern. A non-homogeneous hypointense zone is seen lying between the articular cartilage and the hemispherical growth zone of the ossification centre (Jaramillo and Hoffer 1992). This hypointense area is thought to represent a broad zone of poorly organized chondrocytes. With increasing age, the growth plate is reduced to a narrow zone between the epiphysis and the metaphysis, which shows an intermediate signal on T_1 and T_2 images and a high signal on T_2* and fat-suppressed images (Harcke *et al*. 1992) (Fig. 26).

Physiological closure of the growth plate begins centrally. Gadolinium-enhanced substraction MRI depicts the increased vascularization and perfusion of the metaphyseal growth zone (Debaert *et al*. 1992; Saifuddin *et al*. 1994) (Fig. 28).

The distribution of fatty and haematopoietic marrow changes in the paediatric skeletal system from infancy to adulthood (Moore and Sebag 1990). The high fat content of fatty marrow results in an increased signal on T_1-weighted sequences and a lower signal on T_2-weighted sequences. In contrast, the higher water and protein content of haematopoietic marrow results in a signal on T_1-weighted sequences that is low in neonates and slightly increased in the older child, together with a signal that is intermediate or increased on T_2-weighted sequence. After injection of contrast material, haematopoietic marrow shows a greater enhancement than fatty marrow, especially in small children and infants (Sze *et al*. 1991).

Starting in infancy, the haematopoietic marrow is converted progressively to a fatty marrow. This begins in the distal appendicular skeleton and progresses to the axial skeleton. In an individual bone, the conversion begins in the diaphysis and then extends to the metaphysis, reaching the adult pattern by 25 years. Knowledge of the type of marrow that is normally present in a given bone at a given age is essential to recognize any local or diffuse marrow abnormality in children (Sebag *et al*. 1993). T_1-weighted spin-echo sequences are optimal for showing the status of marrow conversion, while STIR and T_2-weighted sequences are better at distinguishing an abnormal marrow (Moore and Sebag 1990) (Fig. 29).

The articular surfaces are smooth and show up well on T_2 and T_2* sequences, with an arthrographic-like appearance (bright joint fluid) (Fig. 26 and Fig. 27).

The synovium is not seen in the normal joint. Without contrast enhancement, proliferative synovial tissue (i.e pannus) has a variable signal and cannot be distinguished reliably from joint effusion and cartilage (Hervé-Somma *et al*. 1992) (Fig. 27). After injection of a paramagnetic agent (gadolinium), proliferative synovium which is extremely vascular may be enhanced. Enhancement is indicative of disease activity and allows a precise assessment of pannus extension, joint effusion, and cartilage loss (Hervé-Somma *et al*. 1992).

Contrast enhanced MR is the most sensitive modality to determine an arthritis is present, but rarely helps in establishing a specific diagnosis (Fig. 30).

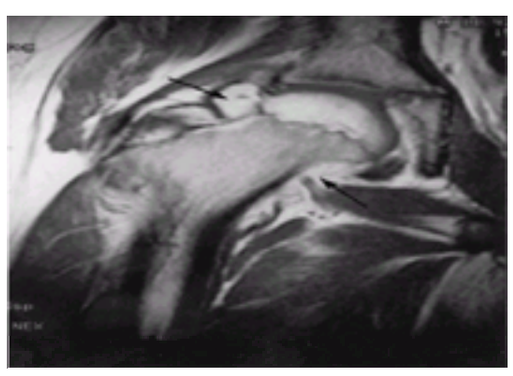


Fig. 30 Tuberculosis of the hip in an 11-year-old boy. Enhanced T_1-weighted MR image shows a non-specific intense enhancement of a dramatic synovial proliferation in the right hip (arrow).

In haemarthrosis and lipohaemarthosis, MR imaging can document the presence of haemorrhage owing to a predictable sequence in the chemical degradation of haemoglobin in extravasated blood (Fig. 27). Decreased signal intensity resulting from haemosiderin deposition can be demonstrated within an abnormal synovial proliferation (Fig. 27) However, although this finding is suggestive it is not specific and can be seen in haemophilia, bleeding disorders, pigmented villonodular synovitis, and intra-articular neoplasms such as haemangiomas (Kottal *et al*. 1987; Jelinek *et al*. 1989) (Fig. 31).

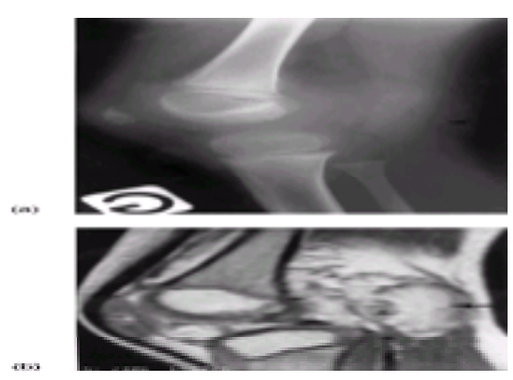


Fig. 31 Synovial haemangioma affecting the knee in a 5-year-old boy. (a) Plain film demonstrates a phlebolithis and a popliteal mass. (b) Enhanced T_1-weighted MR

image shows the extension of the synovial haemangioma.

Role of MRI in paediatric musculoskeletal system

MRI is an extremely sensitive modality for bone marrow imaging; however the pattern of MR changes is non-specific and the diagnosis may be dependent on clinical information (Bonnerot *et al*. 1994).

Joints are ideally suited for investigation by MRI in children (Recht and Resnick 1994). In many centres arthrography and CT arthrotomography have been replaced by MRI, which is currently used to evaluate congenital, traumatic, and inflammatory processes involving various joints. While the knees and hips are the joints most commonly examined with this modality, MRI of ankles, shoulders, wrist, elbow, and C1 to C2 are being performed with increasing frequency. Within the knee, MRI is ideal for determining internal derangement; discoid menisci and meniscal tears are easily visualized (Zobel *et al*. 1994) (Fig. 21). In osteochondritis dissecans, the extent of involvement can be evaluated as well as the status of the overlying cartilage (Rogers and Poznanski 1994).

In Legg–Calvé–Perthes' disease, MRI allows the detection and definition of the extent of necrotic bone, which is essential for the diagnosis, prognosis, and successful treatment of this condition (Ducou le Pointe *et al*. 1994) (Fig. 26 and Fig. 28).

Within the hip, MRI allows detection of vascular changes and assessment of femoral head incongruity. It is useful to evaluate femoral head containment within the acetabulum even in abduction (Fig. 26). The role of MRI compared with that of scintigraphy is still under study (Henderson *et al*. 1990; Conway 1993; Uno *et al*. 1995). MRI also allows early detection of osteonecrosis in children at risk and, especially, complications of steroid therapy (Mulliken *et al*. 1994). MRI is also useful in demonstrating late changes of congenital hip dislocation including deformity of the femoral head, acetabular dysplasia, infolding of the labrum, hypertrophy of the pulvinar, invagination of the ilio psoas tendon, and superimposed osteonecrosis (Johnson *et al*. 1988; Bos *et al*. 1988; Bos *et al*. 1991).

MRI is essential for preoperative planning in congenital limb deficiency by showing the unossified structures connecting the hypoplastic bone.

MRI is the most satisfactory modality for diagnosing possible growth plate abnormalities prior to the osseous abnormality (Norton *et al*. 1991; Rogers and Poznanski 1994).

The early detection of any damage (most often acute or chronic trauma) is essential to avoid severe growth deformity and long-term sequella (Jaramillo *et al*. 1990). MRI is very useful in the evaluation of sports injuries, showing occult fractures and trabecular bone impaction, which are seen increasingly in children and adolescent (Kapelov *et al*. 1993).

Inflammatory and synovial diseases in children are increasingly being studied by MRI, though its roles in these conditions have not yet been defined (Resnick 1988; Kaye 1990). In juvenile chronic arthritis, clinical evaluation of symptomatic joints is frequently supplemented with plain radiographs; but radiographic changes represent late and indirect signs of synovial disease.

Contrast-enhanced MRI can be used to evaluate precisely the extension of synovial proliferation, the status of articular cartilage, joint effusion, and bone erosion in juvenile chronic arthritis (Hervé-Somma *et al*. 1992) and haemophilic arthropathy (Fig. 32).

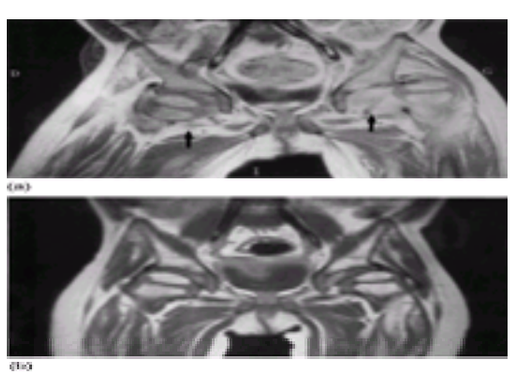


Fig. 32 Juvenile chronic arthritis. Systemic onset disease with hip involvement in a 3-year-old girl. Enhanced T_1-weighted MR image: (a) Prior to intra-articular injection of steroid: bilateral, active enhancing pannus is seen (arrow) with left hip dislocation. (b) After intra-articular injection of steroid: bilateral decrease in volume and enhancement of pannus with normal femoral head containment.

Cartilage loss is detected better on MRI than on plain films, especially in young children with thick, growing cartilage and at the early stage of the disease (Senac *et al*. 1988; Hervé-Somma *et al*. 1992).

Contrast-enhanced MRI is promising in the evaluation of the effectiveness of intra-articular therapy in juvenile chronic arthritis (Hervé-Somma *et al*. 1991; Eich *et al*. 1994). Enhanced MRI allows better evaluation of residual anatomic lesions after treatment. MRI may be used to predict subsequent therapeutic failure (fluid loculations) or early relapse (persistence of pannus enhancement) (Fig. 32). It is particularly valuable for hip appraisal in which clinical evaluation is difficult.

MRI is now the preferred approach for evaluating children for lesions of the spine and spinal cord (either congenital, traumatic, inflammatory, or tumorous).

Finally, as in the adult, MRI has proved to be essential in the evaluation of :

- extent of soft tissue tumours and bone tumours (Norton et al.1991)
- evaluation of the response to treatment of musculoskeletal tumours (Debaert *et al*. 1992)
- musculoskeletal infection (acute and chronic osteomyelitis)

In osteomyelitis, MRI should be performed in cases that may require surgical drainage including infections of the spine or pelvis, infections extending into the growth plates of long bones, and infections that fail to respond to antibiotics (Dangman *et al*. 1992; Jaramillo *et al*. 1995) (Fig. 1).

In conclusion

Selection of the appropriate sequence of imaging studies for a given child and a given diagnostic problem has become a complex undertaking. Investigations have to be tailored to the individual need of the child, and radiologists and paediatricians have to work together to ensure that the needs of all are best matched. Clinical examination and/or plain films remain the initial investigations (see Boxes 1, 2, 3, and 4). The need for radiation protection must remain a priority in children. Finally, in this era of advancing imaging technology, a knowledge of the relative value of available imaging techniques is necessary to optimize the management of children with musculoskeletal disease.

Chapter References

Ansell, B.M. (1990). Classification and nomenclature. In: *Update in paediatric rheumatology*, (ed. P. Woo, P. White, and B.M. Ansell), pp. 3–5. Oxford University Press.

Bellah, R.D. *et al*. (1991). Low back pain in adolescent athletes: detection of stress injury to the pars interarticularis with SPECT. *Radiology*; **180**, 509–12.

Blumhagen, J.D. (1994). The child with a limp. In: *Practical pediatric radiology* (ed. S.W. Hilton, D.K. Edwards, and J.W. Hilton), pp. 391–442. W.B. Saunders, Philadelphia.

Bonnerot, V. *et al*. (1994). Dota-gadolinium enhanced MRI of painful osseous crisis in children with sickle cell anemia. *Pediatric Radiology*, **24**, 92–5.

Bos, C.F.A., Bloem, J.L., Obermann, W.R., and Rozing, P.M. (1988). Magnetic resonance imaging in congenital dislocation of the hip. *Journal of Bone and Joint Surgery*, **70B**, 174–8.

Bos, C., Bloem, J.L., and Verbout, A.J. (1991). Magnetic resonance imaging in acetabular residual dysplasia. *Clinical Orthopaedics*, **265**, 207–17.

Conway, J.J. (1993). A scintigraphic classification of Legg–Calvé–Perthes disease. *Seminars in Nuclear Medicine*, **23** (4), 274–5.

Dale, K., Paus, A.C., and Laires, K. (1994). Radiographic classification system in juvenile rheumatoid arthritis applied to the knee. *European Radiology*, **4**, 27–32.

Dangman, B.C. *et al*. (1992). Osteomyelitis in children: gadolinium-enhanced MR imaging. *Radiology*, **182**, 743–7.

Debaert, T., Vanel, D., Shapeero, L.G., Charpentier, A., Terrier, P., and Di Paola, M. (1992). Osteosarcoma after chemotherapy: evaluation with contrast material-enhanced substraction MR imaging. *Radiology*, **185**, 587–92.

Ducou Le Pointe, H., Haddad, S., Silberman, B., Filipe, G., Monroc, M., and Montagne, P. (1994). Legg–Calvé–Perthes' disease: staging by MRI using gadolinium. *Pediatric Radiology*, **24**, 888–91.

Eich, G.F., Halle, F., Hodler, J., Seger, R., and Willi, U.V. (1994). Juvenile chronic arthritis: imaging of the knees and hips before and after intraarticular steroid infection. *Pediatric Radiology*, **24**, 558–63.

Englaro, E.E., Gelfand, M.J., and Paltiel, H.J. (1992). Bone scintigraphy in preschool children with lower extremity pain of unknown origin. *Journal of Nuclear Medicine*, **33**, 351–4.

Georgy, B.A. and Hessenlink, J.R. (1994). MR imaging of the spine: recent advances in pulse sequences and spatial techniques. *American Journal of Roentgenology*, **162**, 923–34.

Gilsanz, V. *et al*. (1988). Vertebral bone density in children: effect of puberty. *Radiology*, **166**, 847–50.

Greulich, W.W. and Pyle, S.J. (1959). *Radiographic atlas of skeletal development of the hand and wrist*. Stanford University Press, Stanford.

Harcke, H.T., Synder, M., Caro, P.A., and Bowen, J.R. (1992). Growth plate of normal knee: evaluation with MR imaging. *Radiology*, **183**, 119–23.

Henderson, R.C., Renner, J.B., Sturdivant, M.C., and Greene, W.B. (1990). Evaluation of magnetic resonance imaging in Legg–Calvé–Perthes' disease: a propective, blinded study. *Journal of Pediatric Orthopedics*, **10**, 289–97.

Hernandez, R.J. (1985). Visualization of small sequestra by computerized tomography: report of 6 cases. *Pediatric Radiology*, **15**, 238–41.

Hernandez, R.J., Tachdjian, M.O., Poznanski, A.K., and Dias, L.S. (1981). CT determination of femoral torsion. *American Journal of Roentgenology*, **137**, 97–101.

Hervé-Somma, C., Sebag, G., Lallemand, D., Touzet, P., and Prieur, A.M. (1991). Gadolinium-enhanced MR imaging of juvenile chronic arthritis before and after intraarticular therapy. *Radiology*, **181**, 303.

Hervé-Somma, C., Sebag, G.H., Prieur, A.M., Bonnerot, V., and Lallemand, D. (1992). Juvenile rheumatoid arthritis of the knee: MR evaluation with Gd–DOTA. *Radiology*, **182**, 93–8.

Jacobs, J.C. (1982). *Pediatric rheumatology for the practitioner* (2nd edn). Springer-Verlag, Berlin.

Jaramillo, D. and Hoffer, F.A. (1992). Cartilaginous epiphysis and growth plate: normal and abnormal MR imaging findings. *American Journal of Roentgenology*, **158**, 1105–10.

Jaramillo, D., Hoffer, F.A., Shapiro, F., and Rand, F. (1990). MR imaging of fractures of the growth plate. *American Journal of Roentgenology*, **155**, 1261–5.

Jaramillo, D., Treves, T.S., Kasser, J.R., Harper, M., Sundel, R., and Laor, T. (1995). Osteomyelitis and septic arthritis in children: appropriate use of imaging to guide treatment. *American Journal of Roentgenology*, **165**, 399–403.

Jelinek, J.S. *et al*. (1989). Imaging of pigmented villonodular synovitis with emphasis on MR imaging. *American Journal of Roentgenology*, **152**, 337–42.

Johnson, N.D., Wood, B.P., and Jackman, K.V. (1988). Complex infantile and congenital hip dislocation: assessment with MR imaging. *Radiology*, **168**, 151–6.

Kalifa, G., Koifman, P., Demange, P.H., Sellier, N., Fraconal, C., and Seringe, R. (1991). Apport de l'échographie dans le dépistage et l'aide au diagnostic de la luxation congénitale de hanche. *Revue d'Imagerie Médicale*, **3**, 93–7.

Kapelov, S.R. *et al*. (1993). Bone contusions of the knee: increased lesion detection with fast spin-echo MR imaging spectroscopic fat saturation. *Radiology*, **189**, 901–4.

Kaye, J.J. (1990). Arthritis: roles of radiography and other imaging techniques in evaluation. *Radiology*, **177**, 601–8.

Keats, T.E. (1991). *An atlas of normal variants that may simulate disease* (5th edn). Year Book Medical Publishers, Chicago.

Keller, M.S., Weltin, G.G., Rattner, Z., Taylor, K.J.W., and Rosenfield, N.S. (1988) Normal instability of the hip in the neonate: US standards. *Radiology*, **169**, 733–8.

Kottal, R.A., Vogler, J.B., Matamoros, A., Alexander, A.H., and Cookson, J.L. (1987). Pigmented villonodular synovitis: a report of MR imaging in two cases. *Radiology*, **163**, 551–3.

Kushner, D.C. *et al*. (1986). Radiation dose reduction in the evaluation of scoliosis: an application of digital radiography. *Radiology*, **161**, 175–81.

Lawson, J.P., Keller, M.S., and Rattner, Z. (1994). Recent advances in pediatric musculoskeletal imaging. In *Advances in musculoskeletal imaging* (ed. D.J. Sartoris), **32** (2), pp. 353–75. W.B. Saunders, Philadephia.

Magid, D. *et al*. (1988). 2D and 3D computed tomography of the pediatric hip. *Radiographics*, **8**, 901–33.

Majundar, S., Genant, H.K., Grampp, S., Jergas, M.D., Newitt, D.C., and Gies, A.A. (1994). Analysis of trabecular bone structure in the distal radius using high-resolution MRI. *European Radiology*, **4**, 517–24.

Mandell, G.A., Harcke, T.H., and Kumar, S.J. (1990). Strategies and technique. In *Imaging strategies in pediatric orthopaedics* (ed. M.E. Kricun), pp. 1–17. Aspen Publication, Gaithersburg, Maryland.

Marchal, G.J. *et al*. (1987). Transient synovitis of the hip in children: role of US. *Radiology*, **162**, 825–8.

Moore, S.G. and Sebag, G.H. (1990). Primary disorders of bone marrow. In *Magnetic resonance imaging of children* (ed. M.D. Cohen and M.K. Edwards), pp. 824–65. B.C. Decker, Philadelphia.

Mulliken, B.D., Renfrew, D.L., Brand, RA., and Whitten, C.G. (1994). Prevalence of previously undetected osteonecrosis of the femoral head in renal transplant recipients. *Radiology*, **192**, 831–4.

Norton, K.I., Hermann, G., Abdelwahab, I.F., Klein, M.J., Granowetter, L.F., and Rabinowitz, J.G. (1991). Epiphyseal involvement in osteosarcoma. *Radiology*, **180**, 813–16.

Oyen, W.J. *et al*. (1992). Diagnosis of bone joint and joint prosthesis infections with in-111-labeled nonspecific human immunoglobulin G scintigraphy. *Radiology*, **182**, 195–9.

Ozonoff, M.B. (ed.) (1992). Special radiologic techniques measurements and calculations. In *Pediatric orthopedic radiology* (2nd edn), pp. 765–85. W.B. Saunders, Philadelphia.

Peterfy, C.G. *et al*. (1994). Quantification of articular cartilage in the knee with pulsed saturation, transfer saturation and fat suppressed MR imaging: optimization and validation. *Radiology*, **192**, 485–91.

Pettersson, H. and Ringertz, H. (1991). *Measurements in pediatric radiology*, pp. 20–98. Springer-Verlag, London.

Pettersson, H. and Rydholm, V. (1984). Radiologic classification of knee joint destruction in juvenile chronic arthritis. *Pediatric Radiology*, **14**, 419–21.

Pettersson, H., Ahlberg, G.A., and Nilsson, I.M. (1980). A radiologic classification of hemophilic arthropathy. *Clinical Orthopaedics*, **149**, 153–9.

Poznanski, A.K. (1976). *Practical approaches to pediatric radiology*. Year Book Medical Publishers, Chicago.

Recht, M.P. and Resnick, D. (1994). MR imaging of articular cartilage: current status and future directions. *American Journal of Roentgenology*, **163**, 283–90.

Resnick, D. (1988). Common disorders of synovium-lined joints: pathogenesis, imaging abnormalities, and complications. *American Journal of Roentgenology*, **151**, 1079–92.

Risser, J.C. (1948). Important practical facts in the treatment of scoliosis. *American Academic Orthopedic Surgery*, **5**, 248–60.

Rogers, L.F. and Poznanski, A.K. (1994). Imaging of epiphyseal injuries. *Radiology*, **191**, 297–8.

Rubin, D.A. and Kneeland, J.B. (1994). MR imaging of the musculoskeletal system: technical considerations for enhancing image quality and diagnostic yield. *American Journal of Roentgenology*, **163**, 1155–63.

Saifuddin, A., Bann, K., Ridgway, J.P., and Butt, W.P. (1994). Bone-marrow blood supply in gadolinium-enhanced magnetic resonance imaging. *Skeletal Radiology*, **23**, 455–7.

Sebag, G.H. and Moore, S.G. (1990). Effect of trabecular bone on the appearance of marrow in gradient-echo imaging of the appendicular skeleton. *Radiology*, **174**, 855–9.

Sebag, G., Dubois, J., Tabet, M., and Lallemand, D. (1993). Spinal bone marrow in children: assessment of normal age related changes in MR appearance. *Pediatric Radiology*, **23**, 515–18.

Senac, M.O. *et al*. (1988). MR imaging juvenile rheumatoid arthritis. *American Journal of Roentgenology*, **150**, 873–8.

Steinbrocker, O., Trager, G.H., and Butterman, R.C. (1949). Therapeutic criteria in rheumatoid arthritis. *Journal of the American Medical Association*, **140**, 659–62.

Sullivan, D.C. (1980). Problems in the scintigraphic detection of osteomyelitis in children. *Radiology*, **135**, 731–6.

Sureda, D., Quiroga, S., Arnal, C., Boronat, M., Andreu, J., and Casas, L. (1994). Juvenile rheumatoid arthritis of the knee. Evaluation with US. *Radiology*, **190**, 403–6.

Swann, M. (1993). Total hip replacement for juvenile chronic arthritis. *Clinical and Experimental Rheumatology*, **11**, 47–50.

Swischuk, L.E. (1994). The extremities. In *Emergency imaging of the acutely ill or injured child* (3rd edn), pp. 361–567. Williams and Wilkins, Baltimore.

Swischuk, L.E. and John, S.D. (1995). Bones and soft tissues. In *Differential diagnosis in pediatric radiology* (2nd edn), pp. 187–342. Williams and Wilkins, Baltimore.

Swischuk, L.E., Swischuk, P.N., and John, S.D. (1993). Wedging of C-3 in infants and children: usually a normal findings and not a fracture. *Radiology*; **188**, 523–6.

Sze, G., Bravo, S., Baierl, P., and Shimkin, P.M. (1991). Developing spinal column: gadolinium enhanced MR imaging. *Radiology*, **180**, 497–502.

Thomas, K.A., Cook, S.D., Bennet, J.T., Whitecloud III, T.S., and Rice, J.C. (1991). Femoral neck and lumbar spine bone mineral densities in a normal population 3–20 years of age. *Journal of Pediatric Orthopedics*, **11**, 48–58.

Thomas, P.S., Renton, P., Hall, C., Kalifa, G., Dubousset, J., and Lalande, G. (1994). The musculo skeletal system. In *Imaging children* (ed. H. Carty, F. Brunelle, D. Shaw, and B. Kendall), pp. 845–1291. Churchill Livingstone, London.

Uno, A., Hattori, T., Noritake, K., and Suda, H. (1995). Legg–Calvé–Perthes' disease in its evolutionary period: comparison of MRI with bone scintigraphy. *Journal of Pediatric Orthopedics*, **15**, 362–7.

Van de Streck, P.R., Carretta, R.F., and Weiland, F.L. (1994). Nuclear medicine approaches to musculoskeletal disease: current status. In *Advances in musculo skeletal imaging*, (ed. D.J. Sartoris). *Radiologic Clinics of North America Imaging*, **32** (2), 227–54.

Wolf, S.D. and Balaban, R.S. (1994). Magnetization transfer imaging: practical aspects and clinical applications. *Radiology*, **192**, 593–9.

Zieger, M.M., Dorr, U., and Schulz, R.D. (1987). Ultrasonography of hip joint effusions. *Skeletal Radiology*, **16**, 607–11.

Zobel, M.S., Borrello, J.A., Siegel, M.J., and Stewart, N.R. (1994). Pediatric knee MR imaging: pattern of injuries in the immature skeleton. *Radiology*, **190**, 397–401.

4.10 Histopathology

Patrick J. Gallagher and Janice R. Anderson

Most chronic rheumatological disorders affect multiple tissues and organ systems (Fig. 1 and Fig. 2). In turn, this generates a wide variety of different histological biopsies. Here, the problems associated with the handling, processing, and interpretation of the more specialized rheumatological biopsies are considered. The value of skin, gastrointestinal, renal, hepatic, pulmonary, marrow, and other biopsies have been considered in various chapters in Section 1.

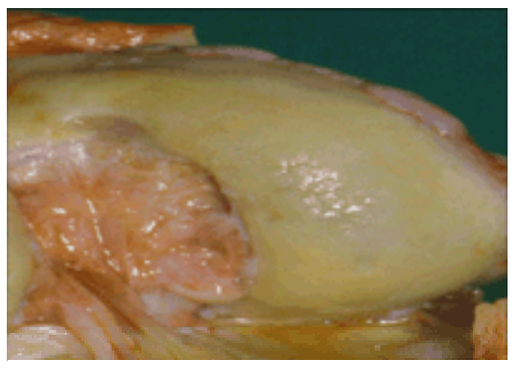


Fig. 1 Early rheumatoid disease. There is a synovial hyperplasia in the intertrochanteric area and a small erosion on the medial aspect of the condyle.

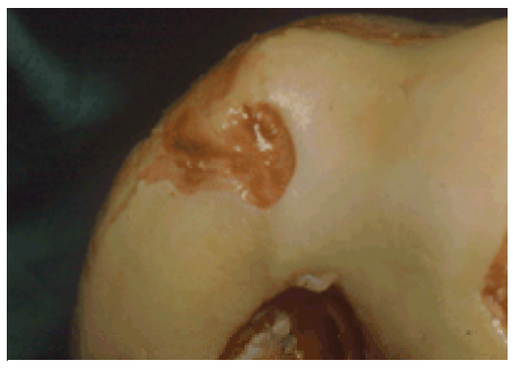


Fig. 2 Rheumatoid disease. An erosive lesion with a base of pannus.

Synovial biopsies

Methods of closed synovial biopsy were described over 50 years ago and, even before the introduction of the narrow-gauge biopsy needle, there were comprehensive accounts of the histological changes in a variety of rheumatic disorders. In everyday practice most synovial biopsies are reported non-specifically in descriptive terms. This is unsatisfactory for both rheumatologists and pathologists, and has generated mutual scepticism as to the value of the procedure. In response to this, computerized methods have been used to assess the importance of many different histological features in large series of synovial biopsies (Soren 1978; Rosenberger *et al*. 1981). These detailed studies have clarified exactly which alterations are most common in each disorder. However, as the range of histological changes is limited, the precision with which diagnoses can be made in individual patients has not necessarily increased.

Technical aspects

Most biopsies obtained with a needle or during arthroscopy are 5 to 10 mm in maximum dimension and ideally three or more fragments are required. It has been

calculated that at least 2.5 mm^2 of synovial tissue are needed to reflect the variation in cellular density expected in rheumatoid synovium (Kennedy *et al*. 1988a).

Formalin fixation is satisfactory for routine histology and an increasing number of simple immunohistochemical techniques. However, many monoclonal antibodies directed against leucocyte antigens only give good results in frozen sections and these may be required in experimental studies. If this is the case, one fragment should be snap frozen in liquid nitrogen and kept at as low a temperature as possible. Urate may be dissolved out by formalin and alcohol is the ideal fixative if gout is suspected.

No special processing techniques are required but multiple levels should be examined and spare sections reserved for additional staining methods.

Evaluation of biopsies

Pathologists who do not work closely with rheumatological units may have difficulties in interpreting synovial biopsies. This is usually ascribed to the limited range of changes that develop in synovium and the paucity of diagnostic features. At least 10 per cent of biopsies are unsatisfactory, either because the material retrieved is too small for evaluation or only collagenous joint capsule has been biopsied. Comparatively large biopsies of synovium can be obtained during joint replacement operations and are an invaluable source of material for investigation. As more 're-do' replacements are performed an increasing number of abnormal synovial samples are submitted for evaluation (see below). Although stereological methods of quantification are used increasingly in assessing bone biopsies, they have not been routinely applied to synovial biopsies (Artacho-Perula *et al*. 1994).

Synovial lining cells

In a normal biopsy the synovial lining is a thin monolayer of flattened or cuboidal basophilic cells, closely applied to the underlying fibroadipose tissue (Fig. 3). Although ultrastructural studies have defined at least two distinct cell types, these cannot be distinguished at the light microscopic level, at least in routine needle-biopsy material. The larger A, or M for macrophage-like, cells have a phagocytic function. The B cells (not B lymphocytes) resemble fibroblasts and have ultrastructural appearances that indicate active synthesis.

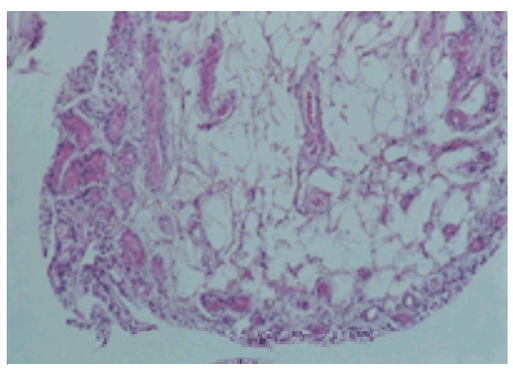


Fig. 3 A normal biopsy. The bulk of the biopsy is loose fibroadipose tissue. Blood vessels are prominet but the appearances are within normal films.

Synovial hyperplasia is diagnosed easily when the lining cells form a clear multilayer but if the villous pattern is accentuated, even hyperplastic synovium may have a monolayered appearance (Fig. 4). Normal synovial lining cells are only slightly larger than underlying fibrocytic nuclei but in most reactive conditions they enlarge and may show considerable variation in size and shape. Areas of synovial ulceration and fibrin deposition are common features in many different forms of acute synovitis (Fig. 5 and Fig. 6).

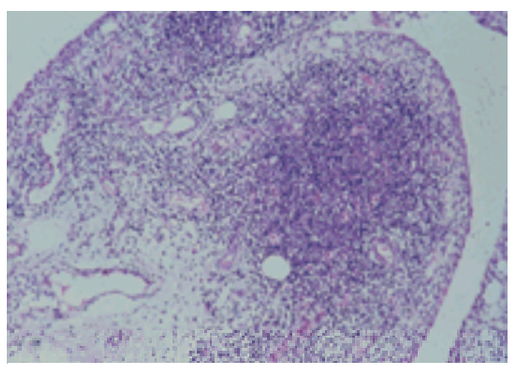


Fig. 4 Chronic rheumatoid arthritis. The synovium shows villous hypertrophy and there is a prominent lymphoid aggregate

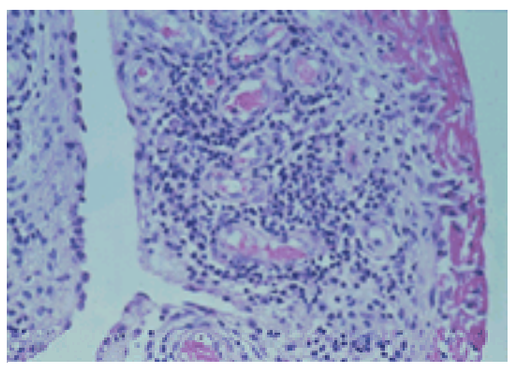


Fig. 5 Chronic rheumatoid arthritis. There is a diffuse lymphocytic infilltration, the synovial membrane is ulcerated, and there is a superficial layer of fibrin (right hand side)

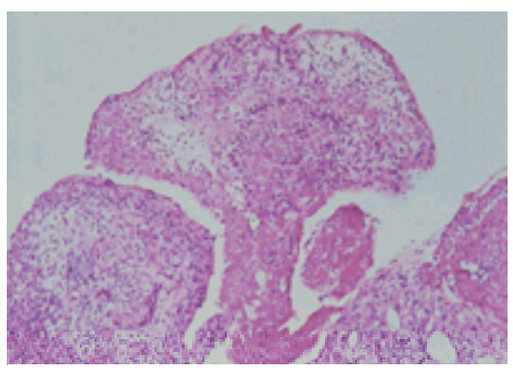


Fig. 6 Early Rheumatoid arthritis. There is a mass of fibrin in the centre and the underlying synovium is infiltrated with acute and chronic inflammatory cells.

Inflammation

Inflammatory infiltration and increased vascularity are the most important histological features that should be evaluated in synovial biopsies. Occasional groups of chronic inflammatory cells may be seen in synovial biopsies from healthy young subjects or early in the clinical course of post-traumatic synovitis. Immunohistochemical studies have shown that most of these cells are T lymphocytes and there is clearly an overlap with mild chronic synovitis (Lindblad and Hedfors 1987). More extensive inflammation must be regarded as pathological, especially if there is an admixture of cells or the pattern of infiltration is diffuse. Large nodular aggregates of lymphocytes and macrophages are commonly seen, especially in rheumatoid disease and related disorders. In contrast, true granulomatous lesions are rare and usually indicate a specific disorder such as sarcoidosis.

Synovial vasculature

In acute synovitis, increased vascularity with perivascular margination and emigration of leucocytes may be prominent. Although hyperaemia is a major clinical and pathological feature of acute synovitis, its histological expression is highly variable. The normal synovial membrane has a rich blood supply and it can be very difficult to assess a marginal increase in vascularity (see Fig. 3). In the most acute stages the appearance resembles granulation tissue, with associated haemorrhage and fibrin deposition. The capillaries are often distended and engorged, and there may be focal endothelial necrosis. In chronic inflammatory synovitis, small and medium-sized arteries and veins may be prominent and show fibrous intimal thickening. A uniform, eosinophilic or hyaline appearance in the vessel wall should suggest amyloidosis but we stain routinely all synovial biopsies with Congo red.

Histological changes in individual disorders

Although histopathologists rarely make a definitive diagnosis on a synovial biopsy, there are some general features that suggest particular disorders and a few changes that indicate specific diseases (Table 1).

Frequent changes, no specific diagnostic value
Hyperplasia of synovial lining cells
Villous synovial hyperplasia
Synovial ulceration and fibrin deposition
Acute and chronic inflammatory cell inflammation
Chronic inflammatory aggregates, with germinal centre formation
Less frequent changes, possible diagnostic value
Granulomatous lesions in joint capsule (rheumatoid arthritis) or synovium (sarcoidosis)
Dense acute inflammatory infiltration with micro-organisms (infective arthritis) or urate crystals (gout)
Subsynovial deposits of amyloid
Dense haemosiderin deposits (repeated haemorrhage, e.g. haemophilia)

Table 1 Histopathological changes in synovial biopsies

Rheumatoid disease

The histological changes in rheumatoid arthritis have been described in numerous reports but there are no specific diagnostic features. Non-committal histological reports may be annoying to rheumatologists but there are many reasons why these must be the norm. The knee is usually biopsied but this may not be the most severely affected joint. Within the knee itself there is some variation in the pattern of histological change (Hutton *et al*. 1987), though in general terms there is a good association between the macroscopic and microscopic features of inflammation (Lindblad and Hedfors 1985). In individual joints the clinical severity of synovitis can be correlated with the degree of histological change, particularly in the early stages of disease (Rooney *et al*. 1988). Pathologists should be told what treatments have been given, especially if second-line drugs were used before the biopsy was taken.

The chief microscopic features in classical rheumatoid arthritis are mononuclear cell infiltration, villous synovial hyperplasia (Fig. 7), synovial ulceration and fibrin deposition, fibrosis, and increased vascularity (Fig. 5 and Fig. 6). Although there are no features that allow a specific diagnosis of rheumatoid disease, there is some value in grading the density of synovial hyperplasia, vascularity, and inflammation. Patients with substantial abnormalities early in their clinical course tend to develop persistent disease (Gallagher *et al*. 1985). Disease activity can be correlated with the intensity of ferritin staining within synovial macrophages. The amount of ferric iron, as shown by Perls' staining, is some guide as to how the disease might progress (Blake *et al*. 1984).

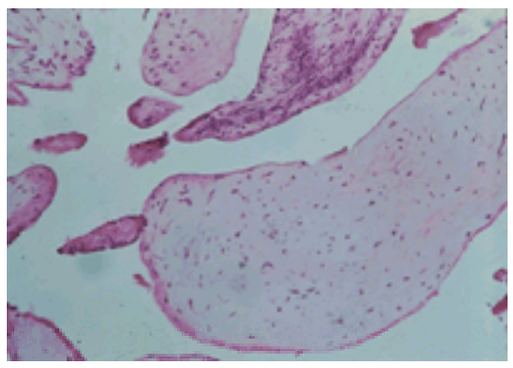


Fig. 7 Vilous synovial hyperplasia in long-standing rheumatoid arthritis. In this field there is no associated inflammation.

Studies with monoclonal antibodies have shown that T-helper lymphocytes are the most numerous of the mononuclear cells in active rheumatoid disease. The density of the inflammatory infiltrates and the ratio of helper to suppressor lymphocytes is reduced in patients who respond to treatment (Rooney *et al*. 1989). There is some evidence that mast cells have a role in the pathogenesis of inflammatory joint disease. They are difficult to recognize in routine sections but paraffin sections stain well with antibodies to mast cell tryptase.

As immunohistochemistry is not of proven diagnostic value, frozen sections are probably superfluous in routine practice. However, if abundant tissue is available, a small sample should be kept frozen. Recent experimental studies of rheumatoid synovium have clarified the ways in which inflammatory cells and regulatory peptides interact in rheumatoid disease (Edwards and Cambridge 1995) (see Chapter 3.1). At present there is no role for specialized molecular investigations in routine synovial biopsies. This may change as instruments are developed that allow combined techniques, such as polymerase chain reaction (PCR) and *in situ* hybridization, to be performed on tissue sections.

Osteoarthritis

Although osteoarthritis is primarily a degenerative disorder there can be substantial inflammatory cell infiltration in severely affected joints (Fig. 8). When osteoarthritic and rheumatoid synovium are compared morphometrically, marked inflammation can be seen in both groups of patients but well-developed lymphoid follicles and significant numbers of plasma cells are restricted to rheumatoid disease. Nevertheless, in one study, 40 per cent of osteoarthritic samples could not be differentiated from rheumatoid biopsies and half of all osteoarthritic synovia had moderate or marked inflammation (Goldenberg *et al*. 1982). Immunohistochemical studies suggest that, in contrast to rheumatoid disease, the numbers of subsynovial CD4 and CD8 cells are about equal (Kennedy *et al*. 1988b). In needle biopsies of osteoarthritis associated with chondrocalcinosis there are no histological features that reliably distinguish them from needle biopsies obtained from cases of uncomplicated osteoarthritis.

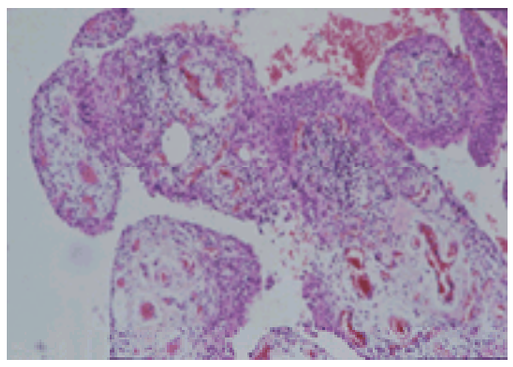


Fig. 8 Villous synovial hyperplasia and moderate associated chronic inflammation in osteoarthritis. Inflammation of this degree is not uncommon.

Other arthropathies

Non-specific inflammatory changes are seen in most of the less common forms of chronic joint disease such as psoriatic or enteropathic arthropathy, juvenile chronic arthritis, Reiter's disease, Behçet's syndrome, or ankylosing spondylitis. There is no single histological feature that is in any way diagnostic or even suggestive of an individual disorder. In both adults and children the density of the inflammatory infiltrates is usually less than in typical, long-standing rheumatoid disease. In a comparative study of patients with ankylosing spondylitis and rheumatoid arthritis the histological features were similar but the lymphocyte helper to suppressor ratios were only increased in patients with rheumatoid disease (Kidd *et al*. 1989).

Acute purulent inflammation is seen in infective arthritis, gout, some reactions to intra-articular foreign bodies, and the early stages of rheumatoid disease (see Fig. 6). In unusual disorders, such as sarcoidosis or amyloidosis, a specific histological diagnosis is sometimes possible.

Prosthetic joints

Metallic debris and fragments of polyethylene or methacrylate bone cement are phagocytosed by synovial macrophages and may accumulate in the synovium. An exuberant reaction may be seen in synovial biopsies taken during 're-do' joint replacements. Abraded polyethylene fragments may stimulate a marked macrophage and foreign body giant cell reaction. The material is strongly birefringent.

Salivary gland biopsy

Inflammatory cell infiltration of the lacrimal and salivary glands may occur as an isolated event or in association with a variety of autoimmune diseases. Together with the associated fibrosis and loss of glandular tissue, production of saliva and tears is reduced. This causes dryness of the mouth and eyes and in turn may progress to a sicca syndrome with keratoconjunctivitis.

The typical, well-developed lesions in major salivary glands in Sjögren's syndrome have been termed 'benign lymphoepithelial lesions' or 'myoepithelial sialadenitis'(MESA). Although glandular tissue is lost, larger salivary ducts may be preserved and the subsequent proliferation of their lining epithelium and myoepithelial covering produces typical nodular masses—'myoepithelial islands'. There is considerable interest in the nature of the associated lymphocytic infiltrates, which range from patchy and histologically innocuous aggregates to malignant lymphoma. Patients with Sjögren's syndrome and MESA have about a 40-fold risk of developing malignant lymphoma, either in affected salivary glands or other nodal or extranodal sites. Furthermore, they have an increased incidence of lymphoproliferative abnormalities such as lymphocytic interstitial pneumonitis and unusual forms of atypical hyperplasia in lymph nodes. There is also an increased risk of malignant lymphoma in patients with MESA without the clinical features of Sjögren's syndrome (Isaacson and Norton 1994).

Technical aspects

Malignant lymphoma should be suspected in any patient with an autoimmune rheumatic disease who presents with salivary gland enlargement or cervical lymphadenopathy. If biopsy is undertaken it is essential that fresh tissue is sent to an immunopathology laboratory as quickly as possible. Although an increasing range of antibodies give satisfactory immunohistochemical staining in formalin-fixed, paraffin-processed tissues, many important monoclonal antibodies require frozen sections for optimal demonstration of reaction product. Furthermore, fresh tissue is essential for chromosomal studies and may be more suitable for the detection of immunoglobulin and T-cell-receptor gene rearrangements.

In Sjögren's syndrome both major and minor salivary glands are involved (Greenspan *et al*. 1974) and the small glands on the inner aspect of the cheek and lower lip are biopsied easily (Fig. 9). The lower lip is everted and a 15- to 20-mm incision is made under local anaesthesia, parallel to the vermilion border in the labial mucosa. Some four to seven nodules of minor salivary gland are required. In normal subjects the procedure is relatively innocuous; occasionally a feeling of numbness around the biopsy site remains. In patients with severe stomatitis the biopsy site may become infected and ulcerate.

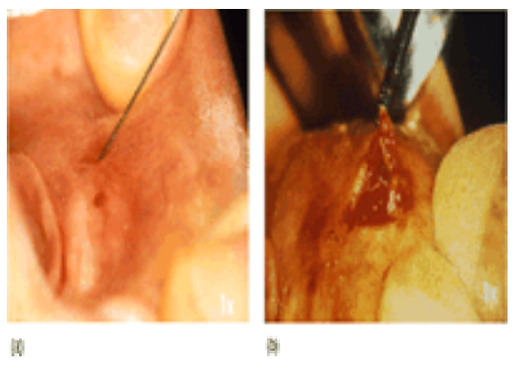


Fig. 9 (a), (b) Biopsy of a minor salivary gland on the inner aspect of the lower lip.

Evaluation of biopsies

Major salivary glands

Ideally, salivary gland biopsies in which there is a suspicion of a lymphoproliferative disorder should be referred to centres with considerable experience of oncological pathology. Specialized laboratory techniques are continually adding to our knowledge of basic aspects of the biology of malignant lymphomas and a general pathologist is rarely able to follow the changes in terminology that these advances necessitate. Furthermore, oncologists are seldom willing to treat these patients until the histology has been reviewed formally by pathologists with whom they have a close working relationship. In the most detailed accounts of the histological and immunohistochemical changes in Sjögren's syndrome, several different groups have emphasized the difficulties encountered in assessing and classifying the changes in myoepithelial sialadenitis (McCurley *et al*. 1990). Immunohistochemical and molecular hybridization techniques have demonstrated the clonal nature of the B-cell infiltrates in many of these lesions (Fishleder *et al*. 1987). Isaacson and his colleagues suggest four histological categories (Hyjek *et al*. 1988):

1. early myoepithelial sialadenitis with focal lymphoid infiltration;
2. established myoepithelial sialadenitis with dense lymphoid infiltration but no evidence of light-chain restriction (i.e. monotypic immunoglobulin staining);
3. established myoepithelial sialadenitis with histological and immunohistological evidence of early lymphoma;
4. frank malignant lymphoma.

Early diagnosis of malignant lymphoma requires careful cytological analysis of the lymphoid infiltrates and good quality immunohistochemical preparations. Virtually all of these are B-cell MALT-type lymphomas, some with a characteristic monocytoid appearance (Shin *et al*. 1991). In the International Lymphoma Study Group Classification (Harris *et al*. 1994) these are grouped under marginal zone B-cell lymphomas.

Minor salivary gland biopsies

Labial gland biopsies from patients with a sicca syndrome show a variety of histological changes, which have been quantified and contrasted with the appearances of glands obtained from control subjects at autopsy (Daniels 1984). Some are entirely free of inflammation and others contain only occasional, small aggregates of 50 or more lymphocytes. For convenience results are sometimes expressed as the number of these aggregates per 4 mm^2. When there are multiple or confluent foci there is usually some loss of glandular tissue. In the most inflamed glands, very little acinar tissue remains and there can be marked duct distension and fibrosis (Fig. 10). Multiple lymphoid aggregates are most unusual in autopsy samples from previously healthy subjects but areas of diffuse inflammation with associated loss of glandular tissue may be seen.

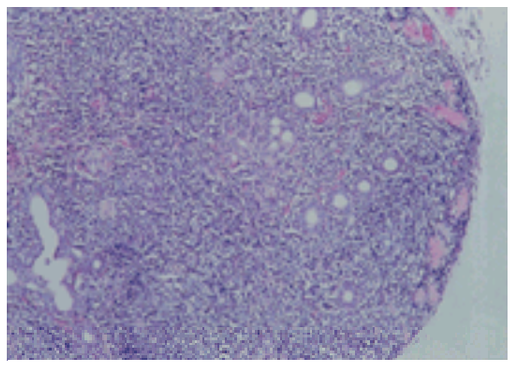


Fig. 10 Labial salivary gland biopsy in a patient with sicca syndrome. The normal architecture of the gland has been effaced by diffuse lymphocytic infiltrate.

Biopsies are most often taken in the investigation of patients with an obscure sicca syndrome. In a study of over 360 cases, Daniels demonstrated that multiple lymphoid aggregates (>1 focus/4 mm^2) in adequate biopsies of labial salivary glands were better indicators of a Sjögren's syndrome than reduced salivary flow from the parotid or subjective symptoms of a dry mouth (Daniels 1984). The technique has also been used in the diagnosis of sarcoidosis. In a study of 40 consecutive biopsies from patients with rheumatological disorders, 10 were entirely normal, 9 had advanced inflammation and fibrosis, and the remainder slight or moderate changes (Harvey *et al*. 1989). There was no evidence of malignancy or granuloma formation. In contrast, non-caseating granulomas (Fig. 11) were present in 5 of 25 patients in whom there was a strong clinical suspicion of sarcoidosis.

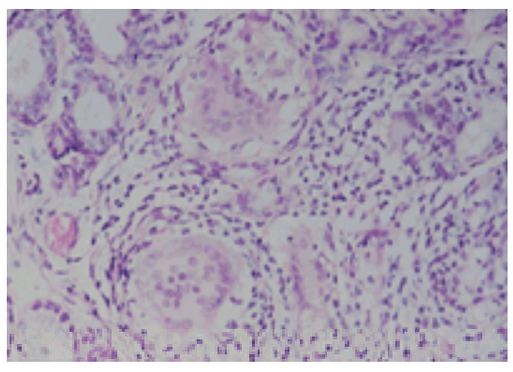


Fig. 11 A non-caseating granuloma, with prominent giant cells, in the labial salivary gland of a patient with sarcoidosis.

Labial salivary gland biopsy could be used more widely. There is no reason why the newer, molecular techniques should not be applied to material extracted from these glands, as comparatively little tissue is required for extraction of enough DNA. A recent study used *in situ* hybridization to demonstrate that 13 of 70 of patients with Sjögren's syndrome had light chain restriction in labial gland biopsies. Four of these 13 patients later developed extra salivary lymphomas (Jordan *et al*. 1995).

Bone biopsy

In systemic bone disorders, such as osteomalacia, renal osteodystrophy, and hyperparathyroidism, bone biopsy can be a valuable aid in establishing a diagnosis and in assessing the severity of the disease process and the response to treatment. The role of biopsy in the common forms of osteoporosis is less clear but it certainly has a place in the investigation of patients less than about 50 years old and in those with associated abnormalities of calcium metabolism (Eriksen *et al*. 1994). Biopsies should only be referred to pathologists whose laboratories have the technical expertise to process and section undecalcified bone and who have sufficient experience to evaluate and quantify the histological changes.

Technical aspects

Eriksen and his colleagues have produced an excellent short guide to bone biopsy which is required reading for pathologists and rheumatologists with an interest in metabolic bone disease (Eriksen *et al*. 1994). The terminology can be confusing but an international group has attempted to standardize the nomenclature and specialist journals, such as *The Journal of Bone and Mineral Research*, give succinct guidelines in their instructions to authors. Virtually all biopsies are from the ilium,

usually within 20 to 30 mm of the anterior superior iliac spine. The iliac crest should be avoided as the demarcation between cortical and trabecular bone can be indistinct (Vigorita 1984). A core should be taken in a horizontal plane through the full thickness of the gluteal surface. An ideal biopsy should, therefore, include the inner and outer cortical plates and the full width of trabecular bone. A 5-mm Jamshidi needle produces a more than adequate biopsy. Cores of 10 mm are less likely to fragment but in our experience this advantage is offset by the added difficulties in processing and sectioning.

Tetracycline is a naturally fluorescent antibiotic that binds to immature bone at the 'calcification front'—the interface between osteoid and mineralizing bone. Tetracycline fluorescence can therefore be used to estimate how much of the bone surface is undergoing mineralization. If two doses are given at a fixed interval the rate of bone calcification can also be calculated (Fallon and Teitelbaum 1982). Two separate 3-day courses of oxytetracycline are given (250 mg, four times a day) at least 10 days apart. The biopsy should be taken no less than 3 days after the end of the second course. The results are unpredictable, especially in severe osteomalacia, when it may be difficult to identify two clear lines of fluorescence. If there is strong clinical evidence of osteomalacia or renal bone disease, it is advisable to increase or prolong the doses of tetracycline and to extend the intervals between the two courses of antibiotic, and between the second dose and the biopsy, as described below. When tetracycline has been given the biopsies should be fixed in 70 per cent alcohol rather than formalin.

Bone that has not been decalcified is difficult to section especially if it has been routinely processed into paraffin wax. Trabecular bone tends to tear and shatter during cutting and the alignment of marrow with the bony trabeculae can be lost. Bone biopsies can be dehydrated satisfactorily in modern tissue processors and we use a kit (JB4; Polysciences Inc.) to embed the samples in 2-hydroxyethylmethacrylate. Sections are then cut with glass knives at 3 µm. It is sometimes possible to distinguish osteoid from mineralized bone in haematoxylin and eosin stains but better definition is obtained with Goldner's trichrome method or the Van Kossa stain for calcium phosphate.

Histopathological assessment

Although simple histological assessment can allow a confident diagnosis in florid cases of osteomalacia (Fig. 12) and renal osteodystrophy, quantitative microscopy is required when osteoporosis is suspected or the changes of osteomalacia are more subtle. Almost all histological features can be measured with a suitable range of eye-piece graticules and highly reproducible results can be obtained by experienced observers. Most large laboratories now have some form of computerized measuring system and even the simplest of these can be adapted for bone histomorphometry. At low magnifications we use a drawing arm to produce a map of the trabecular architecture and shade in any important additional features. The drawings are then placed on a graphics tablet and the relevant areas outlined with a cursor. Well-focused, low-power photomicrographs with good contrast are notoriously difficult to produce and they have few advantages over good line drawings. If a darkroom is available, a projecting microscope can be used to shine a high-power image directly on to the graphics tablet. This is especially useful for measuring osteoid and osteoblastic or resorptive surfaces, and for counting osteblasts or osteoclasts. Fully automated image analysis systems have been adapted for bone histomorphometry but their cost is many times greater than an interactive system.

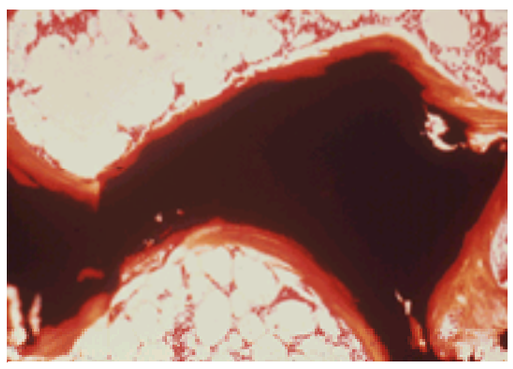


Fig. 12 Bone biopsy from a patient with osteomalacia. The material was not decalcified and was stained by the von Kossa method. An abnormally thick rim of non-calcified osteoid (orange-yellow colour) surrounds the mineralized trabeculum (black). The relative osteoid volume in this patient was 11.2 percent (normal range approximately 1.5–3.2 per cent).

Most pathologists assume that the information obtained from a two-dimensional image is a reflection of the results that would be obtained from the full three-dimensional structure. This is not necessarily true but recent reviews have outlined mathematical methods by which three-dimensional measurements can be estimated from a small number of sections taken at known intervals in a carefully defined plane (Cruz-Orive and Weibel 1990). These methods can be applied to bony tissue (Gundersen *et al*. 1988) and are especially useful for estimating trabecular bone volume.

Many different measurements can be made in bone biopsies (Melsen *et al*. 1978) but comparatively few of these are used in the diagnosis of osteoporosis and the various forms of osteomalacia (Table 2). The histological features of some variables, such as trabecular bone volume, are clear cut, at least if the sections are free of artefactual tears or fissures. If the measurements are made from a hard copy, usually a drawing or photograph, even an inexperienced pathologist should obtain reproducible results. There is much more variation when subtle microscopic changes are evaluated. For example trabecular resorptive surfaces are defined on the basis of their irregular or scalloped surfaces but the identification of these is highly subjective.

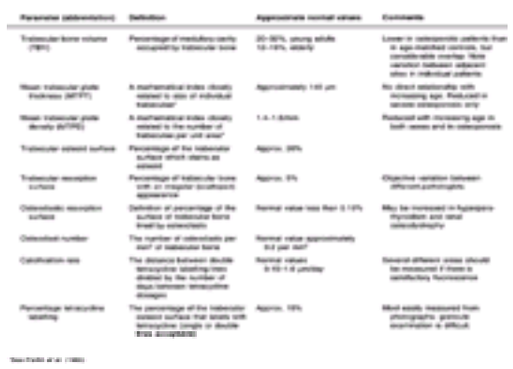


Table 2 Histological measurements in bone biopsies

Osteoporosis

As there are no reliable, qualitative features that distinguish osteoporotic from normal bone, some form of quantitation is essential for diagnosis. The comprehensive approach used in recent long-term studies of bisphosphonate treatment sets an ideal standard but is impractical in most routine laboratories (Storm *et al*. 1993). As a first step we measure trabecular bone volume and osteoid volume. Trabecular bone volume is defined as the percentage of the medullary cavity occupied by trabecular bone. In most studies of this volume there are differences between groups of patients with osteoporosis and healthy controls, but in elderly people there may be a considerable overlap. Furthermore, measurements of trabecular bone volume in biopsies taken from different sites in the same patient can vary by as much as 80 per cent (Chavassieux *et al*. 1985). Healthy young adults have trabecular bone volumes in the range 20 to 32 per cent but mean results in elderly subjects have been quoted as 11.9 and 19.9 per cent (normal) and 8.7, 9.6, and 13.3 per cent (osteoporotic patients). Patients with trabecular bone volumes of 20 per cent or greater are most unlikely to have osteoporosis, but volumes as low as 6 per cent can occasionally be recorded in elderly subjects with no clinical or pathological evidence of osteoporosis (Ashton-Key and Gallagher 1992).

Mean trabecular-plate thickness and mean trabecular-plate density can be derived from the trabecular bone volume using simple formulas (Parfitt *et al*. 1983). The

plate thickness is an indication of the size, and the density an index of the number of trabecular plates per unit area. By using these three measurements together it is possible to determine whether bone loss has resulted from loss of complete trabeculae, by thinning of trabeculae, or a combination of both. In non-osteoporotic controls there is an almost linear decline in both trabecular bone volume and mean trabecular-plate density with increasing age, and this is accentuated in patients with vertebral fractures. It is now accepted that the underlying cellular mechanisms in age-related and postmenopausal osteoporosis are different (Dempster and Lindsay 1993) (see also Chapter 5.17.1). In postmenopausal disease enhanced osteoclastic activity during the erosive phase of bone remodelling may cause deep perforations of trabeculae with a generalized loss of 'connectivity'. In contrast, in age-related osteoporosis there is a reduction in trabecular plate thickness but the trabeculae retain their 'connections' with each other. At least two methods have been described to assess this—the marrow star volume (Eriksen *et al*. 1994) and node-strut analysis. The value of these measurements is illustrated clearly in a recent report of hyperparathyroidism in postmenopausal women (Pariesen *et al*. 1995). The principles of these methods are illustrated in Fig. 13.

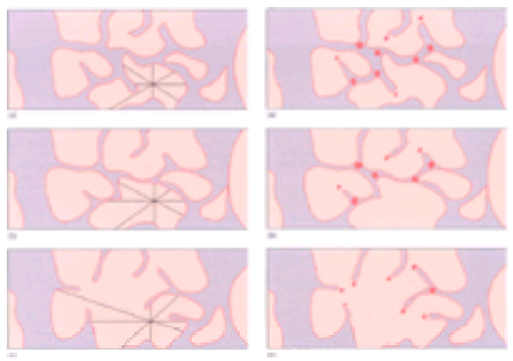


Fig. 13 Histopathalogical assessment of osteoporosis. The top panels (a) depict normal bone. Interconnecting bony trabeculae are in the centre and cortical bone at the edges. For illustrative purposes the proportion of cortical bone has been exaggerated. (b), in the centre, is age-related osteoporosis. The trabeculae are thinned out but remain connected. The trabecular bone volume (TBV) would be reduced. The lower panels (c) are severe postmenopausal osteoporosis. The trabeculae are thinned and, because they have been deeply eroded, have lost connectivity. The star volume (left hand panels) is obtained by measuring and adding the distances from a random point to the nearest bony trabeculum. It is slightly increased in age-related osteoporosis (b, left) but markedly increased when connectivity has been lost (c, left). In practice, several random histological levels would be examined and the results would be expressed volumetrically (see text). Node-strut analysis is illustrated in the right hand panels. A node (red circles) is a junction between three (or more) trabeculae, a strutis a connection between two nodes, and a terminus (red dot) is a blindly endingtrabeculum. In severe postmenopausal osteoporosis the number of nodes and struts are decreased and termini are increased. For full details refer to Pariesen *et al*.(Pariesen *et al*. 1995).

There is no evidence that osteoid volume is increased in either the 'postmenopausal' or 'senile' form of osteoporosis. Furthermore, most reports indicate that osteoid volume is not affected by increasing age. If a high osteoid volume is recorded in an osteoporotic patient, an additional form of metabolic bone disease should be suspected.

Osteomalacia

The histological appearances of florid cases of osteomalacia are characteristic (see Fig. 12) and histomorphometry is not strictly necessary for immediate diagnosis. Trabecular bone volume, osteoid volume, and osteoid surface should be measured in all cases, along with an index of hyperparathyroidism such as percentage trabecular or osteoclastic resorption surface. When tetracycline has been given it is usually possible to calculate the calcification rate and percentage tetracycline labelling. In practice a perfect pattern of fluorescence is seldom obtained. If severe osteomalacia or renal osteodystrophy is suspected the interval between the two courses of antibiotic should be increased and the biopsy taken 1 or 2 weeks after the last dose. Even if this is possible the number of true double lines may be small and the real value of this technique may be doubtful. Nevertheless, we use tetracycline in all adults in whom there is a suspicion of osteomalacia or renal osteodystrophy. It should not, of course, be used in children. Confocal laser microscopy can produce excellent fluorescent images of tetracycline-labelled bone biopsies but the instruments are expensive and not generally available to diagnostic histopathologists.

Biopsies from patients with renal bone disease or hyperparathyroidism should be handled in much the same way. Reabsorption surfaces should be measured. Some form of 'hard copy'—good drawings or well-prepared photomicrographs—makes these measurements easier. Nevertheless, with practice, consistent results can be obtained by direct microscopy. Osteoblasts and osteoclasts should be counted when there is any suspicion of hyperparathyroidism. Their distribution is not random in these disorders, nor even in normal bone. Because of this a set of measurements made from a single section may be imprecise. Gundersen and his colleagues have described simple techniques that allow more accurate estimation of irregularly distributed histological features (Gundersen *et al*. 1988). By examining a comparatively small number of sections it is possible to volumetrically, rather than in two dimensional terms.

Vascular disorders (see also Chapter 5.11.1, Chapter 5.11.2, Chapter 5.11.3, Chapter 5.11.4 Chapter 5.11.5, Chapter 5.11.6, Chapter 5.11.7, Chapter 5.11.8 and Chapter 5.12.1)

Temporal artery biopsy

Despite many detailed clinical, pathological, and epidemiological studies, very little is known about the cause of cranial ('giant-cell' or 'temporal') arteritis (see Chapter 5.11.6). Although medium-sized arteries of the head and neck are frequently involved, almost any area of the body, including the aorta, may be affected. It is disease of the posterior ciliary and ophthalmic arteries that leads to blindness but biopsies are usually taken from the superficial temporal branch of the external carotid or the terminal branches of the occipital artery.

Technical aspects

Many patients with the classical, clinical features of temporal arteritis have normal biopsies and this is often ascribed to focal involvement of the artery by disease (Klein *et al*. 1976). To minimize the incidence of these false-negative biopsies, samples removed for histological examination must be as long as possible and should be taken from areas that are tender. Scalp necrosis is a rare complication of cranial arteritis but it is safe to remove a 2- to 3-cm length of the superficial temporal artery. Biopsies must be handled gently and should not be trimmed for processing until they are fully fixed. Histological artefacts can be produced if fresh biopsies are squeezed or cut transversely, and in the most curious of these the artery appears to intussuscept on itself. All of the biopsy should be processed, cut at multiple levels, and stained by haematoxylin and eosin and a connective tissue method, such as elastic–van Gieson.

Evaluation of biopsies

The microscopic features of acute and resolving cranial arteritis and the arterial changes associated with ageing are summarized in Table 3. The classical picture of granulomatous acute and chronic inflammation, with multinucleated giant cells in close relationship to an extensively fragmented internal elastic lamella, is seen in approximately 60 per cent of typical clinical cases (Table 4, Fig. 14 and Fig. 15)). Giant cells are not a prerequisite for diagnosis. However, there must be evidence of inflammation within the muscular wall, for adventitial chronic inflammatory infiltrates are a feature of common disorders such as atherosclerosis. In some biopsies from patients with acute cranial arteritis there is florid intimal oedema (and it may be the rapid resolution of this that underlies a prompt response to anti-inflammatory drugs).

'Ageing changes' (arteriosclerosis)	Active cranial arteritis	Subacute or healed cranial arteritis
Concentric intimal and inner medial fibrosis	Acute and chronic transmural inflammation, often in relationship to small branches	Focal aggregates of lymphocytes and macrophages in the media
Focal fragmentation and apparent reduplication of internal elastic lamella	Prominent giant cells in relation to destructed internal elastic lamella	Bizarre, irregularly arranged intimal fibrous tissue
Focal calcification, especially around elastic lamellae and adjacent media	Marked intimal fibrosis and oedema	Fragmentation of elastic lamella, usually involving more than one-quarter of the arterial circumference
Hyalinization of muscular media	Focal areas of intimal and medial necrosis; not extensive necrotizing change with fibrin deposition and luminal thrombus formation unusual	Medial scarring with ingrowth of new blood vessels (neovascularization)
Aggregates of lymphocytes and macrophages in adventitia, especially in association with atherosclerosis		

Table 3 Histological changes in temporal artery biopsies

	[illegible]	[illegible]
No. of biopsies suitable for study*	130	91
No. of patients (%) with clear clinical evidence of cranial arteritis	57 (43%)	39 (43%)
Histological results (on review)		
Typical cranial arteritis	23 (17%)	21 (23%)
Atypical arteritis†	5 (3%)	2 (2.2%)
Healed arteritis	2 (1.5%)	3 (3.3%)
Arteriosclerosis	91 (67.5%)	55 (60.5%)
Normal	9 (4.5%)	0 (0.0%)
Percentage of patients with clear clinical cranial arteritis who had positive biopsies	58%	49%
Histological errors		
No. of patients (%) with arteriosclerosis erroneously diagnosed as healed or atypical arteritis	8 (5.0%)	5 (5.5%)
No. of cases in which a histological diagnosis of arteritis was missed	0	1

*Cases in which both histology and clinical records were available for study.
†No giant cells but inflammatory lysis of cranial arteritis.
[illegible] (1991) and Ashton-Key and Gallagher (1991)

Table 4 Histological results in temporal artery biopsies, Southampton University Hospitals

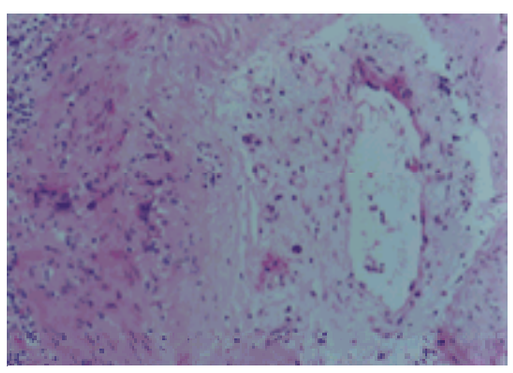


Fig. 14 A positive temporal artery biopsy. Note the florid intimal oedema and mixed chronic inflammatory infiltrate.

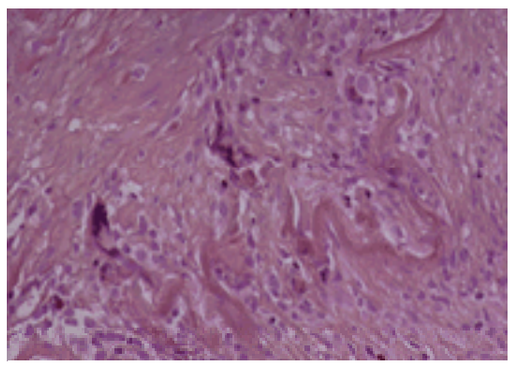


Fig. 15 Cranial arteritis. Note the close relationship between the internal elastic lamella and giant cells. Not all positive biopsies contain giant cells.

Surgical pathologists must be familiar with the full range of ageing changes that occur in the aorta and muscular arteries. These alterations can be misinterpreted and, in reviewing 91 biopsies taken over a 5-year period, we identified five cases reported as healed arteritis that showed only ageing change (Ashton-Key and Gallagher 1991). These alterations, sometimes termed arteriosclerosis, include intimal and inner medial fibrosis, concentric intimal thickening by bands of collagenous tissue resembling the internal elastic lamella, focal fragmentation and calcification of elastic tissue, hyalinization of the medial muscle, and small aggregates of lymphocytes and macrophages in the intima and media. In contrast, in genuine cases of healed arteritis there are extensive breaks in the internal elastic lamella and areas of irregular or coarse scarring involving much of the thickness of the media. On occasions a bizarre pattern of intimal fibrosis is recognized (Fig. 16). If these changes are present, a diagnosis of healed arteritis can be made even in the absence of inflammation.

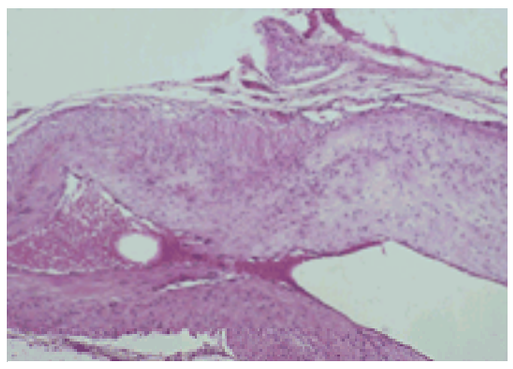


Fig. 16 Healing temporal arteritis. On the right there is irregular inimal thickening and oedema but at the left the structure of the arterial wall is almost normal.

Indications for biopsy

About 20 per cent of temporal artery biopsies show acute, subacute, or healed arteritis but between 10 and 20 per cent of patients with negative biopsies have classical signs and symptoms of cranial arteritis and respond well to steroid treatment (Hall *et al*. 1983; Ashton-Key and Gallagher 1991). Some rheumatologists and ophthalmologists do not routinely take a biopsy from patients who they intend to treat with steroids and the results shown in Table 5 provide some support for this practice (Allison and Gallagher 1984). In one group of patients we studied, 18 per cent of those with a clinical diagnosis of cranial arteritis developed a complication associated with long-term steroid treatment and in cases such as these a positive biopsy report is invaluable.

	No steroids before biopsy (n=41)	Steroids for 7 days or less (n=31)	Steroids for more than 7 days (n=22)
Positive biopsy			
Typical cranial arteritis	42 (59%)	26 (52%)	2 (10%)
Atypical arteritis	8 (14%)	4 (8%)	–
Healed arteritis†	–	–	2 (10%)
Negative biopsy	11 (19%)	21 (40%)	18 (80%)

*These results demonstrate that the chance of a positive biopsy diminishes after steroid therapy is started. Adapted from Allison and Gallagher (1984).
†Giant cells absent but otherwise typical of cranial arteritis.

Table 5 Temporal artery biopsy and the duration of steroid treatment before biopsy in patients with good clinical evidence of cranial arteritis*

Larger arteries are involved in a proportion of cases. Disease of the proximal aorta may be associated with dilatation of the proximal aorta, aortic incompetence, or aortic dissection. In some cases the diagnosis is only made when cardiac surgeons biopsy an unusually thick aortic wall. As in the temporal artery, the histological changes in the aorta may be focal. Any inflammation within the aortic media, as opposed to the adventitia, is pathological and must not be dismissed. The erythrocyte sedimentation rate and acute phase reactants should be measured preoperatively in all patients with proximal aortic disease.

There is no doubt that steroid treatment produces a rapid improvement in clinical symptoms and appears to be associated with a higher incidence of 'false negative' biopsies. It is unlikely that complete resolution of histological changes can occur in steroid-treated patients in days or even weeks. The low incidence of positive biopsies may be the result of sampling error—and when steroid treatment has already been started it is particularly important to take biopsies from areas of tenderness. Not all patients with genuine temporal arteritis have a raised erythrocyte sedimentation rate or plasma viscosity, and normal values (Wong and Korn 1986) are not a contraindication to biopsy.

Granulomatous arterial diseases

Three other chronic inflammatory diseases of large vessels have superficial similarities to cranial arteritis but each is a distinct clinical and pathological entity. Granulomatous angiitis of the central nervous system affects small and medium-sized arteries of the meninges and cerebrum. As in cranial arteritis, giant cells may be prominent but, if anything, the inflammatory infiltrates are denser and there may be extensive necrosis of the muscular wall (Fig. 17). Patients present with a variety of focal or diffuse neurological symptoms but visual symptoms are not especially common and the scalp vessels do not appear to be affected. If cerebral angiography is performed, a characteristic beaded appearance may be seen; on occasions the diagnosis has been confirmed by meningeal biopsy. Some patients have responded to cyclophosphamide and corticosteroids (Frayne *et al*. 1986). The cause of this disorder is unknown but there is an association with a form of amyloid angiopathy distinct from the usual pattern seen in Alzheimer's disease (Shintaku *et al*. 1986).

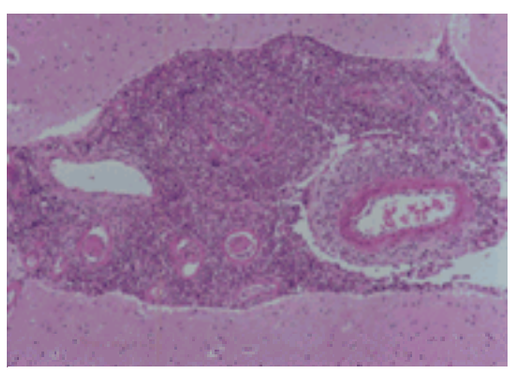


Fig. 17 Granulomatous angiitis of the central nervous system. There is florid transmural chronic inflammation with prominent giant cells. In contrast to cranial arteritis there is extensive necrosis of the vessel wall.

Takayasu's arteritis affects the aorta and its larger branches and, at least in the early stages, it is a granulomatous chronic inflammatory disease (Hall *et al*. 1985). The usual presenting symptoms are upper limb claudication or hypertension but the diagnosis is often difficult and aortography is usually required. Most biopsies are taken from completely obstructed arteries and these may show only intimal fibrosis, medial scarring, and patchy chronic inflammation. In active cases the histological appearances closely resemble those of cranial arteritis but the clinical features are so different that this is seldom a source of confusion. The same may be said of Buerger's disease, a distinctive peripheral vascular disorder, in which there may be granulomatous inflammation of the lower limb arteries. Histological confirmation of the diagnosis is usually made in amputation specimens but cerebral and gastrointestinal arteries can be affected (Rosen *et al*. 1985). The disease has no important rheumatological complications.

Histological diagnosis of systemic vasculitis

The term 'vasculitis' is applied loosely to non-infectious inflammatory disorders of blood vessels. There is some direct and some circumstantial evidence that the inflammation is a result of an immunological reaction at the endothelial surface or in the wall of the affected vessel. In the early stages the clinical diagnosis may be uncertain. Screening for autoantibodies is an essential investigation and circulating autoantibodies to neutrophil cytoplasmic antigens are strongly associated with the development of several forms of systemic vasculitis (Savage *et al*. 1991). A wide variety of tissue biopsies may be submitted in order to establish a histological diagnosis (Table 6).

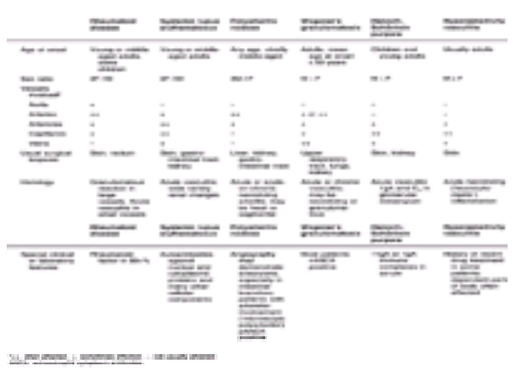


Table 6 Histological and immunopathological features of vasculitis

Evaluation of biopsies

No single microscopic feature is diagnostic of a particular disorder and there is considerable overlap between the appearances in the different forms of vasculitis. A careful assessment of the nature of the inflammatory reaction, the type and size of the diseased vessel, and the site of the tissue biopsy may provide a correct diagnosis. The most characteristic histological features of a true vasculitis are necrosis of the media of the vessel wall, transmural and adventitial inflammation, and a granulomatous reaction (Travis *et al*. 1991) (Fig. 17 and Fig. 18). In contrast, in other chronic inflammatory processes, inflammatory cells are usually restricted to the adventitia. Although the intima and media may undergo reactive thickening ('endarteritis obliterans'), there is no significant necrosis or inflammation of the wall itself. The inflammatory changes are often more marked in veins than in arteries and this would be unusual in a true vasculitis. These are useful guidelines but in practice there are many exceptions. Not all biopsies from undoubted cases of vasculitis will show full-thickness inflammation or necrosis of the vessel wall. On the other hand, arterial haemorrhage is a well-recognized complication of disorders such as peptic ulceration or tuberculosis and this could not occur without extensive medial necrosis.

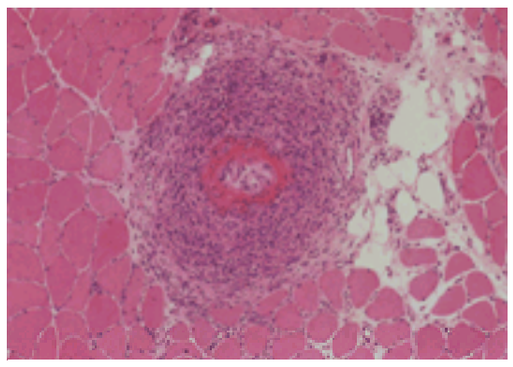


Fig. 18 Polyarteritis nodosa in a muscle biopsy. There is fibrinoid necrosis of the inner media and the inflammatory infiltrate extends through the full thickness of the wall.

Immunohistochemical methods have not found a place in the routine diagnosis of systemic vasculitis. In small vessels, such as arterioles and capillaries, it is comparatively easy to demonstrate pathological deposits of immunoglobulins. This is especially true of skin, where simple immunofluorescent methods with frozen sections are at least as satisfactory as more modern methods in paraffin-embedded material.

There are substantial technical problems in larger vessels, where the reaction product may be obscured by heavy background staining (Gallagher 1991). Although it is assumed that many forms of vasculitis are the result of immune complex deposition, in most instances the nature of the antigenic component is uncertain. Nevertheless, if substantial deposits of immunoglobulin can be demonstrated in diseased vessels the process is more likely to be true vasculitis than a reactive.

Muscle biopsy in the rheumatic disorders

Technique

Muscle biopsy, obtained by either needle or open biopsy, is a minor procedure that can be done under local anaesthesia in almost all patients. An open biopsy requires a small skin incision but has the advantage that a larger specimen is obtained, often giving more diagnostic information in a patchy inflammatory disorder such as polymyositis. In comparison, a needle biopsy provides a smaller sample but several biopsies in different directions can be obtained through the same small incision, and it is much less likely to leave a scar. In addition, the patient is more likely to tolerate repeat biopsies. With either method, careful handling and freezing of the specimen is essential to obtain maximum histological information. It is best to orientate the specimen before freezing and then to snap-freeze in isopentane cooled in liquid nitrogen. This rapid freezing eliminates ice-crystal artefact, which can grossly distort the muscle fibres and obscure pathological changes. A simple technique for freezing biopsy specimens is illustrated (Fig. 19). It is important to keep the specimen completely frozen and to avoid contact with fingers or metal instruments at room temperature, as these may cause thawing at the periphery. For diagnostic purposes, transverse sections generally yield the most information. Fibre atrophy is one of the more common pathological changes in the rheumatic disorders and fibre size can be accurately assessed from fibre diameter, but only in genuine transverse sections (Fig. 20). If there is enough material, it is always sensible to store a small frozen sample for biochemical analysis and to put a tiny strip into glutaraldehyde for electron microscopy, lest either of these investigations prove necessary.

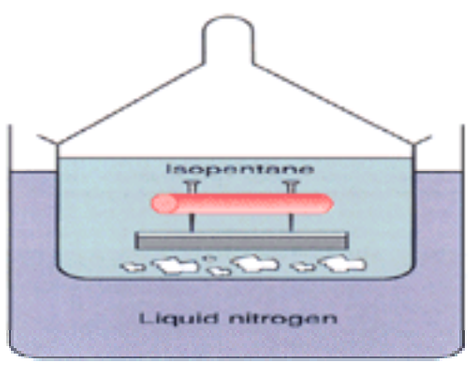


Fig. 19 A small cylinder of muscle under slight tension is pinned to a piece of rubber and prevented from sticking to it by a layer of plastic. When dropped into isopentane, cooled by liquid nitrogen until crystals appear on the bottom of the beaker, the specimen freezes instantly.

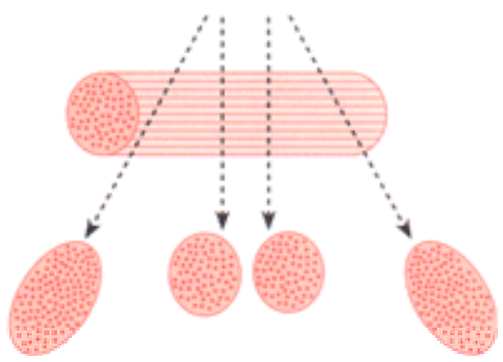


Fig. 20 Fibre size can be reliably assessed from fibre diameter or area, but only in true transverse sections. Poor orientation results in larger elliptical sections and

erroneous measurements of fibre size.

Site

A positive muscle biopsy is an important diagnostic criterion of polymyositis and dermatomyositis (see Chapter 5.9.1), where skeletal muscle is the main and often sole pathological target. In other systemic rheumatic disorders, muscle biopsy may assist in primary diagnosis, but more often is done to elucidate the cause of muscle symptoms, particularly to detect an active myositis. The site of biopsy is best determined by the clinical picture. A severely wasted muscle in a patient with a chronic disorder should be avoided. A tender muscle may yield a positive result in inflammatory disorders; electromyographic abnormalities on one side in symmetrical muscle disease may assist localization of an appropriate muscle to biopsy in the opposite limb.

Muscle fibre types and staining reactions

A wide variety of enzyme histochemical and tinctorial stains is available, but only a few are necessary for initial screening (Fig. 21). Differences in physiological properties, that is differences in twitch speed and fatiguability, correlate with differences in enzyme profile of individual muscle fibres. One fibre category may be selectively involved in disease; therefore identification of fibre types and their variation from the normal pattern are important diagnostic criteria. The myosin ATPase histochemical reaction is the best and most readily reproducible method, by which three types can be distinguished: slow-twitch, fatigue-resistant type-1, and fast-twitch type-2 fibres, which are subdivided into 2A with intermediate fatiguability and 2B that fatigue rapidly (Fig. 22). The muscle fibres within a motor unit are of uniform type but fibres from adjacent units are intermingled, creating a normal mosaic pattern in all normal human limb muscles. A fourth type, 2C fibres, are immature fibres found only in fetal, neonatal, or regenerating muscle. Other stains give additional information about sarcoplasmic contents, the arrangement of myofibrils, and other organelles. Immunocytochemistry may be used to identify and type infiltrating lymphocytes and macrophages in inflammatory disease. Electron microscopy has played an important research role in muscle disorders but has little diagnostic value in the rheumatic disorders.

Myosin ATPase *	Fibre typing
Oxidative enzymes *	Intermyofibrillar network
Gomori trichrome	Mitochondria
Acid phosphatase *	Lysosomes, macrophages
Periodic acid Schiff	Glycogen
Oil Red O	Lipid

* *Histochemical reaction*

Fig. 21 Basic panel of special stains for muscle biopsy interpretation

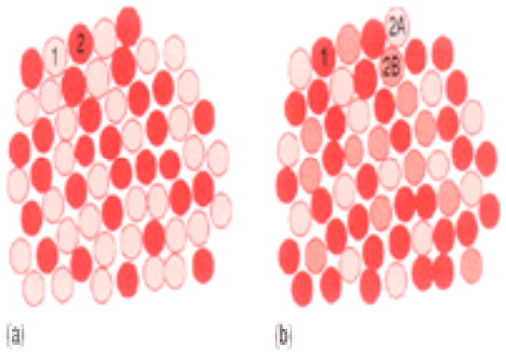


Fig. 22 Fibre types. (a) Myosin ATPase at pH9.4 shows strong reactivity in type 2 fibres (dark), but enzyme activity is inhibited in type 1 fibres (pale). (b) At acid pH 4.6 there is a reversal of this pattern. The type 1 fibres are dark and the type 2 fibre subtypes are revealed, type 2A (pale) and 2B (intermediate). The intermingling of fibres of adjacent motor units creates the normal mosaic pattern.

Clinicopathological correlation

Histological abnormalities in a muscle biopsy are rarely pathognomonic of a single disorder. Rheumatic diseases and conditions with totally different pathogeneses, such as dystrophies or neurogenic atrophy, may not always be distinguished by histological features alone. Close clinicopathological correlation is essential to obtain the maximum diagnostic information from a muscle biopsy. For detailed descriptions of the pathology of neuromuscular disease and biopsy interpretation the reader is referred to larger, comprehensive works (Anderson 1985; Engel and Banker 1986). This section seeks to correlate histopathological abnormalities in muscle with pathogenetic mechanisms in the rheumatic disorders. As with the clinical picture, there is considerable histological overlap in the various rheumatic disorders, and the different diseases are discussed under the headings of key pathological features.

Inflammation

Inflammatory myopathies are disorders in which muscle cell injury is directly or indirectly attributable to the inflammatory reaction and mononuclear chronic inflammatory cells are usually present in the muscle (see also Chapter 5.9.1).

Infectious agents

Infective myositis due to active proliferation of bacterial, fungal, or parasitic organisms in muscle may elicit a vigorous inflammatory response, but it is rare except in the tropics and unlikely to be confused with the immunological disorders. Acute viral myositis is attributable to viral invasion of muscle cells, for example by influenza or enteroviruses, but is a self-limiting disorder. Infectious agents have long been incriminated as trigger factors of the immunologically mediated destruction of muscle in polymyositis and dermatomyositis, but micro-organisms are consistently absent in muscle examined by light and electron microscopy or tissue culture. In addition, a steroid-responsive inflammatory myopathy, indistinguishable from idiopathic polymyositis occurs in patients infected with human immunodeficiency virus (**HIV**) (Calabrese 1989; Wrzolek *et al*. 1990), and retroviral antigens have been identified in sparse macrophages in muscle biopsies (Nordstrom *et al*. 1989; Dalakas 1990).

Mononuclear cells

Whilst there is no absolute distinction, polymyositis and dermatomyositis generally show differences in distribution and character of the inflammatory cells, which probably reflect different immune effector mechanisms (Engel and Arahata 1986). In both disorders the cellular infiltrate is predominantly lymphocytic, but in polymyositis it is chiefly endomysial and focal (Fig. 23(b)). Individual, apparently healthy muscle fibres are surrounded and invaded by lymphocytes, shown by immunocytochemistry to be cytotoxic T cells accompanied by macrophages. This invasion of non-necrotic muscle fibres implicates cell-mediated cytotoxicity in polymyositis. In contrast, in dermatomyositis the infiltrate is dispersed in the perimysial connective tissue surrounding fascicles and frequently forms perivascular aggregates (Fig. 23(a)). Endomysial inflammation is less conspicuous. B cells are more numerous in dermato-than in polymyositis, but the character of the infiltrate

changes from B-cell dominance in the perivascular location to increasing numbers of T cells, mainly CD4, approaching the endomysium. The relative abundance of B cells suggests local humoral immune mechanisms have a more important role in dermatomyositis. Polymyositis and dermatomyositis are the archetypal disorders, but the same patterns of inflammation may occur in any of the autoimmune rheumatic disorders.

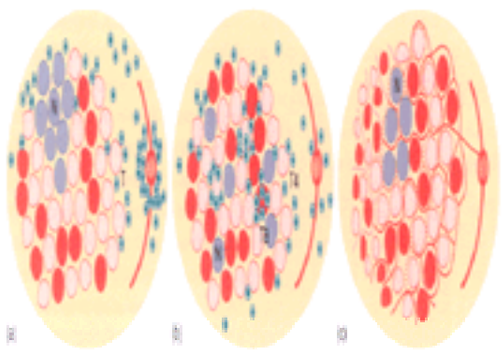


Fig. 23 Immunological injury. (a) Pattern of lymphocytic (T and B cell) infiltration in dermatomytosis. Necrotic fibres (N) in small groups suggesting microinfarcts. (b) Pattern of lymphocytic infiltration in poliomyositis. Invasion of non-necrotic fibres by cytotoxic T cells (T8) and spotty fibre necrosis (N). (c) Membrane attack complex (MAC) deposition of capillaries (dotted line) precedes capillary loss. Reduction of flow in capillary bed causes microinfarcts (N) and atrophy of fibres at the pperiphery of the vascular field.

Invasion by cytotoxic T cells and destruction of non-necrotic muscle cells also occurs in inclusion-body myositis (Engel and Arahata 1986). Histological distinction from polymyositis depends on the presence of numerous vacuoles rimmed by basophilic material, and eosinophilic cytoplasmic inclusions (Lotz *et al*. 1989). The vacuoles are acid phosphatase-positive, lysosomal vacuoles containing cytoplasmic debris. The eosinophilic inclusions are found within or closely associated with vacuoles and electron microscopy reveals they are composed of masses of electron-dense filaments, 14 to 18 mm in diameter, of unknown origin and composition. Despite resemblance to paramyxoviral nucleocapsids, negative findings for viral DNA by *in-situ* hybridization refute the possibility that mumps virus is an aetiological agent (Nishino *et al*. 1989). Furthermore, neither the filaments nor vacuoles are disease specific but occur in a wide variety of chronic wasting neuromuscular disorders. The autophagic lysosomal system plays a central role in muscle cell degradation in all these conditions, but the trigger factors are unknown.

Eosinophils

Skeletal muscle may be involved in the spectrum of hypereosinophilic disorders, in which eosinophil-derived toxins probably contribute to tissue damage. These disorders encompass eosinophilic fasciitis (Shulman's syndrome), where a mixed chronic inflammatory infiltrate containing many eosinophils is present in an oedematous and sclerotic fascia, often spreading deeply into the perimysium (Serratrice *et al*. 1990). In some cases the infiltrate involves the endomysium and is associated with spotty necrosis of muscle fibres, an eosinophilic polymyositis. Indistinguishable eosinophilic inflammation of fascia and muscle may be found in the eosinophilia–myalgia syndrome attributable to tryptophan ingestion (Silver *et al*. 1990).

Granulomatous inflammation

The 'tuberculoid' granuloma, a histological marker of cell-mediated immunity, is a tight focus of epithelioid histocytes and multinucleate Langhans giant cells. Granulomas are rare, inconspicuous components of the predominantly lymphocytic inflammation of muscle fascicles in polymyositis. In contrast, sarcoidosis involving skeletal muscle shows discrete, non-caseating tuberculoid granulomas, chiefly in the interstitium, but without accompanying endomysial inflammation or necrosis of muscle fibres.

Necrosis

All stages of necrosis and regeneration may be seen in a single biopsy in the active phase of the idiopathic inflammatory myopathies. Initially a necrotic segment is pale staining and structureless. Invasion by macrophages that engulf cytoplasmic debris rapidly follows and subsequently small basophilic segments containing large nuclei indicate regeneration from satellite cells. Different patterns of fibre necrosis are observed in polymyositis and dermatomyositis. In polymyositis there is frequently spotty, single-fibre necrosis (see Fig. 23(b)). These necrotic fibres are not surrounded by T lymphocytes but contain the membrane-attack complex, indicating a different method of immunological destruction, probably humoral antibody-dependent, complement-mediated cytotoxicity. In a minority of patients with an idiopathic inflammatory myopathy clinically compatible with polymyositis and with no drug exposure, the muscle biopsy is completely devoid of inflammatory cell infiltration and shows only segmental necrosis. The same picture is seen in a proportion of patients with AIDS-related myopathy (Wrzolek *et al*. 1990) and is recorded in Lyme disease (Schoenen *et al*. 1989).

In dermatomyositis necrotic fibres are frequently in small clumps, with the appearance of microinfarcts (see Fig. 23(a)). Primary immunological attack upon the vasculature may be responsible for these ischaemic foci. Capillary vessels positive for the complement membrane-attack complex and decreased capillary density are consistent findings and present early in the disease (Emslie-Smith and Engel 1990) (see Fig. 23(c)). Microtubular inclusions of uncertain nature revealed by electron microscopy in endothelium in dermatomyositis, and also in systemic lupus erythematosus, may be a cellular response to injury. In addition to microangiopathy, fibrinoid necrosis, attributable to immune complex deposition in vessel walls, is occasionally seen in arterioles and venules in dermatomyositis, particularly the juvenile form.

A necrotizing inflammatory myopathy, histologically indistinguishable from either poly- or dermatomyositis, may occur in association with any of the systemic autoimmune rheumatic diseases, that is in overlap syndromes. In systemic sclerosis, necrosis of muscle fibres is unusual, and inflammatory cell infiltration, predominantly T cell, is largely confined to the perimysium, supporting the likelihood of a cell-mediated immune response directed against connective tissue components rather than muscle cell antigens.

Necrosis of muscle fibres in patients with a rheumatic disorder is occasionally attributable to drug therapy. Many different drugs, with different modes of action, can cause a necrotizing myopathy but D-penicillamine, which induces a wide range of autoimmune phenomena, is probably the most frequently encountered in rheumatological practice. A florid, necrotizing inflammatory myopathy is histologically identical to idiopathic poly- or dermatomyositis. Zidovudine toxicity may contribute to muscle cell necrosis in HIV-positive patients (Dalakas *et al*. 1990).

Atrophy

Muscle fibre atrophy is invariably present in patients with muscle weakness. Different causes and pathogenetic mechanisms give different patterns of atrophy (Fig. 22). Atrophy may be selective, affecting only one fibre type, usually type-2, or involve groups of fibres, or be confined to the periphery of fascicles (Fig. 24). Regenerating fibres are also smaller than normal, but can be recognized by their cytoplasmic basophilia and type-2C ATPase reaction. Selective atrophy of the fast-twitch type-2, particularly type-2B, fibres is a frequent but by no means specific finding in patients with rheumatoid disease and muscle symptoms. Whilst often seen in association with myositis, type-2 atrophy may be the only abnormality. The cause is often multifactorial and selective type-2 atrophy alone is insufficient for diagnosis of active inflammatory myopathy. Disuse atrophy, and hence immobility, associated with joint pain, malignancy, cachexia, and steroid therapy may all induce type-2 atrophy. Muscle biopsy is often used in an attempt to distinguish between steroid-induced myopathy and exacerbation of myositis. If the only abnormality is type-2 atrophy, the many possible contributory factors and the patchy nature of the inflammatory process make distinction impossible.

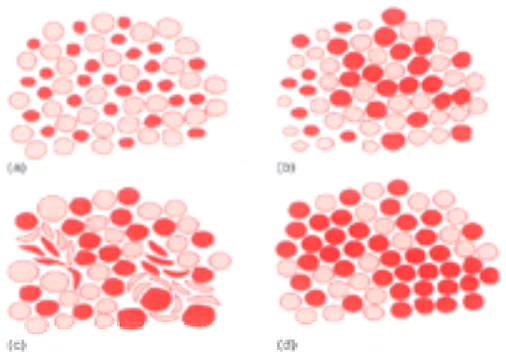


Fig. 24 Patterns of atrophy. For simplicity only two fibre types are illustrated. (a) Selective type-2 fibre atrophy, a non-specific change encountered in many disorders. (b) Perifasicular atrophy involving both fibre types, as in dermatomyositis. (c) Neurogenic atrophy, small group atrophy affecting individual motor units of both fibre types. Some residual innervated fibres show compensatory hypertrophy. (d) Reinnervation reverses atrophy and may result in type grouping, i.e. large groups of uniform fibre type.

Perifascicular atrophy, involving type-1 and type-2 fibres, is particularly characteristic of dermatomyositis, both juvenile and adult (Fig. 24(b)). It is attributed to ischaemia, secondary to the microangiopathy of this myositis. The distribution coincides with the periphery of the vascular field, where effects of reduced blood supply are most severe.

Grouped atrophy, affecting type-1 and type-2 fibres in any proportion, is characteristic of denervation. The muscle fibres of each motor unit are in fairly close proximity within the belly of the muscle. Denervated fibres slowly shrink and the muscle fibres within a denervated motor unit shrink simultaneously, creating the grouped pattern (Fig. 24(c)). Denervated fibres assume an angular outline, as though moulded by adjacent healthy contracting fibres. Grouped atrophy is most evident in the motor-neurone diseases, but denervation may be a component of a rheumatoid disorder, usually as a consequence of ischaemic motor neuropathy secondary to vasculitis. It is also another frequent muscle abnormality in AIDS patients, secondary to HIV infection of peripheral nerves (Gabbai *et al*. 1990).

Denervation atrophy may be accompanied by a change in the fibre type distribution due to reinnervation. Distribution of fibre type is not an intrinsic property of muscle cells, but is determined by the motor nerve. Reinnervation of denervated fibres may induce a change in fibre type and create enlarged motor units, which appear as clumps of fibres of uniform type (Fig. 24(d)).

Hypertrophy

Hypertrophy of muscle fibres is a normal physiological response to physical training and the fibres of healthy active young adults are generally slightly larger than those of inactive elderly persons. Hypertrophy also occurs in pathological conditions, when it is most often a compensatory change in residual, healthy, over-stressed fibres alongside atrophic and damaged fibres. Hypertrophy is not seen in early or acute phases of inflammatory myopathies, nor in muscle of patients disabled and immobilized by arthritis, but it may be found in chronic, insidious or burnt-out disease, when patients remain active. Limb-girdle forms of muscular dystrophy and chronic spinal muscular atrophy may enter the histological differential diagnosis of chronic polymyositis.

Disturbances of internal cell architecture

A variety of alterations in the normal, regular arrangement of myofibrils, the peripheral situation of nuclei, and the normal distribution of other cell organelles indicate sublethal cell injury. The oxidative enzyme reaction and Gomori's trichrome stain are particularly valuable for detecting these changes. These minor abnormalities are not disease specific and on their own serve only to confirm an occult organic basis for the patient's muscle symptoms.

Myofibrillar disarray

Myofibrillar disarray, frequently seen in the inflammatory myopathies, is responsible for a patchy staining reaction, a so-called moth-eaten pattern with the oxidative enzyme reaction. In polymyalgia rheumatica, despite the severity of myalgia, a few moth-eaten fibres and a very mild degree of type-2B atrophy is generally all that is seen on biopsy. Inflammation is conspicuously lacking and there is no vasculitis in muscle.

Vacuoles

Empty cytoplasmic vacuoles, derived from dilated organelles such as sarcoplasmic reticulum or the T-tubule system, are an occasional finding and rarely numerous in inflammatory myopathies. The numerous lysosomal or rimmed vacuoles of inclusion-body myositis have already been described. A few may be found in steroid-sensitive poly- or dermatomyositis, together with the eosinophilic inclusions; it is the very large number that separates inclusion-body myositis. Identical vacuoles are also characteristic of chloroquine-induced myopathy, where in addition, electron microscopy may reveal short, curved, lamellar cytoplasmic bodies, identical to the inclusions in Batten's disease.

Mitochondria

The wide clinical spectrum of myopathies due to mitochondrial enzyme deficiencies encompasses patients with symptoms indistinguishable from polymyositis. A diagnostic biopsy in mitochondrial myopathy reveals numerous 'ragged red' fibres with Gomori's trichrome stain, an appearance created by aggregates of large, structurally abnormal mitochondria. There is no inflammation. In contrast, secondary mitochondrial injury may be responsible for an occasional 'ragged red' fibre amongst the more characteristic changes of the idiopathic inflammatory myopathies, and may be a manifestation of zidovudine myotoxicity in HIV-positive patients.

Fibrosis

Increase in interstitial connective tissue—perimysial and endomysial fibrosis—only occurs in chronic rheumatoid disorders, preceded by considerable atrophy and destruction of myofibres.

Vasculitis in muscle

Vasculitis—inflammation directed against blood vessels of almost any calibre in many different tissues—is a pathological component of most rheumatic disorders and may contribute to fibre atrophy and necrosis in skeletal muscle. Acute necrotizing arteriolitis and arteritis are not uncommon in juvenile dermatomyositis and may on occasion be seen in skeletal muscle in systemic lupus and rheumatoid arthritis. However, in polyarteritis nodosa, fibrinoid necrosis of medium-sized arteries is the essential diagnostic feature; muscle fascicles in the territory of an affected vessel may show ischaemic atrophy, or a pattern of denervation atrophy because the motor nerve is ischaemic, but there is no endomysial inflammation. Muscle infarcts due to arteritis in the absence of associated microangiopathy are very unusual, probably because there is a good collateral supply and the full capacity of the capillary network is only required during strenuous exercise.

The histological aspects of vasculitis are sometimes given undue diagnostic weight. Large-vessel vasculitides, such as polyarteritis nodosa, Wegener's granulomatosis, and rheumatoid arthritis, may all involve arteries in skeletal muscle and cannot be distinguished on this basis. The microscopic appearances are largely determined by the structure and calibre of the vessel, and the time-course of the reaction. Thus, in the acute phase of inflammatory destruction of an artery, which has an intima, internal elastic lamella, and smooth muscle coat, there is a fibrin thrombus, fibrin seepage into the wall, and inflammation is predominantly neutrophilic. At a later stage there is intimal hyperplasia and the inflammation obliterating the wall becomes more granulomatous, with a predominance of lymphocytes, histocytes, and occasional giant cells.

Peripheral nerve biopsy in the rheumatic disorders

Definition of the peripheral nervous system

The peripheral nerves include all the neural structures external to the pial sheath of the central nervous system, that is the axons and sheaths of all cranial nerves except optic and olfactory, all spinal and autonomic nerves, and the dorsal root and autonomic ganglia.

Factors that influence the decision to biopsy

Peripheral nerve biopsy is not an investigation that should be undertaken lightly. Not only is it an invasive procedure, with inevitable scarring and not insignificant postoperative sensory disturbance, but the diagnostic yield is negligible unless there are specialist laboratory facilities for processing the nerve and a pathologist experienced in interpretation. Furthermore, only a tiny, possibly unrepresentative, section of an extensive network can be examined. There are no histological criteria that differentiate motor and sensory fibres, and patterns of degeneration are limited and stereotyped. Only rare disorders, such as leucodystrophies, exhibit disease-specific fine-structural changes. In most peripheral neuropathies, molecular pathogenetic mechanisms are unknown. Morphological clues may incriminate the axon or myelin sheath as the primary target but, inevitably, because of structural and metabolic interdependence, secondary changes will occur. In a single biopsy, the sequence of events is not always easy to reconstruct and diagnostic interpretation must be amplified by close correlation with clinical and electromyographic data.

Indications for biopsy

Biopsy is unnecessary in a large proportion of patients with a peripheral neuropathy because they have a known cause, such as diabetes, or a family history and recognized pattern of hereditary disease. Diseases that selectively involve motor neurones, namely motorneurone disease and hereditary motor neuropathies, manifest as muscle weakness. Muscle biopsy and electromyography, rather than nerve biopsy, are more appropriate methods of investigation. Nerve biopsy is generally only applicable to peripheral neuropathies with a definite sensory component. In the rheumatoid and autoimmune rheumatic diseases, a diagnostic quest for vasculitis, which is not only the most common cause of peripheral neuropathy in these disorders but also potentially treatable, is the foremost indication for nerve biopsy. Entrapment neuropathies and neuropathy due to drug toxicity occur but are usually recognized without biopsy.

Site

Certain factors dictate the site of nerve biopsy.

1. It must be easily accessible through a small skin incision made under local anaesthetic, therefore superficial, but not subject to repetitive trauma in everyday life, as this alone will injure the nerve.
2. Biopsy must not give rise to unacceptable deficit; therefore cutaneous sensory nerve is preferable to a motor or mixed sensory–motor nerve.
3. The nerve selected must be involved and in most polyneuropathies, the lower legs are earliest and most severely affected.

Hence, the sural nerve, which is purely sensory and easily identified subcutaneously at the ankle, is most often appropriate, but the superficial peroneal and radial nerves are useful alternatives. After exposure the proximal end of the nerve should be incised first. The distal end can then be handled without the patient experiencing abnormal sensations. To minimize discomfort during and deficit after surgery, only partial transection of the nerve and excision of several fascicles is necessary (Fig. 25) but, ideally, these should be at least 3 cm long. The success of biopsy relies heavily on careful surgical technique. Artefactual disruption of delicate axons and myelin sheaths is easily produced by stretching or pinching with forceps. The specimen is best transferred immediately, fresh and unfixed, to the laboratory where a skilled technician can subdivide it for the most appropriate methods of study—generally resin-embedded transverse sections, teased fibre preparations, and electron microscopy. Quantitation of internodal length and nerve fibre diameter and density has an important role in the interpretation of the nerve biopsy (Fig. 26 and Fig. 27). Normal values are essential for comparison, and these are best documented for the sural nerve, although some data for other nerves are available (Dyck *et al*. 1984; Vital *et al*. 1990).

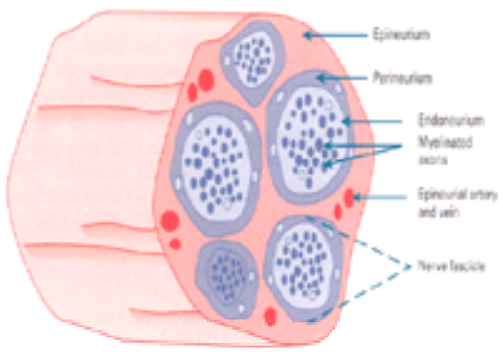


Fig. 25 Diagram of transected nerve to show the arrangement of fascicles bounded by connective tissue sheaths. The sural nerve at the ankle contains five or six fascicles and thus partial transection may be adequate. Full thickness is preferable for detection because arteries, most often invlved, lie within the epineurium.

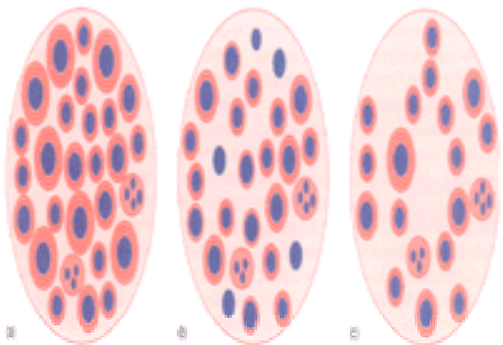


Fig. 26 Transverse sections are required for measurement of nerve fibre density and fibre diameters. (a) The normal nerve has three major populations: (i) large myelinated fibres; (ii) small myelinated fibres; (iii) unmyelinated axons—the smallest fibres. (b) Demyelination and remyelination. No axonal loss. Large fibres with abnormally thin myelin sheaths indicate remyelination. (c) Axonal loss. Decreased fibre density, chiefly due to loss of large diameter fibres.

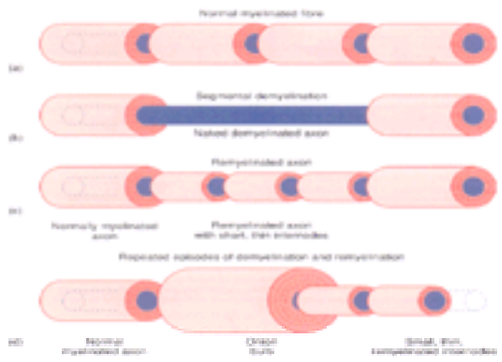


Fig. 27 Teased fibre preparations permit examination of the thickness of several consecutive internodes and are particularly valuable for recognition of demyelination and remyelination. (a) Normal myelinated fibre. (b) Segmental demyelination. (c) Remyelinated axon. (d) Repeated episodes of demyelination and remyelination.

Interpretation of the biopsy

Pathological changes are related to the neuronal components principally involved, with an attempt to explain resulting clinical and electrophysiological abnormalities.

Wallerian degeneration

Wallerian degeneration refers to the chain of events initiated by transection of the axon. Whilst trauma is an obvious cause, the pathology is identical when a segment of nerve is destroyed by ischaemia and infarction, and thus it occurs in vasculitic neuropathies.

The changes are greatest in the distal axon, where total dissolution of the axon and fragmentation of the myelin sheath occur, with formation of myelin ovoids (Fig. 28(a)). The debris is phagocytosed by Schwann cells and invading macrophages, and Schwann cells in the vicinity proliferate. The proximal axonal segment swells, owing to accumulation of cytoplasmic organelles; the myelin sheath is retracted from internodes. The extent of proximal change depends on proximity to the cell body. If injury is very close to the cell body, the neurone may die, but if at a distance, recovery and regrowth of the distal axon may occur by sprouting from the proximal segment. The recently formed Schwann cells remyelinate the new axon (Fig. 28(b)). Although the axon may grow to achieve its original diameter, the internodes of a remyelinated axon remain smaller and thinner than the original and can thereby be recognized in teased fibre preparations (Fig. 27(a-c)). The size of each internode appears to be determined by the area of axon in contact with the Schwann cell (Thomas and Ochoa 1984). More Schwann cells make contact with new axon than with the original, therefore the territory of each is smaller and the internodes are smaller.

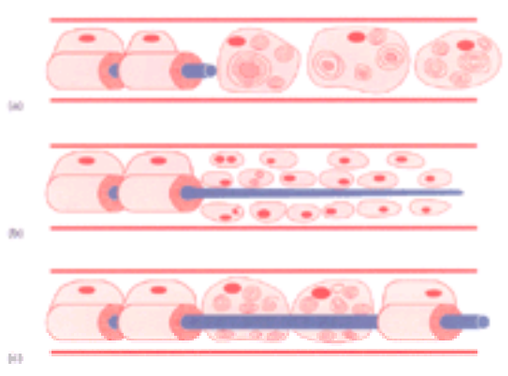


Fig. 28 The role of Schwann cells. (a) Wallerian degredation—breakdown of distal axon and myelin sheath—formation of myelin ovoids within original Schwann cell basement membrane. (b) Axonal regeneration—axonal sprouting from proximal stump accompanied by proliferation of Schwann cells that will remyelinate new axon. (c) Segmental demyelination—myelin debris engulfed by Schwann cells. Schwann cell proliferation and remyelination follows as above.

The neurone

The cell body is the metabolic power house of the cell. Normal structure and function of the axon are maintained by continuous supply of factors transported centrifugally; if the mechanism is impaired the distal ends of the longest axons are the most vulnerable. Consequently, morphological changes of neuronal injury frequently appear first in the distal axon. Distal glove-and-stocking sensory impairment, together with muscle weakness of extremities, is characteristic of many polyneuropathies.

The severity of pathological insult probably determines the time course of events. In acute axonal degeneration attributed to metabolic failure, of which uraemic neuropathy is an example, changes in the distal axon are indistinguishable from Wallerian degeneration. In more chronic insidious forms of degeneration, such as alcoholic–nutritional neuropathy, the distal axon shrinks, but axonal atrophy is accompanied by wrinkling of the myelin sheath and paranodal retraction. This is followed by demyelination, that is fragmentation of the myelin sheath and remyelination to form a thinner sheath around the atrophied axon. Complete axonal degeneration may eventually occur. Fine structural changes often provide evidence of decreased axonal transport in the form of axonal spheroids—an accumulation of neurofilaments and other organelles which create a focal axonal swelling.

Myelin sheath and Schwann cells

Demyelination is destruction of the myelin sheath with preservation of the axon. This may be a primary event or secondary to axonal atrophy, as described above. Provided the axon survives, demyelination is followed by remyelination and the formation of smaller, thinner myelin internodes (Fig. 27(c)). Chronic disorders, exemplified by the hereditary motor and sensory neuropathies, are characterized by repeated episodes of segmental demyelination and remyelination. The repetitive focal proliferation of Schwann cells and formation of new basement membranes creates localized, circumferential thickening of nerves, histologically likened to an onion bulb (Fig. 27(d) and Fig. 29). These thickened nerves may become clinically palpable, hence the term 'hypertrophic neuropathy'.

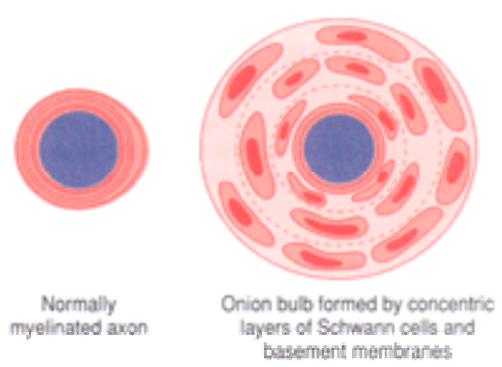


Fig. 29 Onion bulb—the result of repetitive remyelination as seen in transverse section.

The principal pathogenetic mechanisms of primary demyelination are inflammatory, immunological, and toxic. The Guillain–Barré syndrome is the prototype of acute inflammatory demyelinating polyneuropathy (**AIDP**). Small numbers of sensitized, cytotoxic T lymphocytes infiltrating the myelin sheaths effect patchy segmental myelin breakdown, particularly involving nerve roots; this corresponds clinically with asymmetrical, sensorimotor radiculopathy. With severe, widespread acute demyelination, the physical signs can be almost symmetrical. However, in this apparently random attack on peripheral nerves, a clear anatomopathological explanation for the frequent predominance of motor deficits is lacking. Myelin sheaths fragment and are degraded within macrophages (Fig. 28(c)), but although some are shrunken, the axons are largely preserved. Conduction velocity is proportional to the thickness of the myelin sheath, and thus extensive demyelination results in slowing of conduction velocity or complete conduction block, but electromyographic changes may be absent at early stages, before fragmentation has begun. AIDP is seemingly triggered by infection and is a form of neuropathy associated with HIV infection and may coincide with seroconversion (Chaunu *et al*. 1989). The virus has occasionally been cultured from peripheral nerve but is not detectable by any microscopical technique, suggesting that very few virions are present (Monte *et al*. 1988).

Chronic inflammatory demyelinating polyneuropathy (**CIDP**) is due to repeated episodes of demyelination that give relapsing or progressive neurological deficit. Histologically, mononuclear cell infiltration of the nerve is accompanied by evidence of remyelination and often onion-bulb formation. CIDP also occurs in AIDS (Leger *et al*. 1989), where it is associated with a greater degree of axonal loss than in the idiopathic cases.

In peripheral neuropathy associated with paraproteinaemias, segmental demyelination occurs in the presence of circulating antibody, most often of IgM class, specifically directed against myelin-associated glycoprotein (Gosselin *et al*. 1991). Antibody deposition in the myelin sheath is demonstrable by immunoelectron microscopy (Mata *et al*. 1988) and myelin lamellas may be widely spaced (Takatsu *et al*. 1985). Nevertheless, there is poor correlation between level of antibody and severity of neuropathy, and a primary tissue damaging role for the antibody has yet to be established (Gosselin *et al*. 1991).

Chloroquine-induced neuropathy may be encountered in patients with autoimmune rheumatic disease. Biopsy shows segmental demyelination and remyelination, and distinctive lamellar phospholipid lysosomal inclusions can be found in Schwann cells by electron microscopy (Tegner *et al*. 1988). The lipid-combining property of the drug may impede lipid degradation and promote its accumulation.

Interstitium

Ischaemia due to vascular disease is responsible for peripheral neuropathy in a variety of disorders. The clinical picture is governed by the territory and calibre of vessels involved, irrespective of the nature of the vascular pathology. Large-vessel, arterial occlusion invariably causes focal infarction and distal Wallerian degeneration, clinically evident as abrupt onset of mononeuritis and mononeuritis multiplex (Said *et al*. 1988). Arteriolar occlusion, frequently due to vasculitis, may have similar clinicopathological manifestations, but a high proportion of patients present with symmetrical distal polyneuropathy (Harati and Niakan 1986; Kissel *et al*. 1989). Most probably, this represents summation of the effects of small, multifocal, ischaemic lesions. Very mild ischaemia may only damage the myelin sheath of the distal axon. Electrophysiological studies reflect this range of changes. Segmental infarction can cause conduction block (Ropert and Metral 1990); on the other hand, provided some large myelinated fibres are spared, conduction velocity may be maintained. Inevitably, severe axonal loss will lead to conduction slowing and the compound action potential, a summation of all the impulses transmitted, is diminished.

Vascular disease (artery and arteriole)

Mononeuritis occurring in any of the autoimmune rheumatic disorders, not only polyarteritis nodosa, but also, for example, in rheumatoid disease (Said *et al*. 1989), temporal arteritis (Caselli *et al*. 1988), Sjögren's syndrome (Mellgren *et al*. 1989), systemic lupus (Omdal *et al*. 1991), suggests large-vessel vasculitis. Necrotizing arteriolitis is partly responsible for cryoglobulinaemic neuropathy (Garcia-Bragado *et al*. 1988) and the multifactorial pathology of neuropathies associated with HIV infection (Chaunu *et al*. 1989). It is one cause of carcinomatous neuropathy and a probable cause of peripheral neuropathy in Lyme disease (Meier *et al*. 1989). Nerve biopsy is not a very fruitful diagnostic test in sarcoidosis, but vasculitis and epineurial epithelioid granulomata may be found in symptomatic patients.

Large-vessel vasculitic lesions are sought in the epineurium (Fig. 25) and therefore a full-thickness rather than fascicular nerve biopsy will give a greater yield. Although florid fibrinoid necrosis may not be detected, perivascular lymphocytes and plasma cells are positive clues to adjacent foci; haemosiderin deposition, perivascular fibrosis, and recanalization of thrombus are all gravestones of previous episodes. In the absence of any vascular pathology, focal axonal degeneration, with asymmetry between adjacent fascicles, strongly suggests proximal ischaemia. In all these disorders, neuronal vasculitic injury is attributed to an immune mechanism. Immune complex deposition has been detected, but a T-cell mediated attack is also invoked (Kissel *et al*. 1989). In contrast, mononeuropathy in diabetes mellitus is due to the combination of arterial atherosclerosis and diabetic microangiopathy.

Microvascular disease (capillary and venule)

The only vessels to penetrate the endoneurium are capillaries and venules. Leucocytoclastic vasculitis frequently occurs in skin and other organs in many autoimmune rheumatic diseases, but is rare within the endoneurium (Kissel *et al*. 1989). Perivascular lymphocytes may be associated with adjacent florid epineurial vascular disease but not endoneurial vasculitis. Occlusion of the capillary venules results from other pathologies. Small-vessel occlusion in cryoglobulinaemia may be due to sludging of the paraprotein and endothelial cell swelling (Vital *et al*. 1988). In diabetes mellitus, diversion of glucose into non-insulin dependent metabolic pathways leads to capillary basement thickening. Whilst not the sole cause, this microangiopathy undoubtedly contributes to the axonal atrophy that underlies the distal sensory and autonomic neuropathy of diabetes.

Amyloid (see also Chapter 5.13.1)

Intraneural amyloid deposition is a cause of rare hereditary neuropathies and may occur in myeloma. Large endoneurial deposits undoubtedly have a direct pressure effect on fragile myelin sheaths, but ischaemia resulting from deposition in walls of small vessels is a contributory mechanism. Despite the high incidence of secondary systemic amyloidosis in chronic rheumatoid disease, it is not a cause of peripheral neuropathy. This serum AA protein-derived amyloid is rarely, if ever, deposited in peripheral nerves.

Eosinophils

The endoneurial capillaries maintain the blood–nerve barrier. In inflammatory demyelinating and vasculitic neuropathies, endoneurial oedema is a subsidiary event reflecting increased permeability. Breakdown of the barrier is probably a key event in the neuropathy of the hypereosinophilic syndrome. Nerve biopsy reveals endoneurial oedema and distal Wallerian-type degeneration, but no eosinophil infiltration. Eosinophils do not penetrate the endoneurium, but substances released by their degranulation both impair the integrity of the endoneurial barrier and exert a neurotoxic effect (Sunohara *et al*. 1989).

In conclusion, the role of nerve biopsy is limited. It should not be performed without clear objectives and, whilst obvious, it cannot be overstated that prior liaison between physician, surgeon, and pathologist is a fundamental prerequisite of success.

Chapter References

Allison, M.C. and Gallagher, P.J. (1984). Temporal artery biopsy and steroid treatment. *Annals of the Rheumatic Diseases*, **43**, 416–17.

Allsop, C.J. and Gallagher, P.J. (1981). Temporal artery biopsy in giant cell arteritis: a reappraisal. *American Journal of Surgical Pathology*, **5**, 317–23.

Anderson, J.R. (1985). *Atlas of skeletal muscle pathology*. MTP Press, Lancaster.

Artacho-Perula, E., Roldan-Villalobos, R., Collantes-Esteves, E., and Lopez-Beltran, A. (1994). Stereological analysis of the synovial membrane in rheumatic disorders: diagnostic value of volume weighted mean nuclear volume estimation. *Histopathology*, **25**, 357–63.

Ashton-Key, M.R. and Gallagher, P.J. (1991). The pathology of temporal arteritis and polymyalgia rheumatica. *Clinical Rheumatology*, **5**, 387–404.

Ashton-Key, M.R. and Gallagher, P.J. (1992). The value of simple morphometric techniques in the diagnosis of osteoporosis. *Pathology Research and Practice*, **188**, 616–19.

Blake, D.R., Gallagher, P.J., Potter, A.R., Bell, M.J., and Bacon, P.A. (1984). The effect of synovial iron on the progression of rheumatoid disease. *Arthritis and Rheumatism*, **27**, 495–501.

Calabrese, L.H. (1989). The rheumatic manifestations of infection with the human immunodeficiency virus. *Seminars in Arthritis and Rheumatism*, **18**, 225–39.

Caselli, R.J., Daube, J.R., Hunder, G.G., and Whisnant, J.P. (1988). Peripheral neuropathic syndromes in giant cell (temporal) arteritis. *Neurology*, **38**, 685–9.

Chaunu, M.P. *et al*. (1989). The spectrum of changes on 20 nerve biopsies in patients with HIV infection. *Muscle and Nerve*, **12**, 452–9.

Chavassieux, P.M., Arlot, M.E., and Meunier, P.J. (1985). Intersample variation in bone histomorphometry, comparison between parameter values measured on two contiguous transiliac bone biopsies. *Calcified Tissue International*, **37**, 345–50.

Cruz-Orive, L.M. and Weibel, E.R. (1990). Recent stereological methods for cell biology: a brief survey. *American Journal of Physiology*, **258**, L148–56.

Dalakas, M. (1990). Immunopathologic features of HIV and AZT-associated myopathies. *Journal of the Neurological Sciences*, **98**, (Suppl.), 109–10.

Dalakas, M.C., Isabel, I., Pezeshkpour, G.H., Laukaitis, J.P., Cohen B., and Griffin, J.L. (1990). Mitochondrial myopathy caused by long term zidovudine therapy. *New England Journal of Medicine*, **322**, 1098–105.

Daniels, T.E. (1984). Labial salivary gland biopsy in Sjögren's syndrome. Assessment as a diagnostic criterion in 362 suspected cases. *Arthritis and Rheumatism*, **27**, 147–56.

Dempster, D.W. and Lindsay, R. (1993). Pathogenesis of osteoporosis. *Lancet*, **341**, 797–801.

Dyck, P.J., Karne, J., Lais, A., and Clarke Stevens, J. (1984). Pathologic alterations of the peripheral nervous system of humans. In *Peripheral neuropathy*, Vol. 1, (2nd edn) (ed. P.J. Dyck, P.K. Thomas, E.H. Lambert, and R. Bunge), pp. 760–870. W. B. Saunders, Philadelphia.

Edwards, J.C.W. and Cambridge, G. (1995). Is rheumatoid disease a failure of B cell death in the synovium. *Annals of the Rheumatic Diseases*, **54**, 696–700.

Emslie-Smith, A.M. and Engel, A.G. (1990). Microvascular changes in early and advanced dermatomyositis. *Annals of Neurology*, **27**, 343–56.

Engel, A.G. and Arahata, K. (1986). Mononuclear cells in myopathies. *Human Pathology*, **17**, 704–21.

Engel, A.G. and Banker, B.Q. (1986). *Myology*. McGraw-Hill, New York.

Eriksen, E.F., Axelrod D.W., and Melsen F. (1994). *Bone histomorphometry*. Raven Press, New York.

Fallon, M.D. and Teitelbaum, S.L. (1982). The interpretation of fluorescent tetracycline markers in the diagnosis of metabolic disease. *Human Pathology*, **13**, 416–17.

Fishleder, A. *et al*. (1987). Uniform detection of immunoglobulin gene rearrangement in benign lymphoepithelial lesions. *New England Journal of Medicine*, **316**, 1118–21.

Frayne, J.H., Gieligan, B.S., and Essex, W.B. (1986). Granulomatous angiitis of the central nervous system. *Medical Journal of Australia*, **145**, 410–12.

Gabbai, A.A., Schmidt, B., Castelo, A., Oliveira, A.S.B., and Lima, J.G.C. (1990). Muscle biopsy in AIDS and ARC: analysis of 50 patients. *Muscle and Nerve*, **13**, 541–4.

Gallagher, P.J. (1991). Blood vessels. In *Histology for pathologists* (ed. S.S. Sternberg), pp. 195–213. Raven Press, New York.

Gallagher, P.J., Blake, D.R., and Lever, J.V. (1985). Audit of closed synovial biopsy in the diagnosis of inflammatory joint disease. *Scandinavian Journal of Rheumatology*, **14**, 307–14.

Garcia-Bragado, F., Fernandez, J.M., Navarro, C., Villar, M., and Bonaventura, I. (1988). Peripheral neuropathy in essential mixed cryoglobulinemia. *Archives of Neurology*, **45**, 1210–14.

Goldenberg, D.L., Egan, M.S., and Cohen A.S. (1982). Inflammatory synovitis in inflammatory joint disease. *Journal of Rheumatology*, **9**, 204–9.

Gosselin, S., Kyle, R.A., and Dyck, P.J. (1991). Neuropathy associated with monoclonal gammopathies of undetermined significance. *Annals of Neurology*, **30**, 54–61.

Greenspan, J.S., Daniels, T.E., Tallal, N., and Sylevester, R.A. (1974). The histopathology of Sjögren's syndrome in labial salivary gland biopsies. *Oral Surgery, Oral Medicine, Oral Pathology*, **37**, 227–9.

Gundersen, J.J.G. *et al*. (1988). Some new, simple and efficient stereological methods and their use in pathological research and diagnosis. *Acta Pathologica, Microbiologica, et Immunologica Scandinavica*, **96**, 379–94.

Hall, S., Lie, J.T., Kurland, L.T., Persellin, S., O'Brien, P.C., and Hunder, G.G. (1983). The therapeutic impact of temporal artery biopsy. *Lancet*, **ii**, 1217–20.

Hall S., Barr, W., Lie, J.T., Stanson, A.W., Kazmer, F.J., and Hunder, G.G. (1985). Takayasu's arteritis. A study of 32 North American patients. *Medicine*, **64**, 89–99.

Harati, Y. and Niakan, E. (1986). The clinical spectrum of inflammatory-angiopathic neuropathy. *Journal of Neurology, Neurosurgery and Psychiatry*, **49**, 1313–16.

Harris, N.L., Jaffe, E.S., Stein, H., *et al*. (1994). A revised European–American classification of lymphoid neoplasms: a proposal from the international lymphoma study group. *Blood*, **84**, 1361–92.

Harvey, J., Catoggio, L., Gallagher, P.J., and Maddison, P.J. (1989). Salivary gland biopsy in sarcoidosis. *Sarcoidosis*, **6**, 47–50.

Hutton, C.W., Hinton, C., and Dieppe, P.A. (1987). Intra-articular variations of synovial changes in knee arthritis: biopsy study comparing changes in patellofemoral synovium and the medial tibio-femoral synovium. *British Journal of Rheumatology*, **26**, 5–8.

Hyjek, E., Smith, W.J., and Isaacson, P.G. (1988). Primary B-cell lymphoma of salivary glands and its relationship to myoepithelial sialadenitis. *Human Pathology*, **19**, 766–76.

Isaacson, P.G. and Norton, A.J. (1994). In *Extranodal lymphomas*, pp. 69–78. Churchill Livingstone, London.

Jordan, R.C.K., Pringle, J.H., and Speight, P.M. (1995). High frequency of light chain restriction in labial gland biopsies of Sjögren's syndrome detected by *in situ* hydridization. *Journal of Pathology*, **177**, 35–40.

Kennedy, T.D., Plater-Zyberk, C., Partridge, T.A., Woodrow, D.F., and Maini, R.N. (1988 *a*). Representative sample of rheumatoid synovium: a morphometric study. *Journal of Clinical Pathology*, **41**, 841–6.

Kennedy, T.D., Plater-Zyberk, C., Partridge, T.A., Woodrow, D.F., and Maini, R.N. (1988 *b*). Morphometric comparison of synovium from patients with osteoarthritis and rheumatoid arthritis. *Journal of Clinical Pathology*, **41**, 847–52.

Kidd, B.L., Moore, K., Walters, M.T., Smith, J.L., and Crawley, M.I.D. (1989). Immunohistochemical features of synovitis in ankylosing spondylitis: a comparison with rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **48**, 92–8.

Kissel, J.R., Riethman, J.L., Omerza, J., Rammohan, K.W., and Mendell, J.R. (1989). Peripheral nerve vasculitis: immune characterisation of the vascular lesions. *Annals of Neurology*, **25**, 291–7.

Klein, R.G., Campbell, R.J., Hunder, G.G., and Carney, J.A. (1976). Skip lesions in temporal arteritis. *Mayo Clinic Proceedings*, **51**, 504–10.

Leger, J.M. *et al*. (1989). The spectrum of polyneuropathies in patients affected with HIV. *Journal of Neurology, Neurosurgery and Psychiatry*, **52**, 1369–74.

Lindblad, S. and Hedfors, E. (1985). Intra-articular variation in synovitis. Local macroscopic and microscopic signs of inflammatory activity are significantly correlated. *Arthritis and Rheumatism*, **28**, 977–88.

Lindblad, S. and Hedfors, E. (1987). The synovial membrane of healthy individuals—immunohistochemical overlap with synovitis. *Clinical and Experimental Immunology*, **69**, 41–7.

Lotz, B.P., Engel, A.G., Nishino, H., Stevens, J.C., and Vichy, W.L. (1989). Inclusion body myositis: observations in 40 patients. *Brain*, **112**, 727–47.

McCurley, T.L., Collins, R.D., and Ball, E. (1990). Nodal and extranodal lymphoproliferative disorders in Sjögren's syndrome. *Human Pathology*, **21**, 482–92.

Mata, M., Kahn, S.N., and Fink, D.J. (1988). A direct electron microscopic immunocytochemical study of IgM paraproteinemic neuropathy. *Archives of Neurology*, **45**, 693–7.

Meier, C., Grahmann, F., Engelhardt, A., and Dumas, M. (1989). Peripheral nerve disorders in Lyme borreliosis. *Acta Neuropathologica*, **79**, 271–8.

Mellgren, S.I., Conn, D.L., Stevens, J.C., and Dyck, P.J. (1989). Peripheral neuropathy in primary Sjögren's syndrome. *Neurology*, **39**, 390–4.

Melsen, F., Melsen, B., Mosekilde, L., and Bergman, S. (1978). Histomophometric analysis of normal bone from the iliac crest. *Acta Pathologica, Microbiologica, et Immunologica Scandinavica*, **86**, 70–81.

Monte, S.M. *et al*. (1988). Peripheral neuropathy in the acquired immunodeficiency syndrome. *Annals of Neurology*, **23**, 485–92.

Nishino, J., Engel, A.G., and Rima, B.K. (1989). Inclusion body myositis. The mumps virus hypothesis. *Annals of Neurology*, **25**, 260–4.

Nordstrom, D.M., Petropolis, A.A., Giorno, R., Gates, R.H., and Reddy, V.B. (1989). Inflammatory myopathy and acquired immunodeficiency syndrome. *Arthritis and Rheumatism*, **32**, 475–9.

Omdal, R., Henriksen, O.A., Mellgren, S.I., and Husby, G. (1991). Peripheral neuropathy in systemic lupus erythematosus. *Neurology*, **41**, 808–11.

Parfitt, A.M., Matthews, C.H.E., Villanueva, A.R., Kleerekoper, M., Frame, B., and Rao, D.S. (1983). Relationships between surface, volume, and thickness of iliac trabecular bone in ageing and osteoporosis. *Journal of Clinical Investigation*, **72**, 1396–409.

Pariesen, M., Cosman, F., Mellish, R.W.E., *et al*. (1995). Bone structure in postmenopausal hyperparathyroid, osteoporotic and normal women. *Journal of Bone and Mineral Research*, **10**, 1393–9.

Rooney, M. *et al*. (1988). Analysis of the histologic variation of synovitis in rheumatoid arthritis. *Arthritis and Rheumatism*, **31**, 956–63.

Rooney, M., Whelan, A., Feighery, C., and Bresnihan, B. (1989). Changes in lymphocyte infiltration of the synovial membrane and the clinical course of rheumatoid arthritis. *Arthritis and Rheumatism*, **32**, 361–9.

Ropert, A. and Metral, S. (1990). Conduction block in neuropathies with necrotizing vasculitis. *Muscle and Nerve*, **13**, 102–5.

Rosen, N., Sommer, I., and Knobel, B. (1985). Intestinal Buerger's disease. *Archives of Pathology and Laboratory Medicine*, **109**, 962–3.

Rosenberger, J.L., Cooper, N.S., Soren, A., and McEwen, C. (1981). A statistical approach to the histopathologic diagnosis of synovitis. *Human Pathology*, **12**, 329–37.

Said, G., Lacroix-Ciaudo, C., Fujimura, H., Blas, C., and Faux, N. (1988). The peripheral neuropathy of necrotizing arteritis: a clinicopathological study. *Annals of Neurology*, **23**, 461–5.

Savage, C.O.S., Pottinger, B.E., Gaskin, G., Lockwood, C.M., and Pusey, C.D. (1991). Vascular damage in Wegener's granulomatosis and microscopic polgarteritis: presence of anti-endothelial antibodies and their relation to anti-neutrophic oytopeason antibodies. *Clinical and Experimental Immunology*, **85**, 14–19.

Schoenen, J., Sinard-Gainko, J., Carpentier, M., and Reznik, M. (1989). Myositis during *Borrelia burgdorferi* infection (Lyme disease). *Journal of Neurology, Neurosurgery and Psychiatry*, **52**, 1002–8.

Serratrice, G., Pellisier, J.F., Roux, H., and Quilichini, P. (1990). Fasciitis, perimyositis, myositis, polymyositis and eosinophilia. *Muscle and Nerve*, **13**, 385–95.

Shin, S.S. *et al*. (1991). Monocytoid B-cell lymphoma in patients with Sjögren's syndrome. A clinicopathological study of 13 patients. *Human Pathology*, **22**, 422–30.

Shintaku, M., Osawa, K., Toki, J., Maeda, R., and Nishiyama, T. (1986). A case of amyloid angiopathy of the central nervous system associations with amyloid angiopathy. *Acta Neuropathologica*, **70**, 340–2.

Silver, R.M., Heyes, M.P., Maize, J.C., Quearry, B., Vionnet-Fuasset, M., and Sternberg, E.M. (1990). Scleroderma, fasciitis and eosinophilia associated with the ingestion of tryptophan. *New England Journal of Medicine*, **322**, 874–81.

Soren, A. (1978). *Histodiagnosis and clinical correlations of rheumatoid and other synovitis*. Thieme, Stuttgart.

Storm, T., Steiniche, T., Thamsborg, G., and Melsen, F. (1993). Changes in bone morphometry after long term treatment with intermittent, cyclic etridonate for post menopausal osteoporosis. *Journal of Bone and Mineral Research*, **8**, 199–208.

Sunohara, N., Furukawa, S., Nishio, T., Mukoyama, M., and Satoyoshi, E. (1989). Neurotoxicity of human eosinophils towards peripheral nerves. *Journal of Neurological Sciences*, **92**, 1–7.

Takatsu, M. *et al*. (1985). Immunofluorescence study of patients with neuropathy and IgM M proteins. *Annals of Neurology*, **18**, 173–81.

Tegner, R., Tomé, F.M.S., Godeau, P., L'hermitte, F., and Fardeau, M. (1988). Morphological study of peripheral nerve changes induced by chloroquine treatment. *Acta Neuropathologica*, **75**, 253–60.

Thomas, P.K. and Ochoa, J. (1984). Microscopic anatomy of peripheral nerve fibres. In *Peripheral neuropathy*, Vol. 1, (2nd edn.) (ed. P.J. Dyck, P.K. Thomas, E.H. Lambert, and R. Bunge), pp. 39–96. W. B. Saunders, Philadelphia.

Travis, W.D., Hoffman, G.S., Leavitt, R.Y., Pass, H.I., and Fauci, A.S. (1991). Surgical pathology of the lung in Wegener's granulomatosis. *American Journal of Surgical Pathology*, **15**, 315–33.

Vigorita, V.J. (1984). The bone biopsy protocol for evaluating osteoporosis and osteomalacia. *American Journal of Surgical Pathology*, **8**, 925–30.

Vital, C. *et al*. (1988). Peripheral neuropathy with essential mixed cryoglobulinemia: biopsies from 5 cases. *Acta Neuropathologica*, **75**, 605–10.

Vital, A., Viral, C., Rigal, B., Decamps, A., Emeriau, J.P., and Galley, P. (1990). Morphological study of the ageing human peripheral nerve. *Clinical Neuropathology*, **9**, 10–15.

Wong, R.L. and Korn, J.H. (1986). Temporal arteritis without and elevated sedimentation rate. *American Journal of Medicine*, **80**, 959–64.

Wrzolek, M.A., Sher, J.H., Kozlowski, P.B., and Rao, C. (1990). Skeletal muscle pathology in AIDS: an autopsy study. *Muscle and Nerve*, **13**, 508–15.

4.11 Electrophysiology

Adam Young

Modern nerve conduction studies and electromyography provide unique quantitative information about the function of nerves and muscles. These investigations can be an invaluable aid to both the diagnosis and management of musculoskeletal and neuromuscular disorders. The general aims of electrical tests in these disorders are to:

1. detect and distinguish between disorders of:
   a. anterior horn cells
   b. nerve roots
   c. plexi
   d. peripheral nerves
   e. neuromuscular junctions
   f. muscles
2. define location, extent, and severity;
3. relate neurophysiological abnormalities to clinical features;
4. infer pathology if possible.

The principal application of electrical studies to rheumatic disease is in the diagnostic and prognostic assessment of the neurological and myopathic features that complicate several of the inflammatory arthropathies and autoimmune rheumatic disorders and some of the mechanical and soft tissue conditions. The usual features are muscle weakness or wasting, sensory symptoms, pain or deformity, or a combination of these, in a patient with an established autoimmune rheumatic disorder or presenting for the first time with a rheumatic complaint for diagnosis. The most frequently encountered indications, possible pathophysiological processes and some examples for electrodiagnosis in rheumatology are shown in Table 1.

Clinical feature	Pathophysiology	Examples
Muscle weakness or wasting	Atrophic	Rheumatoid joint
	Myopathic	Polymyositis
	Neuropathic	Rheumatoid or entrapment neuropathy
Sensory symptoms	Peripheral neuropathy	Rheumatoid arthritis
	Entrapment neuropathy	Ulnar neuropathy
Pain	Entrapment neuropathy	Carpal tunnel syndrome
	Nerve injury	Causalgia
Deformity	Hereditary or rare neuromyopathies	Muscular dystrophy
	Entrapment neuropathy	Ulnar nerve claw hand

Table 1 Indications for electrodiagnosis in rheumatic disorders

There are several and varied procedures available and good clinical information is essential for planning an individual examination. The answer to a specific clinical question depends on the interpretation of the appropriate choice of a number of laboratory, radiological and electrical investigations and the requesting clinician should consider carefully the diagnostic and prognostic yield of the electrical study with respect to inevitable discomfort to the patient during this procedure. It is, however, a safe investigation and, except when needle sampling of muscles is required, is non-invasive.

The following points should be considered when an electrical test is planned:

1. The optimum diagnostic yield depends on thorough clinical evaluation of the patient's symptoms and signs.
2. It only provides physiological information but will complement other studies of pathological (e.g. biopsy) or structural disturbances (e.g. radiology).
3. Biopsy and electromyography must not be done in the same muscle, and muscle enzymes (e.g. creatine phosphokinase) should be assayed either before or after 2 weeks of muscle sampling.
4. It is uncomfortable for the patient.
5. It requires expensive equipment and expertise.

General principles and methods

Technical details of the methods used are outside the scope of this chapter and if required can be found in the specialized literature (Ludin 1980).

Nerve conduction studies

Motor nerve conduction velocity

This is measured by stimulating a motor nerve with a surface (skin) electrode and recording the response from the appropriate muscle with preferably a surface electrode placed over the muscle belly (only occasionally is a concentric needle electrode inserted into the muscle belly required). A supramaximal stimulus is delivered at a minimum of two sites along the peripheral nerve. The onset of the muscle response, the motor action potential (MAP) gives the terminal or distal motor latency (DL). If the distance between the two stimulating sites is measured, the motor nerve conduction velocity (NCV) of that segment of nerve can be calculated from the formula:

Speed (m/s) = distance (cm) / time (ms)

The amplitude, duration and shape of the MAP is examined and all these findings are compared with the contralateral side and other peripheral nerves, and also to the normal values available in standard tables (Ludin 1980) or developed by the user's own laboratory. Figure 1 and Figure 2 show the stimulation and recording sites and traces obtained in an evaluation of the ulnar nerve.

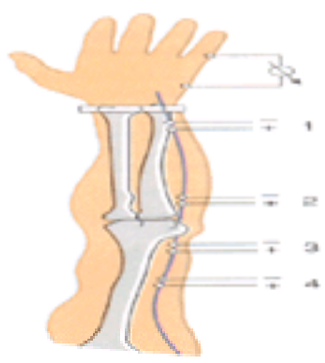


Fig. 1 Stimulation and recording sites of ulnar nerve. The stimulation points of the ulnar nerve are shown by numbers and are at the wrist, below and above the elbow, in the axilla, and at Erb's point in the supraclavicular fossa. Recordings of the muscular response are over the adductor digiti minimi (ADM). The distance between the stimulation and recording sites is measured in centimetres. This examination will demonstrate a localized slowing at a particular part of the ulnar nerve (for example compression by Osborne's band at the elbow) or a more generalized slowing of the ulnar nerve if axonal or wallerian degeneration is present as a result of a more proximal lesion. This is widespread and involves other nerves, it may be part of a peripheral neuropathy.

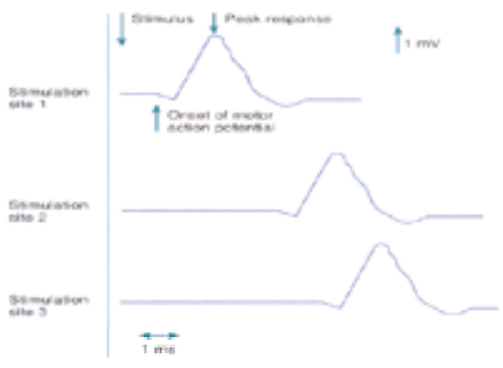


Fig. 2 Recordings of ulnar nerve study. This patient presented with pain, sensory symptoms and slight weakness of the little finger on the right. The referring clinician suspected an ulnar nerve lesion and planned ulnar nerve exploration. MAPs were recorded from the ADM following stimulation of the ulnar nerve at the first three sites previously shown in Fig. 1. The amplitude is calibrated at 1 millivolt (mV) per vertical scale mark, and the latency in 1 millisecond (ms) per horizontal mark. With modern machines cursors are placed at the onset of the responses and the velocities calculated from the formula speed (m/s) equals distance over time. In this example the distal latency recorded in the ADM following stimulation of the ulnar nerve at the wrist was 2.5 ms, and velocity in the forearm and at the elbow was 59.3 m/s and 58.1 m/s respectively. These are normal results for ulnar nerve studies. The findings of normal peripheral ulnar motor nerve conduction studies in the presence of clinical abnormalities suggest a lesion other than in the ulnar nerve itself and indicate performing some of the other procedures descibed further on in this chapter, for example, sensory nerve studies, needle sampling and reflex studies. In this example normal sensory studies but electrical evidence of denervation in the ADM would be highly suggestive of a preganglionic lesion of the T1 nerve root or in the anterior horn cell.

Sensory nerve conduction velocity

This is measured normally by stimulating the sensory nerve and recording the sensory nerve action potential (SNAP) with surface electrodes distally, i.e antidromically, because this is technically easier to do. The responses are of much lower amplitudes (normally measured in microvolts) and more difficult to elicit, and may require averaging techniques. The amplitude, shape and duration are recorded. The velocity is calculated in the normal way, and/or the distal latency standardized for distance. An example of a median nerve sensory study is shown in Fig. 3.

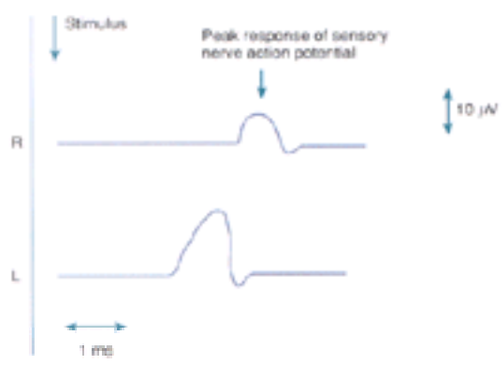


Fig. 3 Sensory nerve conduction study in median nerve. This patient was referred by a clinician who suspected carpal tunnel syndrome which had not responded to conservative measures of therapy. It shows the tracings of right and left median nerves. The recording electrodes are placed over the index finger as shown in Fig. 1 and the median nerve is stimulated at the wrist. The SNAP is measured in microvolts (vertical scale) and DL in milliseconds (horizontal scale). The DL for the right median nerve is delayed and there is a significant difference between the right (symptomatic) and left sides. Although not shown here, in this case the motor conduction studies of the median nerve over the right wrist were abnormal but those performed in the forearm were normal thus confirming the clinician's diagnosis and the patient was referred for surgical decompression.

Mixed nerve potential

The mixed motor and sensory nerves are tested by stimulating a mixed fibre nerve proximally (e.g. in median fossa for median nerve) and recording the amplitude and latency of the compound nerve action potential distally (at the wrist for median nerve).

Reflex latencies

Specific involvement of a nerve root may be indicated from conduction studies of the central segments of peripheral nerves by measuring the latencies of limb reflexes (Malcom 1951). A technique using an electronically triggered patella hammer has made reflex studies easier to perform and more comfortable for the patient. The ankle jerk is elicited in the usual way, the response recorded with a surface electrode over the medial belly of the gastrocnemius. The latency is recorded in milliseconds and compared with the other side. Its absence or a delay of 2 ms or more is found in conditions which clinically affect this reflex. For example it is

particularly useful in patients who present with back pain and sciatica but no evidence of tension signs or abnormal neurology. A delayed ankle jerk latency in the presence of normal conduction of the peripheral segments of sciatic and peroneal nerves is suggestive of an S1 root lesion.

Similarly, the 'H' reflex is mediated through the monosynaptic reflex arc of the first sacral (S1) root and increased latencies in the presence of normal peripheral conduction imply dysfunction of this root (Deschuytere and Rosselle 1973). The posterior tibial nerve is stimulated at below threshold strength in the popliteal fossa and the motor response in the medial head of the gastrocnemius in the calf is recorded with a surface electrode.

When a motor nerve is stimulated impulses travel proximally to the anterior horn cells followed by recurrent conduction back down the nerve where the small muscle response (known as the 'F-wave') can be detected over muscles of the appropriate root distribution. This latency is measured and compared with the other side or height-related normal values. They are sometimes technically difficult to elicit but a delayed F-wave in the presence of normal peripheral conduction implies slowing of proximal motor fibres at plexus or root level.

Neuromuscular function

Repetitive nerve stimulation and recordings of motor action potentials can reveal abnormalities of neuromuscular transmission. In myasthenia gravis a significant decrement occurs in about 60 per cent if one muscle is sampled but this can be improved to 95 per cent if many other muscle groups are included (Oh *et al*. 1982).

Electromyography

Electromyography (EMG) is the recording and study of spontaneous and voluntary electrical activity of muscle. It almost always requires the sampling of muscles with a concentric needle electrode (CNE) and is normally performed in conjunction with the electrical studies described above. The first objective is to distinguish between primary muscle disease and secondary changes as a result of neurogenic disorders. A normal muscle at rest has no detectable electrical activity, apart from the initial insertional activity or if the CNE is placed very close to the motor end plate. If a motor unit loses its nerve supply, sensitivity to acetylcholine increases and spontaneous discharges can be detected by the CNE when amplified and displayed on a cathode ray oscilloscope as small spontaneous fibrillation (SF) potentials (less than 300 microvolts; see Fig. 4). They may be confused with normal motor end plate potentials, an important distinction because spontaneous fibrillation is characteristic of denervation. Positive sharp waves are another electrical sign of denervation. The interpretation of these changes requires considerable skill and experience and is largely a matter of personal judgement, despite the recent improvements in developing objective measurements of the motor unit.

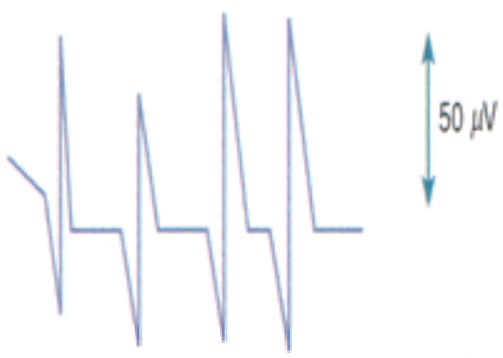


Fig. 4 Electrical signs of spontaneous activity – spontaneous fibrillation. Small units (usually less than 100 μV) seen in muscle at rest often coming in bursts. Several depths and areas of a muscle need to be sampled. When found they are characteristic of denervation.

When the motor unit fires the action potentials which are conducted along the muscle fibre cell membranes are much larger (Fig. 5). If there is loss of motor units (e.g. primary myopathy), or muscle fibres cannot be activated (e.g. myasthenia gravis) the recorded potentials are smaller. In axonal degeneration, sprouting of surviving axons in an attempt to reinnervate nearby denervated muscle fibres results in a larger motor unit which is detected electrically by the appearance of large amplitude motor action potentials (Fig. 5). When the force of contraction is increased more units are activated, and these recruitment patterns vary with different conditions. The analysis of interference pattern is normally subjective and computer based methods are not generally available.

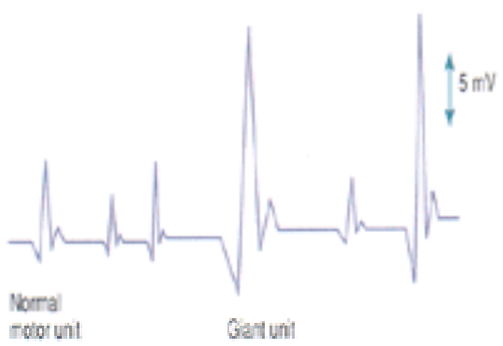


Fig. 5 Electrical signs of voluntary motor units. The amplitude of a normal action potential of a motor unit on voluntary contraction varies from 1 to 8 mV depending on the muscle. Giant motor units exceed this and are a sign of attempted reinnervation.

Various other features are recognized and include the complex repetitive discharges seen in polymyositis, fasciculation potentials characteristic of denervation, and the myotonic discharges of myotonia. Single fibre techniques (Stalberg and Trontelj 1979) can be useful in detailed examination of the motor unit, particularly in the diagnosis of motor end plate disorders and reinnervation. This was initially a very time consuming refined method, but has been simplified to some extent by modern methods. The 'blanket technique' described by Payan (1978) bridges the gap between single-fibre and conventional electromyography.

Sensory evoked potentials

Very small potentials, evoked by repetitive percutaneous stimulation of peripheral nerves can be detected with averaging techniques by surface electrodes over the skin of spinal cord or scalp. These are known as somatosensory evoked potentials (SEPs) and have been used extensively in the investigation of patients with multiple sclerosis, particularly in suspected cases and in revealing clinically silent lesions (Jones 1982). The SEP has also been used in investigations of peripheral neuropathies, and is particularly well suited to the investigation of proximal lesions of the brachial plexus and spinal nerve roots which are so inaccessible to normal recording methods (Jones *et al*. 1981). The dermatomal somatosensory evoked potential (DSEP) is an adaptation of this technique and found a place in the assessment of lumbar spinal disease, especially compressive disc and degenerative lesions (Katifi and Sedgwick 1987).

Clinical application

How can electrical methods assist the clinician in rheumatological practice? The answer to this depends very much on the nature, site and chronicity of the lesion, the clinical information available, and the knowledge of other features which may affect the electrical findings, for example the presence of neurological anomalies or dual pathology. In some cases this investigation can be crucial in decisions regarding surgery, but in others the diagnostic and prognostic yield may be minimal.

Three main pathological processes result in the electrical abnormalities seen in conduction studies:

- demyelination
- axonal degeneration
- metabolic effects

It is customary to use the term demyelination for neuropathies with a primary disorder of the myelin sheath. Prolonged distal latencies and conduction velocities are the characteristic electrical findings, the best known examples being the hereditary neuropathies. The degree of slowing parallels the degree of demyelination, and conduction velocities are usually reduced by at least 30 per cent. Segmental demyelination is found localized to one segment of a nerve in the entrapment neuropathies. More prolonged compression of a peripheral nerve results in further damage to the axons distal to the lesion causing wallerian degeneration (Lewin 1993). Normal or slightly reduced conduction velocities in the presence of reduced amplitudes or absent sensory, motor and/or mixed action potentials are characteristic of axonal degeneration. This is seen typically in amyloidosis, and alcohol and drug abuse. Motor conduction velocities of less than 40 m/s are unusual (McLeod *et al*. 1973). In rheumatological practice axonal degeneration is typical of the vasculitides, for example mononeuritis multiplex seen in polyarteritis nodosa and rheumatoid arthritis. Certain conditions result in a mixed pattern of electrical findings, for example in diabetes mellitus and Guillain–Barré syndrome. The rapid improvement in conduction times after renal transplants is explained by certain metabolic effects.

The degree and site of these changes (Table 2) are an indication of the presence of axonal degeneration or widespread or segmental demyelination and this helps to narrow down the differential diagnosis. Electrical studies help distinguish the relative importance of the coexistence of these processes, for example localized entrapment neuropathies in the presence of generalized neuropathies of rheumatoid arthritis or diabetes. Timing of the investigation may be important, and serial studies may be helpful. For example, because the amplitude of the motor action potential reflects the number of surviving axons, the response in the thenar muscles on stimulating the median nerve at the wrist is a useful prognostic indicator in the early stages of a polyradiculitis (Miller *et al*. 1988). Significant demyelination in the childhood and hereditary neuropathies can be a useful predictor of response to steroid treatment.

Electrical features	Axonal degeneration	Segmental demyelination
Amplitude	Markedly reduced	Normal/slight reduction
Distal latency	Normal/slight delay	Marked delay
Velocity	Normal/slight delay	Marked delay
Examples	Vasculitis	Diabetes

Table 2 Summary of electrical findings of conduction studies in neuropathy

Common clinical conditions

Acute nerve injury

This subject is well covered in surgical texts and the great deal that has been learnt from studying traumatic nerve lesions can be applied to other lesions of the peripheral nerve. The basic principles of damage are taught early in medical school and summarized in Table 3.

Type of injury	Pathophysiology	Recovery
Neurapraxia (contusion)	No axonal degeneration	Days to weeks
Axonotmesis (trauma)	Wallerian degeneration Endoneural tubes survive	Variable Months to years
Neurotmesis (division of nerve)	Complete degeneration of nerve	None without surgery

Table 3 Classification of nerve injury (after Seddon 1972)

A surgical opinion should be sought early in these cases because the timing of surgery can be crucial to recovery. Early referral for rehabilitation may be necessary and the planning of this can be helped by the prognostic information obtained from electrical methods. Wallerian degeneration requires nerve regrowth and if present the timescale of recovery is inevitably longer and dependent on the length of nerve and degree of damage.

Acute compression neuropathies

Sudden but prolonged compression of a peripheral nerve leads to mild focal demyelination affecting large fibres preferentially, and nerve conduction studies reveal motor and sensory conduction block at the damaged site, but normal conduction distally because the nerve is electrically still intact (Ochoa *et al*. 1973). This is followed by gradual electrical and clinical recovery accompanied by wallerian degeneration which may be detected by evidence of electromyography features of denervation. The clinical syndromes most often seen in rheumatological practice are shown in Table 4.

Clinical sign	Nerve	Condition
Wrist drop	Radial nerve palsy	Saturday night palsy Tourniquet palsy
Weak hand	Ulnar nerve	Tourniquet palsy
Foot drop	Common peroneal nerve	Crossed leg palsy Tight plaster cast Tourniquet palsy
	Sciatic nerve	Total hip replacement Haematoma in haemophilia or poor anticoagulation control

Table 4 Acute nerve entrapment syndromes

The so-called 'tourniquet palsy' may in fact be the result of prolonged pressure because of poor positioning of a limb during prolonged anaesthesia. Similar clinical and electrical features are sometimes seen following the reduction of difficult fractures and the removal of tight plaster casts. Bonney (1986) has recently reviewed iatrogenic nerve damage, all of which may be seen in rheumatology clinics and includes traction of peripheral and cervical nerve roots during anaesthesia, damage to the superficial radial nerve following surgery for De Quervain's syndrome, digital nerve damage during surgery for Dupytren's contracture, and sciatic nerve damage during total hip replacement. The changes detected electrically depend critically on when they are performed after injury. Spontaneous fibrillation so characteristic of denervation does not appear for at least ten days. The appearance of this or other electrical signs of axonal degeneration indicate a slow recovery (Bolton and McFarlane 1978). Table 5 summarizes these changes.

Time after injury	Sensory conduction	Motor conduction
Immediate	Absent across injury	None proximal to injury Normal distal response
1–10 days	Progressive decline in SNAP and CNAP	Progressive decline in MAP
	Normal NCV	Normal NCV
10–21 days		Spontaneous fibrillation may be detected on EMG

Table 5 Electrical changes after injury

Chronic or subacute compressive (entrapment) neuropathies

Mononeuropathies developing after prolonged mechanical damage to a nerve at a site of anatomical constriction are common, the most familiar of which is the median nerve in the carpal tunnel. Following localized demyelination, segmental demyelination occurs and progresses to wallerian degeneration. Varying degrees of slowing of nerve conduction velocities are found and in early or mild cases sensory nerves studies are more sensitive (Buchtal and Rosenfalk 1971). Needle sampling may reveal denervation in the relevant muscles in more severe cases. Dawson *et al*. (1983) have described entrapment neuropathies in detail, but the common lesions encountered in rheumatological practice are summarized in Table 6.

Nerve	Site of entrapment
Median	Carpal tunnel at wrist
Ulnar	Palm, wrist or elbow (cubital tunnel)
Radial	Spiral groove in upper arm
Cervical nerve root/plexus	Cervicothoracic outlet
Common peroneal	Head of fibula
Posterior tibial	Tarsal tunnel
Lumbar sacral roots	Lumbosacral spinal canal or foramina

Table 6 Common entrapment neuropathies

Not only can electrical studies determine the exact site and severity of a lesion in these conditions, they may also help to assess prognosis and response to treatment, depending on the nature and site of lesion. However, correlation of clinical and electrical findings are not always straightforward as shown by the lack of uniformity in reported studies of electrical assessment and monitoring of recovery following injury and surgical decompression. These are best discussed on an individual basis later. For diagnosis it is important to demonstrate that delay in conduction distal to the site of compression is disproportionate to any mild proximal slowing (Simpson 1956). Electrodiagnostic demonstration of an isolated motor nerve root lesion can be achieved by finding evidence of denervation in a group of muscles corresponding to the distribution of the spinal segment rather than the peripheral nerve. A preganglionic root lesion will not affect the distal sensory action potential whereas a postganglionic lesion may do so because the integrity of the peripheral nerve fibre may be affected. This may be vital information in differentiating between a lesion of the anterior horn cell (e.g. motor neurone disease) and a localized postganglionic lesion (e.g. in brachial plexus).

Median nerve

Mild cases of carpal tunnel syndrome usually respond initially at least to conservative measures (Le Quesne 1978). Electrodiagnosis is generally indicated if the diagnosis is not secure or the lesion appears progressive as in the following circumstances:

1. If motor neurone disease is suspected because thenar muscle wasting is marked or progressive and sensory symptoms absent or minimal.
2. Osteoarthritis of the hand, especially in the carpometacarpal joint of the thumb, or the presence of reflex sympathetic dystrophy can make a clinical diagnosis uncertain.
3. Dual pathology is suspected, that is both at the cervical spine and at the wrist, the so-called double crush syndrome (Upton and McComus 1982). This information may affect treatment.
4. If surgical exploration and/or decompression is considered.
5. Recurrent symptoms following surgery for median nerve decompression may be because of inadequate surgery which can be better demonstrated electrically in comparison with a preoperative result.

The minimal diagnostic criterion is a prolonged median sensory conduction velocity with a normal ulnar sensory velocity. The distal motor latency to the abductor pollucis brevis is normally also determined. Using these criteria, Boniface *et al*. (1994) reported that nerve conduction studies excluded the clinical diagnosis of carpal tunnel syndrome in 36 per cent, and most of these patients (72 per cent) responded to conservative treatment at follow-up. In this study, it was clear that decisions whether to perform decompressive surgery were influenced greatly by positive electrical findings, and the authors highlighted the importance of the investigation for this clinical condition as part of good clinical practice as well as health service costs.

Postoperative electrical studies are not always easy to interpret. Goodman and Gilliat (1961) reported normalization of conduction velocities within a few months but Melvin *et al*. (1968) reported persistently abnormal sensory latencies in 60 per cent of patients one year after decompression of the median nerve at the wrist. Surgeons still report patients with typical carpal tunnel symptoms who respond to decompressive surgery in the presence of normal electrical findings. Refinements of electrodiagnosis reported by Mills (1985) in which the median nerve is stimulated in the palm and recorded at the wrist should obtain correct results in 95 per cent of cases.

Ulnar nerve

The common indications for ulnar nerve studies are:

1. the degree and exact site of damage at the elbow;
2. whether the lesion is in the ulnar nerve or the 1st thoracic root;
3. whether an ulnar nerve lesion of the hand is proximal or distal to the bifurcation into deep muscle and superficial sensory branches.

Conduction studies of ulnar nerve lesions at the elbow correlate well with clinical severity and localization is correct in 95 per cent of cases (Payan 1969). Mild symptoms and no objective signs can usually be best treated conservatively, but surgery should be considered in patients with moderate symptoms and neurological signs which progress, and electrical tests are often the best way of assessing this. Decompression of the ulnar nerve without transposition is often all that is needed, especially in patients with a short history, associated with mild weakness only, and with mild abnormalities of sensory nerve action potentials (Miller and Hummel 1980). Transposition of the ulnar nerve is still being performed purely on clinical grounds. It should be reserved for patients with objective evidence of ulnar nerve damage. Repair or mobilization of peripheral nerves can produce causalgia and become a long-term treatment problem (Withrington and Wynn Parry 1984). Occasionally the nerve can be trapped in Guyon's canal at the wrist.

Radial nerve

This nerve can be injured in the axilla (e.g. pressure from a crutch), in the spiral groove (Saturday night palsy) or by fracture of the humerus or radius, and thorough clinical examination is sufficient usually for initial diagnosis and monitoring recovery. As with the ulnar nerve it may occur postoperatively from faulty positioning or prolonged tourniquet pressure. If there is any dispute about the cause or timing of such a lesion, electrodiagnosis may be required for medicolegal reasons. Surgery is rarely indicated.

Cervicothoracic nerve roots and brachial plexus

Cervical disc prolapse, cervical ribs and bands in the thoracic outlet syndrome, neuralgic amyotrophy, tumours and irradiation are all relatively uncommon conditions and electrical studies may contribute to diagnosis if there is clinical progression and surgery is considered. Unfortunately the use of F-waves and somatosensory evoked potentials have not helped as much as was hoped originally in the diagnosis of lesions of the proximal segments of these nerves. Electrical studies are still often needed to exclude lesions of these nerves at more peripheral sites and in patients with unusual presentations or poor response to treatment, because the demonstration of dual pathology may alter management. Common clinical situations are carpal tunnel syndrome and cervical nerve root involvement from cervical spine degeneration, rheumatoid involvement of both cervical spine (often silent) and hand or elbow, and multiple nerve entrapment syndromes. The diagnosis and planning of treatment for traction lesions of the brachial plexus has been shown to be helped greatly by expert electrical examination in special centres (Jones *et al*. 1981; Wynn Parry 1988).

Sciatic and common peroneal nerves

The sciatic nerve may be injured following traction during hip surgery. The common peroneal nerve is vulnerable at the level of the fibula head as the nerve enters the peroneus muscle. Electrical studies are indicated when:

1. ecovery is poor;
2. decompressive surgery is considered;
3. here is difficulty distinguishing between an isolated palsy or a lesion of the 5th lumbar nerve root or both.

If the lesion is complete and recovery is not apparent or appears very slow, decisions about the amount of physiotherapy actually required, planning of rehabilitation and type of leg appliances all require prognostic information early which can often only be obtained by electromyography. In distinguishing between an L5 root lesion and peroneal nerve palsy causing a foot drop, it is worth noting that the weakness of foot eversion which is found in peroneal nerve palsies is not always easy to demonstrate. The posterior tibialis is one of the few muscles innervated by the L5 root and not via the peroneal nerve and therefore an important muscle to examine in distinguishing these two lesions. It is a difficult muscle to examine clinically but ideal for needle sampling.

Posterior tibial nerve

The tarsal tunnel syndrome is not common and the electrophysiological findings reported in the limited series are generally disappointing (Kaeser 1970). It is a rather painful area to examine electrically but denervation found on needle sampling proves reliable and treatment can be rewarding.

Back pain and sciatica

Electrical studies do not have a place in the routine work up of acute or chronic mechanical back pain with or without sciatica. Over 90 per cent of prolapsed disc disorders are diagnosed successfully after careful clinical examination followed by radiological studies. These methods have in the past been less successful in correctly diagnosing and localizing chronic nerve root involvement from bony entrapment because water-soluble dyes do not extend far enough laterally into the lateral recesses and exit foramina where impingement of the nerve roots in an often narrowed canal occurs (McNab 1977). Development of electrodiagnostic methods was invaluable in these patients, and successful localization of segmental nerve root damage in 70 per cent or more can be achieved (Young *et al*. 1983). Modern magnetic resonance and computerized scanning methods of the lumbosacral spine have improved greatly the visualization of the spinal canal and its recesses, and electrical investigation is now less often needed. Electrical methods can still be useful in the following circumstances:

1. the problem of failed surgery (Young and Wynn Parry 1988);
2. clinical and radiological signs which do not correspond;
3. in isolating single nerve root involvement by epidural fibrosis in patients with extensive arachnoiditis (Leyshon *et al*. 1981).

Electrical changes remain for more than a year postoperatively both in the peripheral muscles (Young *et al*. 1983) and in the paraspinal muscles (See and Kraft 1975) and must therefore be interpreted carefully.

Other less commonly affected nerves

Anterior and posterior interosseus nerves in forearm

Lesions in these nerves can be confirmed with motor conduction studies and by finding abnormalities in the appropriate muscles on needle sampling.

Facial nerve

Conduction studies are difficult and needle sampling uncomfortable and a problem to interpret. Apart from the method, prognostic information from electromyography is limited. Needle sampling can determine whether loss of motor units is complete or not.

Peripheral neuropathies

Motor, sensory and mixed conduction studies in both arms and legs are essential if a generalized neuropathy is suspected. Sensory investigations are more sensitive than motor studies, and particularly the sural sensory and peroneal mixed action potentials in sensory neuropathies. The differential diagnostic possibilities are narrowed by a careful appraisal of the clinical features, the degree of slowing, and amplitude response, and which sensory and/or motor nerves are affected. Differentiation between the numerous possible causes of peripheral neuropathies on the basis of electrical findings alone is possible in only a few cases. The common causes of peripheral neuropathy such as diabetes, drug and alcohol effects, vitamin deficiencies, metabolic and autoimmune rheumatic disorders, malignancy, and Guillain–Barré syndrome can usually be diagnosed by means of detailed and full clinical, radiological, laboratory and electrical investigations.

Two varieties of peripheral neuropathy are well recognised features of rheumatoid disease (Pallis and Scott 1965). Low amplitudes and slowing of sensory conduction suggest segmental demyelination found in the more common mild distal sensory neuropathy, which has a good prognosis. Widespread denervation in the presence of

relatively normal conduction suggests axonal degeneration found in the more severe sensorimotor neuropathy. This may start as an isolated neuropathy and progress as part of a vasculitic process as in mononeuritis multiplex. Electrical studies are clearly valuable in distinguishing the two and demonstrating the extent of involvement, which may not be apparent clinically because of joint deformity or widespread synovitis.

It is not unusual for normal electrical values to be recorded in patients who present with features of a peripheral neuropathy, and these cases can be explained on the basis of very early mild disease, involvement of small diameter fibres only, central lesions of the nerve root, or non-organic disorders (Payan 1985). Most series record that as many as 50 per cent of cases remain undiagnosed, but in a very detailed study Dyck *et al*. (1981) reported that 42 per cent of patients proved to have an inherited disorder when careful family studies were performed, 21 per cent had an inflammatory demyelinating polyradiculopathy and 13 per cent had other acquired neuropathies.

Myopathies

The electrophysiological diagnosis of many myopathies is based on needle sampling. Conduction studies are required to exclude neuropathic lesions and are usually normal. Electrodiagnosis is most rewarding in the acute phases of the inflammatory myopathies. The classic changes in acute polymyositis are those of myopathic degeneration and acute denervation. Fibrillation potentials have been reported in 74 per cent of polymyositis and 33 per cent of dermatomyositis patients (Bohan *et al*. 1977), but are not an essential prerequisite for the diagnosis of an inflammatory myopathy. Other changes include myopathic motor potentials (low amplitude short duration motor unit potentials) and polyphasic potentials. These changes allow successful diagnosis in about two-thirds of all cases of myopathy, including the other inflammatory arthropathies that can be complicated by myopathy, such as rheumatoid arthritis and systemic lupus erythematosus. The diagnosis depends on the typical clinical features, an elevated creatine phosphokinase (CPK), electromyography abnormalities and histological changes seen on muscle biopsy. In practice, treatment is usually recommended if two of the four features are present (Bohan and Peter 1975). Prognosis has improved with the earlier and judicious use of steroids and immunosuppressive agents, and electromyography can be important in this respect although MRI is being used as a less invasive alternative, especially for children.

The most common drug-induced myopathy is caused by steroid treatment, which is at least partially dose dependent, often subclinical and frequently electromyography and CPK level are normal. Sometimes it is very difficult to distinguish between the changes of polymyositis and the myopathic changes seen in chronic steroid treatment (particularly the fluoridized corticosteroids) which many of these patients are in fact taking for their condition. However, florid spontaneous activity makes an inflammatory aetiology a high probability. A muscle biopsy is usually the only way to make the distinction safely. The other causes of myopathy, for example metabolic myopathies and the muscular dystrophies, require other more detailed investigations in addition to electromyography for a definitive diagnosis (Hudgson 1983). The assessment of treatment of these myopathies is mainly a clinical skill and, although occasionally helpful in some instances, electrodiagnosis has not found a place in the routine management of these cases.

Multifocal pathology

This may be suspected clinically for example in mononeuritis multiplex, but more often is revealed by the appropriate electrical tests. Certain conditions characterized by mild generalized neuropathic changes which are often subclinical, are also prone to localized entrapment neuropathies because of mechanical pressure as in diabetes, hypothyroidism, chronic alcoholism, inflammatory arthritis, and familial pressure palsies. Change in multiple cervical and/or root pathology although mild is not uncommon, particularly in the older patient and may be found during investigation for other conditions. In a series of patients with carpal tunnel syndrome reported by Murray-Leslie and Wright (1976) , 33 per cent also had humeral epicondylitis and cervical spine dimensions were significantly smaller in this group. A full clinical summary is vital to explain unusual electrical findings resulting from dual pathology.

Pitfalls in electrodiagnosis

The normal ranges for motor nerve conduction were first established by Hermann von Helmolz in the summer months of the early 1850s. Subsequent studies by him in the winter demonstrated the significant effect of temperature on nerve conduction (Helmolz and Baxt 1867) and present-day electromyographers are well aware of the necessity to provide standard temperatures (preferably warm ones) for this examination. In some circumstances cold patients need to be warmed up to obtain real results, for example in peripheral neuropathy. Another factor which can influence electrical readings is age, with maximal conduction velocities seen in the teens. Full-term neonates have half the normal adult ranges, and there is a reduction of 0.5 to 1.8 m/s every ten years after the age of 20. Amplitude decay increases with age. Anomalies of neurological arrangements occasionally make the life of an electromyographer difficult, the most common of which is the Martin–Gruber anastomosis between ulnar and median nerves, which probably occurs more frequently than the 6 per cent quoted by Hopf and Hense (1974). Conventional electrical tests will be essentially normal in patients with upper motor neurone lesions or with non-organic signs. In the latter, when the hysteria-conversion reaction or malingering posture makes clinical examination so difficult, electrical testing can be very reassuring, but occasionally reveals a genuine underlying pathological lesion.

Chapter References

Bohan, A. and Peter, J.B. (1975). Polymyositis and dermatomyositis. Parts 1 and 2. *New England Journal of Medicine*, **292**, 344–57, 403–7.

Bohan, A., Peter, J.B., Bowman, R.L., and Pearson, C.M. (1977). A computer assisted analysis of 153 patients with polymyositis and dermatomyositis. *Medicine*, **56**, 255–86.

Bolton, C.F. and McFarlane, R.M. (1978). Human pneumatic tourniquet paralysis. *Neurology*, **28**, 787–93.

Boniface, S.J, Morris, I.M., and Macleod, A. (1994). How does neurophysiological assessment influence the management and outcome of patients withh carpal tunnel syndrome? *British Journal of Rheumatology*, **33**, 1169–70.

Bonney, G. (1986). Iatrogenic injury of nerves. *Journal of Bone and Joint Surgery*, **68B**, 9–13.

Buchtal, F. and Rosenfalck, A. (1971). Sensory potentials in neuropathy. *Brain*, **94**, 241–62.

Dawson, D.M., Hallet, M., and Millender, L.H. (1983). *Entrapment neuropathies*. Little Brown, Boston.

Deschuytere, D. and Rosselle, P. (1973). Reflexes in L5 and S1 root compression. In *New developments in EMG and clinical neurophysiology*, Vol.3, (ed. J.E. Desmedt), pp. 360–66. Karger, Basel.

Dyck, P., Oviatt, K.F., and Lambert, E.H. (1981). Intensive evaluation of referred unclassified neuropathies yields improved diagnosis. *Annals of Neurology*, **10**, 222–6.

Goodman, H.V. and Gilliat, R.J. (1961). The effect of treatment of median nerve conduction in patients with carpal tunnel syndrome. *Annals of Physical Medicine*, **6**, 137.

Helmholz, H. and Baxt, N. (1867). Fortpflanzungs geschwindigkeit der Reizung in den motorischen Nerven des Menschen. *Mber Konig preuss Abhal Wiss*, 238–54.

Hopf, H.C. and Hense, W. (1974). Anomalien der Motorisation Innervation an der hand. *EEG-EMG*, **5**, 220.

Hudgson, P. (1983). Recent advances in diseases of muscle. In *Advances in Rheumatology* (ed. W.L. Dick and M.H. Moll), pp. 117–46. Churchill Livingstone, Edinburgh.

Jones, S.J. (1982). Clinical applications of short latency somatosensory evoked potentials. *Annals of the New York Academy of Science*, **388**, 369–87.

Jones, S.J., Wynn Parry, C.B., and Landi, A. (1981). Diagnosis of brachial plexus traction lesions by sensory nerve action potentials. *Injury*, **12**, 376–82.

Kaeser, H.E. (1970). Nerve conduction velocity measurements. In *Handbook of clinical neurology*, Vol. 7, (ed. P. Vinken and G. Bruyn). North Holland Publishing Company, Amsterdam.

Katifi, H.A. and Sedgwick, E.M. (1987). Evaluation of the dermatomal somatosensory evoked potentials in the diagnosis of lumbo-sacral root compression. *Journal of Neurology, Neurosurgery and Psychiatry*, **50**, 1204–10.

Le Quesne, P. (1978). The carpal tunnel syndrome. *British Journal of Hospital Medicine*, **20**, 155–64.

Levin, K.G. (1993). Common focal mononeuropathies and their electrodiagnosis. *Clinical Neurophysiology*, **10**, 181–9.

Leyshon, A., Kirwan, E. and Wynn Parry, C.B. (1981). Electrical studies in the diagnosis of compression of the lumbar root. *Journal of Bone and Joint Surgery*, **63B**, 71–5.

Ludin, H.P. (1980). *Electromyography in practice*. Georg Thieme Verlag Thieme-Stratton Inc., New York.

McLeod, J.G., Prineas, J.W., and Walsh, J.C. (1973). The relationship of conduction velocity to pathology in peripheral nerves. In *New developments in EMG and clinical neurophysiology*, Vol. 2, (ed. J.E. Desmedt). Karger, Basel.

McNab, I. (1977). *Backache*. Wiliams and Wilkins, Baltimore.

Malcom, D.S. (1951). A method of measuring reflex times applied in sciatica and other conditions due to nerve root compression. *Journal of Neurology, Neurosurgery and Psychiatry*, **14**, 15–24.

Melvin, J.L., Johnson, E.W., and Duran, R. (1968). Electrodiagnosis after surgery for carpal tunnel syndrome. *Archives of Physical Medicine and Rehabilitation*, **49**, 502–7.

Miller, R.G. and Hummel, E.E. (1980). The cubital tunnel syndrome: treatment with simple decompression. *Annals of Neurology*, **7**, 567–9.

Miller, R.G, Petersen, G.W., Daube, J.R., and Albers, J.W. (1988). Prognostic value of electrodiagnosis in Guillain–Barré syndrome. *Muscle and Nerve*, **11**, 769–74.

Mills, K.R. (1985). Orthodromic sensory action potentials from palmar stimulation in the diagnosis of carpal tunnel sydrome. *Journal of Neurology, Neurosurgery and Psychiatry*, **48**, 250–5.

Murray-Leslie, M. and Wright, V. (1976). Carpal tunnel syndrome, humeral epicondylitis and the cervical spine: a study of clinical and dimensional relations. *British Medical Journal*, **1**, 1439–42.

Ochoa, J., Fowler, T.J., and Gilliatt, R.W. (1973). Changes produced by a pneumatic tourniquet. In *New developments in EMG and clinical neurophysiology*, Vol. 2 (ed. J.E. Desmedt). Karger, Basel.

Oh, S.J., Eslami, N., and Nishihara, T. (1982). Electrophysiological and clinical correlates in myasthenia gravis. *Annals of Neurology*, **12**, 348–54.

Pallis, C.A. and Scott, J.T. (1965). Peripheral neuropathy in rheumatoid arthritis. *British Medical Journal*, **1**, 1141.

Payan, P. (1969). Electrophysiological localisation of ulnar nerve lesions. *Journal of Neurology, Neurosurgery and Psychiatry*, **32**, 208–20.

Payan, P. (1978). The blanket principle: a technical note. *Muscle and Nerve*, **1**, 423–6.

Payan, P. (1985). Peripheral neuropathy. In *Contemporary neurology* (ed. M. Harrison), pp. 225–33. Butterworths, London.

Seddon, H.J. (1972). Surgical disorders of the peripheral nerves. Churchill Livingstone, Edinburgh.

See, D.H. and Kraft, G.H. (1975). Electromyography in paraspinal muscles following surgery for root compression. *Archives of Physical Medicine and Rehabilitation*, **56**, 80–3.

Simpson, J.A. (1956). Electrical signs in the diagnosis of carpal tunnel and related syndromes. *Journal of Neurology, Neurosurgery and Psychiatry*, **19**, 275.

Stalberg, E. and Trontelj, J.V. (1979). *Single fibre electromyography*. Miravelle Press, Old Woking.

Upton, A.R.M. and McComus, A.J. (1982). The double crush in nerve entrapment syndromes. *Lancet*, **ii**, 359–62.

Withrington, R.H. and Wynn Parry, C.B. (1984). Management of painful peripheral nerve disorder. *Journal of Hand Surgery*, **9**, 24.

Wynn Parry, C.B. (1988). Update on peripheral nerve injuries. *International Disability Studies*, **10**, 11–20.

Young, A. and Wynn Parry, C.B. (1988). The assessment and management of the failed back. *International Disability Studies*, **10**, 21–5.

Young, A., Getty, J., Jackson, A., Kirwan, E., Sullivan, M., and Wynn Parry, C.B. (1983). Variations in the pattern of muscle innervation by L5 and S1 nerve roots. *Spine*, **8**, 616–24.

5.1.1 Epidemiology and the rheumatic diseases

Alan J. Silman

Introduction

Epidemiology can be defined broadly as the study of the occurrence of diseases in human populations. Occurrence can be considered in terms of the demographic, genetic, and environmental influences and thus, in addition to 'counting', the epidemiological investigation of a disorder aims to uncover risk factors in disease causation. Additionally, and increasingly, epidemiological method is being applied to the study of the natural history of disease and to those factors that influence prognosis.

The term clinical epidemiology refers to a different area of endeavour and describes the application of epidemiology to clinical practice. Areas in this regard of specific relevance to rheumatology are:

1. the derivation of criteria for diagnostic or classification purposes;
2. the evaluation of clinical data in terms of their accuracy (validity) and reliability (reproducibility);
3. the design, conduct, and analysis of intervention studies (clinical trials).

In other areas of medicine (for example, cardiovascular disease), the application of epidemiology to the evaluation of strategies for primary prevention and screening is of much interest. Such activities are, unfortunately, of little relevance in rheumatology, given current knowledge, and will not be considered further. This review will concentrate on the application of epidemiology to understanding the occurrence and causation of disease. Examples are taken, where appropriate, from some of the principal disorders to illustrate the particular points to be made. Detailed epidemiological information of relevance to specific disorders will also be found in the appropriate chapters in this volume.

Criteria for diagnosis

There is no single diagnostic test for most of the major rheumatological disorders. Further, there is considerable overlap, both clinically and pathologically, between the various conditions. As a consequence, sets of criteria covering the main clinical and other facets of each disorder have been formulated with the aim of separating 'cases' from 'non-cases'.

Purpose of criteria

Criteria are useful in so far as they allow like to be compared with like, such that results of clinical and other studies may be compared directly. The American College of Rheumatology has taken an important lead in formulating criteria for most of the principal disorders. It prefers the term 'classification' to 'diagnostic' when applied to the criteria generated for specific disorders, to take account of the atypical case.

Criteria are developed for a number of purposes including population studies of occurrence, aetiological studies, and clinical trials (Fries *et al*. 1994). In trials, there is a need for considerable stringency, i.e. only patients with clearly established disease are eligible for study to prevent dilution of an effect. In that case, the aim of the criteria is maximal specificity (no false positives). By contrast, if the aim is the investigation of familial clustering one would not wish to miss even mild disease in the first-degree relatives of affected probands. Criteria in this situation should opt for maximal sensitivity (few false negatives). In practice, published criteria take no account of the different applications and, as a consequence, they are not always appropriate to the desired task.

Derivation of criteria

Historically, there are three approaches to deriving criteria (Table 1). These are (i) expert consensus, (ii) quasianalytical, and (iii) analytical. The first approach results from the deliberations of a group of 'experts' who reach an agreement between themselves as to what constitutes a case. The second starts as a consensus, but the criteria developed are thereafter tested in patients with and without the disease (determined in an independent manner), and then subjected to subsequent refinement. The third approach is to collect data on the major, potentially diagnostic features and, using a number of statistical approaches (Bloch *et al*. 1990), derive the most discriminatory criteria. The main advantages and disadvantages of these approaches are summarized in Table 1. Diseases for which criteria have been published are listed in Table 2, with their methods of derivation.

Approach	Advantages	Disadvantages
Consensus	Criteria based on clinical good sense and likely to be acceptable	May not maximize discrimination between cases and non-cases
Quasi-analytical	As above but discriminatory power stated	Assumes that appropriate individual items were included in criteria sets; may not maximize discrimination
Analytical	Maximal discrimination between cases and non-cases; excludes bias from preconceived notions	May not reflect clinical sense and thus have difficulty in acceptance

Table 1 Approaches to deriving criteria for rheumatic diseases

Disease	Reference	Approach	Common name
Rheumatoid arthritis	Kellgren (1962)	Consensus	The 'Rome' criteria
	Bennett and Wood (1968)	Consensus	The 'New York' criteria
	Arnett et al. (1988)	Analytical	Revised ARA*
Ankylosing spondylitis	Bennett and Wood (1968)	Consensus	'New York' revision of 'Rome' criteria
	van der Linden et al. (1984)	Quasi-analytical	'Revised New York'
Systemic lupus erythematosus	Cohen et al. (1971)	Consensus	'Preliminary ARA'
	Tan (1982)	Quasi-analytical	'Revised ARA'
Raynaud's syndrome	Wilkins et al. (1982)	Quasi-analytical	
Systemic sclerosis	Masi et al. (1980)	Quasi-analytical	Revised ARA
Osteoarthritis			
Knee	Altman et al. (1986)	Analytical	ARA
Hand	Altman et al. (1990)	Analytical	ARA
Vasculitis (various types)	Hunder et al. (1990)	Analytical	ARA
Fibromyalgia	Wolfe et al. (1990)	Analytical	ARA

*ARA, American Rheumatism Association

Table 2 Publications of criteria for rheumatic diseases

The constitution both of cases and the comparison group used for deriving criteria is of fundamental importance for their interpretation. This is best illustrated by example. The 1987 American Rheumatology Association (**ARA**) revised criteria for rheumatoid arthritis (Arnett *et al*. 1988) allow the following subset of signs to satisfy the criteria: either positive rheumatoid factor test or typical radiological erosion, plus swelling in three or more groups of appropriate joints of which at least one should be either a wrist or metacarpophalangeal joint. Such a subset might not be specific enough for routine clinical practice. The sensitivity of the criteria as a whole were, however, very high in the case series studied because they were all current attenders at a specialist hospital facility and had had disease for a mean of 7.7 years. The application of the same criteria to patients attending with early disease to a non-specialist could give a much reduced sensitivity (Silman and Symmons 1995).

The preliminary criteria for the diagnosis of scleroderma (Masi *et al* 1980) show an extremely high specificity for the major criterion of proximal skin tightness against the comparison group chosen for study from rheumatological practice. Their performance against patients without scleroderma attending a dermatology practice with swollen or tight skin would be substantially different. The purpose of emphasizing this point is not to devalue the criteria themselves, but to argue for caution in their use in circumstances different to those from which they were derived.

Occurrence of disease

The occurrence of diseases is expressed as the number of cases arising in the population 'at risk'. The latter is multiplied by an appropriate power of 10 to yield a sensible number. Thus, it might be appropriate to express the annual incidence of new episodes of back pain as a percentage, whereas the incidence of scleroderma is more conveniently expressed per million population. Frequently there are reports of 'epidemiological' studies that present the relative proportions of diagnoses seen, for example, in a clinical practice, with the interpretation that the proportions represent the relative occurrence of those diseases in the catchment population of the clinic. Such data relying totally on numerator ascertainment, without consideration of the population base, are subject to considerable error. Thus it is difficult to interpret reports from Chinese populations that systemic lupus erythematosus is more frequent than rheumatoid arthritis, suggesting that either the lupus is more frequent than in the West or that the arthritis is less frequent. There may also be selective differences in attendance at hospital between these two disorders.

Measures of occurrence

A number of measures of occurrence are used in the epidemiology of the rheumatic diseases and these are listed in Table 3. Broadly, prevalence, which describes the proportion of existing cases in a population, is less useful than incidence, which describes the rate of occurrence of new cases. Prevalence is dependent on duration, so the longer a disease persists the greater the likelihood that it will be included in a prevalence estimate.

Measure	Description	Appropriate example
Incidence	Rate of occurrence of new cases arising in population conveniently expressed per unit time interval (e.g. per year)	Incidence of rheumatoid arthritis is 0.3 new cases/1000/year
Episode incidence	Rate of occurrence of new episodes of a disease arising in population; usefully applied to conditions where previous episodes may only be weakly related	New episode incidence of painful shoulder is 10/1000/year
Cumulative incidence	Similar to incidence but time interval expressed as fixed period	Cumulative incidence of juvenile chronic arthritis by age 16 is 1/1000
(Point) prevalence	Describes the occurrence of current cases, i.e. those with evidence of disease; implication is based on a particular point in time (notional prevalence day)	Point prevalence of radiographic osteoarthritis of the hip is 1%/1000 (on 1 January 1991)
Period prevalence	Similar to episode incidence in so far as it expresses the rate of individuals displaying evidence of disease in a fixed period (e.g. 1 year) but they may not be continually affected during this period; unlike episode incidence does not require development of new episode during that period	One-year period prevalence of low back pain is 50/1000
Cumulative prevalence	Summation of disease occurrence during fixed period; similar to cumulative incidence but would exclude those who had died from disease prior to investigation	Similar to cumulative incidence

Table 3 Measures of disease occurrence

There are considerable problems in assessing prevalence in rheumatic diseases, because they frequently go into remission without any residual clinical sign of disease. Thus a population prevalence survey can only detect either those with currently active disease (e.g. inflamed joints) or those with evidence of past disease (e.g. joint deformity). Those individuals whose disease resolved without damage will remain undetected. To overcome this investigators may try to assess the existence of past disease, for example by taking a detailed history or reviewing available medical records. This is acceptable, but if such an approach is comprehensive, then it is not point prevalence that is being ascertained (see Table 3) but cumulative prevalence (MacGregor and Silman 1992). Further, the problem with such an approach is that cases who died before being ascertained would obviously be missed. Thus such a cumulative prevalence would underestimate cumulative incidence.

Approaches to estimating disease occurrence in rheumatology

The approach to be used (Table 4) depends both on the frequency of the disease (Safavi *et al*. 1990) and its severity. Population surveys are prohibitively expensive for rare diseases, but necessary for common ones. Thus it would not be appropriate to ascertain back pain by review of hospital attenders as many cases will not seek medical attention. Conversely it would be reasonable to assume that the majority of cases with scleroderma will seek medical care and thus a population survey is not necessary. The complacency of such an approach, however, was displayed by a population survey from South Carolina suggesting that there is a considerable,

unrecognized prevalence in the community of previously unrecognized scleroderma and 'scleroderma spectrum' disorders (Maricq *et al*. 1989) and that expensive population surveys may be necessary.

Measure	Approach
Incidence	(i) Retrospective review of diagnosed cases from medical facilities
	(ii) Prospective notification or registration scheme
	(iii) Measurement of cases occuring in interval between two population surveys
Episode incidence	(i) Retrospective population survey*
	(ii) Prospective notification or registration scheme
Cumulative incidence	(i) Retrospective review of diagnosed cases
	(ii) Retrospective population survey*
Point prevalence	(i) Cross-sectional population survey*
	(ii) Estimate based on current diagnosed clinic attenders
Period prevalence	(i) Retrospective* population survey
	(ii) Prospective population survey

*A retrospective population is based on recall for past events whereas a cross-sectional survey aims to investigate current disease status. In practice, both methods can be adopted concurrently, e.g. 'Have you ever had or do you now have painful joints?'

Table 4 Approaches to measuring disease occurrence

It is virtually impossible to derive incidence data from cross-sectional population surveys, as a single survey will miss cases that went into remission or died before the survey. Lawrence (1977) in his classical work in Leigh and Wensleydale in the North of England, undertook a second survey five years after the first and calculated the rate of development of new cases from those that developed the disease in the interval. Again, this approach would miss those that developed the disease and either died or went into remission between the two surveys.

The only realistic options for assessing incidence (see Table 4) are, first, some form of prospective notification system whereby newly developing cases are continuously and prospectively notified to a central source in a similar fashion to cancer registration. Such a scheme is logistically difficult and expensive, although a recent example of its use in rheumatoid arthritis is encouraging (Symmons *et al*. 1994). Alternatively, all relevant clinical facilities used by the target population may be reviewed to determine the rate of newly diagnosed cases. The disadvantages of this second approach are, first, its retrospective element relying on sufficient information being recorded in the case records to allow a subsequent diagnostic opinion. Second, such an approach requires an excellent information system with few mistakes in diagnostic coding, entry, or retrieval. Such systems are infrequent, with the Rochester Epidemiology Program being a notable exception.

Rochester Epidemiology Program

The Mayo Clinic in Rochester, Minnesota, together with the Olmsted Medical Practice, provides apparently the only source of medical care to the local population, which was approximately 60 000 in 1980. The record system at the Mayo Clinic is almost unique in being able to extract diagnostic data from the records over at least the past 50 years and as a consequence any diagnosis made on a resident of the County can be retrieved. These data have been used to generate estimates of the incidence of many rheumatic diseases such as rheumatoid arthritis (Linos *et al*. 1980), systemic lupus erythematosus (Michet *et al*. 1985), ankylosing spondylitis (Carbone *et al*. 1992) and temporal arteritis (Chuang *et al*. 1982). There are problems, however, in interpreting the results. First, the population is really too small for reliably assessing the incidence of some rare disease. Thus there were no male cases of scleroderma recorded in their incidence series (Michet *et al*. 1985), although males with this disorder do exist. Second, the Olmsted population is unusual both in having heavy over-representation of the professional middle classes and also (compared with the rest of the United States) having a relative excess of those of Scandinavian origin.

In most other parts of the world, it is impossible to rely on one clinical facility to ascertain all cases and an obvious alternative approach is to review as many different sources of data as possible. The greater the degree of overlap between the sources, the greater the confidence that ascertainment is complete. Such an approach has been applied to both scleroderma (Silman *et al*. 1988) and systemic lupus erythematosus (Jonsson *et al*. 1990), both studies using many different sources of data. Statistical methods are available which use the amount of overlap to assess the likely range of the true occurrence (McCarty *et al*. 1992).

Overview of the occurrence of the major rheumatic diseases

Table 5 summarizes the typical available estimates of the occurrence of the rheumatic disorders for Western populations. For some diseases, such as rheumatoid arthritis, there have been many studies and the aim has been to produce a 'ball park' figure.

Disease	Annual incidence per 1000	Point prevalence per 1000
Rheumatoid arthritis	0.5	8.0
Psoriatic arthritis	Not available	0.2
Systemic lupus erythematosus	0.05	0.4
Systemic sclerosis	0.01	0.1
Ankylosing spondylitis	0.07	2.0
Knee osteoarthritis	Not applicable	100*
Juvenile chronic arthritis†	0.1	0.7
Gout	1.0	Not applicable

Rates are based on consensus of most recently available data.
*Rates variable and age dependent; this prevalence in age-range 35–74 years.
†Rates refer to children under 15 years of age.

Table 5 Occurrence of major rheumatic diseases: Western populations aged 15+ years

There are a number of points of note. First, in terms of incidence, the inflammatory arthropathies are very rare disorders. Thus less than one new case of rheumatoid arthritis occurs in an adult population of 1000 every two years. Even the prevalence is relatively low at approximately 0.6 to 0.8 per 100, meaning that population surveys of less than 2000 are unlikely to give robust estimates of occurrence. There are, however, considerable differences in the reported prevalence and incidence of many of the rheumatic diseases and it is tempting to ascribe the differences to underlying differences in the populations studied. As an example, the range of incidence rates observed is shown for giant cell arteritis (temporal arteritis) in Table 6. It seems unlikely that these differences represent true biological variation in geographical or temporal occurrence. It is more likely that there is an artefactual explanation for such differences and some of these are:

Country	Incidence rate per 100 000 (both sexes, age 50+ years)
Israel	0.5
Scotland	4.2
France	9.4
USA	11.7
USA	17.0
Sweden	28.6
England	40.0
USA	54.0
USA	70.0
Denmark	76.6

Source: Silman and Hochberg (1993).

Table 6 Variation in reported incidence of polymyalgia rheumatica/giant cell arteritis

1. small sample size leading to random error;
2. differences in age and sex structure of populations surveyed;
3. high non-response rate with selective differences in occurrence between responders and non-responders;
4. differences in completeness of case ascertainment;
5. differences in case definition;
6. observer variation in use of case definition.

Small-denominator population size increases the risk of denominator error. Differences in the ascertainment procedure for cases are a principal source of variation. Even when the same apparent source is used (for example, referrals to a specialist rheumatologist), there is likely to be large variation between populations in the threshold for referral. Sources of this variation include the availability of health care services and physician and population perception of disease severity. Lack of standardization in case definition is a major problem. Thus, in the data in Table 6 some studies restricted inclusion to those with a positive biopsy, others relied on a broader clinical diagnosis.

Generally in rheumatology, it is perhaps easy to declare that cases satisfied 'the ARA Criteria' for the relevant disease, but there may be considerable differences in observer interpretation of the various individual constituents of the particular criteria used. The timing of a report is of interest. The effect of time trends on disease is considered below, but differences in diagnostic fashion will obviously have an effect on the reported occurrence rates in different years. Thus reported prevalence rates of up to 20 per cent for rheumatoid arthritis in a survey undertaken 25 years ago (Engel 1968) are more likely to represent a perception of what was then rheumatoid arthritis and what now would be considered unacceptable.

Factors affecting occurrence of the major rheumatic diseases

Age and sex

It is impossible to compare the results from different studies without considering the age and sex structure of the population studied. Differences in crude (i.e. all age and sex combined) estimates of occurrence rates between populations may well be explained by differences in demographic structure. There are very striking effects of age and sex in virtually all the rheumatic diseases, as shown in Table 7. Extreme examples are the very low incidence of ankylosing spondylitis at age 65 and the virtual absence of knee osteoarthritis at age 25. In all disorders considered, a 5-year shift in the mean age or a 10 per cent difference in the proportions of the two sexes in the population studied can have a large effect on the crude estimates. One problem in surveys is that males and those at the two extremes of the age distribution are less likely to participate, resulting in a 'skewed' estimate of occurrence.

Disease	Ratio ages 65:25 years	Sex (F/M)
Rheumatoid arthritis	6:1	2.5:1
Ankylosing spondylitis	0[a]	1:3
Gout	2:1	1:6
Systemic sclerosis	3:1	4:1
Systemic lupus erythematosus	1.5:1	3:1→9:1[b]
Juvenile chronic arthritis (pauciarticular)	N/A	2:1→7:1[b]
Knee osteoarthritis (prevalence)	0[a]	2:1

Data are typical
[a]No incident cases at this age
[b]Range given as ratio varies considerably in published studies
Source: Silman and Hochberg (1993).

Table 7 Effect of age and sex on incidence of rheumatic disease

There are two approaches to overcoming this problem. First, the data can be presented separately for each age and sex group (e.g. Table 8), although this is frequently not possible or desirable as the numbers in individual age and sex 'strata' may be too small for precise estimates. The data presented on the prevalence of knee osteoarthritis and on the incidence of rheumatoid arthritis are unusual by virtue of having large sample sizes for the disorders considered. Frequently in publications this is not the case. Alternatively, adjustment may be made to some notional standard population to produce a standardized rate. For example, in the Symmons *et al*. study of rheumatoid arthritis incidence (Symmons *et al*. 1994), male and female crude rates of 14.0 and 35.6 per 100 000 were standardized to the United Kingdom 1991 population to produce standardized rates of 12.7 and 34.3 per 100 000 respectively. The only problem is that other reports do not standardize to the same population and readers rarely make the necessary calculations themselves to obtain comparable results.

Age (years)	Prevalence			
	Males		Females	
	Number studied	%	Number studied	%
40–44	395	0.0	428	0.0
45–49	362	0.3	386	1.6
50–54	312	2.2	298	2.7
55–59	220	1.4	229	0.4
60–64	178	5.6	186	4.8
65–69	116	6.9	183	9.8
70–74	85	7.1	122	16.4
75–79	59	8.5	117	14.5
80+	27	7.4	77	29.9

Source: Van Saase et al. (1988)

Table 8 Age- and sex-specific prevalence rates for grade 3 osteoarthritis of the right knee

With the exception of the HLA B27-associated spondylarthropathies and gout, there is a striking female excess (even after age adjustment) in most of the rheumatic diseases for reasons that are mainly unexplained. Such an excess is always a useful starting point for considering aetiological hypotheses and the effect of hormonal and reproductive factors (e.g. Silman and Black 1988; Brennan and Silman 1994). In disorders with an increased occurrence with age then a degenerative aetiology, perhaps as a consequence of a sustained environmental insult, is likely (e.g. osteoarthritis in the presence of obesity). By contrast, disease with a young age at onset may represent a greater genetic contribution or more short-lived environmental influence (e.g. Reiter's syndrome).

Time trends

It is of considerable interest to monitor the trends over time in the occurrence of diseases. From an aetiological view changes in the incidence of disease may represent changing levels of exposure to putative risk factors. From a public health perspective the appropriate health service provision requires knowledge of both current and future levels of occurrence. Trends in prevalence rates are difficult to interpret because improvement in survival will lead to an increase in prevalence. Thus it is really only appropriate to consider changes in incidence over time, and regrettably, there are few data on incidence at different time periods for most of the major rheumatic diseases, the notable exception being rheumatoid arthritis.

Time trends for rheumatoid arthritis

The earliest reliable data on trends in the incidence of rheumatoid arthritis come from estimates based on cross-sectional surveys of population samples from Leigh and Wensleydale, studied in the 1950s and 1960s. A follow-up survey was undertaken after 5 years in the 620 individuals who had participated in the initial survey in 1954 to 1959. This found that 36 (6 per cent) of the population originally free of rheumatoid arthritis had developed it during the follow-up period (Lawrence 1977). This however, is equivalent to an annual incidence of 12 per 1000, which was six times the rate estimated during the first survey (based on recalled age at onset). It is not likely that the incidence increased sixfold during this period and the data probably seriously overestimate the true occurrence of the disease and probably reflect a case definition of low specificity.

The accurate determination of trends requires the continual monitoring of a population, with access to contemporary medical records; these permitting retrospectively correct diagnostic assignment. Such a system can only work if the monitoring can detect all the cases in a defined population as in the Rochester Epidemiology Program discussed above. Although retrospective examination of medical records will omit those who never seek medical attention for their symptoms and standardization of diagnosis is difficult, the utility of such a system in documenting trends in new cases is clear. Figure 1 shows published results from 1950 to 1975 in the Rochester studies (Linos *et al*. 1980). These show an increasing incidence in the first part of this period, with a marked subsequent decline in females but not in males. Although there have been considerable changes in diagnostic practice over this period, with increasing recognition of the need to separate out the HLA B27-related seronegative arthritides, such an explanation is unlikely, given the patterns observed. Such a diagnostic reassignment would have been expected to alter the trends in both sexes equally. Possible reasons for the difference between the sexes will be considered later in this review. A more recent study from Seattle (Dugowsen *et al*. 1991) aiming to 'capture' all incident cases in women in 1987 in a fixed population, yielded an annual incidence rate of 0.23 per 1000 women aged between 18 and 64 years compared with a rate of 0.46 per 1000 in the same age group in the Rochester population during the period 1950 to 1974, i.e. the rate had halved over the past decade. This decline in the United States is not restricted to white groups. Over the past 25 years there has been a halving in incidence in Pima Indians, a group with one of the world's highest rates (Jacobsson *et al*. 1994).

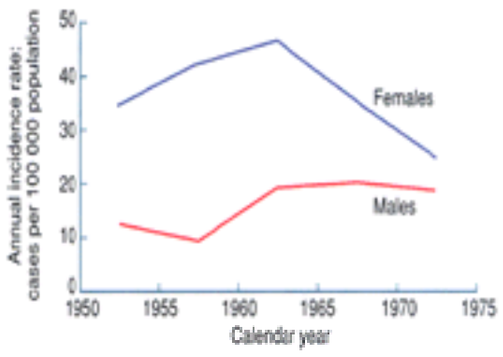


Fig. 1 Published results from 1950 to 1975 in the Rochester studies (reproduced with permission from Linos *et al*. 1980).

There are also supporting data from the United Kingdom. The Royal College of General Practitioners conducted morbidity surveys involving a large group of general practitioners who were required to make a diagnosis for every patient consultation. These surveys in 1955, 1970 to 1972, 1981 to 1982, and 1992 all relied on the recorded diagnosis from the general practitioner and were not standardized. The results comparing the two middle surveys show (Hochberg 1990) (Table 9) that there was no significant decline in incidence between these surveys. The other data from this source are from an ongoing survey of all patients attending one of 30 participating general practitioners, who send back to the 'central' unit a 'weekly return' of the numbers attending with a list of specific diagnoses (Royal College of General Practitioners 1976 to 1987). These include rheumatoid arthritis. Analysis of these data (Fig. 2) (Silman 1988) showed a statistically significant decline of approximately 7.5 per 100 000 per year between 1976 and 1987, equivalent to a halving in the incidence rate during this time. Although most of the participating doctors had remained the same during the period of observation the possibility cannot be excluded that there was a decline in completeness of recording during this period.

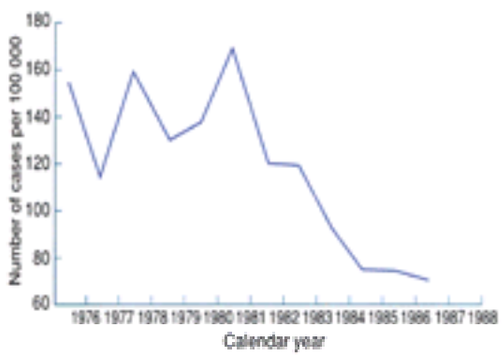


Fig. 2 Time trends in incidence of rheumatoid arthritis (United Kingdom general practice 1967 to 1987). (Reproduced with permission from Silman 1988.)

	1970–1972		1981–1982	
	Males	Females	Males	Females
Incidence rate per thousand (age adjusted	1.3	3.2	1.2	2.6

Source: Hochberg (1990)

Table 9 Incidence rate of rheumatoid arthritis in the United Kingdom

Whole population data are also available from those countries that have a population morbidity register for specific disorders. Interpretation of trends from such sources is difficult as, in addition to the persistent problem of changes in completeness of registration, there may be selective changes in the severity of cases recorded. Such registers are extensively available in Scandinavia, and data from Finland (Isomaki 1989) have recently been published. This source demonstrated an annual incidence of newly registered patients with seropositive rheumatoid arthritis of 0.46 per 1000 in 1980, which was unchanged from that recorded in the early 1970s.

Time trends for other rheumatic diseases

There are fewer data for the other diseases. There are suggestions that symptomatic osteoarthritis is increasing in incidence, based on an increase in consultations to general practitioners from the Royal College of General Practitioners data (Croft 1990) (Fig. 3). Similar data have shown a 30 per cent increase in the incidence of

gout between 1971 and 1981 (Stewart and Silman 1990). There has been a marked increase in the incidence of scleroderma, based on a number of studies covering 35 years of observation of the population of Allegheny County, Pennsylvania (Fig. 4) (reviewed by Williams and Silman 1991). Giant cell arteritis has also doubled in incidence in some countries in the past two decades (Rajala *et al*. 1993). However, there is always a problem with such diseases in that increasing incidence may only represent better case ascertainment as physicians become more confident at making the diagnosis. By contrast, there has been a decline in the incidence of ankylosing spondylitis over six decades in the Olmsted County population (Carbone *et al*. 1992).

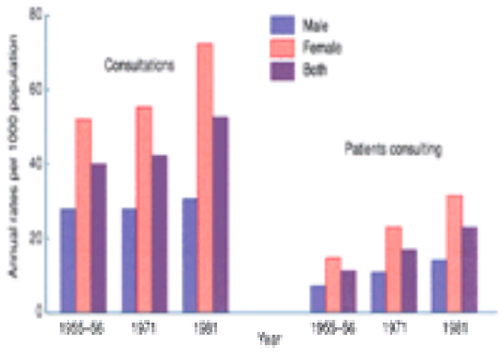


Fig. 3 Incidence of consultations for, and patients consulting with, osteoarthritis and allied conditions in United Kingdom general practice (reproduced with permission from Croft 1990).

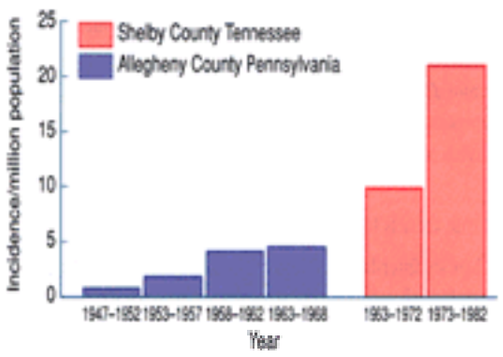


Fig. 4 Time trends in the incidence of systemic sclerosis between 1947 and 1982 in the United States.

Trends in birth cohorts

The onset of many of the rheumatic diseases is difficult to define in time and thus trends in calendar year of presentation or diagnosis may not reflect true temporal patterns of disease. Further, if the presumed environmental exposure which triggers a particular disorder occurs early in life, it might be more appropriate to examine trends in disease in successive generations defined by their cohort of birth. An illustrative example of this in rheumatoid arthritis is that contemporary studies of its incidence demonstrate a marked increase in risk with increasing age (Symmons *et al*. 1994). Although this might represent a true age effect on incidence, an alternative explanation is that the oldest groups in that population were at highest risk based on their year of birth and that follow-up over a long period would confirm that the age-specific incidence rates would fall.

Lawrence (1977) was the first to point out the possibility that the risk of rheumatoid arthritis could be related to the period in which an individual was born. In a prospective study in Oberhörlen (West Germany) in the 1960s the maximal incidence was found to be in those aged 65 years and over, whereas the maximal prevalence of existing cases in the cross-sectional survey in the same population was in the decade below that age.

More interesting perhaps, are the data from the Leigh and Wensleydale population surveys (mentioned above) on the prevalence of rheumatoid factor positivity in relation to period of birth (Lawrence 1977). In brief, there were three observations: (i) in urban populations, positivity increases with age in cross-sectional studies; (ii) over a 10-year follow-up period in that urban population there was a tendency for individuals to have a fall in titre; (iii) on analysis by year of birth, the rate of positivity fell in successive cohorts from those born between 1885 and 1894; indicating perhaps that the 1885 to 1894 groups had a peculiarly high risk. Lawrence suggested that the reduction in titre in the older age groups was consistent with an effect of improvement in atmospheric pollution resulting from the Clean Air Acts in 1956.

Clustering in time

A further approach to considering the influence of time on the onset of disease is to look for clustering of disease, i.e. the non-random distribution in time. The most classical example of this was the initial description of the apparent epidemic of juvenile arthritis in Lyme (Steere *et al*. 1977), where both the parents' and the physicians' feelings were that the number of cases seen was greater than would be expected by normal random distribution of sporadic cases. Although there are complex statistical methods available for confirming the non-randomness, it is normally obvious that such a cluster exists, as was the case in Lyme. The ultimate consequence from the follow-up of this cluster was the incrimination of *Borrelia burgdorferi* as the causative organism.

The reporting of apparent clusters of cases in time has occurred in other rheumatic disorders with no well-defined cause. One interesting example is polymyalgia rheumatica/giant cell arteritis, although the role of increased awareness and changes in diagnostic sensitivity in producing such clusters is unknown. A cluster of five cases in Jerusalem was reported in a 7-week period, where their expected annual referral rate was one case. A study in general practice in the United Kingdom was also suggestive of a cluster with six cases of polymyalgia rheumatica and two of giant cell arteritis presenting in 7 months in a population of 4400. Further, one of the most unusual aspects of the epidemiology of polymyalgia rheumatica/giant cell arteritis is the suggestion of seasonality in incidence. The results are confusing, however. Thus, an initial report of clustering in summer (Kinmont and McCullum 1965) was followed by two of a definite winter peak (Coomes *et al*. 1976; Jonasson *et al*. 1979), one with a marked summer peak (Cimmino *et al*. 1990), and one suggesting both a summer and a winter peak (Mowat and Hazleman. 1974). Two more studies showed no such seasonality (Chuang *et al*. 1982; Omland *et al*. 1986) but the results from the positive studies do suggest that clustering in time and indeed in space might be responsible for some of the variable incidence of this disease.

Racial influences

A summary of the principal racial influences in the rheumatic diseases is given in Table 10. The consideration of racial influences is of interest, although any differences that do emerge are inevitably difficult to interpret. The reasons for this are first that there may be racial differences in symptom perception, physician consultation, and physician bias in diagnosis, all of which could lead to apparent differences in incidence. Second, ethnic groups differ in both genes and environment, the latter both at a macro level—there may be geographical differences due to area of residence—and at a micro level, perhaps due to differences in nutrition or other lifestyle factors. The classical epidemiological method of attempting to distinguish between these explanations is the migrant study. Thus, one aims to compare the incidence of a disease in members of an ethnic group who migrate with that seen in the population who remain in their original environment. One example of this is the study of rheumatoid arthritis in black African populations. Thus a very low prevalence of rheumatoid arthritis was seen in a rural African population in South Africa, but the prevalence approximated to the white persons' rate among those black people who had migrated to an urban environment (Solomon *et al*. 1975). The suggestion is that it is the rural environment rather than the negro race that is protective. However, there has to be caution in interpreting the results of such studies. First the migrants are unlikely to be representative of the 'parent' population in respect of their lifestyle and other factors. Second, the occurrence of disease may lead to migration, thus an individual with rheumatoid arthritis may seek to move nearer to a city for greater access to medical care. One

problem in considering comparative studies of racial differences is that one may not be comparing like with like, and different study methods may be involved.

	Racial groups with rate:	
	Increased*	Decreased
Rheumatoid arthritis	Pima Indians	Black Africans Chinese
Ankylosing spondylitis	Haida Indians	
Systemic lupus erythematosus	American blacks Chinese	
Osteoarthritis	Blackfeet Indians	Caribbean blacks
Juvenile chronic arthritis	Native Americans	Chinese
Gout	Filipinos Tamils Malaysians	Pima Indians

*Compared with 'European' group.
Data from numerous sources; groups mentioned are those with the most extreme results.
Source: Silman and Hochberg (1993).

Table 10 Racial differences in the occurrence of rheumatic diseases

Rheumatoid arthritis is rarer in both developed (Lau *et al*. 1993) and rural (Beasley *et al*. 1983) Chinese groups and in both developed (MacGregor *et al*. 1994a) and rural (Silman *et al*. 1993) black groups than in white populations. Examples of racial differences in a few other selected diseases are considered below.

Osteoarthritis

A single-observer study comparing rural Jamaica with rural England showed a similar overall prevalence of osteoarthritis, although there was a difference in the distribution of affected joints. Jamaicans were more likely to have knee and hand involvement and less likely to have hip and metatarsophalangeal joint involvement (Bremner *et al*. 1968). By contrast, involvement of the distal interphalangeal joint is very common in Pima and Blackfeet Indian groups in North America, with rates of Heberden's nodes as high as 30 per cent in some Blackfeet communities (Bennett and Burch 1968). In North American black populations, radiographic changes in the knee are more common in women but similar in men when compared with a similarly surveyed white population (Felson 1988). Differences in the prevalence of osteoarthritis of the hip are more common, with a relative rarity of disease at this site in Asian Indians, Chinese, and black African populations (Felson 1988).

Juvenile chronic arthritis

The incidence of juvenile chronic arthritis in native Americans in British Columbia, Canada is three times that of a comparable white population (Hochberg 1981). By contrast no cases were found in an extensive survey of 20 000 Chinese children in the same population (Hill 1977). There do not appear to be any differences between black and white children in North America.

Systemic lupus erythematosus

This is up to five times more common in black than in white races, although this excess is more predominant in the United States and West Indian populations than in black African groups (Symmons 1991). South Asians are also at double the risk of white populations (Hopkinson *et al*. 1994). The other at-risk group would appear to be the Chinese, with a number of anecdotal hospital series reporting a relative excess of systemic lupus erythematosus compared with rheumatoid arthritis, in contrast to the normal state in Western Caucasoid populations. There may be environmental factors relevant here however, as the excess is seen in Chinese populations in Malaysia but not in 'Westernized' Chinese for example in San Francisco (Frank 1980). It is likely in this, as in other disorders, that the effect of race is not limited to disease susceptibility but also applies to disease severity.

Geographical influences

The racial distribution of disease can, as was explained above, be interpreted on the basis of geographical differences related to environmental factors. Some of the differences that affect the Third World are considered in Chapter 5.1.2. Geographical, like racial, differences can be useful in framing hypotheses for further testing. Geographical differences may be between countries or even within countries to the level of small, area clusters, in an analogous manner to the clustering in time discussed above for Lyme disease. Clustering without an obvious explanation has been described for connective tissue diseases in general in Georgia, United States (Arnett *et al*. 1990) and for scleroderma in South and West London, where one hypothesis was that this was due to proximity to airports (Silman *et al*. 1990).

Geographical differences between populations for the various diseases have been considered and many of these were well described at a conference nearly 25 years ago (Bennett and Wood 1968), relating particularly to gout and rheumatoid arthritis.

Rheumatoid arthritis

Rheumatoid arthritis has been subject to more population studies than any other rheumatic disease and the breadth of these studies across all five continents is shown in Table 11. The points listed above on methodological explanations for differences in results are relevant in interpreting these studies. Most of the studies are very small and with a disease, such as rheumatoid arthritis, with a less than 1 per cent rate of prevalence, population studies of under 2000 are unlikely to yield useful data. The precision with which the studies were conducted varies, and the published reports do not always allow an accurate appreciation of the methods used, particularly the diagnostic criteria. Despite these caveats, the most striking feature from these numerous studies is the overall similarity in prevalence between populations, with a standard rate of between 0.5 and 1.0 per cent. This consistency is most unusual in human chronic diseases and may reflect the ubiquity of the causative factors, genetic and/or environmental.

Table 11 Geographical variation in prevalence of rheumatoid arthritis

Other rheumatic diseases

Ankylosing spondylitis varies in prevalence according to the underlying population frequency of the HLA-B27 gene. Rates of the disease are substantially higher in Scandinavian than in other European or American Caucasian populations (Rigby 1991). Gout is a disease with a very marked geographical variation in prevalence, probably related to the underlying population levels of serum uric acid. Filipino and Polynesian populations in the Pacific Rim have shown the greatest excess in prevalence compared with that in Caucasoid populations (Decker *et al*. 1968). Interestingly, even within relatively homogeneous populations, there are differences in

prevalence. For example, there is a higher rate of gout in England than in either Wales or Scotland (Currie 1979). Juvenile chronic arthritis has also been extensively studied with no major difference found in Western populations or what was the Soviet Union. There have also been a number of prevalence and incidence studies of polymyalgia rheumatica in different populations, although many of the these studies were from populations of the same ethnic origins, mostly either from Scandinavia or from the largely originally Scandinavian population in Olmsted County served by the Mayo Clinic. There are no consistent differences between the populations studied in Scandinavia or the United States. One study from Israel (Friedman *et al*. 1982) had the lowest recorded incidence (0.5 per 100 000), although it is always difficult to be sure as to the role of under-ascertainment. Similarly studies from Southern Europe—France (Barrier *et al*. 1982) and Italy (Salvarani *et al*. 1987)—yielded incidence estimates substantially lower than those from most of the major studies in Northern European and the United States. The one exception is the study from the Southern United States (Smith *et al*. 1983) which, at 2.4 per 100 000 in the white population, had the lowest incidence rate for giant cell arteritis (apart from the Israeli study).

Aetiological models

Most of the rheumatic diseases have a combination of genetic and environmental factors implicated in their aetiology; it is frequently the role of epidemiology to define the nature of these factors and consider their involvement quantitatively. The problem is not a simple one, however, for two reasons: first, there is not a single aetiological model for the majority of the rheumatic diseases. However, knowledge about genetic factors has increased exponentially as data on gene products is replaced by information from DNA sequencing. Thus it is likely that there is an infinite number of point mutations that explain the occurrence of such wholly genetic disorders as the Ehlers–Danlos syndrome, rather than a single genetic cause.

It is, however, useful to consider models for disease causation according to whether a cause is necessary, i.e. the disease for all practical purposes cannot occur in its absence, and/or sufficient, i.e. the disease can occur in the presence of that risk factor alone, as shown in the diagram below.

		Cause necessary	
		Yes	**No**
Cause sufficient	**Yes**	*Borrelia burgdorferi* and Lyme disease	Organic solvents and scleroderma
	No	HLA B27 and ankylosing spondylitis	HLA DRB1*04 and rheumatoid arthritis

Thus Lyme disease follows infection by *Borrelia burgdorferi* and this organism (probably) alone can cause the disease (Steere *et al*. 1983). By contrast, organic solvent exposure is not necessary for the development of scleroderma but is sufficient for some cases (Black *et al*. 1986). It is (almost) true that HLA B27 is necessary for the development of ankylosing spondylitis but it is not sufficient, as the majority of individuals with this antigen will not develop the disease. The most frequent pattern of causation is, however, the last quadrant (bottom right) where a risk factor is neither necessary nor sufficient. Thus the overwhelming majority of individuals with HLA DR4 (HLA DRB1*04) will not develop rheumatoid arthritis and a significant proportion (30 per cent) of individuals with rheumatoid arthritis will not have HLA DR4 (HLA DRB1*04).

It is obvious to most observers that for most of the rheumatic diseases there is not, as in the Lyme model, a single cause, and one is searching for risk factors that may be relevant but are poor discriminators between those who will and will not develop a particular disease. One useful strategy is to describe clinically or pathologically derived subsets within a clinical entity that may allow the better derivation of causes, the so-called splitting approach. This is particularly useful in relation to juvenile chronic arthritis, where such factors as pattern of joint involvement at onset and the production of antinuclear antibodies have been associated with different immunogenetic backgrounds to an increasingly specific level (Fernandez-Vina *et al*. 1990)

Genetic factors

There are a number of epidemiological approaches to assessing the genetic contribution to disease and these are shown, with their advantages and disadvantages, in Table 12.

Approach	Advantages	Disadvantages
Comparison of occurrence in different ethnic groups	Data may be readily available	Differences between ethnic groups may be more related to environment than genes
Assess whether there is an increased risk in relatives	Families normally willing to participate and provide good data	Environmental factors also cluster within families
		Often difficult to obtain data from correct families
		Need to adjust for age and sex as younger members have not had time to develop disease
Comparison of concordance in identical and non-identical twins	Good matching for environmental factors	Twin pairs not available in rare diseases Identical twins may share closer environment than non-identical (especially if of unlike sex)
Comparison of frequency of genetic marker in probands and controls	Data easy to obtain Accurate definition by DNA techniques of markers may enhance specificity	Such population associations do not prove causation Difficult to obtain strength of genetic contribution or type of inheritance
Studies of linkage between marker and disease in multicase families	Strength of genetic contribution and inheritance are possible to determine	Linkage difficult to show and large number of families needed Results from multicase families may not be applicable to sporadic cases

Table 12 Approaches to studying genetic factors in disease aetiology

Family studies

It is obvious from Table 12 that the study of family members of probands with a disease is an important tool in genetic epidemiology, but there are specific problems in relation to the rheumatic diseases which have been summarized for rheumatoid arthritis (Silman 1986); these are listed below.

1. Categorization of relatives into affected and non-affected is frequently difficult as relatives may show some features of the syndrome under investigation but not sufficient to satisfy the relevant criteria. As an example the relatives of probands with ankylosing spondylitis frequently have either clinical or radiographic evidence of sacroiliitis but not both (Burns and Calin 1983) and to categorize them as affected or unaffected would both be wrong. The answer is to use all the available data and analyse subsets, and thus in family studies not to be bound by rigid criteria.
2. Many of the rheumatic diseases may not become evident until the seventh or subsequent decades and misclassification of those yet to develop the disease may occur. This is different to, for example, the situation with type I diabetes, where development of the disease has normally occurred by the end of the third decade. Some have suggested adjusting for the age of the proband as a suitable method, i.e. if a sibling has not developed the disease by the age at which his or her affected proband developed the disease then he or she can be categorized as non-affected. This is erroneous, as data (for example, in rheumatoid arthritis) have shown that there are considerable differences in the age of onset between affected sibling pairs (Silman *et al*. 1987) and even between concordant monozygotic twins (MacGregor *et al*. 1994b).
3. By contrast for some of the inflammatory joint diseases there is also the problem of misclassification, because the disease has undergone complete remission by the time of the family survey and there is no evidence of past disease either clinically or radiologically. Unless contemporary notes are of high quality, it may be impossible to determine the significance of self-reported past history.
4. It is well recognized that probands derived from hospital series are more likely to have a 'positive' family history than unselected probands chosen from a community survey. This is likely to be due to selection for referral to hospital of an individual with symptoms in the face of a family history, thus explaining the greater familial clustering in hospital series (for example, in rheumatoid arthritis; Wolfe *et al*. 1988). An alternative, but less likely explanation is that hospital cases are more severe, and that the 'genetic contribution' is considered to be one of severity as opposed to susceptibility.

Genetic contribution to specific diseases

In this review it is impossible to provide the detailed data of the genetic contribution to the various rheumatic disorders. A summary of the data from some of the major disorders is shown in Table 13, from which it can be seen that the various strategies outlined above have yielded very different estimates of the role of genetic factors for the different disorders. Recent advances in molecular biology are refining the genetic factors to be considered. An excellent example of this is the 'shared epitope hypothesis' (Gregerson *et al*. 1987), which provides an explanation for the population association between rheumatoid arthritis and different, serologically defined class-II antigens at the DR locus on the basis of the shared possession of a short sequence of amino acids in the third, hypervariable region of the DRB1 gene. Unfortunately for the epidemiologist, although this shared epitope is observed in some 90 per cent of rheumatoid arthritis, it is also found in 50 per cent of the normal population. A quick calculation of the relevant 2 × 2 table gives the results that positive predictive value for the epitope is less than 2 per cent and the specificity is around 50 per cent; in other words, it is a poor discriminator for those destined to develop rheumatoid arthritis.

Disease	Familial clustering (risk× of risk in first-degree relatives to expected population rate)	Monozygotic twin concordance (%)	HLA association (relative risk)
Rheumatoid arthritis	1.7×	12 (population)	DR4 (3×)
Generalized osteoarthritis	2×	43	Inconsistent
Juvenile chronic arthritis	3–4×	40	DR5* (2.3–4.8*)
			DRw8* (10–25×)
Ankylosing spondylitis	15–25×	50	B27 (100×)
Systemic lupus erythematosus	9×	>25	DR2/3 (2–3×)
Systemic sclerosis	6–8×	No data	DR5 (2.5)

Data are based on pooled estimates from many sources. Source: Silman and Hochberg (1993).

Table 13 Genetic features of the major rheumatic diseases

Non-genetic host factors

This term is a slight misnomer but it is a convenient label for the non-environmental factors present in individuals that are not predominantly genetic in origin. A list of such factors examined in the rheumatic diseases is:

1. perinatal problems;
2. family size, position in family;
3. menstrual/menopausal status;
4. sex hormonal status;
5. stature and body form;
6. reproductive history;
7. organic comorbidity;
8. psychiatric comorbidity;
9. biochemical background;
10. height and weight.

Data in relation to some of the factors in the list are considered next.

Family size/position in family

In many instances this would be considered as a surrogate for overcrowding or a similar socioeconomic variable. For instance, in a study of rheumatoid arthritis, there was an increased risk of disease in older than younger siblings, suggesting that being first born carries an increased risk (Hazes *et al*. 1990b). This would be unlikely to have a simple socio-economic explanation. By contrast, a study of Behçet's disease suggested that there was a risk connected with large family size (Cooper *et al*. 1986).

Menstrual, hormonal and reproductive factors

It is appropriate to combine consideration of these variables as suggesting an influence of sex hormones. One of the most interesting aspects of the epidemiology of the major rheumatic diseases (see above) is the marked sex difference in occurrence, which is not easily explained by differences in lifestyle (such as occupation, smoking, or alcohol consumption), as is the case for other chronic disorders. It is thus reasonable to consider whether there are hormonal or reproductive effects that might explain these differences and there has been much recent work on the subject. A summary of some of the findings is shown in Table 14. Many of those factors listed might of course represent the sub- or preclinical effects of the disease rather than the cause, but such factors might be useful in explaining sex differences in occurrence.

Disease	Menstrual, hormonal, and reproductive risk factors
Rheumatoid arthritis	Early menarche, nulliparity, early post-partum period, possibly fetal loss in some groups, low testosterone in males
Osteoarthritis	Menopause, previous hysterectomy for menstrual irregularity, possible oestrogen excess
Systemic sclerosis	Infertility, possible fetal loss
Osteoporosis	Early menopause
Systemic lupus erythematosus	Pregnancy

Table 14 Menstrual, hormonal, and reproductive factors in various rheumatic diseases

Height and weight

There are a number of disorders where anthropometric variables are thought to be of relevance. Tallness may be a risk factor for low back pain. The greatest topic of interest has been the relationship between obesity and, particularly, non-inflammatory joint disease. Although obesity might also be seen as an environmental factor (i.e. excess calorie intake). The mechanical effect of obesity is to increase load on weight-bearing joints, such as the hip and knee, and it is thus of interest to note that obesity is more strongly linked to osteoarthritis of the knee than of the hip (Felson and Radin 1994). Methodologically, it is important to distinguish cause from effect. Thus in the presence of osteoarthritis there may be less physical activity with a greater tendency to obesity. Recent prospective studies have shown conclusively that obesity is a real risk factor not only for osteoarthritis of the knees (Hochberg *et al*. 1995), but also of the hands (Carman *et al*. 1994), suggesting perhaps a 'metabolic' effect.

The effect of comorbidity

Many of the diseases discussed in this volume will be classified as primary or secondary depending on whether there is another underlying pathology for the joint disease, e.g. polycythaemia and gout. The fact that this is true for a (normally small) proportion of clinical cases does not mean that such comorbidity is of relevance for the disease in general. Of greater relevance is an increased risk of joint disease in the presence of other illness, although the latter does not 'cause' the former. There are some data to support the increased coincidence of some disorders: for example, rheumatoid arthritis and autoimmune thyroid disease (Silman *et al*. 1989), scleroderma and cancer (Abu-Shakra *et al*. 1993), and neurosis and back pain (Leino and Magni 1993) are amongst a large list of such associations.

There are some interesting negative associations, i.e. the presence of one disorder 'protecting' against the development of another. Examples include the mutual exclusivity of schizophrenia and rheumatoid arthritis (Spector and Silman 1987), and osteoporosis and osteoarthritis (Dequeker 1986). The underlying explanation for these negative associations is, however, somewhat obscure.

Environmental factors

Infectious agents

The identification of *Borrelia burgdorferi* as the spirochaete responsible for Lyme disease, has acted as a stimulus to search for viral causes of other inflammatory joint diseases, such as rheumatoid arthritis for which a viral background seemed likely. It is of interest to review the epidemiological data suggesting a viral cause for rheumatoid arthritis as the lessons are applicable more widely.

Seroepidemiological studies

In this approach, antibody frequencies in a diseased and a disease-free population are compared. Numerous studies have confirmed the high titres of antibodies against a variety of Epstein–Barr virus (**EBV**)-related antigens in patients with rheumatoid arthritis. However, not all studies have been positive and, more importantly, the differences from a control population have not always been either biologically or statistically significant. Given the almost universal exposure to EBV in Western populations, it may be more relevant to look for a quantitative difference in response but this is methodologically more difficult. Epidemiologically, the geographical distribution of infectious mononucleosis and rheumatoid arthritis are similar, and infectious mononucleosis is apparently unknown in countries with a low prevalence of rheumatoid arthritis (Aho and Raunio 1982). One explanation for these observations in the face of the ubiquity of infection with EBV is the effect of age. In countries with a low prevalence of rheumatoid arthritis, infection with EBV is virtually universal by the age of 3 years and the infection is clinically silent; whereas in countries with a high prevalence, infection occurs at a later age and is more likely to be clinically apparent with, for example, infectious mononucleosis. The increased rate of antinuclear antibodies in rheumatoid arthritis is seen, however, in most populations including Mexican and American Indians and Afghans (Vaughan 1979). The evidence for EBV infection as a cause of rheumatoid arthritis is constrained by the lack of clinical evidence that EBV is arthritogenic, unlike, for example, rubella, hepatitis and mumps virus (Depper and Zvaifler 1981).

The other widely studied agent has been human parvovirus (**HPV**), given that arthritis can follow infection with this agent. A reasonable conclusion is that HPV probably has little relevance (Leading Article 1985) for rheumatoid arthritis, despite two reports from the United Kingdom, the first showing that 19 out of 153 patients with early synovitis had evidence of HPV infection (White *et al*. 1985) and the second describing joint problems in 17 patients after an HPV outbreak (Reid *et al*. 1985). Neither study had appropriate control groups and there were no patients with a persistent arthritis.

Retroviruses have also been investigated, owing to the similarity between rheumatoid arthritis and the arthritis produced by lentiviruses in animals, for example caprine arthritis encephalitis. However, attempts in man to show evidence of retrovirus infection in rheumatoid arthritis have been unsuccessful. There are also very good mycoplasma-induced animal models of rheumatoid arthritis and in these chronicity and severity, as in man, are related to genetic factors. Clinical and epidemiological studies in humans have, however, mainly failed to support a mycoplasma source for rheumatoid arthritis.

Lifestyle

In many chronic diseases there is a considerable wealth of epidemiological data linking aspects of lifestyle, such as diet, exercise, and cigarette and alcohol consumption, with the risk of disease. Studies of such variables in the rheumatic diseases have been mostly non-informative, for example, there have been studies supporting (Heliovaara *et al*. 1993) and refuting (Vessey *et al*. 1987) an increased risk of rheumatoid arthritis from cigarette smoking. Alcohol is well established to be associated with gout and possibly also osteoporosis. Diet is an aspect of lifestyle that is consistently considered by patients to be the explanation behind their disease, although scientific evidence is hard to find. Problems in studying diet include the accurate recall of diet at the appropriate time before disease onset (which might be many years in a disease with a long latency), and within-individual variation in dietary intake for many nutrients, which results in misclassification if an inappropriate dietary methodology is used, such as a food frequency interview or a 24-h recall.

Occupation

The investigation of occupational exposure as a cause of specific rheumatic syndromes is potentially of importance for those who are exposed, but of relatively minor importance at the population level. It is useful to distinguish the two epidemiological concepts of attributable risk and population attributable risk. Attributable risk, in this setting, estimates the proportion of an exposed individual's risk that is directly due to the exposure, whereas by contrast, the population attributable risk provides an estimate of the proportion of cases that arise in the population as a whole owing to that exposure. Thus there are well-established links between being a footballer (soccer) and the subsequent development of osteoarthrosis of the knee. The attributable risk is high for footballers, i.e. the occupation explained most of their increased risk, whereas in the population as a whole, being a footballer will explain only a small proportion of cases. The consequence of this fairly obvious point is that it is virtually impossible to identify most occupational exposures by undertaking retrospective case–control studies of random series of diagnosed cases, the approach frequently used for testing aetiological hypotheses. It is necessary to undertake prospective studies of occupationally derived cohorts.

Occupational exposure as far as the rheumatic diseases are concerned normally reflects either chemical or toxic exposure, or the outcome of mechanical trauma to joints and associated structures from the physical demands of the job. In practice, occupation is only rarely of interest in the inflammatory rheumatic disorders, although there have been suggestions that some cases of rheumatoid arthritis have an occupational cause (e.g. Klockars *et al*. 1987). By contrast there have been a number of toxic exposures linked to the development of scleroderma and indeed in this disease, which predominantly affects women, the development in a man might indicate a chemical exposure. Suggested exposures are:

1. silica dust (coal miners, gold miners, stonemasons);
2. organic chemicals:
    a. aromatic hydrocarbons (toluene, benzene, xylene, aromatic mixes—white spirit, dieselene);
    b. aliphatic hydrocarbons:
        i. chlorinated (vinyl chloride)
        ii. non-chlorinated (naphtha-n-hexane);
3. toxic oil;
4. epoxy resins;
5. biogenic amines—metaphenylenediamine;
6. urea–formaldehyde foam insulations;
7. drugs

It remains an unanswered (and possibly unanswerable) question as to whether silicone breast implants lead to this disease (Gabriel *et al*. 1994).

P>Osteoarthritis provides a useful example of mechanical causes, resulting from occupational exposure, that are linked with the disease; a list of such occupations is shown in Table 15. The scientifically difficult task is to try and combine the study of such occupations so as to achieve some more precise measure of joint stress. From this, overall risk of disease in relation to trauma might be determined, rather than the risk associated with specific occupations. Recent attempts at this exercise for osteoarthritis of the knee, have confirmed the increased risk from jobs with heavy demand and those that involve heavy bending, but the absolute increase in risk is small, less than twofold for the highest compared with the lowest stress group.

Exposure	Site
Sports	
Boxers	Hands
Baseball pitchers	Shoulders
Hurdlers	Hip
Runners	?Knees
Footballers (soccer)	?Knees, ?Hips
Footballers (American)	?Knees
Weight lifters	Knees
Other	
Parachutists	Knees
Jack-hammer operators	Wrists, hands
Cotton-mill workers	Hands
Coal miners	Knee
Dockers/shipyard workers	Knee
Farmers	Knee, hip

[Source: Pelletier (1988)]

Table 15 Mechanical occupational exposures linked to osteoarthritis

Drugs

Pharmaceutical agents are a source of chemical exposure that have been linked to a number of rheumatic diseases. Many of these examples reflect an idiosyncratic response to the drug, or a genetically determined response (e.g. slow or fast acetylators in drug-induced systemic lupus) and are not a clear example of toxicity *per se*. Again, in public health terms, drugs are a relatively weak contributor to overall risk for most rheumatic diseases, systemic lupus being a notable exception. The chief epidemiological interest in relation to the aetiologic effects of drugs has been the hypothesis that the oral contraceptive pill reduces the risk of rheumatoid arthritis, to such an extent that there is a decline in incidence as a result (Vandenbroucke 1983). This has probably been the most investigated area in the epidemiology of rheumatic disease and a clear consensus has now emerged (Silman and Vandenbroucke 1989; Spector and Hochberg 1990). This is that the oral contraceptive pill probably either protects against or postpones the development of severe rheumatoid arthritis, although the mechanism for this is obscure. The studies themselves demonstrate the entire repertoire of investigational methods in epidemiology, and the various problems and biases involved in undertaking this area of research. Table 16 is instructive in so far as it illustrates the variety of approaches used to generate answers to the same question, although no single study is perfect on its own. This list is a fitting description of the current state of epidemiological research into the aetiology of the rheumatic diseases. The conclusions are that there is a lot of interest, that the studies are methodologically difficult but can be completed, that conflicting results can be expected, but that ultimately a useful answer for the basic scientists to follow up may emerge!

First author	Study design	Source of cases	Source of controls	Source of data	Odds ratio for ever used
Wingren (1978)*	Prospective cohort*	General practice records	n/a	General practice	0.09
Vandenbroucke et al. (1982)	Case-control	Hospital clinic	Soft tissue	Postal questionnaire	
Linos et al. (1981)*	Case-control	Diagnostic register	Non-arthritic clinic attenders	Medical record	1.1
Allebeck et al. (1984)	Case-control	Inpatients	Community	Postal questionnaire	0.1
del Junco et al. (1985)*	Case-control	Diagnostic register	Non-arthritic clinic attenders	Medical records	1.1
Vandenbroucke et al. (1986)	Case-control	Hospital clinic	Osteoarthritis	Postal questionnaire	0.57
Vessey et al. (1987)	Prospective, cohort*	Postal questionnaire	n/a	Family-planning clinic	1.13
Oertelt and Ambrosi (1987)	Case-control	Hospital clinic	Non-arthritic clinic attenders	Interview	1.29
Koepsell et al. (1989)	Incident case-control	Hospital clinic	Community	Interview	0.27
Spector et al. (1990)	Case-control	Hospital clinic	a) Community b) Osteoarthritis	Postal questionnaire	0.34 0.64
Medalie et al. (1990)	Case-control	Group health clinic	Pharmacy users	Pharmacy records	2.0
Hernández et al. (1994a)*	Incident case-control	Hospital clinic	Osteoarthritis	Interview	0.38
Hernández-Avila et al. (1993)	Prospective cohort*	Postal questionnaire	n/a	Postal questionnaire	1.0
Hayes et al. (1993)*	Case-control	Hospital clinic	Sisters	Postal questionnaire	0.37

*Data are reports from the cases population.
†Data are cohort studies and study population is chosen by their exposure status at baseline. The controls (comparison group) are the non-rheumatoid users.

Table 16 Studies investigating the possible protective effect of oral contraceptives against the development of rheumatoid arthritis

Chapter References

Abu-Shakra, M., Guillemin, F., and Lee, P. (1993). Cancer in systemic sclerosis. *Arthritis and Rheumatism*, **36**, 460–4.

Aho, K. and Raunio, V. (1982). EB-virus and rheumatoid arthritis: new insights into their interrelation. *Medical Biology*, **60**, 49–52.

Allebeck, P., Ahlbom, A., Ljungstrom, K., and Allander, E. (1984). Do oral contraceptives reduce the incidence of rheumatoid arthritis? *Scandinavian Journal of Rheumatology*, **13**, 140–6.

Altman, R. *et al*.. (1986). Development of criteria for the classification and reporting of osteoarthritis. *Arthritis and Rheumatism*, **29**, 1039–49.

Altman, R. *et al*.. (1990). The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hand. Arthritis and Rheumatism, **33**, 1601–10.

Arnett, F.C. *et al*. (1988). The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis and Rheumatism*, **31**, 315–24.

Arnett, F.C. *et al*. (1990). Connective tissue disease in south east Georgia. A community based study of immunogenetic markers and autoantibodies. *Journal of Rheumatology*, **17**, 1029–35.

Barrier, P., Pion, P., and Massari, R. (1982). Epidemiological approach to Horton's disease. *Revue de la Medicine Interne*, **3**, 13–20.

Beasley, R.P., Bennett, P.H., and Chun, L.C. (1983). Low prevalence of rheumatoid arthritis in Chinese. Prevalence survey in a rural community. *Journal of Rheumatology*, **10** (Suppl.1), 11–15.

Bennett, P.H. and Burch, T.A. (1968). Osteoarthrosis in the Blackfeet and Pima Indians. In *Population studies of the rheumatic diseases*, (ed. P.H. Bennett and P.H.N. Wood), pp. 407–12. Excerpta Medica, Amsterdam.

Bennett, P.H. and Wood, P.H.N. (ed.) (1968). *Population studies of the rhuematic diseases*, p. 477. Excerpta Medica, Amsterdam.

Black, C., Pereira, S., McWhirter, A., Welsh, K., and Laurent, R. (1986). Genetic susceptibility to scleroderma-like syndrome in symptomatic and asymptomatic workers exposed to vinyl chloride. *Journal of Rheumatology*, **13**, 1059–62.

Bloch, D.A., Moses, L.E., and Michel, B.A. (1990). Statistical approaches to classification: methods for developing classification and other criteria rules. *Arthritis and Rheumatism*, **33**, 1137–44.

Bremner, J.M., Lawrence, J.S., and Miall, W.E. (1968). Degenerative joint disease in a Jamaican rural population. *Annals of the Rheumatic Diseases*, **27**, 326–32.

Brennan, P. and Silman, A.J. (1994). An investigation of gene-environmental interaction in the etiology of rheumatoid arthritis. *American Journal of Epidemiology*, **140**, 453–60.

Burns, T. and Calin, A. (1983). The hand radiograph as a diagnostic discriminant between seropositive and seronegative rheumatoid arthritis: a controlled study. *Annals of the Rheumatic Diseases*, **42**, 605–12.

Carbone, L.D., Cooper, C., Michet, C.J., Atkinson, E.J., O'Fallon, W.M., and Melton, L.J. (1992). Ankylosing spondylitis in Rochester, Minnesota 1935–1989. Is the epidemiology changing? *Arthritis and Rheumatism*, **35**, 1476–82.

Carman, W.J., Sowers, M.F., Hawthorne, V.M., and Weissfeld, L.A. (1994). Obesity as a risk factor for osteoarthritis of the hand and wrist. A prospective study. *American Journal of Epidemiology*, **139**, 119–29.

Chuang, T.Y., Hunder, G.G., Ilstrup, D.M., and Kurland, L.T. (1982). Polymyalgia rheumatica. *Annals of Internal Medicine*, **97**, 672–80.

Cimmino, M.A., Caporali, R., Montecucco, C.M., Rovida, S., and Baratelli, E. (1990). A seasonal pattern in the onset of polymyalgia rheumatica. *Annals of the Rheumatic Diseases*, **49**, 521–3.

Cohen, A.S. *et al*. (1971). Preliminary criteria for the classification of systemic lupus erythematosus. *Bulletin of the Rheumatic Diseases*, **21**, 643–8.

Coomes, E.N., Ellis, R.M., and Kay, A.G. (1976). A prospective study of 102 patients with the polymyalgia rheumatica syndrome. *Rheumatic Rehabilitation*, **15**, 270–6.

Cooper, C. *et al*. (1986). Is Behçet's disease triggered by childhood infection? *Annals of the Rheumatic Diseases*, **43**, 421–3.

Croft, P. (1990). Review of UK data on the rheumatic diseases—osteoarthritis. *British Journal of Rheumatology*, **29**, 391–5.

Currie, W.J.C. (1979). Prevalence and incidence of gout in Great Britain. *Annals of the Rheumatic Diseases*, **38**, 101–6.

Darwish, M.J. and Armenian, H.K. (1987). A case–control study of rheumatoid arthritis in Lebanon. *International Journal of Epidemiology*, **16**, 420–4.

Decker, J.L., Healey, L.A., and Skeith, M.D. (1968). Ethnic variations in serum uric acid: Filipino hyperuricemia, the result of hereditary and environmental factors. In *Population studies of the rheumatic diseases* (ed. P.H. Bennett and P.H.N. Woods), pp. 336–43. Excerpta Medica, Amsterdam.

del Junco, D.J. *et al*. (1985). Do oral contraceptives prevent rheumatoid arthritis? *Journal of the American Medical Association*, **254**, 1938–41.

Depper, J.M. and Zvaifler, N.J. (1981). Epstein–Barr-Virus—its relationship to the pathogenesis of rheumatoid arthritis. *Arthritis and Rheumatism*, **24**, 755–61.

Dequeker, J. (1986). The relationship between osteoporosis and osteoarthritis. *Clinics in the Rheumatic Diseases*, **11**, 261–96.

Dugowson, E.C. *et al*. (1991). Rheumatoid arthritis in women: incidence rates in group Health Cooperative Seattle, Washington, 1987–1989. *Arthritis and Rheumatism*, **34**, 1502–7.

Engel, A. (1968). Rheumatoid arthritis in US adults 1960–2. In *Population studies of the rheumatic diseases* (ed. P.H. Bennett and P.H.N. Woods), pp 83–9. Excerpta Medica, Amsterdam.

Felson, D.T. (1988). Epidemiology of osteoarthritis. *Epidemiologic Reviews*, **16**, 499–512.

Felson, D.T. and Radin, E.L. (1994). What causes knee osteoarthrosis? Are different compartments susceptibile to different risk factors? *Journal of Rheumatology*, **21**, 181–3.

Fernandez-Vina, M.A., Fink, C.W., and Stastny, P. (1990). DQA1 and DQB1 alleles in patients with juvenile arthritis. *Human Immunology Abstracts*, 62.

Frank, A.O. (1980). Apparent predisposition to systemic lupus erythematosus in Chinese patients in West Malaysia. *Annals of the Rheumatic Diseases*, **39**, 266–9.

Friedman, G., Friedman, B., and Benbassat, J. (1982). Epidemiology of temporal arteritis in Israel. *Israel Journal of Medical Science*, **18**, 241.

Fries, J.F., Hochberg, M.C., Medsger, T.A., Hunder, G.G., and Bombardier, C. (1994). Criteria for rheumatic disease: different types and different functions. *Arthritis and Rheumatism*, **37**, 454–62.

Gabriel, S.E., O'Fallon, W.M., Kurland, L.T., Beard, C.M., Woods, J.E., and Melton, L.J. (1994). Risk of connective-tissue diseases and other disorders after breast implantation. *New England Journal of Medicine*, **330**, 1697–702.

Gregersen, P.K., Silver, J., and Winchester, R.J. (1987). The shared epitope hypothesis. An approach to understanding the molecular genetics of susceptibility to rheumatoid arthritis. *Arthritis and Rheumatism*, **30**, 1205–13.

Hannaford, P.C., Kay, C.R., and Hirsch, S. (1990). Oral contraceptives and rheumatoid arthritis: new data from the RCGP oral contraceptives study. *Annals of the Rheumatic Diseases*, **49**, 744–6.

Hazes, J.M.W. *et al*. (1990*a*). Reduction of the risk of rheumatoid arthritis among women who take oral contraceptives. *Arthritis and Rheumatism*, **33**, 173–9.

Hazes, J.M.W. *et al*. (1990*b*). Influence of oral contraception on the occurrence of rheumatoid arthritis in female sibs. *Scandinavian Journal of Rheumatology*, **19**, 306–10.

Heliovaara, M., Aho, K., Aromaa, A., Knekt, P., and Reunanen, A. (1993). Smoking and the risk of rheumatoid arthritis. *Journal of Rheumatology*, **20**, 1830–35.

Hernandez-Avila, M. *et al*. (1990) Exogenous sex hormones and the risk of rheumatoid arthritis. *Arthritis and Rheumatism*, **33**, 947–53.

Hill, R. (1977). Juvenile arthritis in various racial groups in British Columbia. *Arthritis and Rheumatism*, **20**, 162.

Hochberg, M.C. (1981). Adult and juvenile rheumatoid arthritis. *Epidemiologic Review*, **3**, 33–44.

Hochberg, M.C. (1990). Changes in the incidence and prevalence of rheumatoid arthritis in England and Wales—1970–1982. *Seminars in Arthritis and Rheumatism*, **19**, 294–302.

Hochberg, M.C., Lethbridge-Cejku, M., Scott, W.W., Reichle, R., Plato, C.C., and Tobin, J.D. (1995). The association of body weight, body fatness, and body fat distribution with osteoarthritis of the knee: data from the Baltimore longitudinal study of aging. *Journal of Rheumatology*, **22**, 488–93.

Hopkinson, N., Doherty, M., and Powell, R.J. (1994). Clinical features and race-specific incidence/prevalence rates of systemic lupus erythematosus in a geographically complete cohort of patients. *Annals of the Rheumatic Diseases*, **53**, 675–80.

Hunder, G.G. *et al*. (1990). The American College of Rheumatology 1990 criteria for the classification of giant cell arteritis. *Arthritis and Rheumatism*, **33**, 1122–8.

Isomaki, H.A. (1989). Rheumatoid arthritis as seen from official data registers: experience in Finland. *Scandinavian Journal of Rheumatology*, **79**, 21–4.

Jacobsson, L.T.H. *et al*. (1994). Decreasing incidence and prevalence of rheumatoid arthritis in Pima Indians over a 25 year period. *Arthritis and Rheumatism*, **37**, 1158–65.

Jonasson, F., Cullen, J.F., and Elton, R.A. (1979). Temporal arteritis. *Scottish Medical Journal*, **24**, 111–17.

Jonsson, H. *et al*. (1990). Estimating the incidence of systemic lupus erythematosus in a defined population using multiple source of retrieval. *British Journal of Rheumatology*, **29**, 185–8.

Kellgren, J.H. (1962). Diagnostic criteria for population studies. *Bulletin of the Rheumatic Diseases*, **13**, 291–2.

Kinmont, P.C. and McCallum, D.I. (1965). The aetiology, pathology and course of giant cell arteritis—the possible role of light sensitivity. *British Journal of Dermatology*, **77**, 193.

Klockars, M., Koskela, R.S., Jarvinen, E., Kolari, P.J., and Rossi, A. (1987). Silica exposure and rheumatoid arthritis: a follow up study of granite workers 1940–81. *British Medical Journal*, **294**, 997–1000

Koepsell, T. *et al*. (1989). Preliminary findings from a case–control study of risk of rheumatoid arthritis in relation to oral contraceptive use. *British Journal of Rheumatology*, **28**, 41.

Lau, E., Symmons, D., Bankhead, C., MacGregor, A., Donnan, S., and Silman, A. (1993). Low prevalence of rheumatoid arthritis in the urbanized Chinese of Hong Kong. *British Journal of Rheumatology*, **20**, 1133–7.

Lawrence, J.S. (1977). *Rheumatism in populations.* Heinemann, London.

Leading Article (1985). Arthritis and parvovirus infection. *Lancet*, **i**, 436–8.

Leino, P. and Magni, G. (1993). Depressive and distress symptoms as predictors of low back pain, neck-shoulder pain, and other musculoskeletal morbidity: a 10-year follow-up of metal industry employees. *Pain*, **53**, 89–94.

Linos, A. *et al*. (1980). The epidemiology of rheumatoid arthritis in Rochester Minnesota: a study of incidence, prevalence, and mortality. *American Journal of Epidemiology*, **111**, 86–98.

Linos, A., O'Fallon, W.M., Worthington, J.W., and Kurland, L.T. (1983). Case–control study of rheumatoid arthritis and prior use of oral contraceptive. *Lancet*, **i**, 1299–1300.

MacGregor, A.J. and Silman, A.J. (1992). A reappraisal of the measurement of disease occurrence in rheumatoid arthritis. *British Journal of Rheumatology*, **19**, 1163–5.

MacGregor, A.J., Riste, L.K., Hazes, J.M.W., and Silman, A.J. (1994 *a*). Low prevalene of rheumatoid arthritis in black Caribbeans compared with whites in inner city Manchester. *Annals of the Rheumatic Diseases*, **53**, 293–7.

MacGregor, A.J., Bamber, S., and Silman, A.J. (1994 *b*). A comparison of the performance of different methods of disease classification for rheumatoid arthritis. Results from a nationwide twin study. *Journal of Rheumatology*, **21**, 1420–26.

Maricq, H.R. *et al*. (1989). Prevalence of scleroderma spectrum disorders in the general population of South Carolina. *Arthritis and Rheumatism*, **32**, 998–1006.

Masi, A.T. *et al*. (1980). Preliminary criteria for the classification of systemic sclerosis (scleroderma). *Arthritis and Rheumatism*, **23**, 581–90.

McCarty, D.J., Kwoh, C.K., and LaPorte, R.E. (1992). The importance of incidence registries for connective tissue diseases. *British Journal of Rheumatology*, **19**, 1–7.

Michet, C.J. *et al*. (1985). Epidemiology of systemic lupus erythematosus and other connective tissue diseases in Rochester, Minnesota, 1950 through 1979. *Mayo Clinic Proceedings*, **60**, 105–13.

Moskowitz, M.A. *et al*. (1990). The relationship of oral contraceptive use to rheumatoid arthritis. *Epidemiology*, **1**, 153–6.

Mowat, A.G. and Hazleman, B.L. (1974). Polymyalgia rheumatica. A clinical study with particular reference to arterial disease. *Journal of Rheumatology*, **1**, 190–202.

Omland, O., Sommer, G., and Elling, H. (1986). Incidence of polymyalgia rheumatica/temporal arteritis in a Danish country. *Ugeskrift fur Laeger*, **143**, 981–3.

Rajala, S.A., Ahvenainen, J.E., Mattila, K.J., and Saarni, M.I. (1993). Incidence and survival rate in cases of biopsy-proven temporal arteritis. *Scandinavian Journal of Rheumatology*, **22**, 289–91.

Reid, D.M., Reid, T.M.S., Brown, T., and Rennie, J.A.N. (1985). Human parvovirus-associated arthritis: a clinical and laboratory description. *Lancet*, **i**, 422–5.

Rigby, A.S. (1991). Review of UK data on the rheumatic diseases—ankylosing spondylitis. *British Journal of Rheumatology*, **30**, 50–3.

Safavi, K.H., Heyse, S.P., and Hochberg, M.C. (1990). Estimating the incidence and prevalence of rare rheumatologic diseases: a review of methodology and available data sources. *Journal of Rheumatology*, **17**, 990–3.

Salvarani, C. *et al*. (1987). Polymyalgia rheumatica and giant cell arteritis: a 5-year epidemiologic and clinical study in Reggio Emilia, Italy. *Clinical and Experimental Rheumatology*, **5**, 205–15.

Silman, A.J. (1986). Epidemiology aspects of family studies in rheumatoid arthritis. *Disease Markers*, **4**, 55–9.

Silman, A.J. (1988). Has the incidence of rheumatoid arthritis declined in the United Kingdom? *British Journal of Rheumatology*, **27**, 77–8.

Silman, A.J. and Black, C. (1988). Increased incidence of spontaneous abortion infertility in women with scleroderma before disease onset. *Annals of the Rheumatic Diseases*, **47**, 441–4.

Silman, A.J. and Hochberg, M.C. (1993). *Epidemiology of the rheumatic diseases*. Oxford University Press, Oxford.

Silman, A.J. and Symmons, D.P.M. (1995). Selection of study population in the development of rheumatic disease criteria: comment on the article by the American College of Rheumatology Diagnostic and Therapeutic Criteria Committee. *Arthritis and Rheumatism*, **38**, 722.

Silman, A.J. and Vandenbroucke, J.P. (1989). Female sex hormones and rheumatoid arthritis. *British Journal of Rheumatology*, **28** (Suppl. 1), 1–73.

Silman, A.J., Ollier, W.E.R., and Currey, H.L.F. (1987). Failure to find disease similarity in sibling pairs with rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **46**, 135–8.

Silman, A.J., Jannini, S., Symmons, D., and Bacon, P. (1988). An epidemiological study of scleroderma in the West Midlands. *British Journal of Rheumatology*, **27**, 286–90.

Silman, A.J., Ollier, W.E.R., and Bubel, M.A. (1989). Auto-immune thyroid disease and thyroid autoantibodies in rheumatoid arthritis patients and their families. *British Journal of Rheumatology*, **23**, 18–21.

Silman, A.J., Hicklin, A.J., and Black, C. (1990). Geographical clustering of scleroderma in South and West London. *British Journal of Rheumatology*, **29**, 92–6.

Silman, A.J. *et al*. (1993). Absence of rheumatoid arthritis in a rural Nigerian population. *Journal of Rheumatology*, **16**, 618–22.

Smith, C.A., Fidler, W.J., and Pinals, R.S. (1983). The epidemiology of giant cell arteritis. *Arthritis and Rheumatism*, **26**, 1214–19.

Solomon, L., Robin, G., and Valkenburg, H.A. (1975). Rheumatoid arthritis in an urban South African Negro population. *Annals of the Rheumatic Diseases*, **34**, 128–35.

Spector, T.D. and Hochberg, M.C. (1990). The protective effect of the oral contraceptive pill on rheumatoid arthritis. *Journal of Clinical Epidemiology*, **43**, 1221–30.

Spector, T.D. and Silman, A.J. (1987). Does the negative association between rheumatoid arthritis and schizophrenia provide clues to the aetiology of rheumatoid arthritis? *British Journal of Rheumatology*, **26**, 307–10.

Spector, T.D., Roman, E., and Silman, A.J. (1990). The pill, parity, and rheumatoid arthritis. *Arthritis and Rheumatism*, **33**, 782–9.

Steere, A.C. *et al*. (1977). An epidemic of oligoarticular arthritis in children and adults in three Connecticut communities. *Arthritis and Rheumatism*, **20**, 7–37.

Steere, A.C. *et al*. (1983). The spirochetal etiology of Lyme disease. *New England Journal of Medicine*, **308**, 733–9.

Stewart, O.J. and Silman, A.J. (1990). Review of UK data on the rheumatic diseases—gout. *British Journal of Rheumatology*, **29**, 485–8.

Symmons, D.P.M. (1991). Review of UK data on the rheumatic diseases—systemic lupus erythematosus. *British Journal of Rheumatology*, **30**, 288–90.

Symmons, D.P.M., Barrett, E.M., Bankhead, C.R., Scott, D.G.I., and Silman, A.J. (1994). The incidence of rheumatoid arthritis in the United Kingdom: results from the Norfolk Arthritis Register. *British Journal of Rheumatology*, **33**, 735–9.

Tan, E.M. (1982). The 1982 revised criteria for the classification of systemic lupus erythematosus. *Arthritis and Rheumatism*, **25**, 1271–7.

van der Linden, S., Valkenburg, H.A., and Catas, A. (1984). Evaluation of diagnostic criteria for ankylosing spondylitis. *Arthritis and Rheumatism*, **27**, 361–8.

Vandenbroucke, J.P. (1983). Oral contaceptives and rheumatoid arthritis. *Lancet*, **ii**, 1282.

Vandenbroucke, J.P. *et al*. (1982). Oral contraceptives and rheumatoid arthritis: further evidence for a preventive effect. *Lancet*, **ii**, 839–42.

Vandenbroucke, J.P. *et al*. (1986). Non-contraceptive hormones and rheumatoid arthritis in premenopausal and postmenopausal women. *Journal of the American Medical Association*, **255**, 1299–303.

van Saase, J.L.C.M. *et al*. (1989). Epidemiology osteoarthrits: Zoetermeer survey. Comparison of radiological osteoarthritis in a Dutch population with that in 10 other populations. *Annals of the Rheumatic Diseases*, **43**, 271–80.

Vaughan, J.H. (1979). Rheumatoid arthritis, rheumatoid factor, and the Epstein–Barr-Virus. *Journal of Rheumatology*, **6**, 381–8.

Vessey, M.P., Villard-Mackintosh, L., and Yeates, D. (1987). Oral contraceptives, cigarette smoking, and other factors in relation to arthritis. *Contraception*, **35**, 457–65.

White, D.G. *et al*. (1985). Human parvovirus arthropathy. *Lancet*, **i**, 419–21.

Wilkins, R.F. *et al*. (1981). Reiter's syndrome: evaluation of a preliminary criteria for definite disease. *Arthritis and Rheumatism*, **24**, 844–9.

Williams, G.H. and Silman, A.J. (1991). Review of UK data on the rheumatic diseases—scleroderma. *British Journal of Rheumatology*, **30**, 365–7.

Wingrave, S.J. (1978). Reduction in incidence of rheumatoid arthritis associated with oral contraceptives. *Lancet*, **i**, 569–71.

Wolfe, F., Kleinheksel, S.M., and Khan, M.A. (1988). Prevalence of familial occurrence in patients with rheumatoid arthritis. *British Journal of Rheumatology*, **27**, 150–52.

Wolfe, F. *et al*. (1990). The American College of Rheumatology 1990 criteria for the classification of fibromyalgia. *Arthritis and Rheumatism*, **33**, 160–72.

5.1.2 Epidemiology of rheumatic diseases in selected non-European populations

Patricia A. Fraser

Epidemiology is the study of the distribution, transmission, and control of disease (Masi 1984). The primary focus of this chapter is the variation in frequencies of rheumatic diseases among several ethnic groups. Ethnicity-specific disease expression also provides clues to aetiological factors for rheumatic diseases. Comparative analysis of the frequencies of specific disease manifestations is a secondary focus of this chapter. The gaps in our knowledge of the causes and the mechanisms of transmission of rheumatic diseases preclude discussion about control and interventions. Observed variations in incidence, prevalence, and clinical and laboratory manifestations of systemic rheumatic diseases between ethnic groups may result from environmental factors. The correlation of exposure to *Mycoplasma pneumoniae* respiratory infections and cyclic variation in incidence of juvenile rheumatoid arthritis in one Canadian province (Oen *et al*. 1995) emphasizes the importance of region-specific microbial exposure. Inherited factors contribute to susceptibility to rheumatic diseases. These genetic factors may be linked to, or regulated by, multiple loci throughout the human genome. If the genes of interest are polymorphic, their allelic frequencies may vary between ethnic groups. Such genetic variation may be the basis for ethnic differences in susceptibility and disease expression. Survival is one important prognostic feature of disease expression. Ethnic differences in disease mortality may be multifactorial. Studies of the prognosis for systemic lupus erythematosus in African Americans emphasize the importance of socioeconomic factors (Karlson *et al*. 1995). Future epidemiological studies of the rheumatic diseases should be designed to obtain data on multiple variables such as genetic markers at multiple loci, climate, microbial exposure, cultural or behavioural factors (e.g. dietary habits, tobacco and alcohol use) (Adebajo 1995), and the conundrum of 'socioeconomic status'.

The occurrence and manifestations of juvenile arthritis, rheumatoid arthritis, systemic lupus erythematosus, and the seronegative spondylarthropathies in selected non-European populations are presented in this chapter. Limited data on Caucasians of European descent will be presented for the purpose of comparison. Similarly, features of arthritic disorders will be compared between Native Americans and Eskimos and between Native Americans and Mexicans because of the common origin of Native Americans and Eskimos and because of the significant contribution of Native Americans to the gene pool of Mexicans (Williams *et al*. 1985).

Interethnic comparisons of the epidemiology of rheumatic diseases are hampered by the lack of uniform sampling and disease definitions, and also by limited or missing data on many ethnic groups. Clinical case series have been included for several ethnic groups when formal epidemiological studies of prevalence or incidence are not available. These data are not equivalent in their scientific rigor to formal epidemiological studies. They are presented for the purpose of completeness of information on rheumatic diseases worldwide.

Juvenile arthritis

Definition

Descriptive terms and definitions of chronic inflammatory arthritis in childhood are numerous. North American diagnostic criteria for juvenile rheumatoid arthritis (Brewer *et al*. 1977) and European criteria for juvenile chronic arthritis (Ansell 1978) provide the foundation for an evolving worldwide nomenclature and classification of juvenile arthritis shown in Table 1 (Ansell 1990) and Table 2 (Fink and Fernandez-Vina 1995).

America [juvenile rheumatoid arthritis (JRA)]

Europe [juvenile chronic arthritis (JCA)]

Systemic JRA/JCA

Pauciarticular JRA/JCA:
Type I—ANA-positive (young girls)
Type II—HLA-B27-positive (older boys)

Polyarticular JRA/JCA:
Rheumatoid factor-negative (young children)
Rheumatoid factor-positive (older children)

[1]From Ansell (1990)
ANA, antinuclear antibody.

Table 1 Existing terminology/classification of childhood arthritis [1]

Systemic arthritis

Polyarthritis: rheumatoid factor-negative

Polyarthritis: rheumatoid factor-positive

Oligoarthritis

Extended oligoarthritis

Enthesitis-related arthritis

Psoriatic arthritis

Proposed classification from the Task Force of the Pediatric Standing Committee of the International League of Associations for Rheumatology (from Fink and Fernandez-Vina 1995)

Table 2 Idiopathic arthritides of childhood

Occurrence

Juvenile arthritis is the most common, childhood, chronic systemic rheumatic disease in the United States (Cassidy and Nelson 1988). This statistic reflects the frequency of juvenile arthritis in the largest ethnic group, Caucasians, and in Native Americans. Data on the occurrence of juvenile arthritis in Africa, Australia, and Asia are sparse.

Several survey methods have been used to estimate the occurrence of juvenile arthritis and other rheumatic diseases (Gewanter *et al*. 1983). The field population survey is the most accurate method because it includes actual physical examination of the population under study by trained observers. It is also the most costly. Estimates are also based on the number of cases presenting to a central clinic. The size of the source population must be known and it is assumed that all cases from this population base would present to this centre. The review of a medical records database for the diagnoses of interest has been used extensively. This approach also assumes the database captures virtually all of the source population. The practitioner survey interviews practitioners by questionnaire, which may be a preferable method in a population with a decentralized health-care system and has been used effectively to estimate frequencies for childhood rheumatic diseases in the United States (Gewanter *et al*. 1983).

Estimates of prevalence reveal an excess risk of juvenile arthritis among Native Americans when compared to Caucasians in various parts of North America. Juvenile arthritis is twice as common among Native Americans than among Caucasians in western Canada. Susceptibility to juvenile arthritis also varies between Native American tribal groups. In this analysis, data from Eskimo ethnic groups are compared to those from other Naive American tribal groups. The twofold difference in the prevalence of juvenile arthritis between Native Americans in Manitoba and British Columbia (Rosenberg *et al*. 1982) may be due to the high prevalence of juvenile-onset spondylarthropathy among the Native Americans of Manitoba (Hochberg 1984). Similarly designed studies reveal a variable prevalence for juvenile arthritis in several samples of Eskimos and Alaskan Native Americans (Oen *et al*. 1986; Boyer *et al*. 1988; Boyer *et al*. 1990; Boyer *et al*. 1991). The Inuit and Native Americans from the south-east coast of Alaska have the two highest reported prevalence rates for juvenile arthritis (126/100 000 and 83/100 000, respectively) (Oen *et al*. 1986; Boyer *et al*. 1991). Figure 1 includes estimates of the period prevalence of juvenile arthritis in Native Americans, Eskimos (Rosenberg *et al*. 1982; Oen *et al*. 1986; Boyer *et al*. 1988; Boyer *et al*. 1990; Boyer *et al*. 1991), and African Americans in Baltimore (Hochberg *et al*. 1983), and the frequency of juvenile arthritis in Caucasians in the United States (Lawrence *et al*. 1989) and Arabs in Kuwait, another Caucasian population (Khuffash and Majeed 1988).

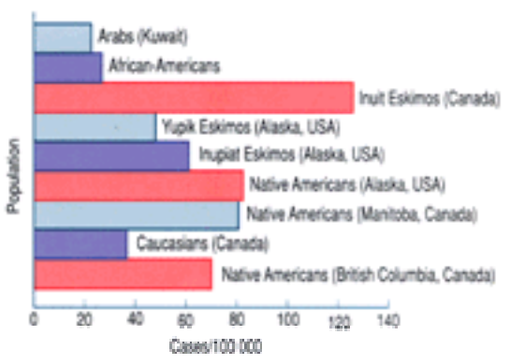


Fig. 1 Prevalence of juvenile arthritis.

Cases of juvenile arthritis have been observed in Chinese in Hawaii (Hicks 1977) and California (Hanson *et al*. 1977), but a formal estimate of the incidence of juvenile arthritis in British Columbia indicates that it is rare in North American Chinese (Hill 1976; Kelsey 1982). The incidence of juvenile arthritis in African Americans approximates that observed in Caucasians in the United States (Lawrence *et al*. 1989). Differences in method may account for the disparity in incidence estimates between United States and Canadian Caucasians. Incidence rates by race and ethnicity are listed in Table 3.

Population	Rate
Chinese (Canada)	0.00
Caucasian (British Columbia, Canada)	2.20
Native American (British Columbia)	7.20
Native Americans (S.E. Alaska, USA)	38.6
Inupiat Eskimo (Alaska, USA)	28.0
Yupik Eskimo (Alaska, USA)	42.5
Inuit Eskimos (Canada)	23.6
African American	6.20
Caucasian (USA)	9.20

Table 3 Annual incidence rate per 100 000 of juvenile arthritis

Gender differences and laboratory manifestations

The distribution of the different onset types of juvenile arthritis varies by race, ethnicity, and region. For example, onset types varied among three predominantly Caucasian samples in the United States (Hanson *et al*. 1977; Jacobs 1982; Aaron *et al*. 1985). In view of the strong genetic component in predisposition to juvenile arthritis and the known variation in the frequency of genetic markers within an ethnic group, the observed difference in the distribution of onset types in United States Caucasians may reflect the particular ethnic composition (e.g. ratio of individuals of northern versus southern European origin) in these samples. Smaller samples of individuals with juvenile arthritis in other ethnic groups also demonstrate variability in the distribution of onset types. Polyarticular-onset juvenile arthritis was twice as common among Canadian Native Americans as Canadian Caucasians (Rosenberg *et al*. 1982). Comparison of three samples of patients with juvenile arthritis among black Africans revealed different proportions of onset types and no evidence for a regional trend (Kanyerezi and Mbidde 1980; Gupta *et al*. 1981; Haffejee *et al*. 1984). The studies of juvenile arthritis among Kuwaiti Arabs (Khuffash and Majeed 1988), Ugandans (Kanyerezi and Mbidde 1980), Zambians (Gupta *et al*. 1981), Black and Indian South Africans (Haffejee *et al*. 1984), and Mexicans (Martinez-Cairo Cueto *et al*. 1978) are presented in Table 4.

Race/ethnicity/author and/or locale	Number	Systemic (%)	Polyarticular (%)	Pauciarticular (%)
Caucasians[a] (USA; Jacobs 1982)	200	9.0	16.0	75.0
Caucasians[a] (USA; Hanson *et al*. 1977)	500	43.0	23.0	34.0
Caucasians[a] (USA; Aaron *et al*. 1985)	327	15.0	27.0	58.0
Native Americans	34	11.8	58.8	29.4
Arabs (Kuwait)	41	39.0	39.0	22.0
Blacks (Uganda)	41	54.0	37.0	9.0
Blacks (Zambia)	8	37.5	37.5	25.0
Blacks (South Africa)	40	15.0	50.0	35.0
Indians (South Africa)	18	16.6	44.4	39.0
Mexicans	46	28.3	45.6	26.1

[a]Sample composition e.g. ~ 80% US Caucasian

Table 4 Juvenile arthritis onset types by author, race, or ethnicity

Equal sex ratios among Black and Indian South African (Haffejee *et al*. 1984) and Mexican individuals with juvenile arthritis (Martinez-Cairo Cueto *et al*. 1978) contrast with the female predominance in samples of Native American and North American Caucasians with juvenile arthritis (Rosenberg *et al*. 1982, Aaron *et al*. 1985) (Table 5). With the exception of two studies on sex ratios (Hanson *et al*. 1977; Aaron *et al*. 1985), data were not analysed to demonstrate the age dependence of sex ratios in pauciarticular juvenile arthritis (i.e. girls more often affected than boys in early onset, boys more than girls in late onset).

Race or ethnicity	Systemic	Polyarticular	Pauciarticular	All JA
Native Americans (British Columbia)				4.1:1
Caucasians (British Columbia)				5.8:1
Blacks/Indians (South Africa)	1.25:1	1.1:1	0.83:1	1:1
Mexicans	0.63:1	2:1	0.7:1	1.1:1
Arabs	2.2:1	33:1	8:1	1.3:1
US Caucasians (Hanson *et al*. 1977)	1.1:1	3.8:1	3.2:1	2:1
US Caucasians (Aaron *et al*. 1985)	1.6:1	5.3:1	2:1	2.4:1
Inuit				1:1

Table 5 Female:male ratio by onset type or all juvenile arthritis (JA)

The frequencies of antinuclear antibodies for individuals with juvenile arthritis in three racial groups in Table 6 are much lower than a recent report from the United States. Szer *et al*. (1991) noted antinuclear antibody positivity of 83, 42, and 39 per cent for patients with pauciarticular, polyarticular, and systemic juvenile arthritis, respectively. This discrepancy may be due to difficulty in comparing clinical case series and epidemiological studies, or to differences in laboratory methods, or both.

Race or ethnic group	Systemic	Polyarticular	Pauciarticular	JA
Native Americans (British Columbia)				53.3
Caucasians (British Columbia)				28.8
Blacks/Indian (South Africa)				6.8

Table 6 Frequency (per cent) of antinuclear antibody by race or ethnicity by onset type or juvenile arthritis (JA) unclassified

Case series must be utilized when we consider comparisons of rheumatoid factor positivity. With the exception of black Ugandans, rheumatoid factor-positive juvenile arthritis is threefold greater in the non-European populations included, although the proportion of polyarticular-onset juvenile arthritis is not significantly different among these populations. The disparity between the proportions of polyarticular-onset juvenile arthritis and seropositivity may indicate the presence of other stimuli for rheumatoid factor production such as malaria or other recurrent parasitic infections and tuberculosis (Haffejee *et al*. 1984). Estimates of the seroprevalence of rheumatoid factor in juvenile arthritis are presented in Table 7.

Race or ethnicity	Systemic	Polyarticular	Pauciarticular	JA
Native Americans (Alaska)				32.0
Native Americans (British Columbia)				36.0
Caucasians (British Columbia)				9.1
Black (South Africa)				36.0
Indian (South Africa)				33.0
Mexican	30.7	42.8	16.6	33.0
Black (Uganda)				10.0

Table 7 Frequency (per cent) of rheumatoid factor by race or ethnicity by onset subtype or unclassified juvenile arthritis (JA)

Juvenile spondylarthropathies

Juvenile-onset seronegative spondylarthropathies, which include ankylosing spondylitis, Reiter's syndrome, and seronegative enthesopathy and arthropathy syndrome (Rosenberg and Petty 1982), account for at least half of the arthritides of childhood in native North American populations. Specific criteria for childhood-onset ankylosing spondylitis are necessary since the application of adult criteria have limited value (Singsen 1990) and may result in underestimation of this condition. The demographic, clinical, and laboratory features of juvenile-onset spondylarthropathies (Rosenberg *et al*. 1982; Oen *et al*. 1986; Boyer *et al*. 1988; Boyer *et al*. 1990; Boyer *et al*. 1991) are summarized in Table 8.

Population	Diagnosis	Number of cases	Prevalence	Incidence	SPA:JA	HLA-B27
Native Americans (British Columbia)	All SPA	14	29.4		0.82:1	71
	SEA	5				
	AS	3				
	RS	4				
Native Americans (Alaska)	All SPA	9			3.75:1	
	SEA	7				
	RS	2				
Yupik Eskimo	All SPA	21			7:1	
	SEA	17				
	RS/AS	4				
Inupiat Eskimo	All SPA	5		47.4	5:1	
	SEA	3				
	RS	3				
	AS	1				
Inuit	All SPA	11	367	105.6	4.5:1	

AS, ankylosing spondylitis; RS, Reiter's syndrome; SEA, seronegative enthesopathy and arthropathy.

Table 8 Juvenile spondylarthropathies (SPA): prevalence (cases/100 000), incidence (cases/100 000 per year), ratio of SPA:juvenile arthritis (JA), and frequency of

HLA-B27 (per cent)

Rheumatoid arthritis

Several variables affect comparisons of the prevalence estimates for rheumatoid arthritis.

Definition

The case definition of rheumatoid arthritis varies between studies. This non-uniform disease definition resulted from the evolution of disease criteria for rheumatoid arthritis over time—the Manchester grading system, the 1958 and 1962 American Rheumatism Association criteria, the New York criteria, 1987 revised, for rheumatoid arthritis [(Ropes *et al*. (1958), Lawrence (1961), Kellgren (1962), Lawrence and Wood (1963), and Arnett *et al*. (1988), respectively]. The earliest prevalence estimates to be presented in this section, from New Zealand Maoris (Rose and Prior 1963), rural Japan (Shichikawa *et al*. 1981), Jamaica (Lawrence *et al*. 1966), Puerto Rico (Mendez-Bryan *et al*. 1964), Liberia and Nigeria (Muller 1970), and among the South African Bantu (Solomon *et al*. 1975; Beighton *et al*. 1975), included probable and definite cases of rheumatoid arthritis. In later studies, prevalence rates were based on classical and definite cases of rheumatoid arthritis.

Occurrence

Rheumatoid arthritis does not occur at the same frequency at all ages. Its age distribution may vary between populations. Crude estimates of prevalence in populations with a younger age structure will be lower than age-adjusted prevalence rates since rheumatoid arthritis is a disease of middle age. This potential source of bias is one explanation offered for the very low prevalence of rheumatoid arthritis in several studies from different Asian and African, and African-derived population samples (Mijiyawa 1995).

Population sampling may also affect the comparability of data. Population-based epidemiological studies are preferred, when feasible.

Rheumatoid arthritis occurs in India, Pakistan, Oman, and Iraq at frequencies (0.75–1.98 per cent) similar to that observed in the United Kingdom and the United States (Al-Rawi *et al*. 1978; Pountain 1991; Malaviya *et al*. 1993; Hameed *et al*. 1995). The common methods used to estimate the occurrence of rheumatoid arthritis among samples of Native American facilitate the comparison of prevalence estimates among these groups. It is noteworthy that rheumatoid arthritis was confined to women in an Inuit sample, although records of both sexes were reviewed (Oen *et al*. 1986). In contrast, a field population survey among the Yakima was limited to women (Beasley *et al*. 1973). Variability between ethnic groups was observed between tribal groups and by locale. The majority (76 per cent) of those with rheumatoid arthritis among south-eastern Alaskan Native Americans were of Tlingit ancestry (Boyer *et al*. 1991). The prevalence of rheumatoid arthritis varied among the groups of Native Americans studied (Chippewa, Haida, Nootka and Pima, Yupik, and Inupiat Eskimos) from 0.6 to 7.1 per cent, with the highest rates among the Chippewa and the Pima (7.1 and 5.3 per cent, respectively) (Gofton *et al*. 1964; Harvey *et al*. 1983; Atkins *et al*. 1988; Del Puente *et al*. 1989; Boyer *et al*. 1990; Boyer *et al*. 1991).

Rheumatoid arthritis occurs at a lower frequency in populations of African ancestry when compared to Caucasians. This is best exemplified in a study in Manchester, England, where the age-adjusted prevalence of rheumatoid arthritis was significantly lower in Afro-Caribbeans (Afro-Caribbean: Caucasian 0.36:1) (MacGregor *et al*. 1994), and in a study that failed to detect any cases of rheumatoid arthritis among 2000 rural Nigerians (Silman *et al*. 1993).

Comparison of the prevalence rates of rheumatoid arthritis among populations of black African descent is more problematical and necessitates review of the raw data. This is necessary because several disease definitions were used in the studies of interest. Intercontinental differences in population age structure may also contribute to the observed differences in the prevalence of rheumatoid arthritis. For example, the age distribution in the United States is significantly older than that in most sub-Saharan African nations (Hall 1991). Since rheumatoid arthritis occurs most frequently in young and middle-aged adults, a country such as the United States may have a higher prevalence of rheumatoid arthritis on the basis of the age distribution of its population. The combined prevalence of definite and classical rheumatoid arthritis has been estimated for African Americans (Lawrence *et al*. 1989), while prevalence data for probable and definite rheumatoid arthritis are available from Jamaica and several black African samples. Variability is observed if the prevalence of definite rheumatoid arthritis among samples of black African descent is compared. Although the prevalence of definite rheumatoid arthritis is similar for African Americans, rural Jamaicans, and urban South African Bantu (Lawrence *et al*. 1966; Solomon *et al*. 1975; Lawrence *et al*. 1989), estimates from rural Africa are significantly lower (Beighton *et al*. 1975; Moolenburgh *et al*. 1984; Brighton *et al*. 1988). With the exception of the Jamaican sample, there is a striking difference between urban and rural black populations. This is best illustrated in southern Africa, where rural and urban estimates among the Tswana reveal a more than threefold difference in the prevalence of rheumatoid arthritis (Beighton *et al*. 1975; Solomon *et al*. 1975). It is difficult to draw any conclusions about ethnic differences in predisposition in black African populations. A study of 39 patients with rheumatoid arthritis admitted to a central hospital in Uganda did not show any ethnic predisposition among the Baganda, Banyarwanda, and Banyankole, the major tribal groups in that region (Kanyerezi 1969). A similar investigation in western Nigeria also failed to show heightened susceptibility in any ethnic group (Greenwood 1969).

The hypotheses that rheumatoid arthritis is a relatively new disease in Africa (Adebajo 1991) and that it is due to infection with a slow virus (Buchanan and Murdoch 1979) may explain the difference in urban and rural prevalence estimates for rheumatoid arthritis in Africa in populations of Bantu ancestry. Presumably, urban Bantu dwellers had greater exposure to the putative viral infection than did rural inhabitants. Similarly, the very low prevalence of the disease (0.2–0.5 per cent) among rural and urban Chinese in the People's Republic of China, Taiwan, Hong Kong, three Japanese samples, and Indonesia (Kato *et al*. 1971, Shichikawa *et al*. 1981; Beasley *et al*. 1983; Lau *et al*. 1993; Darmawan *et al*. 1993; Chou *et al*. 1994; Wigley *et al*. 1994) may indicate lack of exposure to the rheumatoid arthritis-associated virus. These hypotheses do not explain the very low prevalence of rheumatoid arthritis (0.1 per cent) in Chile, which is possibly due to the age distribution of the population (Gomez-Carpio *et al*. 1966). Such theories of differential susceptibility or aetiological agents can neither be confirmed nor refuted until standardized disease definitions are used to estimate the age-adjusted prevalence of rheumatoid arthritis in population-based studies. Table 9 lists 45 prevalence estimates for rheumatoid arthritis.

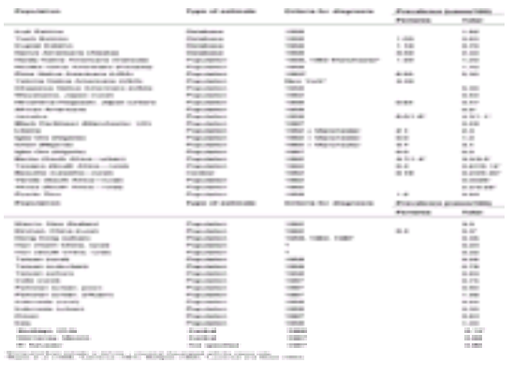


Table 9 Prevalence rates of rheumatoid arthritis

Comparative analyses of seropositivity in rheumatoid and non-rheumatoid arthritis show rates of seropositivity in non-rheumatoid from 2 to 20 per cent and between 9 to 94 per cent in rheumatoid (Table 10). It is interesting that two studies from Nigeria made more than 20 years apart provide very different estimates of the seroprevalence of rheumatoid factor (Greenwood 1969; Adebajo *et al*. 1993). A longitudinal study of rheumatoid factor and rheumatoid arthritis in the Pima indicated that the titre of rheumatoid factor predicts the risk of developing rheumatoid arthritis in this population (Del Puente *et al*. 1988). This observation may also explain the high prevalence rates of rheumatoid arthritis and of seropositive rheumatoid arthritis (78–94 per cent) among other Native American groups.

Population/locale	Type of sample	Total	Non-RA	RA
USA	Population		2.6	32.0
Uganda	Hospitalized			51.3[a]
Western Nigeria	Hospitalized		11.4	13.0[a]
Ibadan, Nigeria	Central clinic		4.0	47.0[a]
Jamaica	Population	1.6		20.0
Tswana (rural South Africa)	Central clinic			6.9[a]
Tswana (urban South Africa)	Central clinic			12.1[a]
Xhosa (rural South Africa)	Central clinic			17.0[a]
Cape Town, South Africa	Central clinic			85.0[a]
Xhosa (rural South Africa)	Population	17.0	14.7	50.0[a]
Basotho (Lesotho)	Population		12–19	58.8
Chippewa (USA)	Population		10.0	92
Yakima (USA)	Population		3.0	94
Haida (Canada)	Population	3.5	2.0	50
Yupik (Alaska)	Population			78
Inuit (Alaska)	Population			83
Japan (urban)	Population	8.9	7.4	75
China (rural)	Population		1.7	28

[a]Kuryenov (1968); [b]Greenwood (1968); [c]Acheson *et al*. (1965); [d]Bruce *et al*. (1963); [e]Moore *et al*. (1968); [f]Moore *et al*. (1963); [g]Kelly and Moore (1968); [h]Moore *et al*. (1977); [i]Vandenbergh *et al*. (1986).

Table 10 Frequency (per cent) of rheumatoid factor seropositivity in selected normal and rheumatoid arthritis (RA) samples

Systemic lupus erythematosus

Definition

The analysis of Jamaicans with systemic lupus erythematosus (Wilson and Hughes 1979) utilized 1971 criteria to define the disease (Cohen *et al*. 1971). The remainder of the studies used the 1982 revised criteria (Tan *et al*. 1982).

H4>*Occurrence*

Although there is little evidence for geographic clustering of systemic lupus erythematosus (Wallace and Quismorio 1995), there are abundant data to support the concept of differential susceptibility to lupus associated with ethnicity. The prevalence of systemic lupus erythematosus among African Americans is more than threefold greater than among Caucasians in the same regions (Siegel and Lee 1973; Fessel 1974). The available rates for systemic lupus in African Americans are stratified by sex and age, with peak prevalence between 15 and 64 years. United States incidence data are two- to threefold greater in African American than Caucasian women (McCarty *et al*. 1995). The same trends are seen in data from Afro-Caribbeans in Birmingham, UK (Johnson *et al*. 1995). The excess risk of systemic lupus erythematosus in populations of African ancestry residing outside continental Africa are unexplained. Genetic admixture of the lupus-prone individuals has been offered as a possible explanation for the excess susceptibility (Symmons 1995). The only study to address admixture in systemic lupus found no difference in its level (approx. 28 per cent) between African Americans patients in Baltimore and local controls (Bias *et al*. 1992). In several samples, age of onset in African-Americans and Afro-Caribbeans is significantly younger than locale-matched Caucasian lupus patients (Johnson *et al*. 1995; McCarty *et al*. 1995). Estimates of systemic lupus erythematosus occurrence in Curacao, a Caribbean island largely populated by descendants of enslaved Africans, differ slightly from other samples of African ancestry. Although the peak incidence occurs among women 45 to 65 years exceeds that for the 15 to 44 age group (186.6 versus 102.9/100 000) (Nossent 1992). The minor discordance between these two rates are not due to differences in mortality rates. Figure 2 summarizes age-specific population, prevalence, incidence, and mortality data for the Curacao sample.

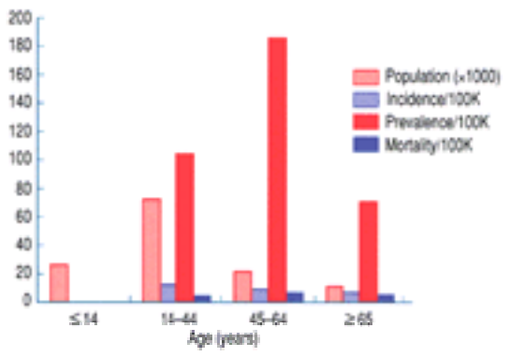


Fig. 2 Age-specific population, prevalence, incidence, and mortality in Curacao.

Among Native American women in Alaska aged 15 to 64 years, the prevalence of systemic lupus erythematosus is 280/100 000, which is intermediate between estimates for United States Caucasians and African Americans (140/100 000 and 410/100 000, respectively) (Boyer *et al*. 1991). The Nootka Native Americans in British Colombia have the highest reported prevalence of systemic lupus (500 cases/100 000 population) (Atkins *et al*. 1988), while the Sioux, Crow, and Arapaho exhibited the highest incidence rates of systemic lupus in a survey of 75 Native American peoples (Morton *et al*. 1976). The estimates for the Nootka, Crow, and Arapaho were derived from small populations (2000–3000) and it is likely that the incidence estimate for the Sioux is the most accurate (Hochberg 1990) because they were the most numerous group studied (population = 30 000).

Black South Africans were over-represented among inpatients with systemic lupus erythematosus in Cape Town (Jessop and Meyers 1973; Fessel 1988). Since admission to hospital with systemic lupus is highly correlated with disease severity, the disproportionate number of Blacks may reflect more severe disease rather than excess risk of lupus. This conclusion in supported by estimates from a central clinic in Johannesburg where the highest prevalence and incidence of systemic lupus was among Asian Indians, with decrements in the frequency among Coloured or mixed-race individuals and Caucasians, and the lowest frequencies in Blacks (Morrison *et al*. 1990).

Higher prevalence rates of systemic lupus erythematosus have been observed among Chinese, Japanese, and Filipinos in Hawaii, and Polynesians in New Zealand, when compared to Caucasians in either locale (Serdula and Rhoads 1979; Hart *et al*. 1983; Catalano and Hoffmeier 1989), but for several reasons we cannot determine whether so-called mongoloid ethnic groups originating in Asia are more susceptible to systemic lupus than are Caucasians. Estimates from China and Japan give rates similar to those computed in Hawaii but there were no local Caucasians available for comparison in those countries. Furthermore, comparisons of international rates among Caucasians also reveal significant variability (Hochberg 1990). The application of similar methods (Hochberg 1990) and an assessment of racial heterogeneity (Fessel 1988) are necessary to determine whether these ethnic groups exhibit a significant predisposition to systemic lupus. Prevalence and incidence estimates appear in Table 11.

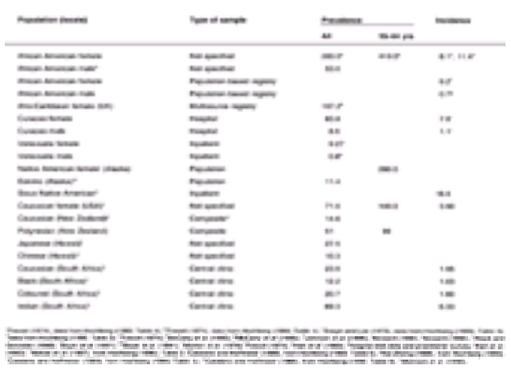


Table 11 Prevalence (cases/100 000) and incidence (new cases/100 000 per year) of systemic lupus

Clinical and laboratory manifestations

The ratio of females to males with systemic lupus erythematosus in the Unites States is 8–10:1 (Siegel and Lee 1973; Fessel 1974) and approaches 12:1 in Venezuela (Abadi and Gonzalez 1990). Case series in Latin America may indicate higher female to male ratios than among United States Caucasians (Alarcon 1986). Populations of black African descent in Jamaica and the United States demonstrate an earlier age of onset than Caucasians (Wilson and Hughes 1979; Ballou *et al*. 1982; Hochberg *et al*. 1985; Ward and Studenski 1990; Smikle *et al*. 1995), more frequent and more severe renal involvement, and greater mortalty (Harris *et al*. 1989; Williams *et al*. 1990; Smikle *et al*. 1995) Earlier age of onset is associated with HLA-DR8 in African Americans (Reveille *et al*. 1989). Among African Americans, discoid lupus (Hochberg *et al*. 1985; Ward and Studenski 1990), lupus pneumonitis, serositis, nephritis, hypocomplementaemia, hyperglobulinaemia (Ballou *et al*. 1982; Hochberg *et al*. 1985) and anti-Sm and antiribonucleoprotein (Ward and Studenski 1990) are significantly more common than among United States Caucasians, while photosensitivity is significantly less frequent (Hochberg *et al*. 1985; Ward and Studenski 1990). Photosensitivity occurs with a similar frequency among African Americans residing in a temperate climate and Jamaicans who are exposed to a subtropical climate. Interestingly, photosensitivity is twice as common in black South Africans who also reside in a temperate zone as among African Americans. The propensity for renal involvement in Afro-Caribbeans is best exemplified by the cohort analysis in Curacao. Lupus nephritis occurred at some time during the follow-up period in 78 per cent of patients (Figure 3) and was associated with greater mortality (Nossent 1993a). When multiple prognostic variables were examined, the SLEDAI score which incorporates data on active renal disease was more powerful than the variable for renal involvement alone as a predictor of mortality (Nossent 1993b). The frequencies of clinical and laboratory features of systemic lupus erythematosus in African American, Jamaican, and black South African and Zimbabwean samples are listed in Table 12 (Wilson and Hughes 1979; Hochberg *et al*. 1985; Taylor and Stein 1986; Dessein *et al*. 1988; Morrison *et al*. 1990; Ward and Studenski 1990; Nossent 1993b; Smikle *et al*. 1995).

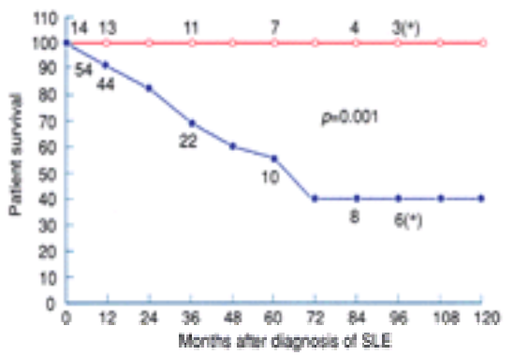


Fig. 3 Survival of lupus patients in Curacao stratified by renal involvement. *Number of patients still in study.

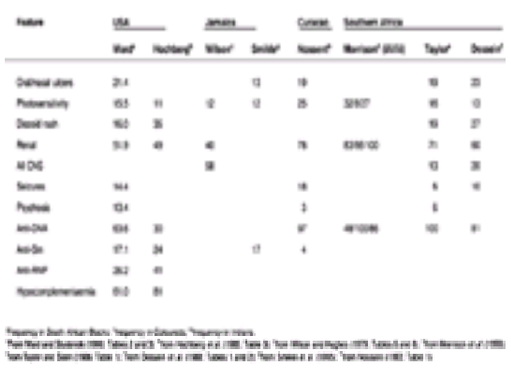


Table 12 Frequency (per cent) of clinical and laboratory manifestations of systemic lupus in selected samples

Seronegative spondylarthropathies

Definition

The seronegative spondylarthropathies are a group of chronic rheumatic diseases characterized by variable combinations of asymmetrical peripheral arthritis, enthesopathy, involvement of the axial skeleton, mucocutaneous symptoms, higher than expected frequencies of the *HLA-B27* allele, and the absence of rheumatoid factor. Ankylosing spondylitis, Reiter's syndrome/'reactive arthritis', psoriatic arthritis/spondylitis, and inflammatory bowel disease-associated spondylitis are included within the spondylarthropathy category. Also to be included are syndromes that have several features of spondylarthropathies but do not meet the criteria for any one. The term undifferentiated spond(lyo)arthropathy has been designated for these conditions (Boyer *et al*. 1988).

The Rome criteria (Kellgren 1962) and the New York criteria (Bennett and Wood 1968) have been developed to assess the frequency of ankylosing spondylitis in populations. The Rome criteria are more suited to field studies since they do not require radiographs to make the diagnosis of definite ankylosing spondylitis. The criteria for Reiter's syndrome (Willkens *et al*. 1981) and the undifferentiated spondylarthropathies (Boyer *et al*. 1988) have not been used as extensively as those for ankylosing spondylitis. Also problematical is the relation of sacroiliitis to the groups of spondylarthropathy. Is asymptomatic sacroiliitis distinct from clinical spondylarthropathy or does it represent one end of the spectrum of spondylarthropathic disease expression?

Occurrence

Prevalence estimates for spondylarthropathies for several Native American and Eskimo groups and the Chinese (Morse *et al*. 1980; Beasley *et al*. 1983; Oen *et al*. 1986; Boyer *et al*. 1988; Boyer *et al*. 1990; Boyer *et al*. 1991) appear in Table 13. HLA-B27 and spondylarthropathies occur at higher frequencies in several Native American groups than in Caucasians, with the highest recorded rates among the Haida in Canada. The published criteria for Reiter's syndrome were not used in the Navajo (the study preceded the publication of the criteria) nor in one of the Inuit studies, so the accuracy and comparability of these estimates is uncertain.

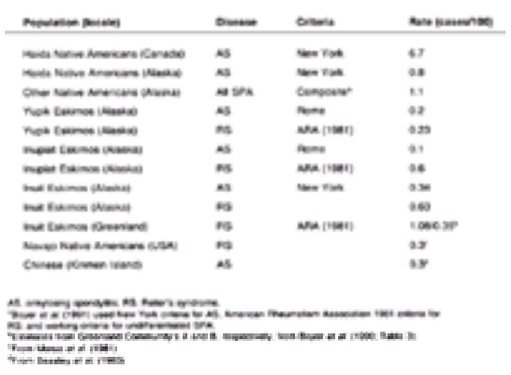


Table 13 Prevalence of seronegative spondylarthropathies (SPA)

In sharp contrast to the Native American groups, spondylarthropathy is diagnosed infrequently in African Americans and black Africans (Baum and Ziff 1971). The frequency of spondylarthropathies in African Americans is not known. The ratio of Caucasians to African Americans with ankylosing spondylitis (range 20:1–6.5:1) admitted to United States Veterans' hospitals serves as an indirect estimate of the prevalence of ankylosing spondylitis in African Americans (Baum and Ziff 1971). The rarity of spondylarthropathies in West Africans, who are racially heterogeneous and are one of the source populations for African Americans, has been attributed to the low frequency of HLA-B27. For example, the *HLA-B*2703* allele has been found in Gambians but ankylosing spondylitis has not been observed (Hill *et al*. 1991). Because of the low overall frequency of HLA-B27 (2.6 per cent) in Gambians, it is possible that one might not see cases of ankylosing spondylitis since only a proportion of HLA-B27-positive individuals develop ankylosing spondylitis. Alternatively, the absence of ankylosing spondylitis in a population with HLA-B27 may indicate the absence of environmental factors that interact with HLA-B27 to trigger spondylarthropathies, or that the HLA-B27 molecules in the population do not confer susceptibility to spondylarthropathies (Hill *et al*. 1991). In support of this latter theory, *HLA-B*2703* differs from other *B27* alleles at amino acid position 59 in the a_1-domain. This observation may explain the results of experiments with alloreactive T-cell clones with this allele (Calvo *et al*. 1990) and may indicate that peptide binding for *HLA-B*2703* is different from other B27 subtypes (Woodrow 1991). HLA-B27 frequencies in normal individuals and those with ankylosing spondylitis appear in Table 14.

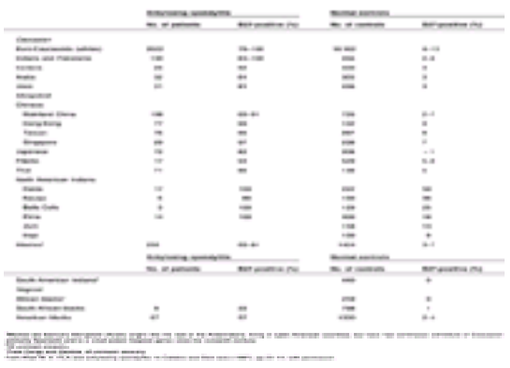


Table 14 HLA-B27 frequency in caucasoid, mongoloid, and negroid populations

Distribution of spondylarthropathy subsets and sex ratios

The relative frequencies of the spondylarthropathy subtypes vary within and between racial groups. As shown in Table 15, ankylosing spondylitis is the predominant spondylarthropathy in Caucasians of Northern European descent (Edmunds *et al*. 1991). Equal frequencies of ankylosing spondylitis and Reiter's syndrome are seen in Alaskan Native Americans and Eskimos (Boyer *et al*. 1990, 1991) and the Shona cultural group in Zimbabwe (Stein *et al*. 1990). Cases of ankylosing spondylitis comprise only 11 per cent of spondylarthropathies among the Alaskan Inuit, where the frequencies of Reiter's syndrome and undifferentiated spondylarthropathy are equal. With the exception of Native Americans in Alaska (Reiter's syndrome and undifferentiated spondylarthropathies are more common in females), males outnumber females with spondylarthropathies (Table 15).

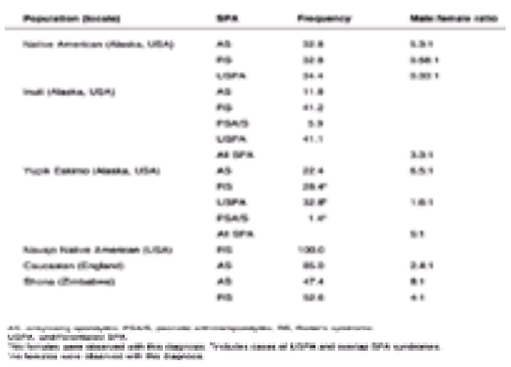


Table 15 Relative frequencies (per cent) of spondylarthropathy (SPA) subsets and sex ratios in several populations

Chapter References

Aaron, S. *et al*. (1985). Sex ratio and sibship in juvenile rheumatoid arthritis kindreds. *Arthritis and Rheumatism*, **28**, 753–8.

Abadi, I. and Gonzalez, N. (1990). Epidemiologia de lupus eritematosos sistemico en Venezuela. Sindrome clinico e immunologico. In *Ediciones Anno Acta Medica Colombiana* (ed. M. Sanchez, M. Diaz, F. Rondon, and G. Ucros), pp. 17–22.

Adebajo, A. O. (1991). Rheumatoid arthritis: a twentieth century disease in Africa? *Arthritis and Rheumatism*, **34**, 248–9.

Adebajo, A. O. (1995). Tropical rheumatology. Epidemiology and community studies: Africa. *Baillières Clinical Rheumatology*, **9**, 21–30.

Adebajo, A. O., Charles, P. J., Hazleman, B. L., and Maini, R. N. (1993). Serological profile of rheumatoid arthritis in West Africa. *Rheumatology International*, **12**, 235–8.

Alarcon, G. S. (1986). Epidemiologia de las enfermedades reumaticas en America Latina. *Bol Oficina Sanit Panam*, **101**, 309–27.

Al-Rawi, Z. S., Alazzawi, A. J., Alajili, F. M., and Alwakil, R. (1978). Rheumatoid arthritis in population samples in Iraq. *Annals of the Rheumatic Diseases*, **37**, 73–5.

Ansell, B. M. (1978). Chronic arthritis in childhood. *Annals of the Rheumatic Diseases*, **37**,107–20.

Ansell, B. M. (1990). Classification and nomenclature. In *Paediatric rheumatology update* (ed. P. Woo, P. H. White, and B. M. Ansell), pp. 3–5. Oxford University Press.

Arnett, F. C. *et al*. (1988). The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis and Rheumatism*, **31**, 315–24.

Atkins, C. *et al*. (1988). Rheumatic disease in the Nuu-Chah-Nulth native Indians of the Pacific Northwest. *Journal of Rheumatology*, **15**, 684–90.

Ballou, S. P., Khan, M. A., and Kushner, I. (1982). Clinical features of systemic lupus erythematosus. *Arthritis and Rheumatism*, **25**, 55–60.

Baum, J. and Ziff, M. (1971). The rarity of ankylosing spondylitis in the black race. *Arthritis and Rheumatism*, **14**, 12–18.

Beasley, R. P., Willkens, R. F., and Bennett, P. H. (1973). High prevalence of rheumatoid arthritis in Yakima Indians. *Arthritis and Rheumatism*, **16**, 743–8.

Beasley, R. P., Bennett, P. H., and Lin, C. C. (1983). Low prevalence of rheumatoid arthritis in Chinese. *Journal of Rheumatology*, **10** (Suppl.),11–15.

Beighton, P., Solomon, L., and Valkenburg, H. A. (1975). Rheumatoid arthritis in a rural South African Negro population. *Annals of the Rheumatic Diseases*, **34**, 136–41.

Bennett, P. H. and Wood, P. H. N. (1968). Report from the subcommittee on diagnostic criteria for ankylosing spondylitis. In *Population studies of the rheumatic diseases* (Proceedings of the Third

International Symposium, New York), (ed. P. H. Bennett and P. H. N. Wood), p 314. Excerpta Medica, Amsterdam.

Bias, W. B., Hochberg, M. C., McLean, R. H., and Machan, C. (1992). Systemic lupus erythematosus joint report. In *HLA 1991: Proceedings of the Eleventh International Histocompatibility Workshop and Conference*, Vol. 1, (ed. K. Tsuji, M. Aizawa, and T. Sasazuki), pp. 740–5. Oxford University Press.

Boyer, G. S., Lanier, A. P., and Templin, D. W. (1988). Prevalence rates of spondyloarthropathies, rheumatoid arthritis, and other rheumatic diseases in an Alaskan Inupiat Eskimo population. *Journal of Rheumatology*, **15**, 678–83.

Boyer, G. S. *et al*. (1990). Spondyloarthropathy and rheumatoid arthritis in Alaskan Yupik Eskimos. *Journal of Rheumatology*, **17**, 489–96.

Boyer, G. S., Templin, D. W., and Lanier, A. P. (1991). Rheumatic diseases in Alaskan Indians of the southeast coast: high prevalence of rheumatoid arthritis and systemic lupus erythematosus. *Journal of Rheumatology*, **18**, 1477–84.

Brewer, E. J. *et al*. (1977). Current proposed revision of JRA criteria. *Arthritis and Rheumatism*, **20** (Suppl.), 195–9.

Brighton, S. W. *et al*. (1988). The prevalence of rheumatoid arthritis in a rural African population. *Journal of Rheumatology*, **15**, 405–8.

Buchanan, W. W. and Murdoch, R. M. (1979). Hypothesis: that rheumatoid arthritis will disappear. *Journal of Rheumatology*, **6**, 324–9.

Calabro, J. J. and Dick, C. (ed.). (1987). *Ankylosing spondylitis. New clinical applications in rheumatology*, Vol. 1. MTP Press, Lancaster.

Calvo, V. *et al*. (1990). Structure and diversity of HLA-B27-specific T cell epitopes: analysis with site-directed mutants mimicking HLA-B27 subtype polymorphism. *Journal of Immunology*, **144**, 4038–45.

Cassidy, J. T. and Nelson, A. M. (1988). The frequency of juvenile arthritis. *Journal of Rheumatology*, **15**, 535–6.

Catalano, M. A. and Hoffmeier, M. (1989). Frequency of systemic lupus erythematosus (SLE) among ethnic groups of Hawaii. *Arthritis and Rheumatism*, **32** (Suppl. 4), S30.

Chou, C. T., Pei, L., Chang, D. M., Lee, C. F., Schumacher, H. R., and Liang, M. H. (1994). Prevalence of rheumatic diseases in Taiwan: A population study of urban, suburban, rural differences. *Journal of Rheumatology*, **21**, 302–6.

Cohen, A. S. *et al*. (1971). Preliminary criteria for the classification of SLE. *Bulletin of the Rheumatic Diseases*, **21**, 643–8.

Darmawan, J., Muirden, K. D., Valkenburg, H. A., and Wigley, R. D. (1993). The epidemiology of rheumatoid arthritis in Indonesia. *British Journal of Rheumatology*, **32**, 537–40.

Del Puente, A. *et al*. (1988). The incidence of rheumatoid arthritis is predicted by rheumatoid factor titer in a longitudinal population study. *Arthritis and Rheumatism*, **31**, 1239–44.

Del Puente, A. *et al*. (1989). High incidence and prevalence of rheumatoid arthritis in Pima Indians. *American Journal of Epidemiology*, **129**, 1170–8.

Dessein, P. H. M. C., Gledhill, R. F., and Rossouw, D. S. (1988). Systemic lupus erythematosus in black South Africans. *South Africa Medical Journal*, **74**, 387–9.

Edmunds. J. E., Kennedy, L. G., and Calin, A. (1991). Primary ankylosing spondylitis, psoriatic and enteropathic spondyloarthropathy: a controlled analysis. *Journal of Rheumatology*, **18**, 696–8.

Fessel, W. J. (1974). Systemic lupus erythematosus in the community incidence, prevalence, outcome and first symptoms: the high prevalence in black women. *Archives of Internal Medicine*, **134**, 1027–35.

Fessel, W. J. (1988) Epidemiology of systemic lupus erythematosus. *Rheumatic Disease Clinics of North America*, **14**, 15–23.

Fink, C. W. and Fernandez-Vina, M. (1995). Classification of juvenile arthritis. *Bulletin of the Rheumatic Diseases*, **14**, 3–5.

Gewanter, H. L., Roghman, K. J., and Baum, J. (1983). The prevalence of juvenile arthritis. *Arthritis and Rheumatism*, **26**, 599–601.

Gofton, J. P., Robinson, H. S., and Price, G. E. (1964). A study of rheumatic disease in a Canadian Indian population. II. Rheumatoid arthritis in the Haida Indians. *Annals of the Rheumatic Diseases*, **23**, 364–71.

Gomez-Carpio, M. *et al*. (1966). Epidemiologia de la artritis reumatoidea en Chile. *Revista Medica de Chile*, **94**, 315–22.

Greenwood, B. M. (1969). Polyarthritis in Western Nigeria. *Annals of the Rheumatic Diseases*, **28**, 489–96.

Gupta, K., Chintu, C., and Raghu, M. B. (1981). Juvenile rheumatoid arthritis in Zambian children. *East Africa Medical Journal*, **58**, 344–5.

Haffejee, I. E., Raga, J., and Coovadia, H. M. (1984). Juvenile chronic arthritis in Black and Indian South African children. *South Africa Medical Journal*, **65**, 510–14.

Hall, P. (1991). Rheumatoid arthritis in sub-Saharan Africa. *Proceedings First AFLAR Congress of Rheumatology* (Abstract 17).

Hameed, K., Gibson, T., Kadir, M., Sultana, S., Fatima, Z., and Syed, A. (1995). The prevalence of rheumatoid arthritis in affluent and poor urban communities of Pakistan. *British Journal of Rheumatology*, **34**, 252–6.

Hanson, V. *et al*. (1977). Three subtypes of juvenile rheumatoid arthritis. *Arthritis and Rheumatism*, **20** (Suppl.),184–6.

Harris, E.N., Williams, E., Shah, D.J., and De Ceular, K. (1989). Mortality of Jamaican patients with systemic lupus erythematosus. *British Journal of Rheumatology*, **28**, 113–17.

Hart, H. H., Grigor, R. R., and Caughey, D. E. (1983). Ethnic difference in the prevalence of systemic lupus erythematosus. *Annals of the Rheumatic Diseases*, **42**, 529–32.

Harvey, J. *et al*. (1983). Rheumatoid arthritis in a Chippewa band. II. Field study with clinical serologic and HLA-D correlations. *Journal of Rheumatology*, **10**, 28–32.

Hill, A. S. V. *et al*. (1991). HLA class I typing by PCR: HLA-B27 and an African HLA-B27 subtype. *Lancet*, **337**, 640–2.

Hill, R. (1976). Juvenile arthritis in various racial groups in British Columbia. *Arthritis and Rheumatism*, **20** (Suppl.), 162–4.

Hochberg, M. C. (1984). *The epidemiology of juvenile rheumatoid arthritis: review of current status and approaches for future research*. In *Epidemiology of the rheumatic diseases*, (Proceedings of the Fourth International Conference, National Institutes of Health) (ed. R. C. Lawrence and L. E. Shulman), pp. 220–30. Gower Medical, London.

Hochberg, M. C. (1990). Systemic lupus erythematosus. *Rheumatic Disease Clinics of North America*, **16**, 617–39.

Hochberg, M. C., Linet, M. S., and Sills, E. M. (1983). The prevalence and incidence of juvenile rheumatoid arthritis in an urban black population. *American Journal of Public Health*, **73**, 1202–3.

Hochberg, M. C. *et al*. (1985). Systemic lupus erythematosus: a review of clinico-laboratory features and immunogenetic markers in 150 patients with emphasis on demographic subsets. *Medicine*, **64**, 285–95.

Jacobs, J. C. (1982). *Pediatric rheumatology for the practitioner*. Springer-Verlag, New York.

Jessop, S. and Meyers, O. L. (1973). Systemic lupus erythematosus in Cape Town. *South Africa Medical Journal*, **47**, 222–5.

Johnson, A. E., Gordon, C., Palmer, R. G., and Bacon, P. A. (1995). The prevalence and incidence of systemic lupus erythematosus in Birmingham, England. *Arthritis and Rheumatism*, **38**, 551–8.

Kanyerezi, B. R. (1969). Rheumatoid arthritis in Uganda. *East Africa Medical Journal*, **46**, 71–5.

Kanyerezi, B. R. and Mbidde, E. (1980). Juvenile chronic polyarthritis in Ugandan African children. *East Africa Medical Journal*, **57**, 484–9.

Karlson, E. W. *et al*. (1995). The independence and stability of socioeconomic predictors of morbidity in systemic lupus erythematosus. *Arthritis and Rheumatism*, **38**, 267–73.

Kato, H. *et al*. (1971). Rheumatoid arthritis and gout in Hiroshima and Nagasaki, Japan: a prevalence and incidence study. *Journal of Chronic Diseases*, **23**, 659–79.

Kellgren, J. H. (1962). Diagnostic criteria for population studies. *Bulletin of the Rheumatic Diseases*, **13**, 291–2.

Kelsey, K. L. (1982). Juvenile rheumatoid arthritis. In *Epidemiology of musculoskeletal disorders*, pp. 94–7. Oxford University Press.

Khuffash, F. A. and Majeed, H. A. (1988). Juvenile rheumatoid arthritis among Arab children. *Scandinavian Journal of Rheumatology*, **17**, 393–5.

Lau, E., Symmons, D., Bankhead, C., MacGregor, A., Donnan, S., and Silman, A. (1993). Low prevalence of rheumatoid arthritis in the urbanized Chinese of Hong Kong. *Journal of Rheumatology*, **20**, 1133–7.

Lawrence, J. S. (1961). Prevalence of rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **20**, 11–17.

Lawrence, J. S. and Wood, P. H. N. (1963). Criteria for rheumatoid arthritis in populations samples. In *Population studies of the rheumatic diseases* (ed. P. H. Bennett and P. H. N. Wood), pp 164–74. Excerpta Medica, Amsterdam,

Lawrence, J. S. *et al*. (1966). Rheumatoid arthritis in a subtropical population. *Annals of the Rheumatic Diseases*, **25**, 59–66.

Lawrence, R. C. *et al*. (1989). Estimates of the prevalence of selected arthritic and musculoskeletal disease in the United States. *Journal of Rheumatology*, **16**, 427–41.

McCarty, D. J., Thomas, S. M., Medsger, T. A., Ramsey-Goldman, R., LaPorte, R. E., and Kentworth, C. (1995). Incidence of systemic lupus erythematosus. *Arthritis and Rheumatism*, **38**, 1260–70

MacGregor, A. J., Riste, L. K., Hazes, J. M., and Silman, A. J. (1994). Low prevalence of rheumatoid arthritis in black-Caribbeans compared with whites in inner city Manchester. *Annals of the Rheumatic Diseases*, **53**, 293–7.

Malaviya, A. N., Kapoor, S. K., Singh, R. R., Kumar, A., and Pande, I. (1993). Prevalence of rheumatoid arthritis in the adult Indian population. *Rheumatology International*, **13**, 131–4.

Martinez-Cairo, C. S., Antonio, O. D., and Frati, A. (1978). Arthritis reumatoide juvenil. Estudio de 46 casos. *Boletin Medico del Hospital Infantil de Mexico*, **35**, 711–17.

Masi, A. T. (1984) Clinical and epidemiologic perspective of systemic lupus erythematosus. In *Epidemiology of the rheumatic diseases* (Proceedings of the Fourth International Conference, National Institutes of Health) (ed. R. C. Lawrence and L. E. Shulman), pp. 145–63. Gower Medical, London.

Mendez-Bryan, R., Gonzalez-Alcover, R., and Roger, L. (1964). Rheumatoid arthritis: prevalence in a tropical area. *Arthritis and Rheumatism*, **7**, 171–6.

Meyers, O. L., Daynes, G., and Beighton, P. (1977). Rheumatoid arthritis in a tribal Xhosa population in the Transkei, Southern Africa. *Annals of the Rheumatic Diseases*, **36**, 62–5.

Meyers, O. L. *et al*. (1983). Rheuma—eine Zivilisationskrankheit? *South Africa Medical Journal*, **101**, 1224–30.

Mijiyawa, M. (1995). Epidemiology and semiology of rheumatoid arthritis in third world countries. *Expansion Scientifique Francaise*, **62**, 121–6.

Mody, G. M. and Meyers, O. L. (1989). Rheumatoid arthritis in blacks in South Africa. *Annals of the Rheumatic Diseases*, **48**, 69–72.

Moolenburgh, J. D., Moore, S., Valkenburg, H. A., and Erasmus, M. G. (1984) Rheumatoid arthritis in Lesotho. *Annals of the Rheumatic Diseases*, **43**, 40–3.

Morrison, R. C. A. *et al*. (1990). Differences in systemic lupus erythematosus manifestations among racial groups in South Africa. *Arthritis and Rheumatism*, **33** (Suppl. B192), S104.

Morse, H. G. *et al*. (1980). High frequency of HLA-B27 and Reiter's syndrome in Navajo Indians. *Journal of Rheumatology*, **7**, 900–2.

Morton, R. O. *et al*. (1976). The incidence of systemic lupus erythematosus in North American Indians. *Journal of Rheumatology*, **3**, 186–9.

Muller, A. S. (1970). Population studies on the prevalence of rheumatic diseases in Liberia and Nigeria. Unpublished doctoral thesis, University of Leiden.

Nossent, J. C. (1992). Systemic lupus erythematosus in the Caribbean island of Curacao: an epidemiological investigation. *Annals of the Rheumatic Diseases*, **51**, 1197–201.

Nossent, J. C. (1993*a*). Clinical renal involvement in Afro-Caribbean lupus patients. *Lupus*, **2**, 173–6.

Nossent, J. C. (1993*b*). Course and prognostic value of systemic lupus erythematosus disease activity index in Black Caribbean patients. *Seminars in Arthritis and Rheumatism*, **23**, 16–21.

Oen, K. *et al*. (1986). Rheumatic disease in an Inuit population. *Arthritis and Rheumatism*, **29**, 65–74.

Oen, K., Fast, M., and Postl, B. (1995). Epidemiology of juvenile rheumatoid arthritis in Manitoba, Canada, 1975–92: cycles in incidence. *Journal of Rheumatology*, **22**, 745–50.

Pountain, G. (1991). The prevalence of rheumatoid arthritis in the Sultanate of Oman. *British Journal of Rheumatology*, **30**, 24–8.

Reveille, J. D. *et al*. (1989). DNA analysis of HLA-DR and DQ genes in American blacks with systemic lupus erythematosus. *Arthritis and Rheumatism*, **32**, 1243–51.

Ropes, M. W. *et al*. (1958). Revision of diagnostic criteria for rheumatoid arthritis. *Bulletin of the Rheumatic Diseases*, **9**, 175–6.

Rose, B. S. and Prior, I. A. M. (1963). A survey of rheumatism in a rural New Zealand Maori community. *Annals of the Rheumatic Diseases*, **22**, 410–15.

Rosenberg, A. M. and Petty, R. E. (1982). A syndrome of seronegative enthesopathy and arthropathy (SEA syndrome). *Arthritis and Rheumatism*, **25**, 1041–7.

Rosenberg, A. M. *et al*. (1982). Rheumatic diseases in Western Canadian Indian children. *Journal of Rheumatology*, **9**, 589–92.

Serdula, M. K. and Rhoads, G. G. (1979). Frequency of systemic lupus erythematosus in different ethnic groups in Hawaii. *Arthritis and Rheumatism*, **22**, 328–33.

Shichikawa, K. *et al*. (1981). A longitudinal population survey of rheumatoid arthritis in a rural district in Wakayama. *The Ryumachi*, **21** (Suppl.), 35–43.

Siegel, M. and Lee, S. L. (1973). The epidemiology of systemic lupus erythematosus. *Seminars in Arthritis and Rheumatism*, **3**, 1–54.

Silman, A. J. *et al*. (1993). Absence of rheumatoid arthritis in a rural Nigerian population. *Journal of Rheumatology*, **20**, 618–22

Singsen, B. H. (1990). Rheumatic disease of childhood. *Rheumatic Disease Clinics of North America*, **16**, 581–99.

Smikle, M. F., James, O. B., Barton, E. N., and De Ceulaer, K. (1995). Letter to the Editor. *Tropical and Geographical Medicine*, **47**, 231.

Solomon, L., Robin, G., and Valkenburg, H. A. (1975). Rheumatoid arthritis in an urban South African Negro population. *Annals of the Rheumatic Diseases*, **34**, 128–35.

Stein, M. *et al*. (1990). The spondyloarthropathies in Zimbabwe: a clinical and immunogenetic profile. *Journal of Rheumatology*, **17**, 1337–9.

Symmons, D. P. M. (1995). Occasional series: lupus around the world; frequency of lupus in people of African origin. *Lupus*, **4**, 176–8.

Szer, W., Sierakowska, H., and Szer, I. S. (1991). Antinuclear antibody profile in juvenile rheumatoid arthritis. *Journal of Rheumatology*, **18**, 401–8.

Tan, E. M. *et al*. (1982). The 1982 revised criteria for the classification of systemic lupus erythematosus. *Arthritis and Rheumatism*, **25**, 1271–7.

Taylor, H. G. and Stein, C. M. (1986). Systemic lupus erythematosus in Zimbabwe. *Annals of the Rheumatic Diseases*, **45**, 645–8.

Wallace, D. G. and Quismorio, F. P. Jr (1995). The elusive search for geographic clusters of systemic lupus erythematosus. *Arthritis and Rheumatism*, **38**, 1564–7.

Ward, M. M. and Studenski, S. (1990). Clinical manifestations of systemic lupus erythematosus. *Archives of Internal Medicine*, **150**, 849–53.

Wigley, R. D. *et al*. (1994). Rheumatic diseases in China: ILAR–China study comparing the prevalence of rheumatic symptoms in northern and southern rural populations. *Journal of Rheumatology*, **21**, 1484–90.

Williams, R. C. *et al*. (1985). GM allotypes in native Americans: evidence for three distinct migrations across the Bering land bridge. *American Journal of Physical Anthropology*, **66**, 1–19.

Williams, W. and Shah, D. (1990). Lupus nephritis at the University Hospital of the West Indies, Kingston, Jamaica: A 10-year experience. *Renal Failure*, **12**, 25–33.

Willkens, R. F. *et al*. (1981). Reiter's syndrome: evaluation of preliminary criteria for definite disease. *Arthritis and Rheumatism*, **24**, 844–9.

Wilson, W. A. and Hughes, G. R. V. (1979). Rheumatic disease in Jamaica. *Annals of the Rheumatic Diseases*, **38**, 320–5.

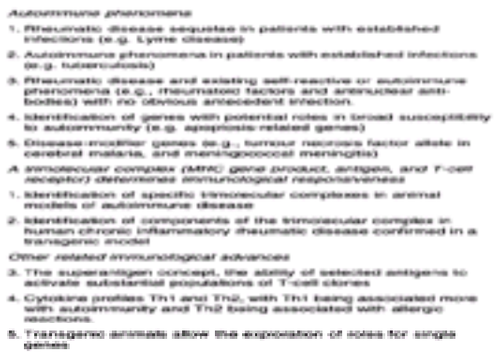
Woodrow, J. (1991). Genetics of the spondyloarthropathies. *Current Science*, **3**, 586–91.

5.2 Nosology of the chronic inflammatory rheumatic diseases

Patricia Woo and David N. Glass

Because the causes and pathological mechanisms of the various chronic inflammatory rheumatic diseases are incompletely understood, their classification has traditionally relied on their clinical evolution. Synovitis, a feature of many rheumatic diseases, is not helpful in differentiating between them; it is so common a response to different stimuli that it does not necessarily signal the presence of any given disease. However, features such as the pattern of arthritis (symmetrical or asymmetrical), age at onset, and presence or absence of systemic features, including specific organ involvement, may be used for classification criteria. In juvenile chronic arthritis (juvenile rheumatoid arthritis in North America), for example, the number of joints involved during the first 3 months of illness is an important element in classification (Brewer *et al*. 1977). Aided by the development of a few supportive but generally non-specific laboratory tests, investigators have subdivided the chronic inflammatory rheumatic illnesses in a manner that has proved to be of practical help in patient management. This approach is not so useful in the planning of clinical and laboratory-based investigations, nor in the exchange of information between investigators and clinicians. It necessitates the considerable use of exclusions based on the absence of the development of other diseases. The proper classification of a group of patients can take many years to accomplish as their individual diseases evolve.

Advances in a variety of disciplines, both clinical and experimental, have greatly improved our understanding of the immune response and of relations between infection, inflammation, and autoimmunity (Table 1). This new information can be used to classify the chronic inflammatory rheumatic diseases using a mechanistic/pathological approach (Vaughan 1989).



Autoimmune phenomena

1. Rheumatic disease sequelae in patients with established infections (e.g. Lyme disease)
2. Autoimmune phenomena in patients with established infections (e.g. tuberculosis)
3. Rheumatic disease and existing self-reactive or autoimmune phenomena (e.g., rheumatoid factors and antinuclear antibodies) with no obvious antecedent infection
4. Identification of genes with potential roles in broad susceptibility to autoimmunity (e.g. apoptosis related genes)
5. Disease-modifier genes (e.g., tumour necrosis factor allele in cerebral malaria, and meningococcal meningitis)

A trimolecular complex (MHC gene product, antigen, and T-cell receptor) determines immunological responsiveness

1. Identification of specific trimolecular complexes in animal models of autoimmune disease
2. Identification of components of the trimolecular complex in human chronic inflammatory rheumatic disease confirmed in a transgenic model

Other related immunological advances

3. The superantigen concept, the ability of selected antigens to activate substantial populations of T-cell clones
4. Cytokine profiles Th1 and Th2, with Th1 being associated more with autoimmunity and Th2 being associated with allergic reactions.
5. Transgenic animals allow the exploration of roles for single genes.

Table 1 Advances in knowledge on which the hypothesis for infectious–immunological causes is based

Relating chronic inflammatory rheumatic diseases to infection, immune responses, and autoimmunity

It has become evident, especially within the last decade, that a few infections may lead directly to chronic rheumatic disease. Prominent examples include Lyme arthritis (an infection with the spirochaete *Borrelia burgdorferi*), rheumatic fever with group A b-haemolytic streptococcal infection, and Reiter's syndrome, which is associated with a range of gastrointestinal and urogenital infections that are primarily bacterial.

Immunological phenomena, long believed to be involved in most chronic inflammatory rheumatic diseases, are now better documented and understood. Immunologically competent cells such as lymphocytes and plasma cells are universally present in diseased tissues, and immune complexes are often formed locally and within the vascular compartment. The local release of cytokines involved in generating these immune responses is important, as is the profile of cytokines produced (Th1/Th2 from CD4 cells, and Tc1 and Tc2 from CD8 cells). This profile may well reflect the nature of the pathogen and the route of immunization, as well as host factors (Marrack and Kappler 1994; Simon *et al*. 1994; Mossman and Sad 1996). It is also noteworthy that the selective inhibition of cytokines can alter the manifestations of the disease: for example, treatment with antibody to tumour necrosis factor-a in rheumatoid arthritis (Elliott *et al*. 1993).

Many of the immunological changes are autoimmune (self-reactive) in nature. At about the same time as rheumatoid factors and antinuclear antibodies with a wide range of specificities were discovered in patients with chronic rheumatic diseases (e.g., rheumatoid arthritis and systemic lupus erythematosus), similar autoimmune phenomena were noted in individuals with established, non-rheumatic, infectious diseases (e.g., chronic forms of bacterial endocarditis and leprosy). This discovery supported the notion that the autoimmune phenomena in patients with chronic rheumatic diseases may have originated from an infectious process.

Meanwhile, understanding of the basic mechanisms of the immune response has greatly improved. Particularly important is the concept of a central role for a trimolecular complex consisting of a T-cell receptor responsive to antigen presentation by proteins of the major histocompatibility complex (**MHC**) genes, which determine the host's response (Davis and Bjorkman 1988). This finding implies that at least two immunological components, HLA and T-cell receptors, underlie susceptibility to many autoimmune rheumatic diseases.

Substantial progress has been made in evaluating this trimolecular complex in animal models of autoimmune disease, most notably the model of allergic encephalomyelitis in which the specific determinants (molecular binding sites) on the antigen (myelin basic protein), the MHC genes, and T-cell receptor have all been identified (Acha-Orbea *et al*. 1988). The component of the trimolecular complex most extensively studied in human disease is HLA (Todd *et al*. 1988; see also Chapter 3.3). The direct role of at least one MHC gene (*HLA-B27*) in the spondylarthropathies has been elegantly confirmed in a transgenic model using rats raised in germ-free environments (Hammer *et al*. 1990; Taurog *et al*. 1994). Because certain HLA genes are associated with specific rheumatic diseases, HLA typing of patients can be used to aid the classification of their disease.

The study of the T-cell receptor in human autoimmune disease is just beginning, and is likely to prove complex. A limited T-cell repertoire has been documented in rheumatoid arthritis and T-cell receptor genes have been associated with the disease (McDermott *et al*. 1995). The role of the thymus in regulating the T-cell repertoire (and hence the formation of trimolecular complexes) has also been recognized; further study of this process will clarify the roles that the components of the trimolecular complex have in autoimmune disease. Expansion of certain V_b-chains of the T-cell receptor can be induced peripherally by superantigens. These are thought to make external contacts with the MHC class II molecule and the V_b portion of a T-cell receptor, thereby stimulating entire families of T cells. Bacterial superantigens in man are heat-resistant enterotoxins, such as *Staphylococcus aureus* enterotoxin, causing toxic-shock syndromes and septic arthritis as a result of the production of large quantities of inflammatory cytokines (Bremell and Tarkowski 1995). Murine mycoplasma arthritides superantigen (MAM) can trigger and exacerbate murine collagen-induced arthritis, thus providing a model for autoimmunity (Cole and Griffiths 1993).

Another mechanism of 'disturbance' in the T-cell repertoire is the alteration in programmed death (apoptosis) of cells of the immune system. In three strains of lupus-prone mice there are spontaneous mutations of Fas apoptosis antigen, its receptor or its ligand, leading to persistence of autoreactive cells (reviewed in Mountz *et al*. 1994). Fas mutations have been described in human lymphoproliferative syndrome, which children also have features of autoimmune haemolytic anaemia (Fisher *et al*. 1995; Rieux-Laucat *et al*. 1995). In human systemic lupus erythematosus, expression of Fas antigen and Bcl-2 have been studied, but there is no consensus as yet as to their potential roles in immunopathogenesis.

These conceptual advances, discussed more fully in earlier chapters, suggest that many of the chronic rheumatic diseases have infectious causes with complicating immunological responses. In devising a classification that proposes or infers infectious causes complicated by a range of immune responses, several other factors need to be taken into account. While a substantial amount is known about the HLA genes and infectious agents in some arthritides, there is a paucity of knowledge concerning T-cell receptor genes and general autoimmune-predisposing genes. Any classification based on these variables must, therefore, be considered tentative. During the evolution of the disease, the immune response may be directed initially at antigen-recognition sites (epitopes) on the intact infecting organism. Later, antigenic fragments of the organism may persist, providing continuing immunopathological stimulation; the immune response may also become directed against self-antigens. Thus any classification must recognize that clinical stages of the disease may change as a result of changing immune response. Routes of inoculation as well as host genetic factors are responsible for determining whether the immune response will become chronic and the disease progressive. Because the evolution of disease, from a straightforward infection to a complex series of immune responses associated with a range of pathologies, has much to do with genetically defined host variables, genetic studies of patients can help the process of disease classification. The above considerations make it possible to devise classifications based on infectious/immunological causes, even when the likelihood of infectious processes can only be inferred from parallel circumstances in which the presence of infection is well established.

Classification of disease

For the purposes of this discussion, the chronic inflammatory rheumatic diseases can be divided into overlapping categories (see Table 2). The first category comprises the relatively rare diseases in which the products of infection-induced inflammation directly cause tissue damage. The harmful products of inflammation include proteolytic enzymes that are associated with the degradation of synovium and cartilage. Such diseases usually take an acute rather than a chronic course. Pyogenic arthritis, in which the products of the inflammatory process cause the destruction of substantial amounts of synovial and cartilaginous tissue, is an example. Other arthropathies caused by infectious agents include those of rubella and parvovirus, mycobacterial and mycoplasma infections.

Table 2 Categories in the classification of chronic inflammatory rheumatic diseases

Because pyogenic organisms do not normally seed to joints, unusual conditions conducive to rheumatic disease must be present. Local injury or disease may make individual joints especially vulnerable and systemic factors may affect the host's susceptibility to a given pathogen. Some form of mechanical problem in a joint, such as osteoarthritis, is common among these patients. Patients with rheumatoid arthritis, for example, have a greater risk of infection than those with osteoarthritis. Their disease, rheumatoid arthritis, is associated with impaired immune responses, which treatment may aggravate. In addition, arthroplasties carry a particular risk of infection, especially in immunocompromised patients. Therefore, in the patient with rheumatoid arthritis who undergoes an arthroplasty, both local and systemic factors may favour the development of bacterial arthritis (Goldenberg and Reed 1985). Systemic illnesses, for example diabetes mellitus, appear to confer a greater susceptibility (see Chapter 5.3.1). The majority of individuals with septic joints have at least one of these predisposing factors. Less common, but more specific, predisposing factors are host difficulties in handling infectious organisms, as in complement deficiency, hypogammaglobulinaemia, and deficiency of lymphocyte function antigen 1 (Atkinson 1995).

In the second category of rheumatic diseases an immune response directed primarily against a pathogenic organism or fragments of that organism causes inflammation that results in disease. Much of the tissue injury results from immune complexes. Cross-reactivity to self-antigens is not critical. Some rheumatic syndromes associated with streptococcal infection and with hepatitis B virus are typical of this category of illness. The intact infecting organism may sometimes be isolated from the involved joint and provide a source of initial tissue damage, as in category 1, but in other instances the organisms cannot be isolated. In these conditions an immune response directed specifically at antigenic fragments of the organism is likely to be responsible for some components of the pathology of the disease. Such organism-specific responses are well documented. For example, in post-streptococcal syndromes such as glomerulonephritis and vasculitis, immune responses in the form of immune-complex deposition contribute to the inflammatory process. Similarly, in hepatitis syndromes the increased amounts of immune complexes and reduced amounts of complement argue for their inclusion in the second category of diseases, as may the presence of hepatitis antigen in the immune complexes of some patients with vasculitis. Even if persistence of the intact organism is not essential, antigenic fragments may be important in generating a chronic immune response.

Characteristics of the host that affect his or her susceptibility to this category of disease are, at least as presently understood, not organism-specific, and have been identified in a small proportion of individuals at risk. They include some complement, immunoglobulin, and T-cell receptor deficiencies as well as defects of immune-complex clearance.

In a third category of disease, components of the first and second are present, that is, infection causes damage directly and results in an organism-specific immune response. In addition, autoimmunity develops in the form of reactivity with self-antigens including HLA. The best-documented example is probably rheumatic fever, in which the immune response to streptococcal infection becomes directed against host tissue antigens in cardiac muscle (see Chapter 5.3.12). Other diseases likely to be included are the spondylarthropathies, in which it is probable that initial infections with Gram-negative organisms and their plasmids are followed by an immune response directed against the self-antigen HLA-B27 in addition to components of the organisms involved; that is, molecular mimicry (Granfors *et al*. 1989) where the two antigens have common epitopes. We suggest that this theory of the development of HLA-B27-associated rheumatic disease will serve as a model for diseases in which infection is suspected as an initiating event but the organisms are as yet unknown (see below). Explanations other than cross-reactivity or molecular mimicry for the HLA-B27 association with spondylarthropathies are conceivable (Sieper and Braun 1995). Because certain class I or class II HLA genes are associated with each disease in the third category, people who do not carry the implicated gene are not susceptible to the diseases. The particular genes are necessary for the disease to develop but probably do not initiate the disease process.

In an extension of this category are diseases in which infectious agents have not been identified but a strong, autoimmune disease phenotype probably results from the infectious process. The more common chronic inflammatory rheumatic diseases, including rheumatoid arthritis, systemic lupus erythematosus, dermatomyositis and polymyositis, fall into this category. The major laboratory findings in patients in this fourth category are immune responses, and most of the antigenic determinants identified to date are self-antigens. The earliest documented examples of these responses were rheumatoid factors found in rheumatoid arthritis and antinuclear antibodies found in systemic lupus erythematosus. The antibodies may contribute to the pathogenesis, as in congenital complete heart block in the new-born of some patients with systemic lupus erythematosus (Buyon and Winchester 1990). Their presence in these patients, given the overall hypothesis that many of these diseases are infections, has parallels with the presence of the same autoantibodies in patients with various chronic inflammatory infections such as tuberculosis, leprosy, and bacterial endocarditis (in which rheumatoid factors are particularly prominent).

Efforts to identify the infectious agents believed to precipitate these chronic rheumatic diseases have continued to be made. Candidate organisms include Epstein–Barr virus and *Proteus mirabilis* in rheumatoid arthritis, and coxsackie virus in dermatomyositis. It is likely, however, that several agents can precipitate the disease, given the same genetic susceptibility. The host's MHC gene products are important in pathogenesis. Certain HLA class I and class II genes are associated with these diseases, and MHC class III genes may be involved in systemic lupus erythematosus in some populations. Whether the HLA genes are themselves directly involved or serve as markers for MHC non-HLA genes involved in the pathogenesis is unknown. Additional MHC-region genes have been identified: these include those coding for tumour necrosis factors, heat-shock protein, and intracellular transporter genes, any of which may be critical to the disease (see Chapter 4.7).

The concept of a trimolecular complex, discussed above, indicates a role for T-cell receptor genes. T-cell receptors have been implicated in the collagen type-II model of an autoimmune-induced stimulus (Banerjee *et al*. 1988) and in other forms of autoimmunity induced in animals (Acha-Orbea *et al*. 1988). Currently, however, very few T-cell receptor genes have been associated with human disease, their analysis being more complex than that of HLA.

The identification of defects in apoptotic pathways and their association with autoimmunity in both experimental models and in man adds to the importance of the host element in determining the outcome of an infection in terms of subsequent autoimmunity (Watanabe-Fukunaga *et al*. 1992). It is also noteworthy that single cytokine genes in transgenic animals can create an *in vivo* environment in which arthritis develops (Keffer *et al*. 1991).

Some implications of an infection–immunological classification

Within these broad concepts, an infection would lead to the second and third categories of disease, that is, organism-specific, and then autoimmune responses could develop over time. The occurrence of these diseases in some individuals but not in others needs to be explained. Diseases in categories 2–3 can be regarded as failures of the hosts' initial defence mechanisms. Inherited susceptibility to disease has been suspected for a long time and some of the genetic influences on particular diseases are known, as is clear from the preceding discussion of HLA genes. The multiplicity of genes involved in addition to HLA suggests that autoimmunity is a complex genetic trait, as is the case in insulin-dependent diabetes mellitus (Davies *et al*. 1994). Microsatellite analysis of the human genome and immune models of diabetes mellitus illustrate the complexity of disease-susceptibility genes. It is clear that certain genes are necessary for disease to develop and that others act as 'modifiers,' responsible for the severity of disease, for example the tumour necrosis factor-a $_2$ allele in cerebral malaria (McGuire *et al*. 1994) and meningococcal meningitis (Girrourdin *et al*. 1992). Other unexplored factors are the virulence of the infecting organisms: for example, adhesive properties of bacteria can be involved in determining the outcome of bacterial infections. Exposure to infectious agents when the immune system is developing may be a cause for polyarthritis. For example, recent evidence suggests that an influenza A $_2$ epidemic caused an intrauterine infection that predisposed children to the development of a subtype of juvenile chronic arthritis (Pritchard *et al*. 1988). Such host/pathogen relations need to be investigated for their contributions to the chronic inflammatory rheumatic diseases.

The concept of a primary infectious origin for the chronic rheumatic diseases does not include discussion of other antigens known to generate disease (e.g., procainamide-induced systemic lupus erythematosus). Nevertheless, the infectious concept is applicable to many chronic rheumatic diseases, as well as other chronic (autoimmune) inflammatory diseases affecting other organs systems such as the gastrointestinal and renal tracts Further investigation should be promoted because it can be used to develop testable hypotheses aimed at identifying the types of immune responses, the infecting organisms, and genetic predisposing factors causing the disease. In addition, rationales for new therapeutic strategies are also provided (Table 3).

T-cell receptor immunization V_b17 in rheumatoid arthritis
Induction of oral tolerance (e.g., to collagen II in rheumatoid arthritis)
Use of antibiotics, sulphasalazine/tetracycline therapy in rheumatoid arthritis and spondylarthropathy

Table 3 New therapeutic strategies

Chapter References

Acha-Orbea, H. *et al*. (1988). Limited heterogeneity of T cell receptors from lymphocytes mediating autoimmune encephalitis allows specific immune intervention. *Cell*, **54**, 263–73.

Atkinson, J. P. (1995). Some thoughts on autoimmunity. *Arthritis and Rheumatism*, **38**, 301–5.

Banerjee, S., Haqqi, T. M., Luthra, H. S., Stuart, J. M., and David, C. S. (1988). Possible role of V T cell receptor genes in susceptibility in collagen-induced arthritis in mice. *Journal of Experimental Medicine*, **167**, 832.

Bremell, T. and Tarkowski, A. (1995). Preferential induction of septic arthritis and mortality by superantigen producing staphylococci. *Infection and Immunity*, **63**, 4185–7.

Brewer, E. J. *et al*. (1977). Current proposed revision of JRA criteria. *Arthritis and Rheumatism*, **20**, 195–9.

Buyon, J. P. and Winchester, R. (1990). Congenital complete heart block. A human model of passively acquired autoimmune injury. *Arthritis and Rheumatism*, **33**, 609–14.

Cole, B. C. and Griffiths, M. M. (1993). Triggering and exacerbation of autoimmune arthritis by the mycoplasma arthritides superantigen MAM. *Arthritis and Rheumatism*, **36**, 994–1002.

Davis, M. M. and Bjorkman, P. J. (1988). T-cell antigen receptor genes and T-cell recognition. *Nature*, **334**, 395–402.

Davies, J. L. *et al*. (1994). A genome-wide search for human type I diabetes susceptibility genes. *Nature*, **371**, 130–6.

Elliott, M. J. *et al*. (1993). Treatment of rheumatoid arthritis with chimeric monoclonal antibodies to tumor necrosis factor a. *Arthritis and Rheumatism*, **36**, 1681–90.

Fisher, G. H. *et al*. (1995). Dominant interfering *Fas* gene mutations impaired apoptosis in a human autoimmune lymphoproliferative syndrome. *Cell*, **81**, 935–46.

Girrourdin, E., Roux-Lombard, P., Grau, G. E., Suter, P., Gallati, H., and Dayer, J. M. (1992). Imbalance between tumour necrosis factor alpha and soluble TNF receptor concentrations in severe meningococcaemia. *Immunology*, **76**, 20–3.

Goldenberg, D. L. and Reed, J. I. (1985). Bacterial arthritis. *New England Journal of Medicine*, **312**, 764–71.

Granfors, K. *et al*. (1989). Yersinia antigens in synovial fluid cells from patients with reactive arthritis. *New England Journal of Medicine*, **320**, 216–21.

Hammer, R. E., Malka, S. D., Richardson, J. A., Tang, J. P., and Taurog, J. D. (1990). Spontaneous inflammatory disease in transgenic rats expressing HLA-B27 and human b $_2$M: an animal model of HLA-B27-associated human disorders. *Cell*, **63**, 1099–112.

Keffer, J. *et al*. (1991). Transgenic mice expressing human tumour necrosis factor: a predictive genetic model of arthritis. *EMBO Journal*, **10**, 4025–31.

McDermott, M. *et al*. (1995). The role of T-cell receptor b chain genes in susceptibility to rheumatoid arthritis. *Arthritis and Rheumatism*, **39**, 91–5.

McGuire, W., Hill, A. V., Allsopp, R. E., Greenwood, B. M., and Kwiakowski, D. (1994). Variation in the TNF alpha promoter region associated with susceptibility to cerebral malacia. *Nature*, **371**, 508–10.

Marrack, P. and Kappler, J. (1994). Subversion of the immune system by pathogens. *Cell*, **76**, 323–32.

Mosmann, T. R. and Sad, S. (1996). The expanding universe of T-cell subsets: Th1, Th2 and more. *Immunology Today*, **17**, 138–46.

Mountz, J. D., Wu, J., Cheng, J., and Zhou, T. (1994). Autoimmune disease: a problem of defective apoptosis. *Arthritis and Rheumatism*, **37**, 1415–20.

Pritchard, M. H., Munro, J., and Matthews, N. (1988). Antibodies to influenza A in a cluster of children with juvenile chronic arthritis. *British Journal of Rheumatology*, **27**, 176–80.

Rieux-Laucat, F. *et al*. (1995). Mutations in *fas* associated with human lymphoproliferative syndrome and autoimmunity. *Science*, **268**, 1347–9.

Sieper, J. and Braun, J. (1995). Pathogenesis of spondyloarthropathies. Persistent bacterial antigen, autoimmunity or both? *Arthritis and Rheumatism*, **38**, 1547–54.

Simon, A. K., Seipelt, E., and Sieper, J. (1994). Divergent T-cell cytokine patterns in inflammatory arthritis. *Proceedings of the National Academy of Sciences (USA)*, **91**, 8562–6.

Taurog, J. D. *et al*. (1994). The germ-free state prevents development of gut and joint inflammatory disease in HLA-B27 transgenic rats. *Journal of Experimental Medicine*, **180**, 2359–64.

Todd, J. A. *et al*. (1988). A molecular basis for MHC class II-associated autoimmunity. *Science*, **240**, 1003–9.

Vaughan, J. H. (1989). Infection and rheumatic diseases: a review. *Bulletin on the Rheumatic Diseases*, **39**, 1–7; 1–80.

Watanabe-Fukunaga, R., Brannan, C. I., Copeland, N. G., Jenkins, N. A., and Nagata, S. (1992). Lymphoproliferation disorder in mice explained by defects in Fas antigen that mediates apoptosis. *Nature*, **356**, 314–17.

5.3.1 Pyogenic arthritis in adults

Marc L. Miller

Septic arthritis caused by pyogenic bacteria is a true medical emergency. Prompt recognition of an infected joint and an immediate start to proper treatment are essential for a good outcome. Despite advances in antimicrobial therapy and in surgical approaches to pyogenic arthritis, the death rate in several large series has ranged up to 10 per cent (Rosenthal *et al*. 1980; Dubost *et al*. 1993), and up to one-third of cases are affected by residual functional impairment, persistent pain, or other complications (Goldenberg and Cohen 1976; Sharp *et al*. 1979; Cooper and Cawley 1986). Therefore, it is important to appreciate the clinical presentations of pyogenic arthritis, underlying risk factors and associated diseases, range of infectious agents, diagnostic techniques, and appropriate medical and surgical treatments.

This chapter will discuss pyogenic bacterial infections in adults only. Paediatric infections will be addressed in Chapter 5.3.2.

Pathophysiology

Bacteria can reach the joint by one of three routes:

1. by haematogenous spread from a distant infected site;
2. by direct penetration through the skin to the joint space;
3. by direct spread from a contiguous infected site.

Haematogenous seeding is by far the most common route for pyogenic joint infection. In the setting of bacteraemia, organisms leave the bloodstream and lodge in the synovium. The synovium is a highly vascular tissue without a limiting basement membrane to block bacterial access to the synovial fluid. However, factors other than the mere presence of bacteria in the bloodstream must be involved as most bacteraemic episodes do not result in septic arthritis and some bacteria that are frequent causes of bacteraemia are rare causes of joint infections. One of these factors is structural changes in the joint. Joints affected by a variety of chronic arthritides (including rheumatoid arthritis, osteoarthritis, Charcot's arthropathy, and gout or pseudogout) and prosthetic joints are at increased risk of pyogenic infections. The presence of synovitis, effusion, granulation tissue, or foreign material increase the likelihood of bacterial colonization (Goldenberg and Reed 1985; Youssef and York 1994).

Characteristics of the infecting organisms are also important. Those organisms that most commonly cause pyogenic arthritis, staphylococci and *Neisseria gonorrhoeae*, have an enhanced ability to adhere to synovial tissue or produce toxins that facilitate colonization (Switalski *et al*. 1993). These organisms are able to infect normal joints in normal hosts. Those organisms with low virulence for joint infection, such as Gram-negative bacilli, typically cause infections in the setting of a structurally abnormal joint or an alteration in host immune defences.

A third contributing factor is alteration in host immunity. This may be due to inherited or congenital immunodeficiency states involving granulocyte function, immunoglobulin production or complement deficiency, or to acquired immunodeficiency due to immunosuppressive therapy or to disease states such as cancer, diabetes, chronic liver disease, or human immunodeficiency virus infection (Rivera *et al*. 1992).

Direct entry of bacteria into the joint space can occur as the result of penetrating trauma, arthrocentesis, arthroscopy, or open surgical procedures such as arthroplasty. The incidence of infection following arthrocentesis is very low, less than 0.06 per cent in two large series (Hollander 1969; Gray *et al*. 1981). The risk of joint infection after arthrocentesis may be increased when an area of cellulitis overlies the joint or in the setting of bacteraemia, in which cases bacteria can be carried by the aspirating needle into the joint space. Persistent drainage from surgical portals should always raise the suspicion of postarthroscopic infection.

Direct spread from an adjacent focus of infection occurs in several clinical settings. Pyogenic arthritis may result from an adjacent osteomyelitis in adults, although this complication is rare, owing to the infrequency of untreated osteomyelitis in the antibiotic era. In adults, but not in children, there is an anastomosis between the metaphyseal and synovial vascular beds allowing direct entry of bacteria from an osseous focus of infection into the joint (Atcheson and Ward 1978). Direct extension of an enteric fistula in inflammatory bowel disease, of a psoas abscess, or of a gluteal abscess in a paraplegic can result in pyogenic arthritis of the hip.

Once infection is established bacteria multiply and spread throughout the synovium and eventually into the synovial fluid. Organisms are phagocytosed by polymorphonuclear leucocytes, resulting in the release of chemotactic factors and activation of complement. Additional phagocytic cells are recruited, resulting in a classic inflammatory reaction. Bacterial products and lysosomal proteolytic enzymes released from leucocytes stimulate the degradation of cartilage. In animal models, the proteoglycan content of cartilage is reduced by 40 per cent within 48 h of the induction of joint infection. Within 2 to 3 weeks, significant loss of collagen develops (Riegels-Nielsen *et al*. 1987). This rapid destruction underscores the importance of prompt detection and initiation of treatment, one aspect of which includes the removal of the products of bacterial and leucocyte degradation by draining the joint. Chronic inflammatory tissue, resembling rheumatoid pannus, that further erodes cartilage and subchondral bone may develop in untreated infections.

Non-gonococcal arthritis

Clinical manifestations

The clinical manifestations, epidemiology, and natural history of gonococcal and non-gonococcal arthritis are sufficiently distinct (Table 1) that infections due to *Neisseria* spp. will be discussed separately.

Gonococcal	Non-gonococcal
Most often affects healthy, sexually active, young adults	Most often affects the very young or very old. Underlying joint or other medical conditions
Females more often than males	Males more often than females
Hip uncommonly affected	Hips involved in 29%
Migratory polyarthralgias common	Polyarthralgias uncommon
Rash, tenosynovitis common	Extra-articular manifestations common
Synovial fluid:	Synovial fluid:
Gram's stain positive, 25%	Gram's stain positive, 50–60%
culture positive, 50%	culture positive, 90%
lactate not elevated	lactate elevated
Rapid response to therapy	Response often slow; may require surgical drainage
Full recovery in most cases	10% mortality; one-third with residual joint damage

Table 1 Distinguishing features of gonococcal and non-gonococcal arthritis

The typical patient with non-gonococcal bacterial arthritis presents with acute onset of pain, swelling, erythema, warmth, and tenderness of the infected joint (Baker and Schumacher 1993). There may be a tense effusion. The patient usually guards the joint against any movement. The severity of pain, tenderness, and swelling is generally greater than in other causes of joint inflammation and should immediately raise the suspicion of infection. Very rarely, patients present with a subacute or even chronic course and a paucity of inflammatory signs. This presentation is more commonly associated with less virulent organisms.

Localization of pain and tenderness may be difficult, and erythema, warmth, and swelling are not detectable in deep-seated joints. For instance, hip infections produce pain deep in the buttock, in the anterior thigh, or even in the knee, while pain from the sacroiliac joint may be referred to the buttock or posterior thigh and may mimic sciatica.

Systemic signs and symptoms of infection are common but not invariably present. In one large series, 90 per cent of patients with acute bacterial arthritis presented with a temperature of at least 37.8°C (Goldenberg and Cohen 1976); in another, fever was found in only 56 per cent (Rosenthal *et al*. 1980). Marked fever (greater than 39°C) was found in two series in 25 and 39 per cent, respectively (Newman 1976; Cooper and Cawley 1986). Therefore, while the presence of high fever in a patient with an acute arthritis should always raise the suspicion of a pyogenic infection, the absence of fever does not preclude the diagnosis.

The patient may also show signs and symptoms of an infection at a distant site from the joint, such as the skin, respiratory tract, or urinary tract.

While any joint is potentially susceptible to pyogenic infection, the knee is the site in more than half of cases and the hip in another 20 per cent. The remainder of cases is divided largely among the shoulder, wrist, elbow, and ankle (Manshady *et al*. 1980; Rosenthal *et al*. 1980; Cooper and Cawley 1986). Rarely the small joints of the fingers and toes or the axial joints (sacroiliac, sternoclavicular, sternomanubrial, or pubic symphysis) are involved. Typically, pyogenic infections cause monarticular arthritis, but in approximately 20 per cent of cases bacteraemia can lead to polyarticular involvement (Epstein *et al*. 1986; Dubost *et al*. 1993). Certainly a polyarticular presentation should not deter a search for infection in a patient with acute onset of illness, signs of toxicity, or a distant site of infection.

Laboratory studies

Patients with pyogenic arthritis commonly have non-specific laboratory signs of acute infection and inflammation. About half will have a peripheral leucocyte count greater than 10 000/mm^3. The erythrocyte sedimentation rate is generally elevated. In one series, the erythrocyte sedimentation rate was greater than 20 mm/h in 117 of 118 cases of pyogenic arthritis. In 76 of these, the rate was greater than 50 mm/h and in 30 cases greater than 100 mm/h (Newman 1976). Similarly, other acute phase reactants such as C-reactive protein may be elevated. The response of the acute phase reactants parallels the course of the infection and is a good indicator of response to therapy. Blood cultures are positive in about one-third of cases (Rosenthal *et al*. 1980).

The critical test in the evaluation of a patient with suspected pyogenic arthritis is the analysis of synovial fluid. The aspiration of frank pus immediately suggests bacterial infection, although joint fluid in chylous effusions associated with long-standing rheumatoid arthritis or subchondral fractures may have a similar appearance. Effusions due to gout or pseudogout can also at times appear grossly purulent.

The leucocyte count in synovial fluid is greater than 50 000/mm ^{3} in most cases and often exceeds 100 000/mm^3 (Krey and Bailen 1979). The differential count will show greater than 90 per cent polymorphonuclear leucocytes. When the count is less than 50 000/mm ^{3}, repeat aspiration in the next 24 to 48 h will show an increase to over that figure. Where the infection has been partially treated with antibiotics before aspiration, the count may never exceed 50 000/mm ^{3}. In about half of cases, the glucose in synovial fluid will be decreased to 20 to 40 per cent of that of the simultaneous serum glucose. However, similarly low levels are commonly found in rheumatoid effusions, which diminishes the differential diagnostic value of the synovial fluid glucose in rheumatoid patients (Goldenberg and Reed 1985). A markedly elevated level of lactic acid in the synovial fluid is associated with non-gonococcal pyogenic arthritis but not with non-infectious causes of inflammatory effusions (Riordan *et al*. 1982). The concentration of lactic acid may be of diagnostic help before cultures of synovial fluid are available or in cases of partially treated infections in which synovial fluid cultures remain negative.

Gram's stain of synovial fluid is positive in 50 to 65 per cent of all cases of non-gonococcal arthritis, approaching 75 per cent in staphylococcal infections but only about 50 per cent in infections caused by Gram-negative bacilli (Goldenberg and Cohen 1976; Newman 1976). Bacteria can be cultured from synovial fluid in nearly all cases of non-gonococcal arthritis. The occasional negative results can be attributed to preceding antibiotic therapy, poor handling of specimens, failure to culture for anaerobic organisms, or theoretically to the localization of organisms to the synovium rather than the fluid during the very early stage of infection. The yield from culture of synovial fluid may be improved by inoculating large volumes of it directly into bottles of blood culture media (Von Essen and Holtta 1986).

Imaging studies

Plain radiographs have limited value in evaluation of the suspected infected joint. Most commonly only periarticular soft-tissue swelling or a joint effusion will be present. The primary value of the radiograph is to demonstrate signs of other causes of acute pain and swelling, such as fracture, chondrocalcinosis, gouty erosions, or avascular necrosis, and to look for underlying chronic joint disease or accompanying osteomyelitis. Plain radiographs can be used to assess the extent of joint damage and cartilage loss in cases in which the diagnosis and treatment of infection have been delayed. The radiographic changes indicative of joint damage may lag behind the clinical signs of infection and may continue to appear even after there has been a response to therapy. Conventional tomography may better demonstrate early bone destruction, particularly in hard-to-visualize joints such as the sternoclavicular.

Except in a few clinical situations, radionuclide scanning also has limited value in the assessment of the suspected infected joint. While pyogenic arthritis of peripheral joints is readily detectable on physical examination, infections of the axial joints are often difficult to localize from the patient's description or by physical examination. In these cases, a technetium-99m bone scan can demonstrate increased activity that suggests involvement of a deep-seated joint (Gordon and Kabins 1980; Vyskocil *et al*. 1991). For example, in the evaluation of the patient with signs of infection and complaining of diffuse low back or buttock pain, the technetium scan may identify involvement of the sacroiliac joint. This may help to differentiate joint sepsis from intra-abdominal or retroperitoneal abscess, or from osteomyelitis of the pelvis. The technetium bone scan, however, does not differentiate between infectious and non-infectious inflammatory arthritis and therefore cannot distinguish an infected sacroiliac joint from sacroiliitis due to Reiter's syndrome, or an infected from an uninfected but actively inflamed peripheral joint in rheumatoid arthritis.

The accuracy of detecting joint infections with radionuclide scanning can be improved by scanning with either gallium-67- or indium-111-labelled leucocytes. Gallium scans used in sequence after a positive technetium scan may increase the specificity of technetium scanning. The gallium scan may become positive earlier than the technetium in poorly vascularized joints such as the sacroiliac or sternoclavicular, making it a useful test to consider even when the technetium scan is negative (Lopez-Longo *et al*. 1987). False-positive gallium scans have been reported in rheumatoid arthritis, thereby limiting the usefulness of this technique in evaluating the possibility of joint sepsis in that condition. Gallium and indium scans are of most value in suspected infection of a prosthetic joint, where they may distinguish an infected from an uninfected, painful, loose prosthesis. Although indium scanning does require the additional steps of removing, labelling, and then reinfusing the patient's leucocytes, it is more sensitive and specific than sequential technetium/gallium scanning and results are available at 24 h rather than 72 h as with gallium scanning (Merkel *et al*. 1985).

Conventional fluoroscopy or computed tomographic scanning can be of value in guiding diagnostic aspiration of deep-seated joints such as the hip or sacroiliac joint. While computed tomography and magnetic resonance imaging are seldom necessary in the evaluation of septic arthritis, these imaging techniques can be of great value when an infected joint fails to respond as expected to appropriate treatment because of adjacent osteomyelitis or a persistent collection of undrained pus. These techniques can also demonstrate soft tissue masses adjacent to axial joints, so indicating the need for abscess drainage (Wohlgethan *et al*. 1988).

Bacteriology

A wide variety of organisms has been cultured from infected joints (Table 2). *Staphylococcus aureus* is by far the most common cause of non-gonococcal arthritis, accounting for 40 to 50 per cent of cases in adults. *Staph. epidermis* is found in 10 to 15 per cent, various streptococcal species in 20 per cent, and various Gram-negative bacilli in 15 per cent of cases. The incidence of pneumococcal joint infections has been drastically reduced in the antibiotic era, reflecting the more effective treatment of pneumococcal infections of the respiratory tract and the decreased incidence of pneumococcal bacteraemia. *Haemophilus influenzae*, the organism most commonly found in pyogenic infections in children under the age of 2 years, accounts for only about 2 per cent of cases in adults (Goldenberg and Cohen 1976; Newman 1976; Sharp *et al*. 1979; Rosenthal *et al*. 1980; Goldenberg and Reed 1985; Cooper and Cawley 1986; Mikhail and Alarcon 1993; Youssef and York 1994). Anaerobic infections, often polymicrobial, occur in about 5 per cent of cases. Case reports have documented the unusual occurrence of arthritis due to *Listeria monocytogenes, Streptobacillus moniliformis*, and *Pasteurella multocida*. While *Staph. aureus* is the most common cause of both monarticular and polyarticular presentations, streptococci and *H. influenzae* are more frequently associated with polyarticular infections than would be expected by the overall incidence of joint infections with these organisms (Epstein *et al*. 1986; Dubost *et al*. 1993).

Organism	Percentage of cases
Staphylococcus	
aureus	40–50
epidermis	10–15
Streptococcal species	20
Gram-negative bacteria	15
S. pneumoniae	2
H. influenzae	2
Anaerobes	5

Table 2 Infecting organisms in non-gonococcal arthritis in adults

A variety of streptococcal species, most commonly group A b-haemolytic, but also group B, group C, group G, enterococcus, *Streptococcus milleri* and *Strep. viridans* have been found in infected joints. Group A streptococci are common inhabitants of the skin and respiratory tract and can seed the joints of normal hosts. Group B streptococci, usually associated with neonatal and puerperal sepsis, have been a cause of adult pyogenic arthritis, particularly in patients with prosthetic joints or predisposing medical conditions such as diabetes mellitus. About 40 per cent of reported cases have been diabetics, in whom the death rate was about 15 to 20 per cent. Group B streptococcal joint infections have been associated with significant complications in the form of residual damage and limited range of movement (Small *et al*. 1984; Pischel *et al*. 1985). Similarly, group C and G streptococcal infections appear to affect patients with underlying illnesses, and have been polyarticular in 30 per cent of reported cases (Ike 1990; Burhert and Watanakunakorn 1991). Enterococci are not usually reported as joint pathogens, despite their frequent association with infections of the biliary and urinary tracts, and with bacteraemia. Affected patients generally have either predisposing conditions or pre-existing joint disease.

H. influenzae is an encapsulated organism requiring opsonizing antibody and an effective reticuloendothelial system for optimal clearance. Patients with *H. influenzae* arthritis therefore often have predisposing conditions such as alcoholism, hypogammaglobulinaemia, multiple myeloma, systemic lupus erythematosus, or asplenia. Extra-articular *H. influenzae* infections are frequent, including sinusitis, pneumonia, meningitis, pharyngitis, or cellulitis. Almost half of reported cases have been polyarticular (Borenstein and Simon 1986).

Gram-negative bacilli have emerged as a more frequent cause of septic arthritis in the past two decades. This probably reflects the prolonged survival of patients with serious medical illnesses, the increased use of immunosuppressive drugs, and the increase in intravenous drug abuse. Gram-negative joint infections generally occur in two distinct clinical settings. The first involves elderly patients with underlying systemic illnesses or pre-existing joint disease. Here the presentation is usually abrupt, with frequent fevers and rigors. The most common organism is *Escherichia coli* (Newman *et al*. 1988). Also reported with some frequency are *Pseudomonas aeruginosa, Proteus mirabilis, Serratia marcesens*, and *Klebsiella* '*pneumoniae*'. Scattered case reports document rare joint infections with a number of Gram-negative organisms including *Aeromonas hydrophila, Acinetobacter* spp., *Arizona hinshawii, Eikenella corrodens, Enterobacter* spp., and *Kingella kingae*; these reports indicate the importance of bacteriological identification by culture in patients with suspected Gram-negative infections.

The second group consists of patients with a history of intravenous drug abuse. These patients tend to be younger and to have no pre-existing joint disease. Their presentation is often more insidious, with a longer duration of illness before diagnosis (Bayer *et al*. 1977; Chandrasekar and Narula 1986). There is a predilection for the axial joints and periarticular soft-tissue abscess or osteomyelitis are frequently encountered. *Ps. aeruginosa* and *Serr. 'marcesens*' have been reported in increased frequency in this group in some but not all series from large urban centres.

Salmonellal arthritis is a rare complication of salmonellosis. *Salmonella typhimurium* is the most common serotype cultured from joints, followed by *Sal. choleraesuis*. Patients with systemic lupus erythematosus or sickle-cell anaemia are at increased risk for salmonella arthritis because the chronic salmonella carrier state is more common in these conditions. This is the result of impaired clearance by the reticuloendothelial system, owing to inhibition of Fc receptors by circulating immune complexes in systemic lupus and to the functional hyposplenism that can accompany either condition. The presentation is usually monarticular but polyarticular cases have been reported. In many cases there is no history of a preceding diarrhoeal illness. The diagnosis of salmonellal arthritis may not be suspected after arthrocentesis because the synovial fluid is often non-purulent, and described as turbid, serosanguinous, or straw-coloured, despite a positive culture from it (Cohen *et al*. 1987; Medina *et al*. 1989).

Rat-bite fever caused by *Streptobacillus moniliformis*, a pleomorphic Gram-negative organism, presents as an acute illness with fever, rash, vomiting, and arthritis that is usually polyarticular. The infection is acquired from bites or close contact with infected rodents or from ingestion of contaminated food or beverages. When articular symptoms are prominent the clinical presentation may be confused with rheumatoid arthritis, adult Still's disease, systemic lupus, or gonococcaemia. Joint effusions can be either sterile and inflammatory or purulent with positive cultures of synovial fluid. Isolation and identification of the organism is difficult, owing to its variable staining characteristics and specific growth requirements (Holroyd *et al*. 1988).

Past. multocida is a small, Gram-negative coccobacillus that causes joint infections after animal bites, most commonly from cats or dogs. In approximately one-third of patients there is no history of a bite. Arthritis with this organism has been reported in both normal hosts and in patients with rheumatoid arthritis, chronic liver disease, or immunosuppression (Weber *et al*. 1984).

L. monocytogenes infections in general have been increasing in recent years, owing to the increased survival of immunocompromised patients. The few case reports of *L. monocytogenes* infections of the joints suggest that these occur in immunocompromised patients or in those with previous joint disease (Kurosh and Perednia 1989).

In recent years there has been increasing recognition of the role of anaerobic organisms in pyogenic arthritis, although the incidence of anaerobic bacterial arthritis is still quite low. Anaerobic joint infections typically occur after joint surgery, after penetrating trauma, or in patients with underlying diseases. In the setting of surgery or trauma, bacteria are introduced directly from the skin or from soil contaminated with faecal organisms. Peptococcal and clostridial species, and *Propionibacterium acnes* predominate. In the setting of underlying disease the anaerobic Gram-negative organisms, particularly *Bacteroides fragilis*, predominate. The route of infection can be haematogenous from a distant site such as the bowel or direct extension from an adjacent site such as a deep sacral decubitus. In about half of cases in all settings a single organism will be cultured, while half will be polymicrobial, usually mixed aerobic and anaerobic organisms (Fitzgerald *et al*. 1982; Brook and Frazier

1993).

Anaerobic joint infections are probably underdiagnosed. There are no specific diagnostic signs and therefore synovial fluid must be cultured for anaerobic organisms whenever the clinical setting suggests the possibility of anaerobic infection. Anaerobic cultures should also be obtained in any patient with suspected joint infection in whom initial aerobic cultures of synovial fluid are negative. Any anaerobic organisms isolated from synovial fluid should not be dismissed merely as contaminants if the clinical setting suggests infection.

Joint infections and rheumatoid arthritis

Patients with rheumatoid arthritis are at increased risk of pyogenic arthritis. While the precise prevalence of joint infections in rheumatoid arthritis is not known, in several large series approximately 10 per cent of all cases of non-gonococcal arthritis occurred in patients with that condition (Goldenberg and Cohen 1976; Rosenthal *et al*. 1980; Cooper and Cawley 1986).

Proliferative synovitis and structural changes in the rheumatoid joint may allow bacteria to become sequestered in the joint more readily and to escape normal host-defence mechanisms. The formation of pannus and damage to cartilage may more readily allow the spread of bacteria from the synovium to subchondral bone. There is some controversy about whether patients with rheumatoid arthritis have an increased risk for infections in general, but they are often treated with immunosuppressive medications, which put them at increased risk of infection. Rheumatoid patients with advanced disease may be significantly debilitated, may be bedridden with greater risk of infections of the respiratory or urinary tracts, and may develop skin ulcerations from vasculitis, all of which increase the risk of infection.

Staph. aureus is the infecting organism in 70 per cent of reported cases, followed by streptococci in 9 per cent, Gram-negative bacilli in 8 per cent, pneumococci in 7 per cent, and anaerobes in 3 per cent. Blood cultures are positive in 20 to 25 per cent of cases. The knee is involved in about half the cases. The hip, shoulder, elbow, wrist, and ankle are each involved in about 10 per cent of cases, while infections in the smaller joints of the hands and feet are unusual. Polyarticular infections are reported in about 30 per cent of cases. The source of infection can be identified in half of these patients. The skin is the most common site of distant infection, followed by the lungs, and the urinary and gastrointestinal tracts (Gardner and Weisman 1990).

Identifying joint infections in the patient with rheumatoid arthritis can be difficult clinically because joint inflammation is part of the underlying disease. Furthermore, fewer than one-half of rheumatoid patients with joint infections present with fever and even fewer develop a leucocytosis (Gardner and Weisman 1990). The erythrocyte sedimentation rate is of limited value in distinguishing infection from disease flare. Both infection and exacerbation of the rheumatoid condition can be monarticular or polyarticular. This difficulty in distinguishing between the two often leads to delay in diagnosis. Joint infection should be suspected in those patients whose disease activity changes abruptly from its normal pattern, particularly if only one or a few joints are involved; in those with systemic signs of toxicity; in those with a remote site of active or recent infection; or in patients on immunosuppressive therapy with long-standing, deforming disease or with accompanying, debilitating illnesses. Whenever one of these clinical states is present, synovial fluid should be obtained, Gram stained, and cultured before considering any treatment (such as an intra-articular steroid injection or systemic anti-inflammatory medication) for a flare-up of rheumatoid disease (Blackburn *et al*. 1986; Goldenberg 1989; Soria *et al*. 1992).

Distinguishing infection from disease exacerbation is further complicated by the occasional patient who develops a 'pseudoseptic' arthritis. In these cases there is an abrupt onset of usually monarticular joint pain and effusion, accompanied by fever. The synovial fluid is purulent, with the leucocyte count usually greater than 100 000/mm^3. Gram's stain and culture of the synovial fluid are negative, and the process quickly resolves within a few days without antibiotics (Call *et al*. 1985; Singleton *et al*. 1991). These episodes may be recurrent and identical in presentation in an individual patient, in which case antibiotics may be withheld pending the results of synovial fluid culture. In most instances, however, empirical antibiotic therapy should be started as soon as infection is suspected and cultures obtained. If adequate cultures are negative, then 'pseudoseptic' arthritis should be considered and antibiotics discontinued.

Owing mostly to the delay in establishing the diagnosis, joint infections in rheumatoid patients tend to have a poorer outcome than in other groups with pyogenic arthritis. The death rate in reported series is about 20 per cent, more than doubling in patients with polyarticular involvement (Gardner and Weisman 1990). Only about one-third will recover without any worsening of basic joint function. On occasion osteomyelitis or draining cutaneous fistulas will develop.

Prosthetic joint infections

While the risk of prosthetic joint infection has decreased significantly over the past three decades, infection remains a major cause of failure in arthroplasty. With evaluation and treatment of preoperative infections, improved surgical technique, and perioperative prophylactic antibiotics, infection after hip or knee arthroplasty has been reduced to less than 1 per cent of cases. The incidence of infection for revisions of prosthetic joints is about 5 to 10 times that of primary procedures (Poss *et al*. 1984; Wymenga *et al*. 1992).

A number of factors contribute to the risk of prosthetic joint infection. There is a risk of introducing bacteria whenever the joint space is entered. In addition the presence of foreign material in the joint promotes the growth of bacteria by sequestering organisms from the vascular supply and hence from the host immune-defence mechanisms and antibiotics. Methylmethacrylate cement may inhibit the phagocytic function of polymorphs and promote the production of glycocalyx, a fibrous collection of polysaccharides that enhance bacterial growth (Gristina and Kolkin 1983).

Prosthetic joint infections should be suspected whenever there are systemic or localized signs and symptoms of infection, or in cases of new or increasing prosthetic joint pain, even in the absence of clinical signs of infection. Over 90 per cent of all prosthetic joint infections present with joint pain; fever, and erythema and swelling of the joint occur in less than half of cases. Leucocytosis is unusual while an elevated erythrocyte sedimentation rate is almost always present.

About one-half of infections are detected in the first 12 months after surgery, with one-half of these occurring in the early postoperative period up to 3 months after surgery (Inman *et al*. 1984). Those cases that present in the early postoperative period are usually the result of bacteria introduced at the time of surgery or are due to wound infections. A large proportion of these patients will have or have had wounds complicated by haematoma, stitch abscess, dehiscence, or infection. These patients usually have an acute illness with fever, pain, and local signs of infection, including persistent drainage from the joint. However, infection should still be suspected in the patient with fever, leucocytosis, or elevated erythrocyte sedimentation rate that fails to drop with time, even in the absence of local signs of infection.

Infections occurring between 3 and 12 months postoperatively tend to have a more subacute presentation. Less than half will have local signs of joint infection. Most will present with either joint pain or failure to regain the expected functional level.

The prosthetic joint remains at risk for infection indefinitely. Infections diagnosed more than 1 year after surgery are usually the result of haematogenous seeding from a distant focus such as the urinary tract or skin, from dental work, or from surgical manipulation of the gastrointestinal or urinary tracts. The presentation is usually the insidious onset of pain in the absence of both systemic and local signs and symptoms of infection. There is often a delay of months from onset of pain to diagnosis. Occasionally, late infections present with draining sinus tracts.

Diagnosis of prosthetic joint infection is often difficult. In the early period, infection of the actual joint space must be distinguished from adjacent wound and soft tissue infection. In later infections presenting with only joint pain, the infected prosthesis must be distinguished from the uninfected but painful, loose prosthesis. Plain radiographs are usually normal in early infections. Loosening is defined as greater than 2 mm of lucency at the bone–cement or bone–metal (in cementless prostheses) interface, or any lucency at the metal–cement interface. Other findings in late infections include progressive cortical bone loss around the prosthesis, fractured cement, or a periosteal reaction when the infection has spread through the thickness of cortical bone. Radionuclide scans are useful in the evaluation of prosthetic joint loosening. A negative technetium-99m bone scan is strong evidence against infection and sequential technetium-99m and gallium-67 scanning can help to distinguish infected from non-infected loosening. Similarly, indium-111 uptake favours diagnosis of infection. Joint aspiration is necessary when infection is suspected on a clinical basis or if imaging studies are suggestive of infection or are equivocal. Culturing for anaerobic organisms is imperative. At times, bacteriological diagnosis can only be confirmed by bone biopsy or by culturing material obtained at surgery for removal of the prosthesis.

Staphylococcal species account for 50 to 60 per cent of prosthetic joint infections. *Staph. epidermis* is a more common cause of prosthetic than native joint infections. Various stretococcal species and Gram-negative bacilli each account for 15 to 20 per cent and anaerobes 5 to 10 per cent of cases. Mixed organisms are cultured in 10 to 20 per cent of cases, particularly those occurring in the early postoperative period (Inman *et al*. 1984). Organisms of low virulence that are seldom considered pathogens, such as diphtheroids, propionibacteria, and lactobacilli, are occasional causes of prosthetic joint infections and must not be dismissed as contaminants if

cultured from synovial fluid or biopsy material.

Neisserial arthritis

Gonococcal arthritis

N. gonorrhoeae can cause either localized mucosal infections of the genitourinary tract, rectum, pharynx, or conjunctiva, or disseminated infections of the skin, joints, and less commonly, heart and meninges. Disseminated gonococcal infection complicates approximately 0.1 to 0.3 per cent of localized infections. Seventy-five per cent of cases of disseminated gonococcal infection result from asymptomatic local infections (Eisenstein and Masi 1981). This makes effective preventive measures difficult.

Those strains of *N. gonorrhoeae* associated with disseminated disease differ biologically from those strains associated with symptomatic, localized disease (O'Brien *et al*. 1983). The strains that disseminate are more likely to be resistant to killing by normal human serum, have specific nutritional requirements when grown in culture, have specific cell-surface proteins, and tend to be more sensitive to antibiotics. These characteristics are also shared by those strains that cause asymptomatic localized infections. These strains do not provoke a local inflammatory reaction and thus the organism is not limited to the mucosa, allowing seeding of the bloodstream.

While the incidence of disseminated gonococcal infection has been decreasing in recent years due probably to a decrease in both the prevalence of those strains associated with disseminated disease and to a decline in the overall incidence of gonococcal infections, *N. gonorrhoeae* remains a common cause of bacterial arthritis in young, previously healthy, individuals particularly in urban centres where the largest reservoirs of localized gonococcal infection exist (Rompalo *et al*. 1987). Seventy-five per cent of cases occur between the ages of 15 and 30 years, in parallel with the years of peak incidence for localized gonococcal infections, although any sexually active individual is potentially at risk (Geelhoed-Duyvestijn *et al*. 1986). Women are affected three to five times more often than men. By contrast, non-gonococcal arthritis more commonly affects the very young or the very old, men, and individuals with underlying diseases or risk factors.

Individuals with inherited deficiencies of the terminal complement components, C5 to C9, have an increased risk of neisserial infection (Ross and Densen 1984). Complement deficiencies should be searched for in individuals with recurrent gonococcal or meningococcal infections.

Disseminated gonococcal infection has been classified according to the presence or absence of a purulent effusion (Holmes *et al*. 1971; Masi and Eisenstein 1981; O'Brien *et al*. 1983). About two-thirds of patients present with an acute illness of 3 to 4 days' duration, consisting of fever, rash, tenosynovitis, and migratory polyarthralgias or polyarthritis with non-purulent effusions (group I). The rash, present in 60 to 90 per cent of this group, is a diffuse, maculopapular eruption or a more limited number of vesiculopustular lesions on an erythematous base of approximately 5 mm. in diameter. Less commonly the lesions appear as haemorrhagic pustules or bullas. The rash generally spares the face and scalp but can appear anywhere on the trunk and extremities. The lesions may be painful. Tenosynovitis usually involves multiple sites, particularly at the wrist, fingers, ankles, and toes. It is present in 90 per cent of group I patients. Polyarthralgias and polyarthritis are migratory, with a predilection for upper extremity joints. When effusions are present they are small and cultures of synovial fluid are negative.

Group II patients present with a purulent arthritis that resembles the presentation of non-gonococcal arthritis. Typically, these patients present with an acute illness of 4 to 5 days' duration. Most have a monarticular arthritis, with the knee most commonly involved, followed by the ankle, wrist, and elbow. Involvement of the hip, shoulder, temporomandibular, and axial joints is rare, as is polyarticular involvement. About one-third of patients in this group will also have rash or tenosynovitis and about two-thirds will give a history of polyarthralgia preceding the development of monarthritis.

About 40 per cent of group I patients have positive blood cultures; this rarely if ever occurs in group II patients (O'Brien *et al*. 1983). This strongly negative correlation between positive blood cultures and the presence of purulent joint effusions suggests that group I patients present during the bacteraemic phase of the disseminated infection and group II patients during the phase of joint localization that can follow bacteraemia. Some series have reported progression from group I to II in individual patients (Holmes *et al*. 1971), although others believe that the groups are distinct (O'Brien *et al*. 1983; Koss 1985).

Leucocyte counts in synovial fluid in group II patients range from about 30 000/mm ^{3} to greater than 200 000/mm^3. This range overlaps with that in non-gonococcal arthritis, although counts tend to be slightly lower in gonococcal infections. The synovial fluid lactate is not increased in gonococcal arthritis. Gram's stain is positive in about one-quarter of cases, and *N. gonorrhoeae* can be successfully cultured from only about half of the purulent effusions (Scopelitis and Martinez-Osuna 1993). There are several explanations for this low rate of recovery of organisms even from purulent effusions. First, *N. gonorrhoeae* is difficult to culture, owing to its fastidious growth requirements. Second, organisms may be sequestered in the synovium and may not have entered the synovial fluid. Third, organisms may exist only transiently in the synovial fluid before they are phagocytosed and can no longer be cultured, but the synovitis persists in response to constituents of the gonococcal cell wall such as lipopolysaccharide. In animal studies, both killed gonococci and purified lipopolysaccharide can induce a purulent synovitis after injection into the joint space (Goldenberg *et al*. 1984). Finally, effusions may be caused or maintained by host immunological reactions to *N. gonorrhoeae*. Immune complexes have been demonstrated in both synovial fluid and the blood of patients with disseminated infection.

As positive cultures of synovial fluid are obtained in only a minority of patients, the diagnosis of disseminated gonococcal infection must often be presumed from a suggestive clinical presentation, together with isolation of *N. gonorrhoeae* from a site other than the joint. In patients with disseminated infection, positive cultures can be obtained from the cervix in 80 to 90 per cent of women, the urethra in 50 to 75 per cent of men, and the blood in 25 per cent. The pharynx and the rectum will be positive in a smaller number of cases (Scopelitis and Martinez-Osuna 1993). To culture *N. gonorrhoeae* from the joint, synovial fluid should be immediately spread on a prewarmed, chocolate-agar plate and incubated in a atmosphere enriched with carbon dioxide, or sent to the laboratory on transport media. Antibiotic-enriched culture media such as Thayer–Martin are used for isolating *N. gonorrhoeae* from sites that have a native bacterial flora, such as the pharynx, cervix, urethra, and rectum, but are not necessary and in fact may inhibit growth when culturing synovial fluid. Recently, the identification of *N. gonorrhoeae* DNA in culture-negative synovial fluid by the polymerase chain reaction has been reported. This technique potentially can improve diagnostic accuracy and hasten the time to diagnosis but is not yet readily available (Liebling *et al*. 1994)

The differential diagnosis of disseminated gonococcal infection is broad. Reiter's syndrome and disseminated gonococcal infection both commonly affect young, sexually active individuals. Urethritis, rashes, and inflammatory monarthritis or oligoarthritis can occur in both. The onset of illness in Reiter's syndrome tends to be more gradual than in disseminated gonococcal infection. Psoriasiform skin lesions, sacroiliitis, painless oral ulcers, conjunctivitis, and a recent history of a diarrhoeal illness favour the diagnosis of Reiter's syndrome. Analysis of synovial fluid does not distinguish between the two conditions unless a positive culture for *N. gonorrhoeae* is obtained. Not uncommonly the patient with disseminated gonococcal infection presenting with monarthritis must be distinguished from patients with non-gonococcal bacterial arthritis. Table 1 lists the distinguishing features. Other infectious illnesses may present with fever, rash, and polyarthralgias or polyarthritis, such as the prodrome of hepatitis B, subacute bacterial endocarditis, and bacterial arthritis due to *N. meningitidis*, Group A streptococcus, *H. influenzae*, or *Streptobac. moniliformis.* Acute rheumatic fever may present with fever and migratory arthritis distinguishable from disseminated gonococcal infection by evidence of a recent streptococcal infection, accompanying chorea or typical rash, and by rapid response to aspirin. Systemic lupus erythematosus, hypersensitivity vasculitis, and adult Still's disease may all present with fever, rash, and arthritis that at times might be confused with disseminated gonococcal infection, although the nature of the rash and the involvement of other organ systems should help to distinguish most cases.

The course of the disseminated infection is variable. In the preantibiotic era, some cases underwent spontaneous remission, while in others there was progression to destructive joint changes. In general the progression to irreversible damage in the cartilage and joint is not as rapid as with organisms such as staphylococci, streptococci, or Gram-negative bacilli. Rash, tenosynovitis, and polyarthralgias respond very promptly to antibiotics, while purulent effusions respond somewhat more slowly.

Meningococcal arthritis

Arthritis due to *N. meningitidis* follows one of several clinical patterns and may be confused with disseminated gonococcal infection (Fam *et al*. 1979; Schaad 1980). Joint manifestations complicate about 2 to 10 per cent of cases of acute meningococcaemia and may follow one of three clinical patterns. In the first, patients develop polyarthralgias, polyarthritis, or tenosynovitis during the acute phase of their meningococcal infection. They are acutely ill with fever, meningitis, and/or a diffuse, erythematous, macular rash or diffuse haemorrhagic skin lesions typical of acute meningococcaemia. When effusions occur in this group they are small and non-purulent. Synovial cultures are negative but *N. meningitidis* can be recovered from blood or cerebrospinal fluid. The articular manifestations respond rapidly to antibiotic therapy. When meningitis is absent, these patients resemble those in group I of disseminated gonococcal infection (see above) (Kidd *et al*. 1985).

The second group of patients, which resembles group II in disseminated gonococcal infection, develops purulent joint effusions that are most commonly monarticular or oligoarticular during the acute phase of meningococcaemia. Cultures of synovial fluid can be positive or negative. In some cases these effusions may be slow to resolve, even after appropriate antibiotics are given and the synovial fluid cultures are sterile.

A small number of patients with acute meningococcaemia will develop articular manifestations only after the acute stage of infection has responded to antibiotics and the patient is clearly improving (Jarrett *et al*. 1980). In these cases, monarticular, oligoarticular, or polyarticular effusions develop 1 to 2 weeks after the onset of the acute illness. Cultures of synovial fluid are negative but immune complexes are detectable in blood and synovial fluid, suggesting an immunological basis for the synovitis, which eventually resolves without further antibiotic therapy.

Patients with purulent arthritis, usually monarticular, but without signs of acute meningococcaemia are classified as primary meningococcal arthritis (Andersson and Krook 1987). The knee and ankle are most commonly involved. Cultures of synovial fluid are positive in 90 per cent of cases. Despite prompt and appropriate therapy these patients are at risk for joint damage and at times may require open drainage. Blood cultures are negative in this group.

Finally, joint manifestations may occur in the setting of chronic meningococcaemia, a condition marked by chronic rash and positive blood cultures. These patients generally present with polyarthralgias rather than frank arthritis. When effusions are present, cultures are negative, suggesting an immune-mediated mechanism.

Management

The three basic principles of management of pyogenic joint infections are: (i) prompt diagnosis and institution of therapy; (ii) appropriate antibiotics; and (iii) adequate drainage of the infected joint.

The most important factors in determining outcome are the length of time before beginning treatment and the length of time to sterilization of synovial fluid cultures (Ho and Su 1982). Therefore, the suspicion of joint infection should arise in any patient with an unexplained, painful or swollen joint, especially those with an acute presentation, with monarticular presentation, with systemic or local signs and symptoms of infection, with risk factors (such as underlying joint disease, recent joint trauma or surgery, diabetes, immunosuppression, or history of intravenous drug abuse), or with active or recent extra-articular infection.

In these settings there should be no delay in aspirating the involved joint or joints. The synovial fluid should be analysed for total and differential cell count, lactate, and crystals, and Gram stained and cultured to differentiate pyogenic infection from other causes of acute joint pain and swelling, including crystal-induced arthritis, exacerbation or presentation of an inflammatory arthritis, haemarthrosis, or trauma. The presence of urate or calcium pyrophosphate crystals, however, does not completely exclude the diagnosis of infection. In rare cases, crystals and infection have occurred in the same joint. The infectious process may 'leach' crystals out of the cartilage or synovium. Blood cultures should be obtained and a search for a distant focus of infection should be undertaken as appropriate for the individual patient.

If Gram's stain is positive or the analysis is otherwise suggestive of pyogenic infection, treatment should begin immediately. The patient should be admitted to hospital and given intravenous antibiotics as soon as all cultures have been obtained. Intra-articular antibiotics should be avoided because adequate concentrations in synovial fluid can be achieved via the parenteral route and because there is a risk of chemical synovitis with the intra-articular route. The concentration of antibiotics in synovial fluid parallels that in serum, and adequate levels can be achieved with standard parenteral doses. Penetration of the drug into the synovial fluid is greatest when the inflammation is most active at the start of therapy and falls off as the infection is controlled. Although they may achieve bactericidal levels in synovial fluid, aminoglycosides are sometimes less effective than other agents, possibly due to inhibition by the reduced pH of infected synovial fluid. Aminoglycosides are therefore inadequate for primary treatment of pyogenic arthritis. Erythromycin does not reach adequate levels in synovial fluid and therefore cannot be used as an alternative to penicillin (Esterhai and Gelb 1991).

Most often the initial choice of antibiotics is empirically guided by the results of Gram's stain, when positive, or by the clinical features when Gram's stain is negative. In choosing antibiotics, one must consider the age of the patient, any underlying disease, and the presence of extra-articular infection. Table 3 provides a guide to the empirical choice of antibiotics. When the results of antibiotic sensitivity testing become known, antibiotic coverage sould then be modified to provide a regimen that has the narrowest spectrum, is the least toxic, and is the least expensive.

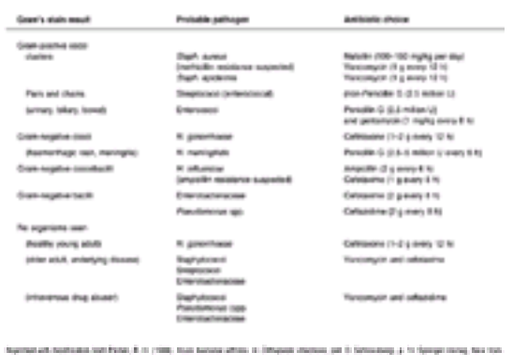


Table 3 Guidelines for initial antibiotic therapy of bacterial arthritis based on Gram's stain

In recent years, methicillin-resistant strains of *Staph. aureus* have become more prevalent in both hospital- and community-acquired infections (Ang-Fonte *et al*. 1985). Vancomycin should be the initial drug of choice for suspected *Staph. aureus* infection in those communities and hospitals where resistant strains exist.

Although those strains of *N. gonorrhoeae* that cause disseminated gonococcal infection are more likely to be antibiotic sensitive than those causing localized infections, strains that have chromosomally-mediated resistance to penicillin or are resistant because they produce plasmid-mediated penicillinase have become more common and widespread, and have been reported in disseminated infection (Wise *et al*. 1994). Therefore, in communities with resistant strains, a third-generation cephalosporin should replace penicillin G until antibiotic sensitivities are reported.

The duration of antibiotic therapy must be tailored to the individual presentation and response to therapy and to the infecting organism. In general, intravenous antibiotics should be continued until signs of joint inflammation have resolved, joint effusions have resolved or significantly decreased, and synovial cultures are sterile. Oral antibiotics can then be used to complete the course of treatment, which is usually about 2 weeks in uncomplicated cases but should be extended to 4 weeks or more where there was a significant delay in starting treatment, or a slow response, where infection was with virulent organisms such as *Staph. aureus* or Gram-negative bacilli, and in immunosuppressed patients (Syrogiannopoulos and Nelson 1988). The development of ambulatory intravenous antibiotic delivery systems has allowed earlier hospital discharge in appropriate patients (Williams *et al*. 1989).

The duration of parenteral antibiotic therapy for disseminated gonococcal infection is generally shorter than for non-gonococcal arthritis. In most patients with gonococcal arthritis with a good response to therapy, parenteral antibiotics can be substituted by oral after about 3 days of intravenous therapy to complete a 2-week course. Patients who have dermatitis/tenosynovitis only, without purulent arthritis, can be treated as outpatients, provided they are compliant and closely followed.

Occasional patients with purulent gonococcal effusions will develop a persistent inflammatory effusion, lasting weeks, after a good response to antibiotics and sterilization of synovial fluid. These effusions eventually resolve without further antibiotics. Recovery may be speeded by giving non-steroidal anti-inflammatory drugs.

Infected joints can be drained by closed needle aspiration, tidal irrigation (Ike 1993), arthroscopy, or arthrotomy. Each procedure has its appropriate place in the treatment of septic arthritis depending on the clinical situation. Which modality is chosen is probably not as critical to the ultimate outcome as the duration of time from

onset of infection to adequate drainage and sterilization of the joint fluid (Goldenberg *et al*. 1975; Lane *et al*. 1990; Ho 1993).

Initial treatment with serial, closed-needle aspiration avoids the expense and surgical morbidity and mortality of the more invasive procedures and is a satisfactory approach as long as the joint can be completely drained and there is clear evidence of improvement as indicated by resolution of systemic signs of toxicity, serial decrease in synovial fluid white blood-cell count and sterilization of the synovial fluid culture within 48 to 72 h (Broy and Schmid 1986).

The presence of risk factors for a poor outcome include delayed diagnosis, virulent organism, polyarticular involvement, or underlying disease, such as rheumatoid arthritis, and might prompt the early use of a more invasive procedure. Arthroscopy has the advantage of good visualization of the joint space and articular cartilage combined with effective drainage and irrigation, while avoiding the postoperative joint stiffness and slow functional improvement sometimes associated with arthrotomy (Thiery 1989). Arthroscopy is best suited for drainage of the knee joint, although, depending on the skill of the arthroscopist, it can be applied to other smaller joints.

Septic arthritis of the hip should be treated primarily by arthrotomy because of the difficulty and uncertainty of achieving adequate joint drainage by closed aspiration, and the increased risk of osteomyelitis and avascular necrosis in an incompletely drained and decompressed hip. Infection of the deep axial joints complicated by abscess formation requires open drainage.

Regardless of the mode of drainage, early mobilization of the infected joint is important in regaining maximal function. Most often the infected joint will be too painful to allow mobilization during the first few days of therapy. However, passive, range-of-motion exercises should be instituted as soon as pain permits. Both animal and human studies have demonstrated the advantages of a continuous, passive-motion device soon after surgical drainage of the knee. Isometric exercises to restore strength to the limb muscles and progressive ambulation to prevent the complications of prolonged bed rest should also be implemented (Esterhai and Gelb 1991; Mikhail and Alarcon 1993).

Management of the infected prosthetic joint involves the added complication of the presence of foreign material. In general, the prosthesis must be removed to treat the infection effectively. Exceptions include some very early postoperative infections in which the wound and joint can be effectively incised and drained, and where antibiotics are started without significant delay. Another exception is the patient who is not considered a candidate for surgery because of poor medical status and who is infected with an organism of low virulence. In this case, chronic suppression of infection by antibiotics has at times been successful in salvaging the prosthesis and preventing further infectious complications.

In nearly all other instances, the prosthesis, cement, and all necrotic bone and soft tissue should be removed as soon as infection is detected or strongly suspected. Appropriate antibiotics should be started, as outlined in Table 3, keeping in mind the increased frequency of *Staph. epidermis*, Gram-negative bacilli, and anaerobic infections. After 6 weeks of intravenous antibiotics a prosthesis can be reimplanted (Fig. 1) with a greater than 90 per cent chance of success and with a better functional outcome and greater patient satisfaction than with arthrodesis of the knee or excision arthroplasty (Girdlestone procedure) of the hip, which should be reserved for those patients unable or unwilling to tolerate a second surgical procedure (Insall *et al*. 1983). Others have reported success with a one-step procedure in which a new prosthesis is implanted at the time of removal of the infected prosthesis (Goksan and Freeman 1992).

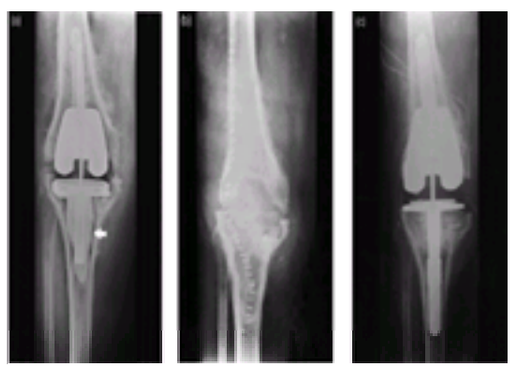


Fig. 1 (a) Infected knee prosthesis. Note the radiolucent zone at the bone–cement interface of the tibial component indicating loosening (arrow); (b) the same knee after complete removal of the infected prosthesis and cement; and (c) the same knee 2 months later after reimplantation of a new prosthesis.

Septic bursitis

The many bursas found throughout the body are lined with a synovial membrane identical to that found in synovial joints. The bursas occupy superficial locations beneath the skin, cushioning the movements of tendons, muscles, and ligaments over bony structures. Bursitis may result from trauma, inflammatory conditions affecting synovial tissues (such as rheumatoid arthritis or gout), or from infection. Most cases of bursal infection are associated with recent trauma to the skin overlying the bursa or with an occupation causing repeated minor trauma to the bursa. Bacteria enter the bursa directly from the skin as a result of trauma, unlike septic arthritis in which the route of infection is most commonly haematogenous seeding. Underlying diseases affecting the bursa, including rheumatoid arthritis and gout, or previous episodes of non-infectious traumatic bursitis, increase the risk of septic bursitis. Immunosuppression, diabetes, and alcoholism may be additional predisposing factors (Canoso and Barza 1993).

Nearly all cases of septic bursitis involve the olecranon or prepatellar bursae. Both of these are closed structures that do not communicate with the adjacent joint. Presentation may be acute, with sudden onset of bursal pain, tenderness, swelling, and erythema leading to diagnosis within a few days of onset. Less commonly, presentation may be subacute or chronic with less dramatic local signs and symptoms present for weeks to months before the diagnosis is established. Involvement of the bursa is readily distinguishable from joint infection by the presence of very superficial, often tense, distension of the bursal sac and preservation of movement in the underlying joint, with little if any pain on motion. Fever is a common but not universal finding. Regional lymphadenopathy and adjacent cellulitis of the forearm or peripatellar skin are frequently present. Desquamation of the skin overlying the infected bursa and spontaneous drainage of the bursa through the skin may occur. Osteomyelitis of the underlying bone may occasionally complicate chronic cases.

The characteristics of synovial fluid from an infected bursa are variable. The gross appearance ranges from clear to frankly purulent. The leucocyte count may range from less than 2000 to several 100 000 cells/mm^3, with the proportion of polymorphs ranging from 50 to nearly 100 per cent. In contrast to pyogenic arthritis, the leucocyte count in synovial fluid is often less than 10 000/mm^3, so that relatively low counts should not deter one from pursuing the diagnosis of infection in the appropriate clinical setting (Ho *et al*. 1978; Canoso and Barza 1993).

In the absence of previous antibiotic therapy, culture of bursal synovial fluid will be positive in every case. Gram's stain is positive in about two-thirds of cases. *Staph. aureus* is the infecting organism in more than 90 per cent of cases. Sporadic cases are due to *Staph. epidermis* or group A streptococci.

Uncomplicated septic bursitis can be treated effectively with serial, closed-needle aspiration and oral antibiotics. Patients with underlying bursal disease, immunosuppressed patients, or those who fail to respond promptly to treatment as outpatients with oral antibiotics should be admitted to hospital and treated with intravenous antibiotics (Ho and Su 1981; Canoso and Barza 1993). Incision and drainage or excision of the infected bursa may be necessary in those cases with extensive involvement or failure to respond to closed drainage. Empirical antibiotic treatment should be with a semisynthetic, penicillinase-resistant penicillin or vancomycin, if methicillin-resistant *Staph. aureus* is a concern.

Chapter References

Andersson, S. and Krook, A. (1987). Primary meningococcal arthritis. *Scandinavian Journal of Infectious Disease*, **19**, 51–4.

Ang-Fonte, G.Z., Rozboril, M.B., and Thompson, G.R. (1985). Changes in non-gonococcal septic arthritis: drug abuse and methicillin-resistant *Staphylococcus aureus*. *Arthritis and Rheumatism*, **28**,

210–3.

Atcheson, S.G. and Ward, J., Jr. (1978). Acute hematogenous osteomyelitis progressing to septic synovitis and eventual pyarthrosis. *Arthritis and Rheumatism*, **21**, 968–71.

Baker, D.G. and Schumacher, H.R., Jr. (1993). Acute monoarthritis. *New England Journal of Medicine*, **329**, 1013–20.

Bayer, A.S., Chow, A.W., Louie, J.S., Nies, K.M., and Guze, L.B. (1977). Gram-negative bacillary septic arthritis: clinical, radiographic, therapeutic, and prognostic features. *Seminars in Arthritis and Rheumatism*, **7**, 123–32.

Blackburn, W.D., Jr., Dunn, T.L., and Alarcon, G.S. (1986). Infection versus disease activity in rheumatoid arthritis: 8 years' experience. *Southern Medical Journal*, **79**, 1238–41.

Borenstein, D.G. and Simon, G.L. (1986). *Hemophilus influenzae* septic arthritis in adults. *Medicine*, **65**, 191–201.

Brook, I. and Frazier, E.H. (1993). Anaerobic osteomyelitis and arthritis in a military hospital: a 10-year experience. *American Journal of Medicine*, **94**, 21–8.

Broy, S.B. and Schmid, F.R. (1986). A comparison of medical drainage (needle aspiration) and surgical drainage (arthrotomy or arthroscopy) in the initial treatment of infected joints. *Clinics in the Rheumatic Diseases*, **12**, 501–22.

Burhert, T. and Watanakunakorn, C. (1991). Group C streptococcus septic arthritis and osteomyelitis: report and literature review. *Journal of Rheumatology*, **18**, 904–7.

Call, R.S., Ward, J.R., and Samuelson, C.O., Jr. (1985). 'Pseudoseptic' arthritis in patients with rheumatoid arthritis. *Western Journal of Medicine*, **143**, 471–3.

Canoso, J.J. and Barza, M. (1993). Soft tissue infections. *Rheumatic Disease Clinics of North America*, **19**, 293–309.

Chandrasekar, P.H. and Narula, A.P. (1986). Bone and joint infections in intravenous drug abusers. *Reviews of Infectious Diseases*, **8**, 904–11.

Cohen, J.I., Bartlett, J.A. and Corey, G.R. (1987). Extra-intestinal manifestations of salmonella infections. *Medicine*, **66**, 349–88.

Cooper, C. and Cawley, M.I.D. (1986). Bacterial arthritis in an English health district: a 10 year review. *Annals of the Rheumatic Diseases*, **45**, 458–63.

Dubost, J. *et al*. (1993). Polyarticular septic arthritis. *Medicine*, **72**, 296–310.

Eisenstein, B.I. and Masi, A.T. (1981). Disseminated gonococcal infection (DGI) and gonococcal arthritis (GCA): I. Bacteriology, epidemiology, host factors, pathologic factors and pathology. *Seminars in Arthritis and Rheumatism*, **10**, 155–72.

Epstein, J.H., Zimmermann III, B., and Ho, G., Jr. (1986). Polyarticular septic arthritis. *Journal of Rheumatology*, **13**, 1105–7.

Esterhai, J.L., Jr. and Gelb, I. (1991). Adult septic arthritis. *Orthopedic Clinics of North America*, **22**, 503–14.

Fam, E.G., Tenenbaum, J., and Stein, J.L. (1979). Clinical forms of meningococcal arthritis. A study of 5 cases. *Journal of Rheumatology*, **6**, 567–73.

Fitzgerald, R.H., Jr., Rosenblatt, J.E., Tenney, J.H., and Bourgault, A. (1982). Anaerobic septic arthritis. *Clinical Orthopaedics and Related Research*, **164**, 141–8.

Gardner, G.C. and Weisman, M.H. (1990). Pyarthrosis in patients with rheumatoid arthritis: a report of 13 cases and a review of the literature from the past 40 years. *American Journal of Medicine*, **88**, 503–11.

Geelhoed-Duyvestijn, P.H.L.M., van der Meer, J.W.M., Lichtendahl-Bernard, A.T., Mulder, L.J., Meyers, K.A.E., and Poolman, J.T. (1986). Disseminated gonococcal infections in elderly patients. *Archives of Internal Medicine*, **146**, 1739–40.

Goksan, S.B. and Freeman, M.A.R. (1992). One-stage reimplantation for infected total knee arthroplasty. *Journal of Bone and Joint Surgery*, **74B**, 78–82.

Goldenberg, D.L. (1989). Infectious arthritis complicating rheumatoid arthritis and other chronic rheumatic diseases. *Arthritis and Rheumatism*, **32**, 496–502.

Goldenberg, D.L. and Cohen, A.S. (1976). Acute infectious arthritis. *American Journal of Medicine*, **60**, 369–77.

Goldenberg, D.L. and Reed, J.I. (1985). Bacterial arthritis. *New England Journal of Medicine*, **312**, 764–71.

Goldenberg, D.L., Brandt, K.D., Cohen, A.S., and Carthcart, E.S. (1975). Treatment of septic arthritis—comparison of needle aspiration and surgery as initial modes of joint drainage. *Arthritis and Rheumatism*, **18**, 83–90.

Goldenberg, D.L., Reed, J.I., and Rice, P.A. (1984). Arthritis in rabbits induced by killed *Neisseria gonorrhoeae* and gonococcal lipopolysccharide. *Journal of Rheumatology*, **11**, 3–8.

Gordon, G. and Kabins, S.A. (1980). Pyogenic sacroiliitis. *American Journal of Medicine*, **69**, 50–6.

Gray, R.G., Tenenbaum, J., and Gottlieb, N.L. (1981). Local corticosteroid injection treatment in rheumatic disorders. *Seminars in Arthritis and Rheumatism*, **10**, 231–54.

Gristina, A.G. and Kolkin, J. (1983). Total joint replacement and sepsis. *Journal of Bone and Joint Surgery*, **65A**, 128–34.

Ho, G., Jr. (1993). How best to drain an infected joint. Will we ever know? *Journal of Rheumatology*, **20**, 2001–3.

Ho, G., Jr. and Su, E.Y. (1981). Antibiotic therapy of septic bursitis: it's implications in the treatment of septic arthritis. *Arthritis and Rheumatism*, **24**, 905–11.

Ho, G., Jr. and Su, E.Y. (1982). Therapy for septic arthritis. *Journal of the American Medical Association*, **247**, 797–800.

Ho, G., Jr., Tice, A.D., and Kaplan, S.R. (1978). Septic bursitis in the prepatellar and olecranon bursae. *Annals of Internal Medicine*, **89**, 21–7.

Hollander, J.L. (1969). Intrasynovial corticosteroid therapy in arthritis. *Maryland State Medical Journal*, **19**, 62– 6.

Holmes, K.K., Counts, G.W., and Besty, H.N. (1971). Disseminated gonococcal infection. *Annals of Internal Medicine*, **74**, 979–93.

Holroyd, K.J., Reiner, A.P., and Dick, J.D. (1988). *Streptobacillus moniliformis* polyarthritis mimicking rheumatoid arthritis: an urban case of rat bite fever. *American Journal of Medicine*, **85**, 711–14

Ike, R.W. (1990). Septic arthritis due to group C streptococcus: report and review of literature. *Journal of Rheumatology*, **17**, 1230–6.

Ike, R.W. (1993). Tidal irrigation in septic arthritis of the knee: A potential alternative to surgical drainage. *Journal of Rheumatology*, **20**, 2104–11.

Inman, R.D., Gallegos, K.V., Brause, B.D., Redecha, P.B., and Christian, C.L. (1984). Clinical and microbial features of prosthetic joint infection. *American Journal of Medicine*, **77**, 47–53.

Insall, J.N., Thompson, F.M., and Brause, B.D. (1983). Two-stage reimplantation for the salvage of infected total knee arthroplasty. *Journal of Bone and Joint Surgery*, **65A**, 1087–98.

Jarrett, M.P., Moses, S., Barland, P., and Miller, M.H. (1980). Articular complications of meningococcal meningitis. An immune complex disorder. *Archives of Internal Medicine*, **140**, 1656–66.

Kidd, B.L., Hart, H.H., and Grigor, R.R. (1985). Clinical features of meningococcal arthritis: a report of 4 cases. *Annals of the Rheumatic Diseases*, **44**, 790–2.

Koss, P.G. (1985). Disseminated gonococcal infection: the tenosynovitis–dermatitis and suppurative arthritis syndromes. *Cleveland Clinic Quarterly*, **52**, 161–73.

Krey, P.R. and Bailen, D. (1979). Synovial fluid leukocytosis. *American Journal of Medicine*, **67**, 436–42.

Kurosh, N.A. and Perednia, D.A. (1989). *Listeria monocytogenes* septic arthritis. *Archives of Internal Medicine*, **149**, 1207–8.

Lane, J.G., Falahee, M.H., Wojtys, E.M., Hankin, F.M., and Kaufer, H. (1990). Pyarthrosis of the knee. Treatment considerations. *Clinical Orthropedics and Related Research*, **252**, 198–204.

Liebling, M.R. *et al*. (1994). Identification of *Neisseria gonorrhoeae* in synovial fluid using the polymerase chain reaction. *Arthritis and Rheumatism*, **37**, 702–9.

Lopez-Longo, F-J., Menard, H-A., Carreno, L., Cosin, J., Ballesteros, R., and Monteagudo, I. (1987). Primary septic arthritis in heroin users: early diagnosis by radioisotopic imaging and geographic

variations in the causative agents. *Journal of Rheumatology*, **14**, 991–4.

Manshady, R.M., Thompson, G.R., and Weiss, J.J. (1980). Septic arthritis in a general hospital 1966–1977. *Journal of Rheumatology*, **7**, 523–30.

Masi, A.T. and Eisenstein, B.I. (1981). Disseminated gonococcal infection (DGI) and gonococcal arthritis (GA): II. Clinical manifestations, diagnosis, complications, treatment and prevention. *Seminars in Arthritis and Rheumatism*, **10**, 173–97.

Medina, F., Fraga, A., and Lavalle, C. (1989). Salmonella septic arthritis in systemic lupus erythematosus. The importance of the chronic carrier state. *Journal of Rheumatology*, **16**, 203–8.

Merkel, K.D., Brown, M.L., Dewanjee, M.K., and Fitzgerald, R.H. (1985). Comparison of indium-111 labeled-leukocyte imaging with sequential technitium–gallium scanning in the diagnosis of low-grade musculoskeletal sepsis. *Journal of Bone and Joint Surgery*, **67A**, 465–76.

Mikhail, I.S. and Alarcon, G.S. (1993). Nongonococcal bacterial arthritis. *Rheumatic Disease Clinics of North America*, **19**, 311–31.

Newman, E.D., Davis, D.E., and Harrington, T.M. (1988). Septic arthritis due to Gram negative bacilli: older patients with good outcome. *Journal of Rheumatology*, **15**, 659–62.

Newman, J.H. (1976). Review of septic arthritis throughout the antibiotic era. *Annals of the Rheumatic Diseases*, **35**, 198–205.

O'Brien, J.P., Goldenberg, D.L., and Rice, P.A. (1983). Disseminated gonococcal infection: a prospective analysis of 49 patients and a review of pathophysiology and immune mechanisms. *Medicine*, **62**, 395–406.

Pischel, K.D., Weisman, M.H., and Cone, R.O. (1985). Unique features of group B streptococcal arthritis in adults. *Archives of Internal Medicine*, **145**, 97–102.

Poss, R., Thornhill, T.S., Ewald, F.C., Thomas, W.H., Batte, N.J., and Sledge, C.B. (1984). Factors influencing the incidence and outcome of infections following total joint arthroplasty. *Clinical Orthopedics and Related Research*, **182**, 117–26.

Riegels-Nielsen, P., Frimodt-Moller, N., and Jensen, J.S. (1987). Rabbit model of septic arthritis. *Acta Orthopaedica Scandinavia*, **58**, 14–9.

Riordan, T., Doyle, D., and Tabaqchali, S. (1982). Synovial fluid lactic acid measurement in the diagnosis and management of septic arthritis. *Journal of Clinical Pathology*, **35**, 390–4.

Rivera, J., Monteagudo, I., Lopez-Longo, J., and Sanchezatrio, A. (1992). Septic arthritis in patients with acquired immunodeficiency syndrome with HIV infection. *Journal of Rheumatology*, **19**, 1960–2.

Rompalo, A.M., Hook, E.W., III, Roberts, P.L., Ramsey, P.G., Handsfield, H., and Holmes, K.K. (1987). The acute arthritis–dermatitis syndrome: the changing importance of *Neisseria gonorrhoeae* and *Neisseria meningitidis*. *Archives of Internal Medicine*, **147**, 281–3.

Rosenthal, J., Boles, G.G., and Robinson, W.D. (1980). Acute non-gonococcal infectious arthritis. *Arthritis and Rheumatism*, **23**, 889–92.

Ross, S.C. and Densen, P. (1984). Complement deficiency states and infections: epidemiology, pathogenesis and consequences of neisserial and other infections in an immune deficiency. *Medicine*, **63**, 243–73.

Schaad, U.B. (1980). Arthritis in disease due to *Neisseria meningitidis*. *Reviews of Infectious Disease*, **2**, 880–8.

Scopelitis, E. and Martinez-Oscina, P. (1993). Gonococcal arthritis. *Rheumatic Disease Clinics of North America*, **19**, 363–77.

Sharp, J.T., Lidsky, M.D., Duffy, J., and Duncan, M.W. (1979). Infectious arthritis. *Archives of Internal Medicine*, **139**, 1125–30.

Singleton, J.D., West, S.G., and Nordstrom, D.M. (1991). 'Pseudoseptic' arthritis complicating rheumatoid arthritis: a report of six cases. *Journal of Rheumatology*, **18**, 1319–22.

Small, C.B., Slater, L.N., Lowy, F.D., Small, R.D., Salvati, E.A., and Casey, J.I. (1984). Group B streptococcal arthritis in adults. *American Journal of Medicine*, **76**, 367–75.

Soria, L.M., Sole, J.M.N., Sacanell, A.R., Garcia, J.V., and Escofet, D.R. (1992). Infectious arthritis in patients with rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **51**, 402–3.

Switalski, L.M., Patti, J.M., Butcher, W., Gristina, A.G., Speziale, P., and Hook, M. (1993). A collagen receptor on *Staphylococcus aureus* strains isolated from patients with septic arthritis mediates adhesion to cartilage. *Molecular Microbiology*, **7**, 99–107.

Syrogiannopoulos, G.A. and Nelson, J.D. (1988). Duration of anti-microbial therapy for acute suppurative osteoarticular infections. *Lancet*, **i**, 37–40.

Thiery, J.A. (1989). Arthroscopic drainage in septic arthritis of the knee: a multicenter study. *Arthroscopy*, **5**, 65–9.

Von Essen, R. and Holtta, A. (1986). Improved method of isolating bacteria from joint fluids by the use of blood culture bottles. *Annals of the Rheumatic Diseases*, **45**, 454–7.

Vyskocil, J.J., McIroy, M.A., Brennan, T.A., and Wilson, F.M. (1991). Pyogenic infection of the sacroiliac joint. *Medicine*, **70**, 188–97.

Weber, D.J., Wolfson, J.S., Swartz, M.N., and Hooper, D.C. (1984). *Pasteurella multocida* infections. Report of 34 cases and review of the literature. *Medicine*, **63**, 133–54.

Williams, D.N., Gibson, J.A., and Bosch, D. (1989). Home intravenous antibiotic therapy using a programmable infusion pump. *Archives of Internal Medicine*, **149**, 1157–60.

Wise, C.M., Morris, C.R., Wasilauskas, B.L., and Salzer, W.L. (1994). Gonococcal arthritis in an era of increasing penicillin resistance. *Archives of Internal Medicine*, **154**, 2690–5.

Wohlgethan, J.R., Newberg, A.H., and Reed, J.I. (1988). The risk of abscess from sterno-clavicular septic arthritis. *Journal of Rheumatology*, **15**, 1302–6.

Wymenga, A.B., van Horne, J.R., Theeuwes, A., Muytjens, H.L., and Slooff, T.J.J.H. (1992). Perioperative factors associated with septic arthritis after arthroplasty. *Acta Orthopaedica Scandinavia*, **63**, 665–71.

Youssef, P.P. and York, J.R. (1994). Septic arthritis: a second decade of experience. *Australian and New Zealand Journal of Medicine*, **24**, 307–11.

5.3.2 Pyogenic arthritis in children

Mary Anne Jackson and Komal B. Desai

Introduction

A successful outcome to suppurative arthritis in infancy and childhood is contingent upon early recognition and timely antimicrobial and surgical therapy. Despite the advent of newer antimicrobials that penetrate readily into infected joints, complications and permanent changes still arise in some cases of septic arthritis. This chapter reviews the epidemiology and pathogenesis of joint infection, and outlines the approach to diagnosis and management. Changes in the spectrum of childhood skeletal infection in the era of the *Haemophilus influenzae* type b (**Hib**) vaccine are highlighted.

Epidemiology

Acute septic arthritis is a relatively uncommon infection. One estimate suggests that two of every 1000 admissions to general hospitals are for septic arthritis (Smith 1974). Our institution serves a large, urban population in the American Midwest where approximately 6 000 children are admitted to the hospital each year. In the last fifteen years, an average of 10 admissions each year have been for septic arthritis. In a 30-year study of paediatric skeletal infection reported by Nelson, 682 cases of suppurative arthritis were described with an occurrence of 25 cases among 10 000 annual admissions (Nelson 1991).

Suppurative joint infection is an important disease of childhood. Among 138 cases treated at our institution since 1980, more than half were less than three years of age (Table 1). Although boys are reported to be affected twice as often as girls, in our series this predominance is seen most in the child older than 5 years; in the younger patient, the sex distribution tends to be more equal.

Age	Total	%
⩽2 months	11	8
3–23 months	58	42
2–5 years	36	26
6–11 years	19	14
>11 years	14	10

Table 1 Suppurative joint infection, The Children's Mercy Hospital Kansas City, Missouri 1980 to 1994

A history of non-penetrating trauma can be elicited in many children with septic arthritis; however, most investigators have questioned its role in the pathogenesis of the infection. In one study (Welkon *et al*. 1986), patients with septic arthritis and sterile cultures more frequently reported a history of trauma than those with culture-confirmed pyarthrosis.

An antecedent infection of the upper respiratory tract or otitis media occurs in 77 per cent of patients with septic arthritis due to *Haemophilus influenzae* type b compared with 18 per cent of those with pyarthrosis due to *Staphylococcus aureus* (Syriopolou and Smith 1987). This observation is consistent with findings in children with meningitis due to *H. influenzae* type b which suggests many have had an antecedent upper respiratory infection or otitis media (Harding *et al*. 1973).

Pathogenesis

In most cases, bacteria enter the joint space haematogenously. Less often, direct inoculation of bacteria into the joint space occurs during an episode of penetrating trauma. Contiguous extension of disease from infected soft tissues is felt to be rare in paediatric practice, and occurs most often in the adult diabetic (Argen *et al*. 1966).

Concomitant osteoarticular infection is a frequent occurrence in the neonate and happens occasionally in the older infant or child.

In neonates, osteoarthritis with destruction of the epiphysis and joint is common because metaphyseal and epiphyseal vessels communicate within the cartilaginous precursor of the ossific nucleus (Trueta 1957). As the ossific nucleus and epiphyseal plate form, separate vessels arise to supply the epiphysis and the communication with metaphyseal vessels disappears at approximately 1 year of age (Ogden 1979).

In children, there are four areas where the bony metaphysis is intracapsular: the proximal femur, proximal humerus, proximal radius, and distal lateral tibia (Morrissy 1989). If bacteria invade the joint from an adjacent osteomyelitis, the clinical presentation is similar to that of isolated joint infection and it is often difficult to differentiate between the two. Data from our institution suggest this type of skeletal infection may occur in up to 20 per cent of cases. In these cases, longer treatment with antibiotics is needed than in primary joint infection. More importantly, these patients have a greater frequency of permanent disability (Jackson *et al*. 1992).

Clinical manifestations

Monoarticular infection of weight-bearing joints is characteristic, with involvement of the knee, hip, or ankle accounting for 70 per cent of cases (Table 2). Usually the child with pyogenic joint infection presents acutely with fever and an exquisitely painful, swollen joint. In some cases, it may be difficult to localize the involved joint,

especially in a febrile, irritable infant. A careful examination after giving a short-acting sedative may be needed to locate the affected joint.

Joint	Total	%
Knee	46	33
Hip	38	28
Ankle	14	10
Elbow	9	7
Shoulder	8	6
Metatarsal	5	4
Sacroiliac	3	2
Wrist	3	2
Multiple	11	9
Interphalangeal	1	–

Table 2 Distribution of involved joints in suppurative joint infection, The Children's Mercy Hospital, Kansas City, Missouri 1980 to 1994

Diagnostic evaluation should be pursued promptly, especially in cases of septic arthritis of the hip where compromise of the vascular supply to the femoral head may result in destruction of the capital epiphysis and growth plate. A complete blood count, erythrocyte sedimentation rate (**ESR**), and radiographs of the joint involved should be obtained. In most cases, leucocytosis and elevation of the ESR in the range of 50 to 90 mm/h will be found; however, the diagnosis should not be excluded even if both of these are normal (Nelson and Koontz 1966). Radiographs may reveal periarticular, soft tissue swelling and joint effusion but also may be non-diagnostic (Morrey *et al*. 1975).

Radioisotope bone scanning is helpful in occasional cases where localization of the diseased joint is difficult. An increase or decrease in uptake on either side of the joint line with limitation to the joint capsule is seen in suppurative arthritis (Tuson *et al*. 1994). Magnetic resonance imaging has been shown to be a sensitive tool in the diagnosis of septic sacroiliitis (Sandrasegaran *et al*. 1994). In the child whose hip is painful on examination but plain radiographs are unrevealing, ultrasonography of the hip is useful in indentifying fluid and in guiding arthrocentesis (Dorr *et al*. 1988).

Aspiration of the joint is imperative to confirm the diagnosis. The fluid is usually frankly purulent and low glucose concentrations may be found. In most cases, the white blood cell count in the synovial fluid is greater than 50 000 and a differential count reveals more than 90 per cent polymorphonuclear leucocytes. A Gram-stained smear of the joint fluid will provide a presumptive identification of the causative agent in approximately one-third of cases. Culture of the joint fluid will reveal the precise organism in 50 to 60 per cent of cases. In 10 per cent blood culture will demonstrate the causative agent when cultures of the joint fluid are sterile. Despite careful culture of blood, joint fluid, and other appropriate sites, a microbiological diagnosis will not be obtained in 25 to 30 per cent of cases. If the clinical diagnosis is deemed to be septic arthritis, lack of a defined pathogen should not change the management of the patient.

General aetiology

The most commonly identified, causative bacteria of suppurative arthritis are presented in Table 3. Although *Haemophilus influenzae* type b has historically caused the majority of cases of suppurative arthritis in infants less than two years of age, we have seen no cases due to this pathogen since 1991 (Jackson, in press). Since the implementation of Hib vaccine in the United States in 1985, this organism has virtually disappeared as a cause of invasive disease. Currently, *Staphylococcus aureus* and streptococci account for close to 70 per cent of culture-confirmed cases of suppurative arthritis. Most streptococcal infections are caused by group A and B (neonates) or pneumococci, but, viridans streptococci and enterococci have also been reported (Fink and Nelson 1986). In teenagers with multiple joint involvement, gonococci are the most common pathogens (Keiser *et al*. 1968). Primary meningococcal arthritis is a rare form of meningococcal disease characterized by polyarticular infection, no involvement of other organs, and excellent prognosis (Schaad *et al*. 1981). *Pseudomonas aeruginosa* may cause infection in cases of puncture wound to the foot; other Gram-negative organisms cause infection infrequently (Fisher *et al*. 1985). Occasionally, salmonella arthritis occurs in the setting of salmonellosis or in the sickleaemic patient (Mallouh and Talab 1985). Anaerobes including *Bacteroides* spp., *Fusobacterium* spp. and *Eikenella corrodens* are most commonly involved in joint infection following a human bite (Resnick *et al*. 1985). *Kingella kingae* has been reported to be a possible aetiological agent in some septic arthritis cases and grew only when the initial synovial fluid specimen was inoculated into a blood culture bottle (Dagan 1993).

Age	Sterile	*S. aureus*	Hib[a]	Cocci[b]	GC[c]	Other[d]
<2 months	9	36	0	36	0	18
2–6 months	40	0	0	60	0	0
7–12 months	26	11	41	15	0	7
13–23 months	36	25	32	0	0	7
24–36 months	29	7	14	36	0	14
3–5 years	35	30	5	20	0	10
6–11 years	35	30	10	10	0	30
>11 years	13	33	0	0	27	26

[a]Haemophilus influenzae type b.
[b]Streptococci including enterococci, group A and B streptococci, viridans streptococci, pneumococci.
[c]Gonococcal.
[d]P. aeruginosa, acinetobacter, enterobacter, klebsiella, coagulase-negative staphylococci, Eikenella corrodens, salmonella, Brucella spp.

Table 3 Aetiology of suppurative joint infection (percentages) by age, The Children's Mercy Hospital, Kansas City, Missouri 1980 to 1994

Specific pathogens in septic arthritis

H. influenzae arthritis

With the advent of conjugate vaccine against *Haemophilus influenzae* type b, the prominence of the pathogen has diminished. Hib vaccine has been used in the United States since 1985, and the more efficacious, conjugate vaccine in infants from 2 months of age since 1990. Between 1986 and 1991, we noted a decline in the incidence of culture-confirmed, septic arthritis from 60 per cent to 20 per cent of infants less than 2 years of age. We have had no cases of Hib arthritis in our institution in the last three years. This changing trend was found in one other study before widespread use of the vaccine; thus the decline in this type of septic arthritis cannot be attributed solely to vaccine-induced immunity, but also to a reduction in nasopharyngeal colonization with this organism (Speiser *et al*. 1985).

Unimmunized infants of less than 2 years or incompletely immunized infants (those infants less than 12 to 15 months of age who have received the appropriate primary vaccine series but no booster) remain susceptible to Hib infection. In such instances, evaluation and management decisions should include Hib as a possible pathogen. As in other cases of disease from invasive *H. influenzae* type b, concomitant meningitis occurs in up to 30 per cent of affected children (Rotbart and Glode 1985). Examination of the cerebrospinal fluid is therefore an essential part of the initial evaluation of the young infant in whom septic arthritis due to *H. influenzae* type b is suspected.

A 1995 report notes the occurrence of the invasive, paediatric disease caused by *H. influenzae* type f, and suggests that reduction in nasopharyngeal Hib carriage may be associated with an increase in carriage of other non-type b strains (Nitta *et al*. 1995). A recent report of skeletal infection due to this organism highlights its ability to cause bone and joint infection (Chusid *et al*. 1992).

Gonococcal arthritis

Suppurative arthritis due to *Neisseria gonorrheae* is one of the manifestations of the disseminated gonococcaemia syndrome (Fig. 1). In more than one-half of cases,

polyarticular disease is found, with joints of the arm, especially wrists and fingers, most frequently involved (Bayer 1980). A history of migratory polyarthralgia in a febrile adolescent with arthritis of the wrist is so clinically distinct that a presumptive diagnosis can be made and appropriate therapy begun. Although analysis and culture of joint fluid is mandatory in all patients, the organism is found in joint fluid in less than one-half of cases. It is essential to culture from the pharynx, rectum and vagina/urethra in all patients with a suspected diagnosis of gonococcal arthritis (Angerine and Hall 1976).

Fig. 1 Typical skin lesions in a teenager with gonococcal arthritis

Septic arthritis of specific joints

Hip

When the hip is involved in septic arthritis, the child will assume a position of comfort with the hip flexed, externally rotated, and abducted. Asymmetry of the skin folds of the buttocks and thighs may be apparent, and significant limitation of passive hip abduction and extension is typical. Pathological subluxation of the hip may occur, and was found in 8 per cent of children with septic arthritis of the hip at our institution.

There is much potential for permanent disability in cases of septic arthritis of the hip. The femoral capital epiphysis is entirely intra-articular and capsular distension may interfere with the vascular supply to the femoral head. Therefore, in all cases where septic hip is suspected, the diagnosis should be promptly confirmed by aspiration of the joint under fluoroscopy. If purulent material is found, immediate drainage of the hip should be performed.

Sacroiliac joint

Infection of the sacroiliac joint is relatively unusual, accounting for 1 to 2 per cent of cases of septic arthritis. These children tend to be older (mean age, 10 years) and usually do not appear systemically ill. Most have acute onset of fever, with pain in the buttocks, groin, hip or abdomen. Often the patient has an antalgic gait but pain is poorly localized. The duration of symptoms before diagnosis averaged 7 days in one study (Patterson 1970), but chronic symptoms of up to 1 month have been reported (Schaad *et al*. 1980). The most common physical finding in patients with sacroiliac arthritis is a positive Fabere sign (pain with flexion, abduction and external rotation of the hip while stressing the sacroiliac joint by placing pressure on the flexed ipsilateral knee) (Hoppenfeld 1976).

As in many cases of septic arthritis, plain radiographs are often normal. A bone scan may localize the involved sacroiliac joint; however, a computed tomographic or magnetic resonance imaging scan is more sensitive and specific (Morgan *et al*. 1981).

Staphyloccocus aureus is the most commonly recognized pathogen and a positive blood culture is often found (Reilly *et al*. 1988). Antistaphylococcal therapy usually results in a prompt clinical response and drainage of the sacroiliac joint is usually unnecessary. *Brucella* spp. have a predilection for sacroiliac involvement and should be considered in individuals who were born or have travelled to endemic areas or have ingested raw milk. Diagnosis of brucella sacroiliitis may be confirmed by culture or serology and therapy should usually include trimethoprim sulphamethoxazole or doxycycline (for patients over 9 years of age) possibly with the addition of rifampicin.

Septic arthritis specific hosts

Neonates (Table 4)

Although suppurative skeletal infection in a newborn infant often reflects systemic infection with *Staphylococcus aureus, Candida albicans* or group B streptococci, systemic symptoms and signs are generally not apparent (Pittard *et al*. 1976). Non-specific symptoms such as irritability, lethargy or poor feeding may occur. However, a non-toxic appearance and absence of fever is found in more than one-half of cases, and a normal white cell count and ESR are generally found (Fox and Sprunt 1978). Local swelling of an extremity with overlying inflammation and/or flexion contracture are the most common clinical findings. Occasionally, a newborn with hip pyarthrosis may present with abdominal distension secondary to rupture into the abdomen of intra-articular pus (Freiberg and Perlman 1976).

Absence of systemic symptoms and signs
Multiple concomitant osteoarticular sites
Delayed diagnosis
Frequent permanent sequelae

Table 4 Characteristics of neonatal skeletal infection

A high index of suspicion is necessary and aspiration of all suspected joints is absolutely essential to confirm the diagnosis. Permanent sequelae are frequent in this age group. Hallel and Salvati (1975) found deformities of the femoral head in 70 per cent of neonates followed up for suppurative arthritis of the hip. Complications from skeletal infection in the neonate can be minimized by a high index of suspicion, detailed evaluation with a careful musculoskeletal examination in any baby with suspected sepsis, and early surgical drainage and antibiotic therapy.

Septic arthritis in chronic joint disease

The diagnosis of suppurative joint infection may be particularly difficult in the patient with underlying chronic disease of the joints, such as those with sickle haemoglobinopathy, haemophilia or systemic juvenile chronic arthritis. Although there is no hallmark for differentiating chronic, active disease from acute suppurative infection of the joint, most infected patients are febrile, have a more toxic appearance, and have marked local manifestations.

In haemophiliacs, septic arthritis has been considered a rare complication of haemarthrosis. Severe joint pain and systemic toxicity that persist despite giving

coagulation factors should be considered an indication for needle aspiration of the affected joint. Pneumococci (which have a predilection for diseased joints) and staphylococci are the most common pathogens of septic arthritis in haemophiliacs (Scott *et al*. 1985). Generally, a single joint, usually the knee, is involved and diagnosis may be delayed when a septic joint is superimposed on haemarthrosis. Recently, septic arthritis complicating pneumococcal, staphylococcal or salmonella bacteraemia has been reported in haemophiliacs infected with human immunodeficiency virus. Polyarticular disease with significant destruction occurred despite appropriate therapy (Ragni and Hawley 1989). However, a more favourable outcome has been reported in four recent cases where non-operative management with prompt antimicrobial therapy was advocated (Merchan *et al*. 1992).

Septic arthritis is uncommon in children with sickle haemoglobinopathy. Localized joint pain and fever are usually found, and symptoms and signs tend to persist for 4 days or longer (Syrogiannopoulos *et al*. 1986). Encapsulated bacteria typically cause infection in the patient with sickle haemoglobinopathy. Pneumococci are the most commonly recognized pathogens in such children with septic arthritis, although occasional disease secondary to *H. influenzae* type b is reported (Mallouh and Talab 1986). Salmonella have a predilection for bone in sicklaemic patients and may be found in cases of suppurative arthritis in association with an adjacent osteomyelitis (Seeler and Jacobs 1977).

Joint infection during rheumatoid arthritis is more common in adults than children. Usually, worsening of pain of an isolated joint occurs during a flare-up of the underlying disease; however, there is polyarticular involvement in almost 40 per cent of cases (Kaufman *et al*. 1976). As in other cases where septic arthritis complicates underlying chronic disease of the joint, a high index of suspicion is necessary to avoid a delay in diagnosis and poor subsequent outcome.

Penetrating injuries and arthritis

Suppurative arthritis following a penetrating injury to a joint is well described, although when first reported was most commonly associated with sewing-needle injury to the knee (Samilson *et al*. 1958). The knee is still the joint usually involved, but sewing needles have rarely been implicated as the agent of trauma in the last two decades. The clinical presentation is similar to other cases of pyarthrosis. In our series, 8 per cent of cases of suppurative arthritis followed a penetrating injury (Table 5). All cases were boys, with a mean age of 6 years. A wide variety of pathogens was found; however, *Pseudomonas aeruginosa* arthritis of the metatarsal joint following nail puncture occurred most frequently.

Age (years)	Sex	Joint	Trauma	Pathogen
2	M	Knee	Toothpick	Viridans streptococci
3	M	Knee	Plastic	Group A streptococci
3	M	Knee	Dog bite	Unknown
1	M	Knee	Human bite	Eikenella corrodens
4	M	Metatarsal	Nail	*Pseudomonas aeruginosa*
12	M	Knee	Nail	Gram-negative rods
4	M	Metatarsal	Nail	*Pseudomonas aeruginosa*
9	M	Metatarsal	Unknown	*Escherichia coli, Enterobacter cloacae*
7	M	Metatarsal	Nail	*Pseudomonas aeruginosa*
9	M	Metatarsal	Nail	*Pseudomonas aeruginosa*
10	M	Metatarsal	Nail	*Pseudomonas aeruginosa*

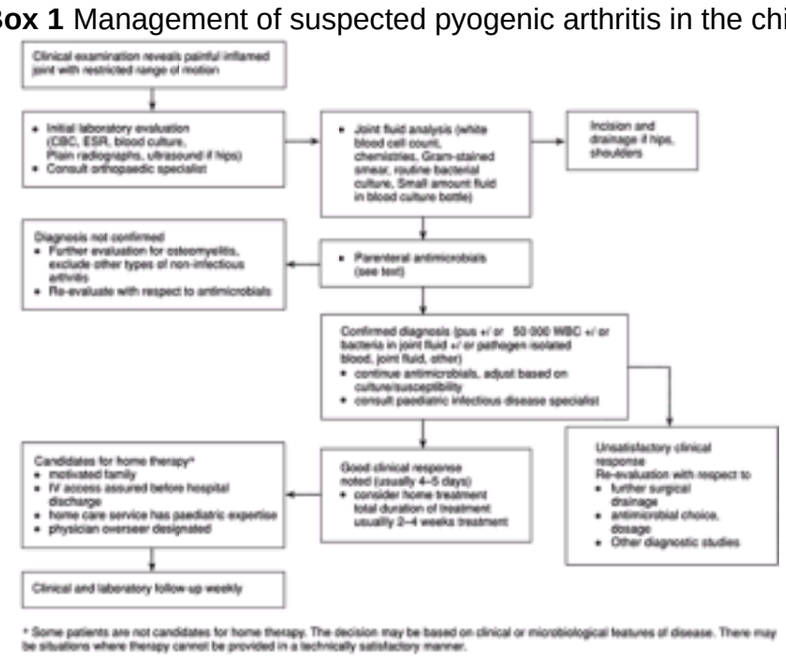
Table 5 Suppurative arthritis following puncture wound

Reactive or so-called traumatic arthritis follows an injury, usually a thorn or splinter puncture, and may be severe and difficult to distinguish from acute suppurative infection. In most cases, symptoms have been present for longer than a week and although there is pain on moving the joint, it is usually not as severe as that associated with a pyarthrosis (Green and Edwards 1987). Exploration of the joint to exclude a foreign body should be considered in cases where traumatic synovitis persists.

Management

The essentials of management of suppurative arthritis in childhood include the combination of an appropriately selected antimicrobial agent and adequate drainage of the pyarthrosis (see Box 1). Simple needle aspiration of the affected joint may be sufficient in some cases. However, emergency surgical drainage of the joint is mandatory in pyarthrosis of the hip and probably shoulder infection (Patterson 1978). Arthrotomy and drainage should also be considered in cases where the pus is thick, or when significant symptoms and signs persist despite initial needle aspiration of pus from the affected joint (Nade 1983).

Box 1 Management of suspected pyogenic arthritis in the child



When a presumptive diagnosis of septic arthritis is made, parenteral antibiotics should be begun promptly. In older children, an antistaphylococcal penicillin is the drug of choice. For infants who have not received a Hib booster vaccine (those less than 12 to 15 months of age), and for infants who have not received Hib vaccine and are less than 2 years old, I add coverage for *H. influenzae* type b in combination with the antistaphylococcal antibiotic. Cefuroxime is appropriate initial coverage as long as concomitant meningitis is not present. In that case, I prefer to combine an antistaphylococcal penicillin with a third-generation cephalosporin (usually cefotaxime or ceftriaxone) until the results of culture are known. Parenteral antimicrobial agents readily penetrate into infected joints, so intra-articular installation of antibiotics is not necessary.

The efficacy of a sequential intravenous–oral antimicrobial regimen has been demonstrated by numerous prospective studies (for example Feigin *et al*. (1975); Jackson and Nelson (1982)). Usually, intravenous therapy is given for 5 to 7 days. After all surgical procedures and when there is definite clinical improvement, a patient may be considered for an oral regimen if (i) an appropriate oral antimicrobial is available, (ii) the patient is able to take and retain oral medication, and (iii) there is a laboratory at hand to analyse the bactericidal activity of serum samples.

The oral dosage of the selected antibiotic is two to three times that used for otitis media or skin and soft tissue infection. A serum specimen for blood levels should be obtained 1 h after the oral dose is given (after the drug has reached steady-state concentrations) and adequate bactericidal activity should be confirmed. A serum bactericidal activity of at least 1:8 is considered adequate except in cases of streptococcal infection where a titre of 1:32 is needed. Adjustment of the antimicrobial dosage is usually required in 15 per cent of cases. In less than 5 per cent of cases, satisfactory bactericidal activity cannot be achieved and a total parenteral regimen must be used for the duration of therapy (Prober and Yeager 1979). Oral antibiotic regimens should not be used outside the hospital setting in a child unless you feel

compliance can be guaranteed.

The optimum duration of antimicrobial therapy for children with suppurative arthritis is dependent upon the duration of symptoms before diagnosis of the pathogen, the response to medical and surgical treatment, and whether or not there is concomitant bone infection. Generally, if the patient's clinical response is good and the ESR returns to normal, 3 weeks of therapy is adequate for staphylococcal and Gram-negative bacillary infection (Syrogiannopoulos and Nelson 1978). Shorter courses have been successful for streptococcal and haemophilus infection, in which a total of 10 to 14 days of therapy may be adequate. In all cases where an oral regimen is used, ESR and serum bactericidal activity should be monitored weekly until therapy is considered complete.

The risk of relapse or recurrent disease is quite small in cases of primary joint infection; however, for the child with pyarthrosis and concomitant bone infection, the potential for relapse or sequelae is significant. Therefore, if compliance cannot be guaranteed, an oral regimen is not appropriate for such patients. Delayed diagnosis, delay in surgical drainage, slow clinical response, and undocumented compliance are all risk factors for chronic disease (Badgley *et al*. 1936; Hallel and Salvati 1975).

Prognosis

In the child over 1 month of age, the prognosis for primary septic arthritis of the knee is good and more than 90 per cent have a satisfactory outcome (Howard *et al*. 1976). However, poor outcome has been estimated to occur in approximately 40 per cent of hip, 20 per cent of ankle and 30 per cent of shoulder infections, particularly if there is an adjacent osteomyelitis (Gillespie 1973; Welkon *et al*. 1986). There should be careful orthopaedic follow-up for at least 6 to 12 months in all children with supporative joint infection.

Chapter References

Angerine, C.D. and Hall, C.B. (1976). A case of gonococcal osteomyelitis. A complication of gonococcal arthritis. *American Journal of Diseases of Childhood*, **130**, 1013–14.

Argen, R.J., Wilson, C.H., and Wood, P. (1966). Suppurative arthritis: clinical features of 42 cases. *Archives of Internal Medicine*, **117**, 661–6.

Badgley, C.E., Yglesias, L., Perham, W.S., and Snyder, C.H. (1936). Study of the end result of 113 cases of septic hips. *Journal of Bone and Joint Surgery*, **18**, 1047–53.

Bayer, A.S. (1980). Gonococcal arthritis syndromes: an update on diagnosis and management. *Postgraduate Medicine*, **67**, 200–8.

Chusid, M.J., Schneider, J.P., Thometz, J.G., and Dunne, W.M. (1992). Osteomyelitis and septic arthritis caused by *Haemophilus influenzae*, type f, in a young girl. *Diagnostic Microbiology and Infectious Diseases*, **15**, 157–9.

Dagan, R. (1993). Management of acute hematogenous osteomyelitis and septic arthritis in the pediatric patient. *Pediatric Infectious Disease Journal*, **12**, 88–93.

Dorr, U., Zieger, M., and Hauke, H. (1988). Ultrasonography of the painful hip. Prospective studies in 204 patients. *Pediatric Radiology*, **19**, 36–40.

Feigin, R.D., Pickering, L.K., Anderson, B., Kooney, R.D., and Shackleford, P.G. (1975). Clindamycin treatment of osteomyelitis and septic arthritis in children. *Pediatrics*, **55**, 213–23.

Fink, C.W. and Nelson, J.D. (1986). Septic arthritis and osteomyelitis in children. *Clinical Rheumatic Disease*, **12**, 423–35.

Fisher, M.C., Goldsmith, J.F., and Gilligan, P.H. (1985). Sneakers as a source of *Pseudomonas aeruginosa* in children with osteomyelitis following puncture wounds. *Journal of Pediatrics*, **106**, 607–9.

Fox, L. and Sprunt, K. (1978). Neonatal osteomyelitis. *Pediatrics*, **62**, 535–42.

Freiberg, J.A. and Perlman, R. (1976). Pelvic abscesses associated with acute purulent infection of the hip joint. *Journal of Bone and Joint Surgery*, **18**, 417–27.

Gillespie, R. (1973). Septic arthritis in childhood. *Clinical Orthopaedics and Related Research*, **96**, 152–9.

Green, N.E. and Edwards, K. (1987). Bone and joint infections in children. *Orthopedic Clinics of North America*, **18**, 555–76.

Hallel, T. and Salvati, E.A. (1975). Septic arthritis of the hip in infancy. End result study. *Clinical Orthopaedics*, **132**, 115–28.

Harding, A.L., Anderson, P., and Howse, V.M. (1973). *Haemophilus influenzae* isolated from children with otitis. In *Haemophilus influenzae* (eds H. Sells and D.T. Karzon), p. 663. University Press, Nashville, TN.

Hoppenfeld, S. (1976). *Physical examination of the spine and extremities*, p. 262. Appleton-Century-Crofts. New York, NY.

Howard, J.B., Highgenboten, C.L., and Nelson, J.D. (1976). Permanent effects of septic arthritis in childhood. *Journal of the American Medical Association*, **236**, 932–5.

Jackson, M.A. and Nelson, J.D. (1982). Etiology and medical management of acute suppurative bone and joint infections in pediatric patients. *Journal of Pediatric Orthopedics*, **2**, 313–23.

Jackson, M.A., Burry, V.F., and Olson, L.C. (1992). Pyogenic arthritis associated with adjacent osteomyelitis: identification of the sequelae-prone child. *Pediatric Infectious Disease Journal*, **11**, 9–13.

Jackson, M.A., Desai, K.B., Burry, V.F., and Olson, L.C. (1995?). The changing face of childhood pyogenic arthritis. *Pediatric Infectious Disease Journal*. (In press.)

Kaufman, C.A., Watanakunakorn, C., and Phair, J.P. (1976). Pneumococcal arthritis. *Journal of Rheumatology*, **3**, 409–19.

Keiser, H., Rubin, F.L., Wolmsky, E., and Kushner, I. (1968). Clinical forms of gonococcal arthritis. *New England Journal of Medicine*, **279**, 234–40.

Mallouh, A. and Talab, Y. (1985). Bone and joint infection in patients with sickle cell disease. *Journal of Pediatric Orthopedics*, **5**, 158–62.

Merchan, E.C.R., Magallon, M., Mauso, F., and Martin-Billar, J. (1992). Septic arthritis in HIV positive haemophiliacs. *International Orthopaedics*, **16**, 302–6.

Morgan, G.J., Schlegelmilch, J.G., and Spiegel, P.K. (1981). Early diagnosis of septic arthritis of the sacroiliac joint by computed tomography. *Journal of Rheumatology*, **8**, 979–82.

Morrey, B.F., Bianco, A.J., and Rhodes, K.H. (1975). Septic arthritis in children. *Orthopedic Clinics of North America*, **6**, 923–34.

Morrissy, R.T. (1989). Bone and joint infection in the neonate. *Pediatric Annals*, **18**, 33–44.

Nade, S. (1983). Acute septic arthritis in infancy and childhood. *Journal of Bone and Joint Surgery*, **65B**, 234–41.

Nelson, J.D. (1991). Skeletal infections in children.. In *Advances in pediatric infecious diseases*, Vol. 6, (eds S.C. Aronoff, W.T. Hughes, S. Kohn, W.T. Speck, and E.R. Wald), pp. 59–76. Mosby Year Book, St. Louis, MO.

Nelson, J.D. and Koontz, W.L. (1966). Septic arthritis in infants and children. *Pediatrics*, **38**, 966–71.

Nitta, D.M., Jackson, M.A., Burry, V.F., and Olson, L.C. (1995). Invasive *Haemophilus Influenzae* type f disease. *The Pediatric Infectious Disease Journal*, **14**, 157–60.

Ogden, J.A. (1979). Pediatric osteomyelitis and septic arthritis: the pathology of neonatal disease. *Yale Journal of Biological Medicine*, **52**, 423–48.

Patterson, D. (1978). Septic arthritis of the hip. *Orthopedic Clinics of North America*, **9**, 135–42.

Patterson, D.C. (1970). Acute suppurative arthritis in infancy and childhood. *Journal of Bone and Joint Surgery*, **52B**, 414–82.

Pittard, W.B., Thullen, J.D., and Fanaroff, A.A. (1976). Neonatal septic arthritis. *Journal of Pediatrics*, **88**, 621–4.

Prober, C.G. and Yeager, A.S. (1979). Use of serum bactericidal titer to assess the adequacy of oral antibiotic therapy in treatment of acute hematogenous osteomyelitis. *Journal of Pediatrics*, **95**, 131–55.

Ragni, M.V. and Hawley, E.N. (1989). Septic arthritis in hemophiliac patients and infection with human immunodeficiency virus. *Annals of Internal Medicine*, **110**, 168–9.

Reilly, J.P., Gross, R.H., Eman, J.B., and Yngve, D.A. (1988). Disorders of the sacroiliac joint in children. *Journal of Bone and Joint Surgery*, **70A**, 31–40.

Resnick, D., Pineda, C.J., Weisman, M.H., and Kerr, R. (1985). Osteomyelitis and septic arthritis of the hand following human bites. *Skeletal Radiology*, **14**, 263–6.

Rotbart, H.A. and Glode, M.P. (1985). *Haemophilus influenzae* type b septic arthritis in children: report of 23 cases. *Pediatrics*, **75**, 254–9.

Samilson, R.L., Bersani, F.A., and Watkins, M.B. (1958). Acute suppurative arthritis in infants and children: the importance of early diagnosis and surgical drainage. *Pediatrics*, **21**, 798–803.

Sandrasegaran, K., Saifuddin, M.B., Coral, A., and Butt, W.P. (1994). Magnetic resonance imaging of septic sacroiliitis. *Skeletal Radiology*, **23**, 289–92.

Schaad, V.B., McCracken, G.H., and Nelson, J.D. (1980). Pyogenic arthritis of the sacroiliac joint in pediatric patients. *Pediatrics*, **66**, 375–9.

Schaad, V.B., Nelson, J.D., and McCracken, G.H. (1981). Primary meningococcal arthritis. *Infection*, **9**, 170–3.

Scott, J.P., Maurer, H.S., and Dias, L. (1985). Septic arthritis in two teenaged hemophiliacs. *Journal of Pediatrics*, **107**, 748–51.

Seeler, R.A. and Jacobs, N.M. (1977). Pyogenic infections in children with sickle hemoglobinopathy. *Journal of Pediatrics*, **90**, 161–2

Smith, T. (1974). On the arthritis of infants. *St. Bartholomew's Hospital Report*, **10**, 189–204.

Speiser, J.C., Moore, T.L., Osborn, T.G., Weiss, T.D., and Zuckner, J. (1985). Changing trends in pediatric septic arthritis. *Seminars in Arthritis and Rheumatism*, **15**, 132–8.

Syriopoulou, V.P. and Smith, A.L. (1987). In septic arthritis. In *Pediatric infectious diseases* (eds R. Feigin and J. Cherry), pp. 773–6. W.B. Saunders, Philadelphia, PA.

Syrogiannopoulos, G.A. and Nelson, J.D. (1978). Duration of antimicrobial therapy for acute suppurative osteoarticular infections. *Lancet*, **i**, 37–40.

Syrogiannopoulos, G.A., McCracken, G.H., and Nelson, J.D. (1986). Osteoarticular infections in children with sickle cell disease. *Pediatrics*, **78**, 1090–6.

Trueta, J. (1957). The normal vascular anatomy of the femoral head. *Journal of Bone and Joint Surgery*, **39B**, 358–94.

Tuson, C.E., Hoffman, E.B., and Mann, M.D. (1994). Osotope bone scanning for acute osteomyelitis and septic arthritis in children. *Journal of Bone and Joint Surgery*, **76B**, 306–10.

Welkon, C.J., Long, S.S., Fisher, M.C., and Alburger, P.D. (1986). Pyogenic arthritis in infants and children: a review of 95 cases. *Pediatric Infectious Disease Journal*, **5**, 669–76.

5.3.3 Osteomyelitis and associated conditions

Emilio Bouza and Patricia Muñoz

Introduction

The term osteomyelitis, although specifically applied to medullary infections involving trabecular bone, is used to describe any infection involving bone and marrow (Brause 1990; Laughlin *et al*. 1994; Laughlin *et al*. 1995).

The appearance of new groups of immunosuppressed patients at risk of osteomyelitis, and the emergence of new patterns of antimicrobial resistance in many micro-organisms responsible for bone and joint infections, is a cause of concern (Poiraudeau *et al*. 1993; Rogeaux *et al*. 1993). At the same time, recent and important changes in the field include the application of more precise diagnostic techniques and the use of aggressive surgery with implantation of new prosthetic devices. Finally, the development of non-toxic, highly efficacious, oral antimicrobial agents frequently permits a long-term approach to these difficult-to-treat infections during which the patient remains ambulatory.

First, general aspects of osteomyelitis, with special reference to chronic infections of the long bones, will be discussed. Other forms of osteomyelitis and arthritis will be considered at the end of the chapter.

Classification

The utility of a classification system depends on its ability to predict relevant prognostic, therapeutic, or aetiological data and to provide uniform criteria for comparative studies (Mader *et al*. 1992). A universally applied classification system for stratifying osteomyelitis and prosthetic joint infection would provide a framework to evaluate the efficacy of medical and surgical treatments in different institutions. To date, no attempt at this has been completely successful and there is no single classification system that is satisfactory. Osteomyelitis must, therefore, be classified on the basis of several characteristics. All of them are relevant to planning management and evaluating outcome. The most common criteria are detailed in Table 1.

Criteria	Classification
1. Age	Children — adults
2. Affected bone	Long — short
3. Host	Drug abusers — immunosuppressed — normal
4. Aetiology	Bacterial — non-bacterial
5. Pathogenesis	Haematogenous — contiguous
6. Existence of fracture	Consolidated — non-unions
7. Risk factors	With/without prosthetic material
8. Blood supply	Adequate — inadequate
9. Evolution	Acute — chronic

Table 1 Classification criteria for osteomyelitis

1. The age of the patient must always be considered, since pathogenesis and aetiology are usually different in neonates, children or adults. Haematogenous osteomyelitis is far more frequent in children. *Staphylococcus aureus, Enterobacteriaceae*, and group A and B b-haemolytic streptococci are the most common aetiological agents in neonates. In children under 4 years of age *Haemophilus influenzae* is the most important pathogen, followed by streptococci and *Staph. aureus*. In children older than 4 years *Staph. aureus*, streptococci, and *H. influenzae* predominate.
2. The affected bone is also an important factor. Long bones, particularly those of the lower limbs, are more susceptible to infection because their blood supply is poorer than in short bones. Also, they bear more weight and have worse venous return. Pelvic and cranial bones are infrequently involved (Sexton *et al*. 1993; Bernier *et al*. 1995; Lentz 1995).
3. The host is also very important and special risk groups must be considered. In intravenous drug abusers, atypical locations are more frequently encountered (pubic bones, clavicle or vertebra) and, besides staphylococci and streptococci, Gram-negative rods, mainly *Pseudomonas aeruginosa*, and yeasts must be considered as potential aetiological agents (Chandrasekar and Narula 1986). Patients with haemoglobinopathies, such as sickle-cell disease, have a higher incidence of infections caused by encapsulated bacteria such as *Strep. pneumoniae, H. influenzae*, or salmonella (Engh *et al*. 1971). Chronic haemodialysis is also a risk factor for osteomyelitis. The most common sites involved are the ribs and thoracic spine.
4. Most osteomyelitis is caused by Gram-positive bacteria, but practically any other micro-organisms may be responsible for bone infection. Gram-negative bacteria are frequently found in post-traumatic and postsurgical osteomyelitis, and anaerobic and mixed infections should also be considered in osteomyelitis associated with poor vascular supply such as in diabetics (Gerding 1995). Brook and Frazier (1993) isolated anaerobes in 33 per cent of cases of osteomyelitis diagnosed at a military institution. *Mycobacteria*, and particularly *Mycobacterium tuberculosis*, are still an important cause of osteomyelitis, and it is occasionally associated with other micro-organisms (*staphylococci*). Finally, fungal infections of the bone have been described in all agents responsible for systemic mycosis both in the normal and the immunocompromised host.
5. Taking into account the mechanism by which the infecting organisms reach the bone, osteomyelitis can be classified as haematogenous, secondary to a contiguous focus of infection, and caused by direct inoculation. Acute haematogenous osteomyelitis is more common in children and will be discussed later. In

adults, haematogenous osteomyelitis frequently involves the spine. This constitutes a diagnostic challenge since clinical and radiological manifestations may be non-specific and rapid evolution may produce significant neurological sequelae (Sapico and Montgomerie 1990; Heary *et al*. 1994). A recent review demonstrates a trend for haematogenous osteomyelitis to have shifted from its known incidence in early age to adulthood, from acute to insidious onset, and from infection by Gram-positive to Gram-negative organisms (Sharma *et al*. 1993).

6. The existence of an underlying fracture, above all with non-union (pseudoarthrosis), is of great importance, since it will influence the surgical approach.
7. The presence of any kind of prosthetic material near the infection site must be carefully sought, since it facilitates persistence of infection, increases the pathogenicity of micro-organisms such as *Staph. epidermidis*, and hinders the efficacy of antimicrobial treatment. Although the withdrawal of prosthetic materials is regularly recommended, it is not always necessary, or possible.
8. Waldvogel *et al*. (1970) classified contiguous osteomyelitis with respect to the integrity of the bone's vascular supply. The most characteristic example of osteomyelitis with poor vascular flow is that associated with infections of the diabetic foot. Neuropathic and vascular changes characteristic of diabetes mellitus put patients at risk for developing chronic foot damage after minor trauma and subsequent bone infection (Gerding 1995).
9. The disease activity in osteomyelitis is classified as acute or chronic. Osteomyelitis is considered acute when the appearance of symptoms is recent, and when there has been no previous therapy. On the other hand, osteomyelitis is chronic when symptoms have been present for 4 to 6 weeks or previous therapy has been given. This classification is important, since therapy and prognosis will be very different. At the present time, less than 6 per cent of haematogenous osteomyelitis becomes chronic but practically all post-traumatic infections are chronic.

Previous classifications have considered these important aspects, the most essential being the presence of vascular compromise, the presence of prosthetic material, and the rate of evolution.

Cierny and Mader (1984) proposed a commendable classification for adult chronic osteomyelitis that takes into account the anatomical status of the bone lesion, the involvement of soft tissues, and the expected host response to infection (Table 2). Accordingly, each case of adult chronic osteomyelitis must be classified with a number and a letter. Nevertheless, the system is imprecise and has not been broadly accepted or generally incorporated into clinical practice. Details of the classification are as follows.

Anatomical	Type I	Medullary
	Type II	Superficial
	Type III	Localized
	Type IV	Diffuse
Physiological	Class A	Normal host
	Class B	Compromised patient
	Class C	Poor-risk candidate

Table 2 Cierny/Mader classification for adult chronic osteomyelitis

Anatomical classification

Type I Intramedullary infection or infected (but consolidated) fractures with an intramedullary rod. Surgical treatment is simple and will not result in bone instability.

Type II Cortical bone infection (pressure sore). Bone excision is easy, although soft-tissue coverage may be more complicated.

Type III Osteomyelitis affecting cortical bone and marrow, but not including the whole bone circumference (sequestrum). Surgical debridement is complicated, although does not necessarily result in instability.

Type IV Osteomyelitis affecting the whole bone circumference. The surgical approach usually requires ablation of a bone segment and may result in bone instability. Type IV includes infected non-unions and infected articular prostheses.

Physiological classification

Class A Patients and tissues with normal response to infection and surgery.

Class B Patients with local or systemic immune deficiencies, that may predispose to infection (Table 3).

Systemic	Local
Malnutrition	Chronic lymphoedema
Renal or hepatic failure	Venous stasis
Immunodeficiency	Large- or small-vessel disease
Neoplasm	Arteritis
Diabetes mellitus	Scars
Extremes of age	Postradiation fibrosis
Smoking (tobacco)	Loss of local sensation
Chronic hypoxia	
Parenteral drug abuse	

Table 3 Local and systemic factors that affect the immunological, metabolic, and vascular response to osteomyelitis (Cierny/Mader)

Class C Patients not considered to be suitable surgical candidates, for whom the morbidity of treatment is greater than the risk of disease or exceeds the expected benefit.

Negative aspects of this classification are that it does not consider the aetiology or the bone involved. Surprisingly, the originators obtained relatively uniform results in all the stages, although compromised patients had a statistically inferior prognosis compared with patients who had apparently normal defence mechanisms (Cierny and Mader 1989).

Risk factors for osteomyelitis

Major risk factors for bone infections are contaminated open fractures, previous surgery, insertion of prosthetic material, delayed postoperative wound healing, and previous infections (Gillespie 1990). In one of the most important series of osteomyelitis, 29 per cent of patients had previous trauma with fracture, 4 per cent trauma without fracture, 42 per cent had recently undergone surgery involving bone, and 11 per cent had a prior open fracture (Guerrero 1989). Pressure sores are closely linked to underlying osteomyelitis, particularly diabetic foot ulcers (Newman *et al*. 1991). The demonstration of underlying bone involvement frequently requires MRI

and indium-111 scintigraphy (Newman *et al*. 1992).

Aetiology

Gram-positive micro-organisms remain the most common causative agents of osteomyelitis at all ages, although there is a trend towards greater involvement of Gram-negative organisms (Gentry and Rodríguez 1990) (Table 4). Gram-negative organisms should particularly be suspected in patients subject to prolonged stays in hospital or previous surgery, or those admitted to intensive care units or with open fractures (Gentry 1990). In contrast to haematogenous osteomyelitis, which is usually caused by a single pathogen, contiguous chronic osteomyelitis may involve multiple organisms.

Micro-organism	1970 (%)	1988 (%)
Staph. aureus	45	27
Other Gram-positive	5	5
Pseudomonas aeruginosa	5	15
Other Gram-negative	8	20
Polymicrobial	37	33

Table 4 Changing profile of micro-organisms isolated from osteomyelitis during the last two decades (per cent) (Gentry 1990)

Staphylococcus aureus is, undoubtedly, the most common aetiological agent of osteomyelitis of any kind (Guerrero 1987). In recent years, the emergence of methicillin-resistant strains (**MRSA**) (strains resistant to all b-lactam drugs) has been described by many hospitals throughout the world (Ish-Horowicz *et al*. 1992). Nevertheless, osteomyelitis caused by MRSA is less common than might be expected.

Staphylococcus epidermidis is one of the most common agents of bone and joint infection in patients with prosthetic materials, and a high proportion of these infections (more than 50 per cent at our institution) are resistant to methicillin. The aetiological role of *Staph. epidermidis* should only be fully accepted when the isolate is obtained from a usually sterile body fluid or tissue, and skin contamination can be reasonably ruled out. *Staphylococcus epidermidis* infection is extremely rare without the presence of underlying prosthetic material (De Wit et al.1993). It is frequently implicated in postsurgical sternal osteomyelitis (Miholic *et al*. 1985).

Although streptococci are of great importance in skin and soft tissue infections, their participation in osteomyelitis is rare (Burkert and Watanakunakorn 1991).

The group B b-haemolytic streptococcus (*Strep. agalactiae*) is exceptional as a cause of osteomyelitis. It is most commonly described in children (Ammari *et al*. 1992; Muñoz *et al*. 1992), or in elderly or immunosuppressed patients (McCarthy and Haber 1987; Elhanan and Raz 1993). Associated bacteraemia may cause high morbidity and mortality (Mateo *et al*. 1993; Farley 1995; Ganapathy and Rissing 1995).

Streptococcus pneumoniae osteomyelitis is also extremely uncommon in adults. It is occasionally described as a single focus of infection in normal children (Jacobs 1991). Cases caused by penicillin-resistant strains have also been reported (Gelfand and Cleveland 1992).

Enterococcus is very rarely implicated in osteomyelitis, and its significance when isolated should be questioned, especially if the biopsy or surgical specimen is not obtained aseptically. The authors have seen two patients with prosthesis-related infections caused by *Enterococcus*, and other cases have been published. In a well-designed study of biopsy confirmed, non-prosthetic osteomyelitis, *Enterococcus* accounted for 3 per cent of the cases (Gentry and Rodríguez-Gomez 1991). Other Gram-positive micro-organisms such as *Listeria* (Housang 1976) or *Bacillus* (Sliman *et al*. 1987; Drobniewski 1993) have only rarely been implicated in osteoarticular infections.

Among the Gram-negative organisms, *Pseudomonas* is one of the most commonly involved in bone infections. It is usually found in osteomyelitis following open fractures or surgical procedures. It is the most frequent cause of calcaneus osteomyelitis following infected puncture wounds of the foot (Dixon and Sydnor 1993; Lavery *et al*. 1994). *Salmonella* osteomyelitis occurs either as a complication of typhoid fever or in patients with various underlying diseases, including immunosuppressed patients and those with sickle-cell disease (Anand and Glatt 1994). Other Enterobacteriaceae have also been described, particularly in infections following open fractures or as a consequence of haematogenous bone involvement in patients with bacteraemia of another origin (Lacour *et al*. 1991; Voss *et al*. 1992).

Conditions predisposing to anaerobic bone infections are vascular disease, bites, a contiguous focus of infection, peripheral neuropathy, haematogenous spread, and trauma. Anaerobes are more frequently detected in osteomyelitis under pressure sores, or in bone infections of the diabetic foot (Hudson 1993). Pigmented *Prevotella* and *Porphyromonas* spp. were mostly isolated in infections of the skull and following bites. Members of the *Bacteroides fragilis* group have been detected in cases of hand and foot infection, and *Fusobacterium* spp. in skull, bite wounds, and haematogenous long-bone infections (Brook and Frazier 1993).

Fungal or nocardial osteomyelitis is found in both normal and immunocompromised hosts (Novak *et al*. 1988; Laurin *et al*. 1991; Pruitt *et al*. 1993; Young 1993; Assaad *et al*. 1994; Straus *et al*. 1994). *Mycobacterium tuberculosis* and, rarely, non-tuberculous mycobacteria, should also be considered as a cause of chronic bone infection (both in normal and immunocompromised hosts) (Cohen and Squires 1992; Jamil *et al*. 1992; Mahan and Jolles 1995 ; Yao and Sartoris 1995). Viruses are only rarely the cause of bone infections (Berman and Jensen 1990; Kain *et al*. 1990).

Diagnostic procedures

The diagnosis of osteomyelitis and prosthetic joint infections is usually made on the basis of clinical, laboratory, and imaging techniques. Although the problem is usually localized, a complete clinical history and examination must be performed (Levine *et al*. 1993). Information on the presence of any kind of prosthetic material, previous surgical and medical therapy, duration of previous antimicrobial courses, and the response must be carefully recorded.

Physical examination should focus on the integrity of involved bone and surrounding tissues, and on evidence of inflammatory signs, pain, bone instability, sinus tracts or neurovascular changes. The nutritional status of the patient should also be considered. The detection of a sinus tract with suppurative drainage will establish the diagnosis in an appropriate clinical setting. Further diagnostic techniques will be required to confirm the diagnosis, if necessary, and to establish the aetiology.

Laboratory investigations

Laboratory data are not essential for the diagnosis of osteoarticular infections. The erythrocyte sedimentation rate is usually high with active infections (92 per cent) and tends to fall after effective therapy. Its accuracy in infected prostheses is variable. In our experience, the erythrocyte sedimentation rate is not useful as an index of either activity or resolution in osteomyelitis.

C-reactive protein is also usually increased, although its measurement is less widely used (Unkila-Kallio *et al*. 1994). Only 35 per cent of patients have leucocytosis at the time of admission.

In summary, no single laboratory measure is reliable enough to be used routinely for the diagnosis of osteomyelitis.

Imaging techniques

Imaging plays an important part in establishing the diagnosis and directing the treatment of osteomyelitis. A variety of imaging methods may be used, including plain radiography, radionuclide imaging, computerized tomography (**CT**), and magnetic resonance imaging (**MRI**). Decisions on the best method can be challenging and should reflect the location of the suspected infection and associated underlying systemic or bone disorder. A brief review of the pathogenesis of osteomyelitis is necessary to help understand when each of imaging techniques is of value (Schauwecker *et al*. 1990; Wegener and Alavi 1991 ; Crim and Seeger 1994).

Once the micro-organism reaches the bone, a suppurative reaction is produced, followed by a marrow oedema, which can only be readily detected by MRI. The next step consists of vascular congestion, thrombosis, and ischaemia. At this time, soft tissue changes may be detected by CT but not by plain radiology. Finally (after at least 2 to 3 weeks), bone reaction begins, with the production of new periosteal bone, sequestrum, decalcification, and new bone formation, which can be detected even with plain films.

Plain films

The detection of acute osteomyelitis on a plain film requires at least a 35 per cent loss of calcium content in the bone lesion. This usually takes a minimum of 15 days. This is not the issue in chronic osteomyelitis, in which the dilemma is to establish whether radiological changes correspond to active infection, surgical sequelae or just trauma.

The detection of a sequestrum (a clearly defined, isolated necrotic area of bone surrounded by a osteopenic zone) or an involucrum (a hyperdense zone of bone under an elevated periosteum) is considered pathognomonic of osteomyelitis. Other signs that should suggest the presence of active infection are the detection of poorly delineated osteolytic areas, periosteal hyperplasia, or irregular periosteal bone extending into adjacent soft tissue (Gómez 1987).

The presence of periprosthetic bone reabsorption or new periosteal bone formation is highly suggestive of infection in the appropriate clinical setting. Considering their simplicity and low price, conventional radiographs should always be obtained if osteoarticular infection is suspected.

Isotope bone scanning

In most cases of chronic osteomyelitis, clinical and radiological data permit an easy diagnosis (Fig. 1, Fig. 2, and Fig. 3). However, sometimes bone changes from other causes and soft tissue infection make laboratory and radiographic signs unreliable as indicators of osteomyelitis. This happens regularly in acute haematogenous osteomyelitis. In this situation, scintigraphic methods can be helpful. Since it provides physiological data, scintigraphy is also useful in the evaluating therapeutic response. However, anatomical definition is inferior to that of CT or MRI. The following is a description of the widely used techniques, although more sophisticated methods are being developed (Fox and Zeiger 1993).

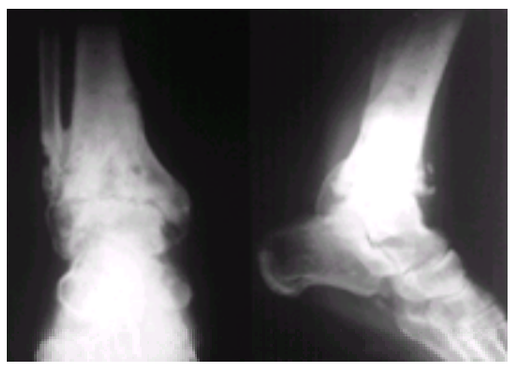


Fig. 1 Plain radiograph: post-traumatic osteomyelitis of the distal tibia.

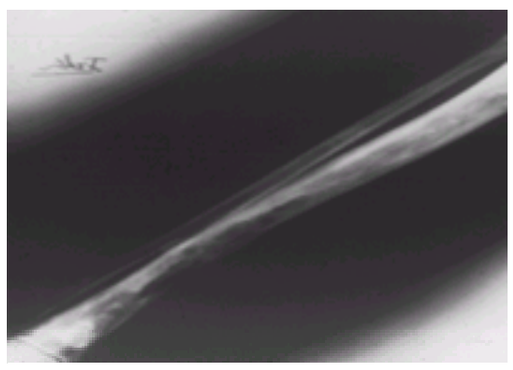


Fig. 2 Haematogenous osteomyelitis of the tibia in an adult with sickle-cell disease.

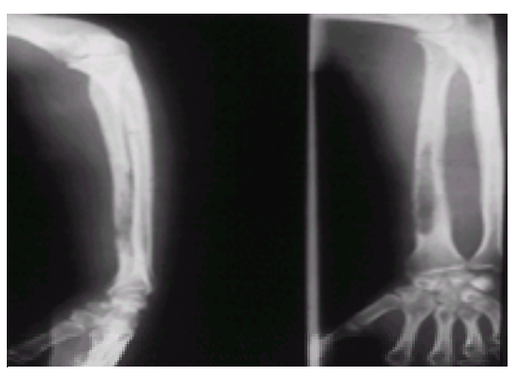


Fig. 3 Haematogenous involvement of bone in a patient with typhoid fever: *Salmonella typhi* osteomyelitis.

Technetium-99m methylene diphosphonate

Technetium-99m methylene diphosphonate (**MDP**) is taken up in areas of increased blood flow or osteoblastic activity. A three-phase bone MDP scan (vascular, pool, and late or bone phase) increases specificity, and is the first-line diagnostic imaging technique after a plain radiograph in evaluating suspected osteomyelitis.

Image quality is fairly good, the dose of irradiation is small, the cost is reasonable, and the technique is available in most centres. Preparation is simple, and the technique provides an earlier and more sensitive diagnosis than plain radiography.

However, specificity is poor and the negative predictive value of this technique is more reliable than positive prediction. False positives may be due to neoplasm,

fractures, heterotopic ossification, arthritis, neuropathic osteopathy, trauma or arthrosis. It is not very useful in children or following recent surgery to bone.

Gallium-67 citrate

Gallium-67 is taken up in areas where leucocytes or bacteria accumulate and provides quantitative information about inflammatory activity. A comparison with the MDP scan is of great value.

Gallium scans become positive earlier than MDP and are sensitive in detecting active bone/joint lesions. A normal gallium scan virtually excludes the presence of an inflammatory process.

False-positive gallium scans in ununited fractures or after recent surgery are common. Images are less precise than MDP scans.

Adult patients with previous bone disorders and possible osteomyelitis or patients with dubious results from MDP scanning should have a gallium scan. Gallium may also be the preferred technique for following response to therapy (Alazraki 1993).

Indium-111 or technetium-99 autologous leucocyte scintigraphy

If the other bone scans are inconclusive, ^{111}In-labelled leucocyte scintigraphy is the next line to take, particularly in adults with other bone disorders.

Patient's leucocytes are obtained and labelled. Afterwards, they are re-infused into the patient and accumulate in the focus of infection where they can be detected.

Indium scans have higher specificity than the previous techniques for the diagnosis of infection, particularly in previously traumatized bone. They provide very good results in osteomyelitis of the diabetic foot, infections associated with delayed or non-union, and prosthetic infections.

Indium scans have higher sensitivity in acute osteomyelitis than in chronic cases (100 per cent vs 60 per cent), probably due to the massive presence of leucocytes in the former (Schauwecker *et al*. 1984). False-positives may occur after trauma, tumours, and other osseous disorders (Nepola *et al*. 1993; Seabold *et al*. 1993). The quality of the image is inferior to that of Tc scintigraphy and it is not very accurate for the axial skeleton. Preparation is complex and long, and the major concerns about this technique are the hazards associated with the handling of blood, and the need to delay imaging for 18 to 24 h, which precludes a rapid result. Potential alternative agents for ^{111}In-labelled leucocytes include labelled immunoglobulin and labelled antigranulocyte antibody reagents.

Indium scans should be reserved for patients with a suspicion of osteomyelitis of the lower extremities not diagnosed with the previous techniques.

None of the types of scintigraphy permits differentiation between septic and non-septic inflammatory processes with sufficient accuracy. For this reason, other techniques are sometimes required.

Computed tomography

This technique accurately detects increased medullary density (typical of the early stages of osteomyelitis), as well as subsequent changes in soft tissues and cortical bone. Its principal advantage is excellent definition of cortical bone, including zones of necrosis, sclerosis, demineralization, periosteal changes, and adjacent soft-tissue swelling. CT does not provide information about the activity of the process and there may be image interference caused by the presence of prosthetic material. Radiation exposure is rather high and it is an expensive technique. CT is especially recommended in chronic osteomyelitis before surgery. It may help to delineate the presence of abscesses, a sinus tract or sequestrum (Maurer *et al*. 1992). It may also be useful for evaluation of infected joint prostheses and osteomyelitis of the spine, pelvis and sternum (Gostishchev *et al*. 1992).

Magnetic resonance tomography (MRI)

Magnetic resonance tomography readily detects the oedema of bone marrow that characterizes the earlier phases of bone infection. Typical features of osteomyelitis on MRI include a low-intensity area in T_1 (less fat) and a hyperintense area in T_2. Bone reaction to fracture or surgery would appear as low marrow intensity in T_1 and normal signal in T_2. Sinus tract and cellulitis would appear as hyperintense areas in T_2.

MRI has proved to be as sensitive as bone scintigraphy in the early detection of osteomyelitis, and, with its superior spatial resolution, it is often more specific than planar scintigraphy in differentiating bone from soft-tissue infection and in separating arthritis, cellulitis, and soft-tissue abscesses from osteomyelitis. In several comparative studies, MRI has been more accurate in detecting the presence and determining the extent of osteomyelitis than scintigraphy, CT scan, and conventional radiography. MRI may facilitate differentiation of acute from chronic osteomyelitis and may help to detect foci of active infection in the presence of chronic inflammation or post-traumatic lesions (Unger *et al*. 1988; Spaeth *et al*. 1991). The patient is not irradiated and newer techniques such as fat-suppressed, contrast-enhanced MRI are significantly more sensitive than scintigraphy, and more specific than non-enhanced MRI or scintigraphy (Morrison *et al*. 1993; Hopkins *et al*. 1995).

MRI also has some drawbacks. It does not provide very precise images of cortical bone and its usefulness decreases in the presence of metallic materials. False-positive results may be obtained in the presence of neoplasm, or intra-/extramedullary inflammation. Experience with MRI is as yet limited and the cost is high. A major indication for MRI is in osteomyelitis of the vertebrae and the foot (Meyers and Wiener 1991), and when diagnosis is still not established after using the previously described techniques.

Microbiological diagnosis

Blood cultures are usually negative in patients with chronic osteomyelitis, and the aetiological diagnosis usually relies on local samples. In acute haematogenous osteomyelitis, blood cultures and/or locally obtained aspirates are the procedures of choice.

It is important to note that only in a small proportion of cases do micro-organisms recovered from sinus tracts reflect the real causative agent present in the bone; *Staph. aureus* is the agent showing the best correlation (Gentry and Rodríguez 1990; Patzakis *et al*. 1994). The overall sensitivity and specificity of sinus-tract cultures for different micro-organisms are summarized in Table 5. Consequently, bone biopsy culture is now the standard method for determining specific antimicrobial therapy (sensitivity and specificity of 87 per cent and 93 per cent, respectively) (Perry *et al*. 1991; Howard *et al*. 1994). Bone biopsy should be taken, using local anaesthesia and imaging control, from the most painful site (Stratton 1989). The specimen must be transported to the laboratory immediately, and stained and cultured for most common pathogens, including aerobes, anaerobes, fungi, and mycobacteria. Quantitative bone cultures have not been effective in differentiating osteomyelitis from infection or colonization of adjacent soft tissue (Darouiche *et al*. 1994).

	Sensitivity (%)	Specificity (%)
Bone biopsy	100	100
Sinus tract:	76	86
Staph. aureus	70	90
Other Gram-positive	65	69
Pseudomonas aeruginosa	82	75
Other Gram-negative	88	96

Table 5 Correlation between cultures from the sinus tract and bone biopsy (Gentry and Rodríguez 1990)

The nature of the infection of a joint prosthesis may be particularly difficult to establish. In these cases, joint aspiration with a fine needle is recommended (sensitivity and specificity of 87 per cent and 95, respectively) (Roberts *et al*. 1992).

Finally, mention must be made of the serum bactericidal test, which is considered to be a predictor of therapeutic efficacy in acute and chronic osteomyelitis. This test is based on determining the ability of progressive dilutions of patient's serum to kill the infecting micro-organism. A multicentre study has shown that, in chronic osteomyelitis, peak and trough levels above 1:16 and 1:4, respectively, are desirable (Weinstein *et al*. 1987).

Histological diagnosis

Histological diagnosis of osteomyelitis is useful in all situations but particularly necessary in cases where there is reasonable doubt about the reliability of cultures. It is required in patients already receiving antimicrobial drugs, in osteomyelitis potentially caused by pathogens that are difficult to grow, and in patients where bone samples have to be taken from infected or colonized soft-tissue structures, which may lead to a false-positive result (McGuire 1989).

Bone histopathology also provides more precise information on the anatomical limits of the infected tissue and reassurance that surgical resection has reached healthy, viable bone.

Treatment

The issue of treatment for osteomyelitis has not been completely resolved. Recent research has provided additional insights into the pathogenesis of bone infection. Advances in pharmacology and in surgical techniques have improved our ability to treat such infections. Despite these advances, successful treatment of post-traumatic chronic tibial osteomyelitis depends on adherence to several basic principles: complete debridement of necrotic and infected tissue, obtaining bone stability, the elimination of dead space, and the provision of durable soft-tissue coverage (Dirschl and Almekinders 1993; Meadows *et al*. 1993). Therapy for osteomyelitis requires a multidisciplinary approach, which ideally should include the collaboration of surgeons, infectious disease physicians, rheumatologists, plastic surgeons, radiologists, and many other specialists. The adequate debridement of necrotic tissue is frequently necessary, and the combined orthopaedic and plastic surgical approach has permitted successful salvage of otherwise severely injured lower limbs.

Acute haematogenous osteomyelitis usually responds to antimicrobial therapy. The presence of an abscess, a metaphyseal cavity in haematogenous osteomyelitis, and evidence of spinal-cord compression in vertebral osteomyelitis require surgical treatment.

Chronic osteomyelitis usually implies that dead bone is present, which requires surgical debridement. Because of the chronic nature of the infection and the various presentations and surgical approaches, antibiotic treatment must be individualized. In general, however, at least 4 weeks of therapy is required (Bamberger 1993).

Systemic antimicrobials

The selection of antimicrobial drugs depends on their *in vitro* activity against the responsible pathogens, their penetration into bone tissue, their pharmacological characteristics, toxicity and cost. Table 6 summarizes some of the characteristics of an 'ideal' antimicrobial agent for the treatment of osteomyelitis. The following is a description of some of the antimicrobial agents most frequently used for the therapy of bone and joint infections.

1. Wide spectrum of antimicrobial activity
2. Bactericidal capacity
3. Good oral absorption
4. Long half-life
5. Good penetration in normal and necrotic bone
6. Good activity at low pH and in anaerobic conditions
7. Low toxicity
8. High resistance to bacterial inactivation mechanisms
9. Good clinical efficacy in animal models and humans
10. Low cost

Table 6 Ideal characteristics of antimicrobial agents for the treatment of osteomyelitis

Quinolones

Quinolones are well suited to the treatment of osteomyelitis, since most of them, including ciprofloxacin, ofloxacin and pefloxacin, reach satisfactory levels in bone tissue and their broad spectrum of activity covers most potential pathogens (Overbeck *et al*. 1995; Suh and Lorber 1995). Activity after oral administration is similar to that with parenteral administration. Combined with rifampin they now constitute a frequently used regimen for the treatment of osteomyelitis (Neu 1993). Several trials have shown that the new oral fluoroquinolones are as effective as parenteral cephalosporins and other broad-spectrum agents in treating osteoarticular infections. Tolerance is excellent and permits prolonged courses in ambulatory patients. Ofloxacin has more predictable absorption, although its activity against *Pseudomonas* is less than that of ciprofloxacin. None of them must be used to treat infections caused by MRSA. The widespread overuse of quinolones (1/44 Americans have received ciprofloxacin) has recently raised an alarm (Guay 1992; Greenfield 1993). However, in our opinion, osteoarticular infections can still be considered one of the most clear indications for these drugs.

Co-trimoxazole

Co-trimoxazole is usually active against *Staph. aureus* (including many methicillin-resistant strains) and against some Gram-negative organisms. The authors have treated 17 patients with complicated chronic osteomyelitis, achieving a satisfactory response in 82 per cent after a follow-up of more than two years (Bouza *et al*. 1992). Co-trimoxazole is, on many occasions, the only oral alternative for the therapy of bone and joint infections caused by MRSA.

Vancomycin

Vancomycin is almost uniformly effective against Gram-positive micro-organisms, but has the drawbacks of toxicity and the need for intravenous administration (Ish-Horowicz *et al*. 1992). It may be administered to patients who require parenteral therapy and in osteomyelitis caused by multiresistant Gram-positive micro-organisms (MRSA, *Staph. epidermidis*, etc.) (Refsahl and Andersen 1991). Bone penetration is poor (14 per cent of the serum concentration) and so simultaneous administration of rifampin is recommended.

Teicoplanin

Teicoplanin is a glycopeptide antibiotic with an antimicrobial spectrum of activity similar to that of vancomycin (Wilson and Gruneberg 1994). Nevertheless, it has a longer half-life that permits once-a-day dosing, which is particularly convenient for patients attending day-care hospitals or having antibacterial therapy at home. It can be administered both by intravenous and intramuscular routes, and it is better tolerated than vancomycin. Its efficacy in Gram-positive osteoarticular infections has recently been assessed. Microbiological eradication was achieved in 86 per cent of patients and clinical cure in more than 80 per cent; toxicity precluded termination of therapeutic courses in 12/60 patients. (LeFrock *et al*. 1992; Weinberg 1993).

Rifampin

Rifampin is a bactericidal antimicrobial agent with excellent activity against many Gram-positive micro-organisms. It is considered the most powerful antistaphylococcal agent. It achieves very good bone tissue concentrations (O'Reilly *et al*. 1992), and we feel that, whenever possible, it should be included in the therapeutic regimen of osteoarticular infections. However, it must never be used alone, since resistance develops rapidly if used as a single agent. Patients musts be warned that rifampin may turn body fluids a reddish colour.

Clindamycin

Clindamycin shows good activity against Gram-positive micro-organisms and some anaerobes, and its penetration into bone tissue is satisfactory (98 per cent of the serum concentration). A well-known side-effect of this antibiotic is *Clostridium difficile*-related colitis, especially in elderly patients.

b-Lactam drugs

Among b-lactam drugs, penicillins continue to have an important role. Penicillins are the drugs of choice in osteomyelitis caused by streptococci and staphylococci that are penicillin sensitive. Nevertheless, one should remember that more than 90 per cent of the staphylococci isolated nowadays are resistant to penicillin.

Amoxycillin is the drug of choice in the rare patients with enterococcal osteomyelitis and also in cases due to *Haemophilus* and *Salmonella*. The combination of amoxycillin and clavulanic acid has considerably increased the spectrum of antimicrobial activity of amoxycillin; it can be used in osteomyelitis caused by different *Enterobacteriaceae* and also in cases with anaerobic bacteria present in either mono- or polymicrobial infections.

Piperacillin and piperacillin with tazobactam are suitable agents for the treatment of osteomyelitis caused by *Pseudomonas*, anaerobic bacteria, and mixed infections.

The isoxazolil penicillins and other penicillins resistant to penicillinases are still the preferred drugs (in competition with first-generation cephalosporins) for treatment of most staphylococcal bone and joint infections.

Cephalosporins of all generations have been used in the treatment of osteomyelitis but those of the first generation remain among the drugs of choice for staphylococcal osteomyelitis. Second- and, more particularly, third- and fourth-generation cephalosporins are useful as substitutes for aminoglycosides in the lengthy treatment of Gram-negative osteomyelitis.

Finally, aztreonam is a useful drug for the treatment of Gram-negative osteomyelitis. Imipenem has a very broad spectrum of in vitro activity, including Gram-positive bacteria, *Enterobacteriaceae, Ps. aeruginosa* and anaerobes, but should be preserved for cases of osteomyelitis that are particularly difficult to treat or whose aetiology is polymicrobial (Jauregui *et al*. 1993).

Parenteral versus oral antimicrobial agents

Traditionally, parenteral agents were the drugs chosen for the treatment of osteomyelitis. Most still believe that at least the initial treatment of osteomyelitis should be with intravenous antibiotics. Others (us included), however, consider that in clinically stable patients with aetiologically well-documented osteomyelitis, some oral drugs are adequate from the very beginning. For maintenance therapy, intravenous antibiotics have been largely replaced by new, orally active, wide-spectrum agents that permit prolonged treatment in the ambulatory patient (Conrad and Marks 1989; Craig and Andes 1995). The drugs most frequently used for staphylococcal infections are combinations of rifampin with either oral fluoroquinolones or co-trimoxazole. Most Gram-negative infections in adults can be treated with an oral quinolone but for others there are no adequate oral drugs (Tice 1993).

Length of therapy

The use of new, orally active antimicrobials such as quinolones has modified the recommended length of therapy in chronic osteomyelitis. The 6-week benchmark, which was determined largely by experience with childhood haematogenous osteomyelitis, may not be applicable to contiguous-focus osteomyelitis after trauma in adults. At the present time, it is known that 6 weeks may not be sufficient for contiguous chronic osteomyelitis and the availability of oral antibiotics with a low toxicity profile permits a much longer duration of treatment and a higher rate of success. Unfortunately, to our knowledge, no well-designed prospective clinical trials have determined the precise duration of antimicrobial therapy for chronic osteomyelitis. We recommend a minimum of 3 months. This may be extended in certain circumstances, such as poor initial response, unsatisfactory coverage of soft tissues, presence of non-unions, and delayed withdrawal of prosthetic material.

Occasionally, patients not considered to be candidates for surgery require prolonged suppression rather than a curative approach.

Local delivery of antimicrobials

Several devices designed to deliver antimicrobials locally (cement, biodegradable microcapsules) have been developed, although the most widely used are gentamicin-impregnated polymethylmethacrylate beads, which are claimed to allow immediate filling of the bone defect and to provide high antibiotic concentrations. In some centres their use has become standard orthopaedic practice. Chains of beads are packed into the infected bone cavity after debridement, with the end projecting from the scar. During the following 4 to 6 weeks, the chain is progressively pulled out. The procedure is painful and sometimes adhesions prevent complete extraction of the beads.

A recent comparative study, involving 384 patients, suggests that local therapy with gentamicin-impregnated beads is as effective as standard systemic therapy for osteomyelitis. The conclusions may be biased by the fact that most patients received combined surgical and medical treatment (local and systemic antimicrobials). However, the data do suggest that both cost of treatment and toxicity are considerably lower in patients who are treated with local antibiotics alone (Blaha *et al*. 1993). Good results are reported in the treatment of infected non-union (Calhoun *et al*. 1993) and contaminated compound fractures (Ostermann *et al*. 1993). The main drawbacks are the potential danger of leaving foreign material in a septic space with connections to the exterior, and that gentamicin, the most commonly used antimicrobial in these devices, does not seem to be the most potent agent, considering that *Staph. aureus* is the major pathogen. New devices impregnated with more effective antimicrobials such as vancomycin, teicoplanin, or ciprofloxacin are being studied (Dacquet *et al*. 1992; Gerhart *et al*. 1993; DiMaio *et al*. 1994).

Surgical therapy

Antibiotic treatment is not a substitute for surgical debridement of infected, devitalized bone. Absence or inadequacy of surgical treatment is clearly associated with higher rates of initial failures or recurrences. We recommend early surgery whenever possible, before the administration of antimicrobial agents. The purpose of the surgical intervention is not only therapeutic but also may serve to establish the precise aetiological diagnosis.

On some occasions, the need for bone consolidation and stability precludes early surgical therapy and surgery has to be postponed until these occur. In this case, the administration of antimicrobials is 'palliative' until the definitive, and potentially curative, surgical procedure is possible.

The different surgical techniques for bone debridement will not be reviewed in detail in this chapter. However, the principle is that only viable bone should be left in place whenever possible.

Bone viability is of fundamental importance in the surgical management of osteomyelitis. The use of laser Doppler flowmetry as an adjunct to surgery, allowing quantitative evaluation of bone vascularity, has recently proved to be of great value (Duwelius and Schmidt 1992). When axial long bones are involved, aggressive debridement is not easy to perform since adequate wound coverage and mechanical stability may be difficult to achieve or maintain. Treatment of infected non-unions is especially difficult. Prognosis has become more optimistic with the development of new orthopaedic methods, such as the Ilizarov limb reconstruction method. This has proved to be cost-effective when compared with amputation (Cattaneo *et al*. 1992; Williams 1994). Occasionally, osteotomies, bone grafts and transports, or muscular flaps are also necessary (Mahan and Jolles 1995; Yajima *et al*. 1995).

Another situation that requires rapid functional reconstruction is osteomyelitis of the sternum associated with mediastinitis. For this, a wide debridement should be followed (at the same time or soon thereafter) by reconstruction of the chest wall by vascularized pectoral-muscle flaps (Banic *et al*. 1995).

A conservative, non-surgical approach is recommended for asymptomatic or mildly symptomatic osteomyelitis in elderly patients.

Finally, it is important to remember that amputation is sometimes the most functional therapeutic alternative in patients with refractory osteomyelitis of the feet or lower limbs (Lerner *et al*. 1993). We feel that early amputation in chronic and relapsing osteomyelitis of the foot, particularly when associated with poor vascular supply or neuropathic disease, is better for the patient than prolonged disability and antimicrobial therapy condemned to failure.

Administration of hyperbaric oxygen

Hyperbaric oxygen has been used successfully in the treatment of air embolism, in radio-osteonecrosis, in intoxication by carbon monoxide, in clostridial myonecrosis, in severely burned patients, and in other soft-tissue infections. Open, non-randomized studies suggest a potential value of hyperbaric oxygen in the treatment of refractory chronic osteomyelitis (Slack *et al*. 1965; Perrins *et al*. 1966; Depenbusch *et al*. 1972; Davis *et al*. 1986; Mader *et al*. 1990).

In an animal experimental model of chronic *Staph. aureus* osteomyelitis, hyperbaric oxygen was compared to treatment with parenteral cephalothin alone and with a combination of both for 4 weeks. Hyperbaric oxygen was as effective as parenteral cephalothin. At the end of treatment, the original *Staph. aureus* could be recovered from 91 per cent of control animals, 36 per cent of those treated with hyperbaric oxygen only, 47 per cent of those treated with cephalothin alone, and from 40 per cent of those treated with the combination.

Hyperbaric oxygen increases the bactericidal activity of phagocytic cells and the bactericidal effect of drugs as vancomycin and aminoglycosides. Collagen production by fibroblasts occurs more efficiently with oxygen tensions greater than 10 mmHg.

The role of hyperbaric oxygen in the treatment of human osteomyelitis requires definition by well-designed, prospective clinical studies.

Prognosis

Osteomyelitis-associated mortality is usually low, but the tendency for osteomyelitis to recur and become chronic is well known. Patients may experience unpredictable episodes of fistulous discharge. Recurrences are most common within a year of the initial episode, but extremely long intervals have been reported.

Malignant transformation arising from a fistula in patients with chronic osteomyelitis is extremely rare, although isolated cases have been described. This emphasizes the importance of histological examination (Mabit *et al*. 1993; Noonan *et al*. 1993).

Patients and doctors should realize that it is impossible to guarantee whether osteomyelitis can ever be 'cured,' since infections become manifest many years after injury or treatment.

Prophylaxis

Antimicrobial prophylaxis is recommended after open fractures, in which the infection rate is high.

For the implantation of articular prostheses both antimicrobial prophylaxis and environmental air cleaning are recommended. For hip, knee, and other major joint prostheses, the recommended antimicrobial prophylaxis is usually one to three doses of cefazolin starting at the induction of anaesthesia and over the next 24 to 48 h. In addition, antimicrobial prophylaxis is specially important, and widely used, when internal fixation of fractures or joint replacements are performed.

As most osteoarticular infections are proved to be acquired in the operating theatre, highly sophisticated, centrifugal laminar-flow systems have been installed, achieving a 10^4-fold reduction in the concentration of environmental micro-organisms and 10-fold reduction in the incidence of infection. At the present time, the incidence of infection has been reduced to 0.5 per cent in hip prostheses (Lidwell 1986; Schutzer and Harris 1988), while knee and shoulder prostheses are associated with a higher incidence of septic complications (1–4 per cent and 4–7 per cent, respectively) (Bengtsson 1993; Malchau *et al*. 1993).

Some orthopaedic surgeons recommend prophylactic antimicrobial agents for prevention of infection in patients with prosthetic joints undergoing certain invasive procedures. Particular attention should be paid to bacteriuria, the implantation of intravascular devices, drainage, and dental manipulations, for which a prophylactic approach similar to that for endocarditis is recommended.

Chronic post-traumatic osteomyelitis

Most information to date comes from experience with chronic post-traumatic osteomyelitis. The risk of infection after an open fracture varies widely with the site, size, and nature of an open fracture, as shown in Table 7, adapted from Gustilo *et al*. (1990).

Type	Wound	Fracture	Infection
Type I	< 1 cm Clean puncture Soft tissue damage +	Simple, transverse Comminution +	0–2%
Type II	> 1 cm Moderate contamination Soft tissue damage ++	Moderate contamination Comminution ++	2–7%
Type III:	Soft tissue damage +++	Instability Comminution +++	
Type III A	Soft tissue coverage: adequate	Contamination ++	7%
Type III B	Loss of soft tissue Exposure of bone or periosteal stripping	Contamination +++	10–50%
Type III C	Arterial injury		25–50%

Table 7 Risk of infection in open fractures (adapted from Gustilo *et al*. 1990)

Chronic osteomyelitis is usually a local disease that only rarely produces systemic manifestations. The most common symptoms are pain and suppurative discharge. As mentioned previously, the infection is characteristically recurrent and resistant to short courses of therapy. Fever is more common in acute infections or in the presence of soft-tissue abscesses.

The presence of an open draining fistula is always a marker of clinical activity of osteomyelitis.

The main principle of management of infection in non-consolidated bone fractures is to retain fracture fixation whenever possible. When infection becomes apparent, ultrasonography may determine the presence or absence of collections requiring drainage. In the absence of a collection or a wound discharge, only antimicrobial treatment should be provided. If discharge or collections are present, surgical drainage and debridement are necessary, the organisms present must be identified, and antimicrobial treatment prescribed.

Implants should be retained unless fracture has occurred, in which case repeated debridement with removal of all devitalized bone and soft tissue is required. After

this, either exchange implants or external fixation are undertaken.

Osteomyelitis in decubitus ulcers

The incidence of osteomyelitis under decubitus ulcers ranges from 28 to 70 per cent (Bohm *et al*. 1988; Bruck *et al*. 1991). Most patients with osteomyelitis have radiological changes but to confirm its activity and aetiology requires needle or surgical bone sampling. Antibiotics are indicated only if osteomyelitis is present. After surgical debridement, soft-tissue coverage is essential. Myocutaneous flaps are superior to skin flaps in securing skin cover.

Acute haematogenous osteomyelitis in children

Haematogenous osteomyelitis may appear as an acute onset of bone pain or limited motion of an extremity, regardless of the presence or absence of signs of infection such as fever, local tenderness, redness, swelling or heat.

Acute haematogenous osteomyelitis in children is becoming infrequent. Nelson (1990) reports that this diagnosis is made only approx. 15 times per year in a large referral unit for paediatric infectious diseases in the United States. Approximately one-third of the cases occur in children under 2 years of age, another one-third between 2 and 5 years, and the remaining one-third in children more than 5 years old (Nelson 1990).

In most children with acute haematogenous osteomyelitis no predisposing factors are apparent but it is well known that sickle-cell anaemia is associated with a higher risk of haematogenous osteomyelitis, particularly due to *Salmonella* (Adeyokunnu and Hendrickse 1980; Syrogiannopoulos *et al*. 1986). and *Staphylococcus* (Sadat-Ali *et al*. 1985).

Pain, pseudoparalysis (voluntary limitation of movement of one extremity), and fever are the most common clinical manifestations. Initially, local inflammatory signs are rarely present. As mentioned earlier, plain radiographs are frequently negative and isotope bone scanning is required (Herndon *et al*. 1985; Demopulos *et al*. 1988). The disease is almost always monostotic and long-bone metaphyses are the sites most frequently involved. Diagnostic confirmation requires needle aspiration or surgical debridement. Blood cultures are, in contrast to chronic osteomyelitis, frequently positive (in at least one-third of the episodes). Both blood and local samples should be always be obtained because, not uncommonly, only one of the samples is positive. *Staphylococcus aureus* is responsible for at least 50 per cent of the cases. The remaining cases are caused by different micro-organisms, among them *Streptococcus* and *Haemophilus* (Nelson 1990).

When pain is severe, needle or surgical decompression should be undertaken. If the infection opens into the synovial cavity, repeated needle aspiration may be sufficient.

When no pain or large collection of pus is present, antimicrobial treatment alone may be enough. Sequential treatment (intravenous drugs followed by oral antimicrobial agents) has proved adequate (Nelson 1982). A total of 10 to 14 days of intravenous antibiotics is followed by 2 to 6 additional weeks of oral agents. These should have an acceptable flavour, good oral absorption, low toxicity profile and be active against Gram-positives and *Haemophilus*. The indication for measuring serum bactericidal concentrations during follow-up still needs to be defined but trough or peak levels greater than 1/8 or 1/32 have been recommended (Prober and Yeager 1979; Nelson 1990).

Nowadays the prognosis of acute haematogenous osteomyelitis in childhood is much better than in the past. Death from this cause is almost non-existent and evolution to chronicity occurs in less than 5 per cent of cases (Dunkle and Brock 1982; Nelson 1990).

Vertebral osteomyelitis

The spinal column is the most common site of haematogenous osteomyelitis in adults, probably due to vertebrae having abundant red marrow and a slow and tortuous blood flow. Bacteria reach the vertebral bodies preferentially by the arteries but occasionally also through the venous plexus (Wiley and Trueta 1959; Waldvogel *et al*. 1970; Adatepe *et al*. 1986; Sapico and Montgomerie 1990).

More than 50 per cent of the episodes of vertebral osteomyelitis are due to *Staph. aureus*, approx. 30 per cent to *Enterobacteriaceae*, and the remaining 10 to 20 per cent to other micro-organisms such as *Brucella* spp, *M. tuberculosis*, and fungi (Sapico and Montgomerie 1979; Sapico and Montgomerie 1980; Sapico and Montgomerie 1990).

Vertebral osteomyelitis is generally monomicrobial but the association of *M. tuberculosis* with other pathogens, particularly *Staph. aureus*, is more than casual.

Pain that increases with spinal movement is the most common clinical manifestation of vertebral osteomyelitis. Other symptoms depend on the site of involvement. In patients with cervical osteomyelitis and prevertebral abscesses, odinophagia and dysphagia may be present.

Fever and/or an increased white blood-cell count is absent in up to half of the cases. An increased erythrocyte sedimentation rate has traditionally been considered to be a very sensitive marker for the presence of active vertebral infection. In our experience this is not the case and the diagnosis of vertebral osteomyelitis should not be rejected in patients with a normal erythrocyte sedimentation rate.

Overall, the lumbar spine is the most frequent location and is involved in more than 50 per cent of episodes of vertebral osteomyelitis, followed by 35 per cent in the dorsal spine, and the remaining 15 per cent in the cervical spine.

Lumbar osteomyelitis occurs occasionally as a septic metastasis of infections of the pelvis and genitourinary tract (Sapico and Montgomerie 1979); cervical bone infection is frequently a complication of parenteral drug abuse and is seen in patients with other predisposing conditions (Sapico and Montgomerie 1986).

Most cases of vertebral osteomyelitis have abnormalities on plain radiographs when first seen but in very acute presentations, pain may be present but the plain film is normal. Erosive, irregular images in the vertebral bodies and the adjacent intervertebral disc are common. Pre- and paravertebral abscesses and vertebral collapse are common complications (Fig. 4 and Fig. 5).

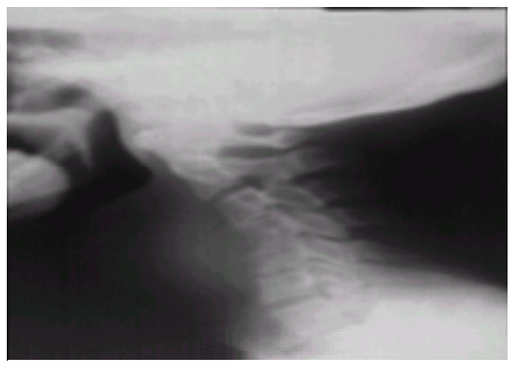


Fig. 4 Cervical vertebral osteomyelitis and prevertebral pyogenic abscess.

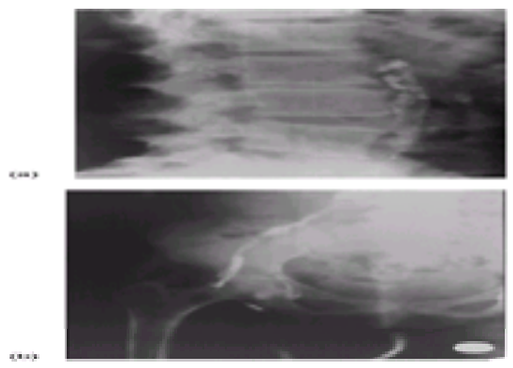


Fig. 5 Tuberculous vertebral osteomyelitis with a fistulous tract draining to the upper thigh.

Isotope imaging is diagnostically very sensitive in patients with normal or equivocal plain films. Both CT and MRI offer early and well-defined images that are very useful for directing aspiration or surgery (Boddicker *et al*. 1980; Modic *et al*. 1985; Sapico and Montgomerie 1990).

Microbiological confirmation requires either positive blood cultures, vertebral material or both. Only one-quarter of the episodes are documented with blood cultures and consequently bone aspiration should be undertaken whenever possible.

Most cases of vertebral osteomyelitis do not require surgical debridement. Surgery should be reserved for patients with spinal instability, neurological impairment, a large, progressive abscess impossible to drain by guided needle aspiration, or cases of unknown aetiology not responding rapidly to antimicrobial therapy.

Considerations governing antimicrobial therapy described previously also apply to vertebral involvement. Immobilization in bed is only required for patients with pain or vertebral instability and casts are not necessary in most cases.

Infected prosthesis (see Chapter 5.3.1)

Infections related to joint prostheses have been divided into those presenting in the first 3 months after surgery (type I), those presenting between 3 months and 1 year postoperatively (type II), and those occurring after 1 year postsurgery (type III) (Coventry 1975). Type I infections are almost exclusively acquired during surgery and may be divided into superficial (involving soft tissues but not the prosthesis), and deep (involving the prosthesis). In group II, most infections are surgically acquired and practically all of them involve the prosthesis. Finally, group III infections are of haematogenous origin and involve the prosthesis. At least 70 per cent of prosthetic joint infections are monomicrobial and the remaining 20 to 30 per cent are either polymicrobial or not documented (Buchholz *et al*. 1981; Buchholz *et al*. 1984). Among the responsible micro-organisms, *Staph. epidermidis* and *Staph. aureus* are far ahead of other bacteria, followed by Gram-negative rods and anaerobes.

Chronic infection of articular prostheses usually begins as continuous pain, frequently without obvious drainage through a sinus tract, prosthetic dysfunction or complete loosening (Fig. 6). Infections associated with prosthetic material are extremely difficult to treat and often require surgical removal of foreign bodies, which may be very radical when artificial joints are involved. Attempts to implant a new prosthesis are usually made after prolonged courses of antimicrobial therapy, leaving the patient with major incapacity for a long time. Recently, Drancourt *et al*. (1993) reported an acceptable rate of success (74 per cent) on treating osteoarticular prostheses infected with staphylococci with prolonged periods (9–12 months) of ofloxacin and replacing the prosthesis, when necessary, in a one-step procedure.

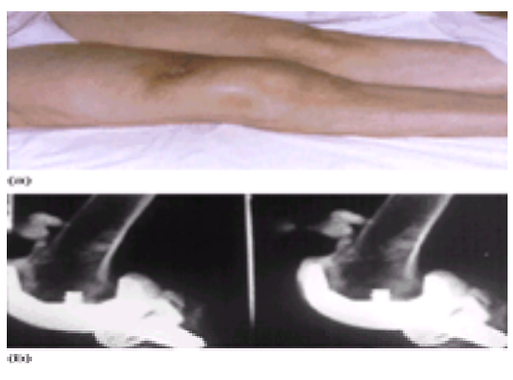


Fig. 6 Infected knee prosthesis: (a) fistulous tract; (b) fistulography.

Type I infections with superficial involvement and no evidence of prosthetic dysfunction should be treated with antimicrobial agents, removal of blood or collections, soft-tissue debridement, and irrigation. A high proportion of these cases (approx. 70 per cent) do not require future prosthetic replacement (Buchholz *et al*. 1984; Fitzgerald 1984; Fitzgerald and Jones 1985). Even in cases with deep infections, antimicrobial therapy and early debridement and irrigation save at least 50 per cent of the prostheses (Drancourt *et al*. 1993).

In patients with type II and III infections the final objective is to preserve a functional, pain-free prosthesis. In patients with no prosthetic dysfunction and with no or minimal pain, a trial of antimicrobial therapy without surgical replacement of the prosthesis can be attempted. We use, whenever possible, oral antimicrobial agents, usually including rifampin in a combination regimen. Treatment is continued for a 6- to 12-month period, provided an adequate response was rapidly obtained. In case of a late relapse, the treatment is individualized depending on many variables such as the age of the patient, functional status, and the risk for surgical replacement.

In patients with pain, prosthetic dysfunction or loosening, surgical replacement is necessary. It may be performed in two stages or in a single step, with a low rate of reinfection (Buchholz *et al*. 1984).

Infectious arthritis

Infectious arthritis must be considered as an emergency, since delays or inadequate treatment may result in great disability. The recognition, diagnosis, and treatment of this condition are therefore very important (Smith 1990).

Infectious arthritis should be suspected in any patient with a swollen joint but especially in children, debilitated patients, immunocompromised persons, those with infection elsewhere (even if on antibiotics), and those with other types of arthritis or a prosthetic joint (Fig. 7). Diagnosis depends on obtaining joint fluid for culture and Gram staining. Initial treatment with appropriate broad-spectrum antibiotics is later best narrowed to suit the individual organism. Treatment requires repeated needle aspiration or surgical drainage plus an antimicrobial course of 2 to 6 weeks (Middleton 1993).

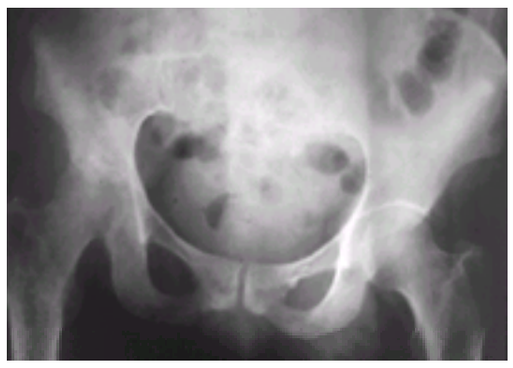


Fig. 7 Right hip infection caused by *Staph. aureus* in an intravenous drug abuser.

This subject is described in more detail in Chapter 5.3.1. Some clinical and epidemiological features related to aetiological agents of infectious arthritis are shown in Table 8.

Risk factor	Micro-organisms
Infants	*Staph. aureus* Enterobacteria Group B streptococci
Children	*Staph. aureus* *H. influenzae* *Streptococcus* spp. Enterobacteria
Adults	*Staph. aureus* Group A streptococci Enterobacteriaceae *N. gonorrhoeae*
Rheumatoid arthritis	*Staph. aureus* *Streptococcus* spp. Enterobacteria
Prosthetic joint, postoperative, intra-articular injection	*Staph. epidermidis* *Staph. aureus* Enterobacteriaceae *Pseudomonas* spp.
Intravenous drug abuse	*Staph. aureus* or *Ps. aeruginosa*
Human bite	*Eikenella corrodens*, oral flora
Cat or dog bite	*Pasteurella multocida, anaerobes*
Rat bite	*Streptobacillus moniliformis*
Risk factor	**Micro-organisms**
Tick exposure	*Borrelia burgdorferi*
Unpasteurized dairy products	*Brucella* spp.
Trauma in aquatic environment	*Mycobacterium marinum*
Quarrelling injury	*Eikenella corrodens*
Abrasions or amputations	*N. gonorrhoeae*

Table 8 Epidemiological and clinical characteristics of aetiological agents of infectious arthritis

Most cases are caused by *Staph. aureus*, although many bacterial and viral pathogens may be involved. Among streptococci, group A b-haemolytic predominates, although group b have been described, mainly in immunosuppressed patients (Barberan *et al*. 1993). *Neisseria gonorrhoeae* is the predominant cause of bacterial arthritis in sexually active adults under 30 years of age with no other risk factor. Transient, disseminated, papular skin lesions precede the appearance of polyarticular symptoms and blood cultures are usually positive. In the rare event of monoarticular infection the micro-organism is only isolated from synovial fluid.

Gram-negative rods usually cause infectious arthritis in patients with underlying conditions, and *Pseudomonas* spp. have been associated sternoclavicular and sacroiliac joint infections in intravenous drug abusers.

Lyme arthritis should be suspected in patients with a history of tick exposure and erythema chronicum migrans. The diagnosis is usually serological.

Chronic monoarticular arthritis is usually caused by mycobacteria or fungus. Infectious arthritis caused by *M. tuberculosis* is infrequent, usually affecting the interphalangeal joints of the hand. *Mycobacterium kansasii* and *M. marinum* arthritis have also been described.

A number of fungi can cause infectious arthritis. *Sporothrix shenckii* infections are usually related to contact with soil and most commonly affect the knee. *Coccidioides immitis* arthritis can occur in endemic areas, in both immunocompetent and immunocompromised patients. Other fungi such as *Candida* (haematogenous spread), *Cryptococcus*, and *Aspergillus* are even rarer agents of infectious arthritis (Christensson *et al*. 1993; Cuende *et al*. 1993).

Many viral infections can produce joint symptoms. Among them, rubella (small joints of the hand), erythema infectiosum (parvovirus B19), and hepatitis B are the most common. Mycoplasmas and ureaplasmas are responsible for about two-fifths of cases of septic arthritis in patients with hypogammaglobulinaemia (Furr *et al*. 1994). Management is dealt with in other chapters.

Prognosis

Bad prognostic factors include: age exceeding 60 years, pre-existing arthritis, infection in hip and shoulder, symptoms for more than 7 days before treatment, involvement of more than four joints, and persistently positive cultures after 7 days of appropriate therapy.

Limitation of motion and persistence of pain are the most common sequel; although a mortality rate between 30 and 100 per cent has been described for polyarticular cases (Bould *et al*. 1993; Pitkin and Eykyn 1993).

Chapter References

Adatepe, M. H. *et al*. (1986). Hematogenous pyogenic vertebral osteomyelitis. Diagnostic value of radionuclide bone imaging. *Journal of Nuclear Medicine*, 27, 1680–5.

Adeyokunnu, A. A. and Hendrickse, R. G. (1980). *Salmonella* osteomyelitis in childhood. *Archives of Disease in Childhood*, **55**, 175.

Alazraki, N. P. (1993). Radionuclide imaging in the evaluation of infections and inflammatory disease. *Radiologic Clinics of North America*, **31**, 783–94.

Ammari, L. K., Offit, P. A., and Campbell, A. B. (1992). Unusual presentation of group B *Streptococcus* osteomyelitis. *Pediatric Infectious Diseases Journal*, **11**,1066–7.

Anand, A. J. and Glatt, A. E. (1994). *Salmonella* osteomyelitis and arthritis in sickle cell disease. *Seminars in Arthritis and Rheumatism*, **24**, 211–21.

Assaad, W., Nuchikat, P. S., Cohen, L., Esguerra, J. V., and Whittier, F. C. (1994). *Aspergillus* discitis with acute disc abscess. *Spine*, **19**, 2226–9.

Bamberger, D. M. (1993). Osteomyelitis. A commonsense approach to antibiotic and surgical treatment. *Postgraduate Medicine*, **94**,177–84.

Banic, A., Ris, H. B., Erni, D., and Striffeler, H. (1995). Free latissimus dorsi flap for chest wall repair after complete resection of infected sternum. *Annals of Thoracic Surgery*, **60**, 1028–32.

Barberan, J., Rodriguez, R., Gomis, M., Quilez, C., Martin-Gamero, C. P., and Atero, F. (1993). *S. agalactiae*-induced septic arthritis in a patient with vulvar carcinoma. (Letter). *Annales Medicina Interna*, **10**, 468–9.

Bengtsson, S. (1993). Prosthetic osteomyelitis with special reference to the knee: risks, treatment and costs. *Annals of Medicine*, **25**, 523–9.

Berman, S. and Jensen, J. (1990). Cytomegalovirus-induced osteomyelitis in a patient with the acquired immunodeficiency syndrome. *Southern Medical Journal*, **83**, 1231–2.

Bernier, S., Clermont, S., Maranda, G., Turcotte, J. Y. (1995). Osteomyelitis of the jaws. *Journal of the Canadian Dental Association*, **61**, 441–2: 445–8.

Blaha, J. D. *et al*. (1993). Comparison of the clinical efficacy and tolerance of gentamicin PMMA beads on surgical wire versus combined and systemic therapy for osteomyelitis. *Clinical Orthopedics*,

295, 8–12.

Boddicker, J. H. *et al*. (1980). Bone and gallium scanning in the evaluation of disseminated coccidioidomycosis. *American Review of Respiratory Diseases*, **122**, 279–87.

Bohm, E., Kuhlmann, I., and Botel, U. (1988). Decubitus ulcer in paraplegic patients: a comparative clinico-pathological study. *Unfallchirurgie*, **14**, 335–42.

Bould, M., Edwards, D., and Villar, R. N. (1993). Arthroscopic diagnosis and treatment of septic arthritis of the hip joint. *Arthroscopy*, **9**, 707–8.

Bouza, E. *et al*. (1992). Eficacia del cotrimoxazol en el tratamiento de infecciones estafilocócicas graves. Abstr. 217 (*Proceedings)* V Congreso de la Sociedad Española de Enfermedades Infecciosas y Microbiología Clínica (SEIMC), Barcelona.

Brause, B. D. (1990). Infections with prostheses in bones and joints. In *Principles and practice of infectious diseases*, (3rd edn) (ed. G. L. Mandell, D. R. Gordon, Jr, and J. E. Bennett), pp. 919–30. Churchill Livingstone, New York.

Brook, I. and Frazier, E. H. (1993). Anaerobic osteomyelitis and arthritis in a military hospital: a 10-year experience. *American Journal of Medicine*, **94**, 21–8.

Bruck, J. C., Buttemeyer, R., Grabosch, A., and Gruhl, L. (1991). More arguments in favour of myocutaneous flaps for the treatment of pelvic pressure sores. *Annals of Plastic Surgery*, **26**, 85–8.

Buchholz, H. W. *et al*. (1981). Management of deep infection of total hip replacement. *Journal of Bone and Joint Surgery (B)*, **63**, 342.

Buchholz, H. W., von Foerster, G., and Heinert, K. (1984). Management of infected prostheses. *Orthopedics*, **7**, 1620.

Burkert, T. and Watanakunakorn, C. (1991). Group G streptococcus septic arthritis and osteomyelitis: report and literature review. *Journal of Rheumatology*, **18**, 904–7.

Calhoun, J. H., Henry, S. L., Anger, D. M., Cobos, J. A., and Mader, J. T. (1993). The treatment of infected nonunions with gentamicin polymethylmethacrylate antibiotic beads. *Clinical Orthopedics*, **295**, 23–7.

Cattaneo, R., Catagni, M., Johnson, E. E. (1992). The treatment of infected nonunions and segmental defects of the tibia by the methods of Ilizarov. *Clinical Orthopedics*, **280**, 143–52.

Chandrasekar, P. H. and Narula, P. (1986). Bone and joint infections in intravenous drug abusers. *Reviews of Infectious Diseases*, **6**, 904–11.

Christensson. B., Ryd, L., Dahlberg, L., and Lohmander, S. (1993). *Candida albicans* arthritis in a nonimmunocompromised patient. Complication of placebo intraarticular injections. *Acta Orthopaedica Scandinavica*, **64**, 695–8.

Cierny, G. and Mader, J. T. (1984). Adult chronic osteomyelitis. *Orthopedics*, **7**, 1557.

Cierny, G. and Mader, J. T. (1989). Adult chronic osteomyelitis: an overview. In *Orthopaedic infections* (ed. R. D. D'Ambrosia and R. L Marier), pp. 31–47. SLACK International, Thorofare, NJ.

Cohen, O. J. and Squires, K. (1992). *Mycobacterium avium* osteomyelitis as the presenting manifestation of AIDS. *Infectious Diseases in Clinical Practice*, **1**, 110–13.

Conrad, D. A. and Marks, M. I. (1989). Oral therapy for orthopaedic infections in children and adults. In *Orthopaedic infections* (ed. R. D. D'Ambrosia and R. L Marier), pp. 345–60. SLACK International, Thorofare, NJ..

Coventry, M. B. (1975). Treatment of infections occurring in total hip surgery. *Orthopedic Clinics of North America*, **6**, 991.

Craig, W. A. and Andes, D. R. (1995). Parenteral versus oral antibiotic therapy. *Medical Clinics of North America*., **79**, 497–508.

Crim, J. R. and Seeger, L. L. (1994). Imaging evaluation of osteomyelitis. *Critical Reviews of Diagnostic Imaging*, **35**, 201–56.

Cuende, E., Barbadillo, C-E., Mazzucchelli, R., Isasi, C., Trujillo, A., and Andreu, J. L. (1993). *Candida* arthritis in adult patients who are not intravenous drug addicts: report of three cases and review of the literature. *Seminars in Arthritis and Rheumatism*, **22**, 224–41.

Dacquet, V. *et al*. (1992). Antibiotic impregnated plaster of Paris beads. Trials with teicoplanin. *Clinical Orthopedics*, **282**, 241–9.

Darouiche, R. O., Landon, G. C., Klima, M., Musher, D. M., and Markowski, J. (1994). Osteomyelitis associated with pressure sores. *Archives of Internal Medicine*, **154**, 753–8.

Davis, J. C. *et al*. (1986). Chronic non-hematogenous osteomyelitis treated with adjuvant hyperbaric oxygen. *Journal of Bone and Joint Surgery (A)*, **68**, 1210–17.

Demopulos, G. A., Bleck, E. E., and McDougall, I. R. (1988). Role of radionuclide imaging in the diagnosis of acute osteomyelitis. *Journal of Pediatric Orthopedics*, **8**, 558.

Depenbusch, F. I., Thompson, R. E., and Hart, G. B. (1972). Use of hyperbaric oxygen in the treatment of refractory osteomyelitis: a preliminary report. *Journal of Trauma*, **12**, 807–12.

De Wit, D., Mulla, R., Cowie, M. R., Mason, J. C., and Davies, K. A. (1993). Vertebral osteomyelitis due to *Staphylococcus epidermidis*. *British Journal of Rheumatology*, **32**, 339–41.

DiMaio, F. R., O'Halloran, J. J., and Quale, J. M. (1994). *In vitro* elution of ciprofloxacin from polymethylmethacrylate cement beads. *Journal of Orthopedic Research*, **12**, 79–82.

Dirschl, D. R. and Almekinders, L. C. (1993). Osteomyelitis. Common causes and treatment recommendations. *Drugs*, **45**, 29–43.

Dixon, R. S. and Sydnor, C. H. (1993). Puncture wound pseudomonal osteomyelitis of the foot. *Journal of Foot and Ankle Surgery*, **32**, 434–42.

Drancourt, M., Stein, A., Argenson, J. N., Zannier, A., Curvale, G., and Raoult, D. (1993). Oral rifampin plus ofloxacin for treatment of *Staphylococcus*-infected orthopedic implants. *Antimicrobial Agents and Chemotherapy*, **37**, 1214–18.

Drobniewski, F. A. (1993). *Bacillus cereus* and related species. *Clinical Microbiology Reviews*, **6**, 324–38.

Dunkle, L. M. and Brock, N. (1982). Long-term follow-up of ambulatory management of osteomyelitis. *Clinical Pediatrics*, **21**, 650.

Duwelius, P. J. and Schmidt, A. H. (1992). Assessment of bone viability in patients with osteomyelitis: preliminary clinical experience with laser Doppler flowmetry. *Journal of Orthopedics and Trauma*, **6**, 327–32.

Elhanan, G. and Raz, R. (1993). Group B streptococcal vertebral osteomyelitis in an adult. *Infection*, **21**, 397–9.

Engh, C. A., Hughes, J. L., Abrams, R. C., and Bowerman, J. W. (1971). Osteomyelitis in the patient with sickle-cell disease. *Journal of Bone and Joint Surgery*, **53**, 1–15.

Farley, M. M. (1995). Group B streptococcal infection in older patients. Spectrum of disease and management strategies. *Drugs and Aging*, **6**, 293–300.

Fitzgerald, R. H. (1984). Antibiotic distribution in normal and osteomyelitic bone. *Orthopedic Clinics of North America*, **15**, 537.

Fitzgerald, R. H. and Jones, D. R. (1985). Hip implant infection. Treatment with resection arthroplasty and late total hip arthroplasty. *American Journal of Medicine*, **78**, S225.

Fox, I. M. and Zeiger, L. (1993). Tc- ^{99}m-HMPAO leukocyte scintigraphy for the diagnosis of osteomyelitis in diabetic foot infections. *Journal of Foot and Ankle Surgery*, **32**, 591–4.

Furr, P. M., Taylor-Robinson, D., and Webster, A. D. (1994). Mycoplasmas and ureaplasmas in patients with hypogammaglobulinaemia and their role in arthritis: microbiological observations over twenty years. *Annals of the Rheumatic Diseases*, **53**, 183–7.

Ganapathy, M. E. and Rissing, J. P. (1995). Group B streptococcal vertebral osteomyelitis with bacteremia. *Southern Medical Journal*, **88**, 350–1.

Gelfand, M. S. and Cleveland, K. O. (1992). Penicillin-resistant pneumococcal vertebral osteomyelitis. *Clinical Infectious Diseases*, **15**, 746–7.

Gentry, L. O. (1990). Antibiotic therapy for osteomyelitis. *Infectious Disease Clinics of North America*, **4**, 485–99.

Gentry, L. O. and Rodríguez, G. G. (1990). Oral ciprofloxacin compared with parenteral antibiotics in the treatment of osteomyelitis. *Antimicrobial Agents and Chemotherapy*, **34**, 40–3.

Gentry, L. O. and Rodríguez-Gomez, G. (1991). Ofloxacin versus parenteral therapy for chronic osteomyelitis. *Antimicrobial Agents and Chemotherapy*, **35**, 538–41.

Gerding, D. N. (1995). Foot infections in diabetic patients: the role of anaerobes. *Clinical Infectious Diseases*, **20** (Suppl. 2), S283–8.

Gerhart, T. N., Roux, R. D., Hanff, P. A., Horowitz, G. L., Renshaw, A. A., and Hayes, W. C. (1993). Antibiotic-loaded biodegradable bone cement for prophylaxis and treatment of experimental osteomyelitis in rats. *Journal of Orthopedic Research*, **11**, 250–5.

Gillespie, W. J. (1990). Epidemiology in bone and joint infection. *Infectious Disease Clinics of North America*, **4**, 361–76.

Gómez, E. (1987). Radiología de la infección ósea. *Enfermedades Infecciosas Microbiología Clinica*, **5**, 525–30.

Gostishchev, V. K., Karmazanovskii, G. G., Shalykova, L. P., Vasilkova, Z. F., and Vavilova, G. S. (1992). Special methods of examination and their role in the selection of the extent and type of surgical intervention in patients with pelvic osteomyelitis. *Khirurgiia (Moskow)*, **7–8**, 17–21.

Greenfield, R. A. (1993). Symposium on antimicrobial therapy. VII. The fluoroquinolones. *Journal of the Oklahoma State Medical Association*, **86**, 166–74.

Guay, D. R. (1992). The role of the fluoroquinolones. *Pharmacotherapy*, **12**, 71–85S.

Guerrero, A. (1987). Estudio etiológico de las osteomielitis bacterianas. *Enfermedades Infecciosas Microbiología Clinica*, **5**, 517–20.

Guerrero, A. (1989). Infecciones del aparato locomotor. In *Monografía en comentarios a la literatura en enfermedades infecciosas*, Vol. 5 (2), (ed. E. Bouza) pp. 9–88. CIMSA, Madrid.

Gustilo, R. B., Merkow, R. L., and Templeman, D. (1990). The management of open fractures. *Journal of Bone and Joint Surgery (A)*, **72**, 299–304.

Heary, R. F., Hunt, C. D., and Wolansky, L. J. (1994). Rapid bony destruction with pyogenic vertebral osteomyelitis. *Surgical Neurology*, **41**, 34–9.

Herndon, W. A. *et al*. (1985). Nuclear imaging for musculo-skeletal infections in children. *Journal of Pediatric Orthopedics*, **5**, 343.

Hopkins, K. L., Li, K. C., and Bergman, G. (1995). Gadolinium-DTPA-enhanced magnetic resonance imaging of musculoskeletal infectious processes. *Skeletal Radiology*, 325–30.

Housang, E. T. (1976). Acute *Listeria monocytogenes* osteomyelitis. *Infection*, **4**, 113.

Howard, C. B., Einhorn, M., Dagan, R., and Yagupski, P. (1994). Fine-needle bone biopsy to diagnose osteomyelitis. *Journal of Bone and Joint Surgery (B)*, **76**, 311.

Hudson, J. W. (1993). Osteomyelitis of the jaws: a 50-year perspective. *Journal of Oral and Maxillofacial Surgery*, **51**, 1294–301.

Ish-Horowicz, M. R., McIntyre, P., and Nade, S. (1992). Bone and joint infections caused by multiply resistant *Staphylococcus aureus* in a neonatal intensive care unit. *Pediatric Infectious Diseases Journal*, **11**, 82–7.

Jacobs, N. M. (1991). Pneumococcal osteomyelitis and arthritis in children. A hospital series and literature review. *American Journal of Diseases of Children*, **145**, 70–4.

Jamil, S., Brennessel, D., Pessah, M., and Hilton, E. (1992). Fluconazole treatment of cryptococcal osteomyelitis. *Infectious Diseases in Clinical Practice*, **1**, 115–17.

Jauregui, L., Matzke, D., Scott, M., Minns, P., and Hageage, G. (1993). Cefepime as treatment for osteomyelitis and other severe bacterial infections. *Journal of Antimicrobials and Chemotherapy*, **32**, 141–9.

Kain, Z., Frogel, M., and Krilov, L. R. (1990). Osteomyelitis associated with varicella infection *Pediatric Infectious Diseases Journal*, **9**, 146–7.

Lacour, M., Duarte, M., Beutler, A., Auckenthaler, R., and Suter, S. (1991). Osteoarticular infections due to *Kingella kingae* in children. *European Journal of Pediatrics*, **150**, 612–18.

Laughlin, R. T., Sinha, A., Calhoun, J. H., and Mader, J. T. (1994). Osteomyelitis. *Current Opinion in Rheumatology*, **6**, 401–7.

Laughlin, R. T., Wright, D. G., Mader, J. T., and Calhoun, J. H. (1995). Osteomyelitis. *Current Opinion in Rheumatology*, **7**, 315–21.

Laurin, J. M., Resnik, C. S., Wheeler, D., and Needleman, B. W. (1991). Vertebral osteomyelitis caused by *Nocardia asteroides*: report and review of the literature. *Journal of Rheumatology*, **18**, 455–8.

Lavery, L. A., Harkless, L. B., Felder-Johnson, K., and Mundine, S. (1994). Bacterial pathogens in infected puncture wounds in adults with diabetes. *Journal of Foot and Ankle Surgery*, **33**, 91–7.

LeFrock, J. L. *et al*. (1992). Teicoplanin in the treatment of bone and joint infections. Teicoplanin Bone and Joint Cooperative Study Group, USA. *European Journal of Surgery*, **567** (Suppl.), 9–13.

Lentz, S. S. (1995). Osteitis pubis: a review. *Obstetrics and Gynecology Survey*, **50**, 310–15.

Lerner, R. K., Esterhai, J. L., Polomano, R. C., and Cheatle, M. D. (1993). Quality of life assessment of patients with posttraumatic fracture nonunion, chronic refractory osteomyelitis, and lower-extremity amputation. *Clinical Orthopedics*, **295**, 28–36.

Levine, S. E., Esterhal, J. L, Heppenstall, R. B, Calhoun, J., and Mader, J. T. (1993). Diagnoses and staging. osteomyelitis and prosthetic joint infections. *Clinical Orthopedics*, **295**, 77–86.

Lidwell, O. M. (1986). The operating environment. *Seminars in Orthopaedics*, **1**, 33.

Mabit, C., Huc, H., Setton, D., Leboutet, M. J,, Arnaud, J. P., and Pecout, C. (1993). Epidermoid carcinoma arising in femoral osteitis. A case. *Revue Chirurgie Orthopédie Reparatrice Apparet Motile*, **79**, 62–5.

McCarthy, J. M. and Haber, J. (1987). Group B streptococcal soft tissue infections beyond the neonatal period. *Western Journal of Medicine*, **147**, 558–60.

McGuire, M. H. (1989). The pathogenesis of adult osteomyelitis. *Orthopedic Reviews*, **18**, 564–70.

Mader, J. T., Adams, K. R., Wallace, W. R., and Calhoun, H. (1990). Hyperbaric oxygen as adjunctive therapy for osteomyelitis. *Infectious Disease Clinics of North America*, **4**, 433–40.

Mader, J. T., Norden, C., Nelson, J. D., and Calandra, G. B. (1992). Evaluation of new anti-infective drugs for the treatment of osteomyelitis in adults. Infectious Diseases Society of America and the Food and Drug Administration. *Clinical Infectious Diseases*, **15**, S155–61.

Mahan, S. and Jolles, P. R. (1995). MAI osteomyelitis. 18-year scintigraphic follow-up. *Clinical Nuclear Medicine*, **20**, 594–8.

Malchau, H., Herberts, P., and Ahnfelt, L. (1993). Prognosis of total hip replacement in Sweden. Follow-up of 92 675 operations performed (1978–1990). *Acta Orthopaedica Scandinavica*, **64**, 497–506.

Mateo, L., Miquel Nolla, J., Rozadilla, A., and del Blanco, J. (1993). Vertebral osteomyelitis caused by *Streptococcus agalactiae*. *Medicina Clinica*, **100**, 398.

Maurer, J., Lehmann-Beckow, D., Vosshenrich, R., Fischer, U., and Grabbe, E. (1992). The place of computed tomography and magnetic resonance tomography in the diagnosis of bone sequestra. *Aktuelle Radiologie*, **2**, 345–9.

Meadows, S. E., Zuckerman, J. D., and Koval, K. J. (1993). Posttraumatic tibial osteomyelitis: diagnosis, classification, and treatment. *Bulletin of the Hospital for Joint Diseases*, **52**, 11–16.

Meyers, S. P. and Wiener, S. N. (1991). Diagnosis of hematogenous pyogenic vertebral osteomyelitis by magnetic resonance imaging. *Archives of Internal Medicine*, **151**, 683–7.

Middleton, D. B. (1993). Infectious arthritis. *Primary Care*, **20**, 943–53.

Miholic, J. *et al*. (1985). Risk factors for severe bacterial infections after valve replacement and aortocoronary bypass operations: analysis of 246 cases by logistic regression. *Annals of Thoracic Surgery*, **40**, 224–8.

Modic, M. T., Feiglin, D. H., and Piraino, D. W. (1985). Vertebral osteomyelitis: assessment using MR. *Radiology*, **157**, 157–66.

Morrison, W. B. *et al*. (1993). Diagnosis of osteomyelitis: utility of fat-suppressed contrast-enhanced MR imaging. *Radiology*, **189**, 251–7.

Muñoz, C., Trujillo, G., Latorre, C., Juncosa, T., and Huget, R. (1992). Osteoarticular infections in children. *Enfermedades Infecciosa Microbiología Clinica*, **10**, 286–9.

Nelson, J. D. (1982). A critical review of the role of oral antibiotics in the management of hematogenous osteomyelitis. In *Current clinical topics in infectious diseases* (ed. J. S. Remington and M. N. Swartz), p. 64. McGraw-Hill, New York.

Nelson, J. D. (1990). Acute osteomyelitis in children. *Infectious Disease Clinics of North America*, **4**, 513–23.

Nepola, J. V., Seabold, J. E., Marsh, J. L., Kirchner, P. T., and el-Khoury, G. Y. (1993). Diagnosis of infection in ununited fractures. Combined imaging with indium-111-labeled leukocytes and technetium-^{99}m methylene diphosphonate. *Journal of Bone and Joint Surgery (A)*, **75**, 1816–22.

Neu, H. C. (1993). Synergy and antagonism of fluoroquinolones with other classes of antimicrobial agents. *Drugs*, **45**, S54–8.

Newman, L. G. *et al*. (1991). Unsuspected osteomyelitis in diabetic foot ulcers. Diagnosis and monitoring by leukocyte scanning with indium in 111 oxyquinoline. *Journal of the American Medical Association*, **266**, 1246–51.

Newman, L. G. *et al*. (1992). Leukocyte scanning with ^{111}In is superior to magnetic resonance imaging in diagnosis of clinically unsuspected osteomyelitis in diabetic foot ulcers. *Diabetes Care*, **15**, 1527–30.

Noonan, K. J., Goetz, D. D., Marsh, J. L., and Peterson, K. K. (1993). Rapidly destructive squamous cell carcinoma as a complication of chronic osteomyelitis. *Orthopedics*, **16**, 1140.

Novak, R. M., Polisky, E. L., Janda, W. M., and Libertin, C. R. (1988). Osteomyelitis caused by *Rhodococcus equi* in a renal transplant recipient. *Infection*, **16**, 186–8.

O'Reilly, T., Kunz, S., Sande, E., Zak, O., Sande, M. A., and Tauber, M. G. (1992). Relationship between antibiotic concentration in bone and efficacy of treatment of staphylococcal osteomyelitis in rats: azithromycin compared with clindamycin and rifampin. *Antimicrobial Agents and Chemotherapy*, **36**, 2693–7.

Ostermann, P. A., Henry, S. L., and Seligson, D. (1993). The role of local antibiotic therapy in the management of compound fractures. *Clinical Orthopaedics*, **295**, 102–11.

Overbeck, J. P., Winckler, S. T., Meffert, R., Tormala, P., Spiegel, H. U., and Brug, E. (1995). Penetration of ciprofloxacin into bone: a new bioabsorbable implant. *Journal of Investigative Surgery*, **8**, 155–62.

Patzakis, M. J., Wilkins, J., Kumar, J., Holtom, P., Greenbaum, B., and Ressler, R. (1994). Comparison of the results of bacterial cultures from multiple sites in chronic osteomyelitis of long bones. *Journal of Bone and Joint Surgery (A)*, **76**, 664–6.

Perrins, J. D. *et al*. (1966). OHP in the management of chronic osteomyelitis. In *Proceeding of the Third International Conference on Hyperbaric Medicine* (ed. I. W. Brown and B. G. Cox), pp. 578–84. National Academy of Sciences, National Research Council, Washington DC.

Perry, C. R., Pearson, R. L., and Miller, G. A. (1991). Accuracy of cultures of material from swabbing of the superficial aspect of the wound and needle biopsy in the preoperative assessment of osteomyelitis. *Journal of Bone and Joint Surgery (A)*, **73**, 745–9.

Pitkin, A. D. and Eykyn, S. J. (1993). Covert multi-focal infective arthritis. *Journal of Infection*, **27**, 297–300.

Poiraudeau, S., Liote, F., Bardin, T., Kuntz, D., and Dryll, A. (1993). Septic arthritis in HIV infection: 5 cases. (Letter). *Annales de Medicine Interne*, **144**, 344–5.

Prober, C. G. and Yeager, A. S. (1979). Use of the serum bactericidal titer to assess the adequacy of oral antibiotic therapy in the treatment of acute hematogenous osteomyelitis. *Journal of Pediatrics*, **95**, 131.

Pruitt, T. C., Hughes, L. O., Blasier, R. D., McCarthy, R. E., Glasier, C. M., and Roloson, G. J. (1993). Atypical mycobacterial vertebral osteomyelitis in a steroid-dependent adolescent. A case report. *Spine*, **18**, 2553–5.

Refsahl, K. and Andersen, B. M. (1991). Coagulase-negative staphylococci-problem bacteria in the hospital. Identification and resistance status. *Nordisk Medicin*, **106**, 228–31.

Roberts, P., Walters, A. J., and McMinn, D. J. W. (1992). Diagnosing infection in hip replacements. The use of fine-needle aspiration and radiometric culture. *Journal of Bone and Joint Surgery (B)*, **74**, 265–9.

Rogeaux, O., Fassin, D., and Gentilini, M. (1993). Prevalence of rheumatic manifestations in human immunodeficiency virus infection. *Annales de Medicine Interne*, **144**, 443–8.

Sadat-Ali, M., Kutty, S., and Kutti, K. (1985). Recent observations on osteomyelitis in sickle-cell disease. *International Orthopedics*, **9**, 97.

Sapico, F. L. and Montgomerie, J. Z. (1979). Pyogenic vertebral osteomyelitis. Report of nine cases and review of the literature. *Reviews of Infectious Diseases*, **1**, 754–76.

Sapico, F. L. and Montgomerie, J. Z. (1980). Vertebral osteomyelitis in intravenous drug abusers. Report of three cases and review of the literature. *Reviews of Infectious Diseases*, **2**, 196–206.

Sapico, F. L. and Montgomerie, J. Z. (1986). Vertebral osteomyelitis. In *Infectious diseases and medical microbiology*, (2nd edn) (ed. A. I. Braude, C. E. Davis, and J. Fierer), pp. 1479–81. Saunders, Philadelphia.

Sapico, F. L. and Montgomerie, J. Z. (1990). Vertebral osteomyelitis. *Infectious Disease Clinics of North America*, **4**, 539–50.

Schauwecker, D. S. *et al*. (1984). Evaluation of complicating osteomyelitis with Tc-^{99m} MDP, In-111 granulocytes and Ga-67 citrate. *Journal of Nuclear Medicine*, **25**, 849.

Schauwecker, D. S., Braunstein, E. M., and Wheat, L. J. (1990). Diagnostic imaging of osteomyelitis. *Infectious Disease Clinics of North America*, **4**, 441–63.

Schutzer, S. F. and Harris, W. H. (1988). Deep wound infection after total hip replacement under contemporary aseptic conditions. *Journal of Bone and Joint Surgery (A)*, **70**, 724.

Seabold, J. E., Ferlic, R. J., Marsh, J. L., and Nepola, J. V. (1993). Periarticular bone sites associated with traumatic injury: false-positive findings with In-111-labeled white blood cell and Tc-^{99m} MDP scintigraphy. *Radiology*, **186**, 845–9.

Sexton, D. J., Heskestad, L., Lambeth, W. R., McCallum, R., Levin, L. S., and Corey, G. R. (1993). Postoperative pubic osteomyelitis misdiagnosed as osteitis pubis: report of four cases and review. *Clinical Infectious Diseases*, **17**, 695–700.

Sharma, S. V., Khare, G. N., Bhalla, N., and Saraf, S. K. (1993). Changing profile of haematogenous osteomyelitis in a teaching hospital. *Indian Journal of Medical Research*, **98**, 92–5.

Slack, W. K., Thomas, D. A., and Perrins, D. J. D. (1965). Hyperbaric oxygenation in chronic osteomyelitis. *Lancet*, **I**, 1093–4.

Sliman, R., Rehm, S., and Shlaes, D. M. (1987). Serious infections caused by *Bacillus* species. *Medicine*, **66**, 218–23.

Smith, J. W. (1990). Infectious arthritis. *Infectious Disease Clinics of North America*, **4**, 523–38.

Spaeth, H. J. *et al*. (1991). Magnetic resonance imaging detection of early experimental periostitis. Comparison of magnetic resonance imaging, computed tomography, and plain radiography with histopathologic correlation. *Investigative Radiology*, **26**, 304–8.

Stratton, C. W. (1989). Microbiologic evaluation of patients with skeletal infections. In *Orthopaedic infections* (ed. R. D. D'Ambrosia and R. L. Marier), pp. 309–20. Slack International Book Distributors, Thorofare, NJ.

Straus, W. L. *et al*. (1994). Clinical and epidemiologic characteristics of *Mycobacterium haemophilum*, an emerging pathogen in immunocompromised patients. *Annals of Internal Medicine*, **120**, 118–25.

Suh, B and Lorber, B. (1995). Quinolones. *Medical Clinics of North America*, **79**, 869–94.

Syrogiannopoulos, G. A., McCracken, G. H., and Nelson, J. D. (1986). Osteoarticular infections in children with sickle cell disease. *Pediatrics*, **78**, 1090.

Tice, A. D. (1993). Osteomyelitis. *Hospital Practice*, **28** (Suppl. 2), 36–9.

Unger, E., Moldofsky, P., Gatenby, R., Hartz, W., and Broder, G. (1988). Diagnosis of osteomyelitis by MR imaging. *American Journal of Radiology*, **150**, 605.

Unkila-Kallio, L., Kallio, M. J., Eskola, J., and Peltola, H. (1994). Serum C-reactive protein, erythrocyte sedimentation rate, and white blood cell count in acute hematogenous osteomyelitis of children. *Pediatrics*, **93**, 59–62.

Voss, L. M., Rhodes, K. H., and Johnson, K. A. (1992). Musculoskeletal and soft tissue *Aeromonas* infection: an environmental disease. *Mayo Clinic Proceedings*, **67**, 422–7.

Waldvogel, F. A., Medoff, G., and Swartz, M. N. (1970). Osteomyelitis: a review of clinical features, therapeutic considerations and unusual aspects. *New England Journal of Medicine*, **198**, 260–316.

Wegener, W. A. and Alavi, A. (1991). Diagnostic imaging of musculoskeletal infection. Roentgenography, gallium, indium-labeled white blood cell, gammaglobulin, bone scintigraphy, and MRI. *Orthopedic Clinics of North America*, **22**, 401–18.

Weinberg, W. G. (1993). Safety and efficacy of teicoplanin for bone and joint infections: results of a community-based trial. *Southern Medical Journal*, **86**, 891–7.

Weinstein, M. P., Stratton, C. W., Hawley, H. B., and Reller, L. B. (1987). Multicenter collaborative evaluation of a standardized serum bactericidal test as a predictor of therapeutic efficacy in acute and chronic osteomyelitis. *American Journal of Medicine*, **83**, 218–22.

Wiley, A. M. and Trueta, J. (1959). The vascular anatomy of the spine and its relationship to pyogenic vertebral osteomyelitis. *Journal of Bone and Joint Surgery*, **41**, 796–809.

Williams, M. O. (1994). Long-term cost comparison of major limb salvage using the Ilizarov method versus amputation. *Clinical Orthopedics*, **301**, 156–8.

Wilson, A. P. and Gruneberg, R. N. (1994). Use of teicoplanin in community medicine. *European Journal of Clinical Microbiology and Infectious Diseases*, **13**, 701–10.

Yajima, H., Tamai, S., Ishida, H., and Fukui, A. (1995). Partial soleus muscle island flap transfer using minor pedicles from the posterior tibial vessels. *Plastic and Reconstructive Surgery*, **96**, 1162–8.

Yao, D. C. and Sartoris, D. J. (1995). Musculoskeletal tuberculosis. *Radiologic Clinics of North America*, **33**, 679–89.

Young, L. S. (1993). Mycobacterial diseases and the compromised host. *Clinical Infectious Diseases*, **17**, S436–41.

5.3.4 Lyme disease

Daniel W. Rahn

Introduction

Lyme disease is a tick-borne infectious disorder caused by a spirochaete, *Borrelia burgdorferi*. The best clinical marker of disease onset is a characteristic expanding skin lesion, erythema migrans. Weeks to months later the nervous system, heart, and joints may be affected with rare involvement of other organ systems. Lyme disease responds to antibiotics throughout its course, but treatment of early disease is the most successful. Endemic foci of Lyme disease have been described in the United States, Europe, and Asia.

History

'Lyme arthritis' was recognized in November 1975 because of an epidemic of arthritis in children in the region of Lyme, Connecticut (Steere *et al*. 1977b). Prospective, community-based study demonstrated the illness to be a multisystem disorder (Lyme disease) (Steere *et al*. 1977b; Reik *et al*. 1979; Steere *et al*. 1979; Steere *et al*. 1980; Steere *et al*. 1983a), occurring at any age and in both sexes, usually beginning with a characteristic expanding skin lesion, erythema chronicum migrans (Afzelius 1910), now shortened by convention to erythema migrans. Erythema migrans had previously been associated with the bite of the sheep tick, *Ixodes ricinus* (Thone 1968), and with tick-borne meningopolyneuritis (Garin-Bujadoux 1922), but not with arthritis (Bannwarth 1944). Field studies in the Lyme region revealed the presence of a closely related tick, *I. scapularis*, which was subsequently implicated as the principal disease vector (Steere *et al*. 1978; Wallis *et al*. 1978; Steere and Malawista 1979).

In 1982, Burgdorfer and associates (Burgdorfer *et al*. 1982) isolated a spirochaete from *I. scapularis* ticks from Shelter Island, New York. Patients with Lyme disease were found to have an elevated serological response to this spirochaete and within months this organism had been cultured from blood, skin, and cerebrospinal fluid of patients with various manifestations of Lyme disease (Benach *et al*. 1983; Steere *et al*. 1983d).

Much has been learned about Lyme disease but many areas of uncertainty remain. There is no clear consensus among practising physicians regarding the geographic range, clinical spectrum, and optimal treatment of this complex disorder. Why some individuals develop chronic symptoms despite apparent elimination of infecting organisms also remains to be elucidated.

Causative agent: *Borrelia burgdorferi*

Lyme disease is caused by *B. burgdorferi*, a spirochaete 10 to 30 µm long and 0.2 to 0.25 µm wide. It can be grown only in a specialized medium (Barbour–Stoner–Kelly medium) that is not available routinely in clinical laboratories (Weber *et al*. 1993). *Borrelia burgdorferi* replicates slowly *in vitro*, with a generation time of around 12 to 20 h. The organism has an outer membrane that surrounds a periplasmic space, multiple flagella, and a protoplasmic cylinder. Its genetic material is contained on a single linear chromosome and both linear and circular plasmids. Immunodominant, species-specific, outer surface proteins are encoded on plasmids. Other non-species-specific antigens of importance are flagellin (41 kDa) and a high molecular-weight heat-shock protein.

Three different genospecies have been described to date: *B. burgdorferi* sensu stricto, *B. garinii*, and *B. afzelii* (Baranton *et al*. 1992; Canica *et al*. 1993). A serotyping system has been developed based on reactivity to different epitopes of OspA (Wilske *et al*. 1993). North American isolates have shown less variability with regard to OspA, than have European isolates. Genetic differences among isolates appear to explain the different virulence patterns seen in European and North American Lyme disease. *B. afzelii*, isolated predominantly in Europe, has a propensity to persist in the skin causing the chronic skin lesion acrodermatitis chronica atrophicans. The other species appear to have a greater tendency for haematogenous dissemination (van Dam *et al*. 1993; Wienecke *et al*. 1994).

Pathogenesis

Lyme disease develops after an infected tick transmits *B. burgdorferi* to a susceptible host. Following a period of latency, organisms spread outward in skin causing the characteristic skin lesion, erythema migrans. Organisms can be readily cultured from biopsy specimens of the primary skin lesion. Histological staining of lesions of erythema migrans may reveal occasional spirochaetes and there is an inflammatory infiltrate consisting of lymphocytes, histiocytes, and plasma cells. In addition to erythema migrans, *B. burgdorferi* causes another acute skin lesion, benign lymphocytoma (Weber *et al*. 1984), seen primarily in Europe. The infection subsequently disseminates to involve secondary skin sites (secondary annular lesions) and other organs (e.g., central nervous system, heart and joints).

Borrelia burgdorferi has been cultured from involved tissues during all stages of the illness. Positive cultures have been obtained from blood (early) (Benach *et al*. 1983; Steere *et al*. 1983d), secondary skin lesions (Åsbrink and Hovmark 1985), cerebrospinal fluid (Steere *et al*. 1983d), joint fluid (Snydman *et al*. 1986), iris, ligamentous tissue, and a long-standing lesion of acrodermatitis chronica atrophicans (Åsbrink and Hovmark 1985). As the disease progresses, however, it becomes progressively more difficult to isolate organisms by culture.

Chronic manifestations have primarily involved the nervous system (Reik *et al*. 1979; Pachner *et al*. 1985; Halperin *et al*. 1987; Halperin *et al*. 1989; Pachner *et al*. 1989; Logigian *et al*. 1990), joints (Steere *et al*. 1987; Steere *et al*. 1990), and skin (Åsbrink and Olsson 1985). There is mounting evidence to explain how *B. burgdorferi* might persist in antibiotic-treated patients. *In vitro* the organism was able to invade and survive within some human cells such as fibroblasts, macrophages and endothelial cells, and thereby evade the action of antibiotics (Comstock and Thomas 1989; Klempner et al.1993; Montgomery et al.1994). Also, *B. burgdorferi* can cross the blood–brain barrier early in the course of infection where routine oral antibiotic regimens do not produce bactericidal levels. Neurological dysfunction resulting from low-grade infection in the central nervous system may only become apparent clinically months to years later (Garcia-Monco *et al*. 1990; Pfister *et al*. 1990; Luft *et al*. 1992).

Other, parainfectious mechanisms may also contribute to the development of chronic neuropathy and arthritis, at least in some individuals. There is a sharp distinction between the response to antibiotics of individuals with early disease and the inconsistent response of those with chronic neurological involvement (see below). Similarly, the likelihood of Lyme arthritis responding to antibiotics is affected by duration of arthritis and immunogenetic factors, particularly HLA-DR4 (Steere *et al*. 1979; Steere *et al*. 1990). Joint inflammation in these individuals may persist after the joint fluid has become negative in the polymerase chain reaction (Nocton *et al*. 1994). Thus, initial infection may trigger persistent inflammation in immunogenetically susceptible individuals. Clinical heterogeneity and inconsistent antibiotic

responsiveness may be explained by differing mechanisms of inflammation at different stages of the illness.

Disabling fatigue has been a particularly troublesome symptom associated with late Lyme disease. Fatigue alone rarely, if ever, responds to antibiotic therapy. Proinflammatory cytokines have been implicated in the pathogenesis of both Lyme arthritis and chronic neurological manifestations, and may account for this pathological fatigue. It is not known whether either live *B. burgdorferi* or retained bacterial products are necessary to stimulate cytokine release, an issue with obvious implications for therapy.

Immune abnormalities are present in all stages of Lyme disease (Steere *et al*. 1977; Steere *et al*. 1979a; Hardin *et al*. 1979a; Hardin *et al*. 1979b). At disease onset, immune complexes are often detectable in serum (Hardin *et al*. 1979a; Hardin *et al*. 1979b) and the serum IgM is often elevated, both of which are associated with disease dissemination. When arthritis is present, immune complexes are uniformly elevated in joint fluid rather than serum (Hardin *et al*. 1979a; Hardin *et al*. 1984). The synovial lesion contains a mixed infiltrate of lymphocytes and plasma cells (Steere *et al*. 1979). Production of anti-*B. burgdorferi* antibody in the cerebrospinal fluid is a hallmark of chronic neurological involvement (Halperin *et al*. 1989; Logigian *et al*. 1990).

As infection spreads from a local (skin) site to involve other organs, an inflammatory response ensues, which includes immune elements and the release of cytokines. Variation in symptoms and clinical course reflects both direct effects of infection and immunological phenomena triggered by the infection. Systemic symptoms result from disseminated infection, immune-mediated inflammation, and release of cytokines. Organ-specific inflammation occurs when localized infection (e.g. joints and nervous system) stimulates a localized immune response. But in immunogenetically susceptible individuals the immunological response may be prolonged, and perhaps even self-perpetuating. Therapeutic advances will require a clearer understanding of the specific disease mechanisms, particularly regarding the question of persistence of spirochaetes or their antigens throughout the entire course of the illness.

Epidemiology

Cases of Lyme disease have been reported from most states in the United States as well as throughout Europe, the former Soviet Union, China, Japan, and (questionably) Australia (Dekonenko *et al*. 1988; Burgdorfer 1989). In the United States, endemic foci are clustered in the north-east from Massachusetts to Maryland, the mid-west in Wisconsin and Minnesota, and the west along the northern California coast (Steere *et al*. 1979; Craven and Dennis 1993). Although Lyme disease was first described only 20 years ago, *B. burgdorferi* has been identified by polymerase chain reaction in preserved mouse tissues collected along the southern New England coast in 1894 (Marshall et al.1994).

A national surveillance case definition was adopted in the United States in 1990 (Table 1). With the use of this case definition, the number of cases reported in the United States has been relatively stable over the past 4 years, with approx. 9000 new cases reported nationally in 1993 (Craven and Dennis 1993). This case count is not an accurate reflection of the true incidence of Lyme disease, however, because most cases are not reported.

Table 1 Lyme disease: United States national surveillance case definition

The highest incidence of Lyme disease is in children under the age of 15 years and middle-aged adults. Illness generally begins between May l and November 30, with the peak in June and July (Steere *et al*. 1983a). Limited studies have shown a significant prevalence of asymptomatic seropositivity in high-risk populations, the significance of which is unknown at present (Steere *et al*. 1986).

The primary vectors of Lyme disease are ixodid ticks. Endemic regions correspond to the distribution of *I. scapularis* (north-east and upper mid-west United States) (Steere *et al*. 1979), *I. pacificus* (California) (Steere *et al*. 1979), *I. ricinus* (Europe and Russia) (Dekonenko *et al*. 1988), and *I. persulcatus* (eastern Russia, China, and Japan) (Burgdorfer 1989). The vector in the north-east United States was previously thought to be a separate species, *I. dammini*, which has recently been shown to be onspecific with *I. scapularis*, its current designation (Oliver *et al*. 1993).

The epidemiology of Lyme disease is explained by the ecology of the tick vectors. The most thoroughly studied, *I. scapularis*, has a three-stage lifecycle (larva, nymph, and adult) spanning 2 years (Fig. 1). In the north-east United States, both the larval and nymphal stages feed on a variety of small mammals, especially the white-footed mouse (*Peromyscus leukopus*) (Wallis *et al*. 1978; Spielman *et al*. 1985; Mather *et al*. 1989). Infected nymphs transmit spirochaetes to mice, which in turn pass the infection on to larval ticks. Humans most often acquire infection from nymphs. Adult ticks feed primarily on larger mammals, especially deer. In endemic regions, 20 to 60 per cent of nymphal *I. scapularis* may be infected with *B. burgdorferi* (Burgdorfer *et al*. 1982; Steere *et al*. 1983d). In contrast the infection rate in *I. pacificus* in endemic areas of California has been found to be 2 per cent or less (Burgdorfer *et al*. 1985), where a complex, two-tick enzootic cycle maintains *B. burgdorferi* in nature (Brown and Lane 1992). An enzootic cycle involving *Neotoma mexicana* (a wood rat) and *I. spinipalpis* has recently been described in northern Colorado (Maupin *et al*. 1994).

Fig. 1 Larva, nymph, adult female and adult male *Ixodes scapularis* ticks (photo M. Fergione).

Birds may serve as hosts for larval and nymphal *Ixodes* ticks and provide a natural means of distributing ticks to new areas (Anderson *et al*. 1985; Anderson *et al*. 1986). Birds do not appear to be a reservoir for *B. burgdorferi*, however, and probably do not contribute significantly to the spread of infection.

Various public-health interventions have been designed to reduce the incidence of Lyme disease. Prompt removal of ticks prevents infection in the vast majority of exposed individuals (Shapiro *et al*. 1992). Eliminating deer can reduce the total tick population but is impractical. Distributing permethrine-impregnated (an acaricide) cotton balls in the nesting environment of mice has been shown to reduce the infestation of mice by ticks, which in turn can reduce the tick infection rate, but this too has obvious practical limitations (Ginsberg 1995). In some regions of the United States the perceived threat of Lyme disease is much greater than objective data can

support. The public-health approach in these regions must be focused on education of the public and health professionals rather than disease control. Adequate control of Lyme disease will require close collaboration between ecologists, epidemiologists, public-health officials, physicians, and an informed public.

Clinical features

Most cases of Lyme disease begin with a characteristic skin lesion (erythema migrans) (stage 1). Within days to weeks, the illness disseminates, with the development of secondary skin lesions, headache, and generalized musculoskeletal symptoms. Weeks later, carditis and acute neurological abnormalities may occur (stage 2). Months later, arthritis and chronic neurological abnormalities appear (stage 3). Chronic skin involvement (acrodermatitis chronica atrophicans) also occurs late in disease, predominantly in Europe. As a guide to therapy, it is best to characterize patients as having early localized (erythema migrans), acute disseminated (neurological or non-neurological), or chronic disease. Clinical stages may overlap or be skipped entirely (Hanrahan *et al*. 1984).

Early Lyme disease

Erythema migrans begins as a red macule or papule at the site of a tick bite (Steere *et al*. 1977c; Steere *et al*. 1983a). After an incubation period of a few days to a month, the lesion expands gradually (0.5–1 cm/day) to a mean diameter of l5 cm (range 3–68 cm). Lesions are flat and non-scaling with a red outer border and partial central clearing (Fig. 2). The centre may be flat, indurated, vesicular or necrotic. The posterior thigh, groin, popliteal fossa, and axilla are particularly common sites. Lesions are warm and minimally tender, and may go unnoticed.

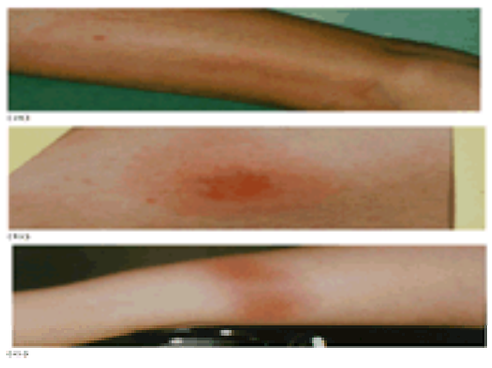


Fig. 2(a)–(c) Erythema migrans, various forms.

Half of patients in the United States develop multiple secondary lesions within days of onset of infection. Lymphocytoma cutis, a purplish nodule often on the nipple, has been reported as a manifestation of primary infection in Europe (Åsbrink and Olsson 1985). Erythema migrans and secondary lesions fade even without treatment in 3 to 4 weeks but may recur. Spirochaetes can be cultured from the skin of untreated patients even after lesions have resolved (Kuiper *et al*. 1994).

Malaise, fatigue, fever and chills, myalgia, arthralgia, headache, and paraesthesias often accompany erythema migrans (Steere *et al*. 1977c; Steere *et al*. 1983a). Recent studies have suggested that headache and paraesthesias may reflect early neurological dissemination (Reik *et al*. 1979; Garcia-Monco *et al*. 1990; Luft *et al*. 1992). Systemic signs and symptoms vary from day to day, and may appear before or after erythema migrans (or without it altogether). In untreated patients, symptoms may recur for months (especially fatigue and lethargy) after skin lesions have disappeared. European erythema migrans often has a prolonged, subacute course; secondary skin lesions, prominent systemic symptoms, laboratory abnormalities and subsequent arthritis are uncommon (Åsbrink and Olsson 1985).

Minor laboratory abnormalities associated with early Lyme disease are an increased erythrocyte sedimentation rate, total serum IgM, and serum glutamic oxaloacetic transaminase (Steere *et al*. 1977b; Steere *et al*. 1977c). Mild anaemia and leucocytosis also are common. Microscopic haematuria and low-grade proteinuria have been reported. Tests for rheumatoid factor or antinuclear antibodies are usually negative.

Disseminated and chronic Lyme disease

Neurological manifestations

Neurological abnormalities develop within several weeks of disease onset in a minority of patients (15 per cent in one series (Reik *et al*. 1979)). Cranial neuropathy (most commonly involving the facial nerve) (Clark *et al*. 1985), meningitis, and radiculoneuropathy alone or in combination (Reik *et al*. 1979; Pachner and Steere 1985; Halperin *et al*. 1987; Halperin *et al*. 1989; Pachner *et al*. 1989). Later in the disease, chorea, demyelinating encephalopathy or myelopathy, chronic encephalopathy, peripheral polyneuropathy, and transverse myelitis may occur (Reik *et al*. 1979; Ackerman *et al*. 1988; Halperin *et al*. 1987; Halperin *et al*. 1989; Logigian *et al*. 1990). Acute and chronic neurological Lyme disease differ in their natural history and response to therapy.

Radiculoneuropathy may affect any dermatome or even several contiguous dermatomes. Symptoms include paraesthesia, pain, sensory deficit and, often, motor weakness. Radiculoneuropathy is often accompanied by meningitis, a complex dubbed Bannwarth's syndrome.

The primary symptom of Lyme meningitis is headache, which may vary in intensity from hour to hour or day to day, and is often accompanied by a mild encephalopathy. Neck stiffness is common but meningismus is not. Spinal fluid contains a lymphocytic pleocytosis and mildly elevated protein, but glucose is normal. Studies with the polymerase chain reaction and rare positive cultures have shown that Lyme meningitis results from direct infection in the cerebrospinal fluid (Keller *et al*. 1992). Both radiculoneuropathy and meningitis eventually remit even without treatment but often only after waxing and waning for months (Reik *et al*. 1979).

Facial palsy, which may be bilateral, is the most common acute neurological manifestation of disseminated, early Lyme disease. Facial paralysis may occur when erythema migrans is present or within a few weeks after its resolution. Even in highly endemic areas, however, facial nerve palsy is usually idiopathic. The course of facial weakness is benign and not distinguishable from that of Bell's palsy; over 95 per cent of individuals experience complete or near complete resolution even without therapy. Fifty per cent of individuals with facial nerve palsy have an associated meningitis, so examination of the cerebrospinal fluid is indicated.

Chronic neurological manifestations emerge months to years after the onset of Lyme disease. Although symptoms vary, memory impairment, and peripheral, sensory polyneuropathy predominate (Reik *et al*. 1985; Halperin *et al*. 1987; Halperin *et al*. 1989; Logigian *et al*. 1990). The cerebrospinal fluid of individuals with chronic encephalopathy usually has a mildly elevated protein but no cells.

Carditis

Carditis develops in less than 5 per cent of cases, generally within several weeks of disease onset (Steere *et al*. 1978). Patients with carditis typically present with palpitations, light-headedness, or syncope due to varying degrees of atrioventricular block. Dilated cardiomyopathy, perhaps reflecting more diffuse myocardial involvement, has also been tentatively linked to Lyme disease (Steere *et al*. 1980; Stanek *et al*. 1990). Carditis remits spontaneously in days to a few weeks and, once resolved, does not recur. Although clinical manifestations are usually limited to the conduction system, subclinical myocardial involvement may be more extensive (McAlister *et al*. 1989). One fatal case involving concurrent *Babesia* infection has been reported (Marcus *et al*. 1985). Carditis should be suspected in individuals in endemic areas who develop heart block without other explanation.

Arthritis

Approximately 60 per cent of untreated individuals develop arthritis from weeks to years after the onset of illness (Steere *et al*. 1977; Steere *et al*. 1979). Frank arthritis may be preceded by months or even years of intermittent migratory myalgias, arthralgias, and periarticular pain. The typical attack begins suddenly with the rapid development of massive swelling of a single large joint, most often the knee. Individual attacks of arthritis usually last a few weeks to a few months and remit

spontaneously. Recurrences are common but attacks decrease in frequency by 10 to 20 per cent per year and eventually cease in most patients, even without antibiotic therapy (Steere *et al*. 1987).

Joint-fluid cell counts average about 25 000 cells/mm^3, with a predominance of polymorphonuclear leucocytes (Steere *et al*. 1977b). Protein is usually elevated; glucose usually normal. Attempts to culture *B. burgdorferi* from joint fluid have been almost invariably unsuccessful (Snydman *et al*. 1986) but the polymerase chain reaction has revealed *B. burgdorferi* DNA in most untreated patients (Bradley *et al*. 1994; Nocton *et al*. 1994). Arthritis in Lyme disease is almost certainly initiated by direct spirochaetal infection of the joint.

Arthritis becomes chronic in approximately 10 per cent of patients, especially individuals who are HLA-DR4 positive in whom pannus may form and erosions develop despite antibiotic therapy that is curative in other patients (Steere *et al*. 1979; Steere *et al*. 1987; Steere *et al*. 1990). The development of a serological immune response against outer surface proteins A and B has been linked to the development of chronic arthritis (Kalish *et al*. 1993). The synovium in chronic Lyme arthritis looks like that of rheumatoid arthritis (Steere *et al*. 1977b; Steere *et al*. 1979). In addition, there may be an obliterative endarteritis and spirochaetes have been seen using a variety of staining techniques. Thus, chronic Lyme arthritis is similar to rheumatoid arthritis, for which it may serve as a model.

Acrodermatitis chronica atrophicans

Acrodermatitis chronica atrophicans (Weber *et al*. 1984; Åsbrink *et al*. 1985) occurs relatively commonly in Europe but is rare in the United States (Kaufman *et al*. 1989). Lesions occur most often on distal extremities and begin as violaceous, infiltrated plaques or nodules that evolve into an atrophic stage. Acrodermatitis chronica atrophicans results from chronic persistence of infection in skin and has been associated with *B. afzelii* (Canica *et al*. 1993; van Dam et al.1993; Wienecke *et al*. 1994).

Diagnostic testing

Culture of *B. burgdorferi* from patients is necessary for definitive diagnosis, but has a significant yield only from affected (erythema migrans) skin (Steere *et al*. 1983; Åsbrink and Hovmark 1985; Berger *et al*. 1985). Spirochaetes are rarely visualized by tissue stains (Steere *et al*. 1983; Berger 1984; Johnston *et al*. 1985). Attempts should be made to isolate the organism by culture when erythema migrans is suspected in a patient from a region not previously known to be endemic; in endemic regions, this is not necessary for diagnosis.

Detection of specific anti-*B. burgdorferi* antibody is the most helpful diagnostic test for confirmation of the diagnosis of Lyme disease. Specific IgM antibody appears first after the onset of the disease and reaches a peak within 3 to 6 weeks. IgG antibody develops more slowly, often reaching a peak months later (Steere *et al*. 1983; Craft *et al*. 1984).

In the past, the performance of many commercially available immunofluorescence tests or enzyme-linked immunosorbent assays for *B. burgdorferi* antibodies has been poor, with unacceptably low sensitivity and/or specificity and lack of reproducibility (Schwartz *et al*. 1989; Luger and Krauss 1990; Magnarelli *et al*. 1990). The major problem at present, however, is not performance, which is comparable to that of many other serological tests, but rather that indiscriminate use of serological testing has set the stage for results with very low positive predictive value (Britton *et al*. 1993; Lightfoot *et al*. 1993). As an example, when the likelihood of Lyme disease is estimated to be 5 per cent before testing, a serological test with sensitivity and specificity of 95 per cent will have a positive predictive value of only 50 per cent. Patient selection has a profound effect on the predictive value of serological testing.

Serological tests may be falsely positive or negative for a variety of reasons. *Borrelia burgdorferi* contains epitopes that are cross-reactive with other spirochaetes, including *Treponema pallidum* and oral treponemes. Some patients with rheumatoid arthritis or systemic lupus erythematosus have low-titre, false-positive serological findings. Seronegative Lyme disease also occurs, but rarely, primarily following incomplete antibiotic therapy for early disease.

Western blotting is available as a confirming test (Dressler *et al*. 1993). This technique can distinguish between true seroreactivity against *B. burgdorferi* and false positivity. Both the technique for Western blotting and criteria for positivity must also be standardized, however. A workshop sponsored by the Centers for Disease Control in 1994 produced recommendations that Western blotting be used to confirm all positive enzyme immunoassay results (Table 2). The most cost-effective approach at present may be to reserve Western blotting for equivocal circumstances.

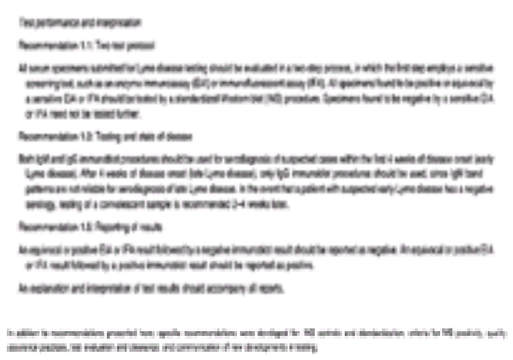


Table 2 Centers for Disease Control and Prevention Workgroup recommendations on laboratory testing for Lyme disease (draft)

Tests have been developed to measure cell-mediated immunity against *B. burgdorferi* but these have not added clinically useful information because of technical difficulties and the rarity with which the cell-mediated immune response differs from the serological response (Dattwyler *et al*. 1989; Zoschke *et al*. 1991).

Lyme disease of the central nervous system is associated with the production of specific antibody in cerebrospinal fluid. This locally driven immune response leads to a measurable increase in the concentration of specific antibody in the cerebrospinal fluid compared with that in serum. Measurement of antibody in cerebrospinal fluid is a useful adjunct in the diagnosis of neurological involvement in Lyme disease (Halperin *et al*. 1986; Halperin *et al*. 1989; Logigian *et al*. 1990). Positive tests for antibody in cerebrospinal fluid confirm *B. burgdorferi* infection of the central nervous system but negative tests do not rule it out, particularly late in the disease.

The polymerase chain reaction is being applied to the study of Lyme disease as a potential means of elucidating sites of active infection (Nocton *et al*. 1994; Keller *et al*. 1992). It has been validated as a sensitive and specific means of identifying *B. burgdorferi* in ticks (Persing *et al*. 1990). Study of both joint and cerebrospinal fluid of individuals has shown that the polymerase chain reaction has the potential to determine who harbours *B. burgdorferi* DNA. It is not clear whether the presence of DNA from the organism can be considered a surrogate for a positive culture, particularly since polymerase chain reaction-positive fluids are routinely culture-negative. A potential explanation for this observation can be based on the observation that the likelihood of a positive polymerase chain reaction seems to be higher if plasmid rather than genomic DNA targets are used (Persing *et al*. 1994). In culture, spirochaetes shed membrane blebs containing only plasmid DNA into the culture medium (Dorward and Garon 1990). It is plausible that *B. burgdorferi* may shed plasmid-rich DNA blebs into joint or cerebrospinal fluids in the virtual absence of intact organisms, which, in turn, could explain the disparate results for culture and the polymerase chain reaction.

The diagnosis of Lyme disease must be approached clinically. In a setting of risk by epidemiological criteria, an individual with clinical manifestations suggesting Lyme disease should undergo confirming serological testing. If negative, the diagnosis is unlikely. If positive, Lyme disease is likely. Serological status can be confirmed by Western blotting but the diagnosis must still be based on clinical criteria, even in individuals with definite seropositivity (Fig. 3).

Fig. 3 Algorithms for diagnosis of Lyme disease.

Differential diagnosis

Although erythema migrans is the most definitive clinical marker of Lyme disease, a cautionary note is warranted because an expanding erythema cannot be considered diagnostic of early Lyme disease in the absence of the right epidemiological setting. When present in its classical form following exposure in an endemic area, it is virtually diagnostic of Lyme disease, but skin lesions from other causes may mimic erythema migrans. Tick bites alone, without infection, may cause small annular areas of erythema. Efforts should be made to culture *B. burgdorferi* from skin lesions in questionable circumstances, particularly in regions not known to be endemic for Lyme disease.

Secondary lesions superficially resemble erythema multiforme. Several features help to distinguish between these entities: blistering, mucosal lesions, and involvement of the palms and soles are not typical of erythema migrans. Lyme disease has occasionally been associated with an urticarial rash, which must be distinguished from other causes of urticaria. Prominent musculoskeletal symptoms (arthralgias, myalgias and fever) may suggest a viral illness, especially when erythema migrans is absent or missed. Upper respiratory symptoms, which are very common with viral disease, are rare in early disseminated Lyme disease.

The headache, stiff neck, and pleocytosis of cerebrospinal fluid associated with Lyme meningitis are similar to the symptoms and signs associated with viral meningitis. Lyme meningitis has a more protracted course, however, often fluctuating in severity for weeks to months, which helps to distinguish it from viral meningitis. Isolated facial nerve palsy mimics idiopathic Bell's palsy; Lyme disease is one of the very few causes of bilateral facial palsy which is rare with idiopathic Bell's palsy. Generalized lymphadenopathy is occasionally seen in disseminated early Lyme disease. The fever and multiple skin lesions usually present at this stage of illness enable one to determine that the adenopathy is due to Lyme disease. Chronic fatigue may be a major and persistent complaint in Lyme disease but the non-specificity of fatigue is so low that it is not helpful diagnostically..

Disseminated and chronic Lyme disease shares clinical features with many other immune-mediated disorders. The pattern of joint inflammation is similar to that seen with reactive arthritis. The relatively brief duration of individual attacks of joint swelling in Lyme disease and absence of mucosal lesions are useful distinguishing features. In children, the attacks of arthritis, although generally shorter, are similar to those associated with the oligoarticular form of juvenile rheumatoid arthritis, but without iridocyclitis.

Late neurological involvement may mimic multiple sclerosis, Guillain–Barré syndrome, a dementing process, brain tumour, or an affective disorder. A history of previous features of Lyme disease helps, as does a complete neurological evaluation. A unique characteristic of neurological Lyme disease is that it may affect multiple levels of the nervous system: peripheral and cranial nerves, meninges, nerve roots, spinal cord and brain itself. Involvement of many different levels of the nervous system either simultaneously or sequentially is rarely seen in other diseases (Pachner and Steere 1985). In the evaluation of chronic fatigue, it is important to note that chronic fatigue alone is not a manifestation of active Lyme disease.

Diagnostic confusion has resulted from poor understanding of the how to distinguish fibromyalgia from Lyme disease. Fibromyalgia may occur as a sequela to Lyme disease, but there is no evidence that it results from persistent infection (Sigal 1990; Hsu *et al*. 1993; Steere *et al*. 1993). Lyme disease appears to be one of many triggers of this common syndrome. Fibromyalgia constitutes up to 25 per cent of presentations in rheumatology practice and affects up to 3 per cent of the normal population. Misdiagnosis of fibromyalgia as Lyme disease has led to extensive overdiagnosis of Lyme disease (Britton *et al*. 1993; Hsu *et al*. 1993; Lightfoot *et al*. 1993; Steere *et al*. 1993). The generalized pain, trigger points, debilitating fatigue, and sleep disturbance that characterize fibromyalgia are distinctly different from the joint and nervous system manifestations of Lyme disease. The treatment of fibromyalgia is not affected by a history of preceding Lyme disease.

Treatment

Antibiotic therapy of Lyme disease has advanced considerably over the past decade. Therapy must be tailored to the extent and duration of disease. The leading reason for antibiotic failure is incorrect diagnosis. Early disease responds readily to a variety of agents but the response of certain late manifestations, particularly neurological, is often incomplete. The appropriate endpoint of treatment for late disease may be difficult to determine because resolution may be slow and lag behind the completion of antibiotic therapy. Although the optimal duration of antibiotic therapy for the various stages of Lyme disease is still being determined, no clinical trials have evaluated treatment courses longer than 4 weeks for any stage of Lyme disease. Longer courses should only be administered in the context of a controlled clinical trial. Current recommendations are presented in Table 3.

Table 3 Treatment recommendations

Early Lyme disease

Early Lyme disease responds promptly to oral antibiotic therapy. Erythema migrans resolves in days and disease progression is halted (Steere *et al*. 1983b; Dattwyler *et al*. 1990; Luft *et al*. 1996). Secondary skin lesions, myalgias, arthralgias, fever, headache, stiff neck, and dysaesthesias indicate more severe infection and have a greater likelihood of incomplete response (Steere *et al*. 1983b; Dattwyler *et al*. 1990). Early treatment is important to prevent dissemination and maximize the likelihood of complete response. Because of the potential for early neurological spread, patients should be carefully evaluated for subtle involvement of the central nervous system. If headache or dys-/paraesthesias are present, formal evaluation with lumbar puncture and electrodiagnostic studies is indicated; if abnormal, intravenous antibiotic therapy should be chosen (Garcia-Monco *et al*. 1990; Pfister *et al*. 1990; Luft *et al*. 1992). If treated early, before any systemic immune challenge, patients may be susceptible to reinfection (Shrestha *et al*. 1985; Aguero-Rosenfeld *et al*. 1993).

Oral doxycycline, 100 mg twice daily, or amoxicillin, 500 mg three times a day (Dattwyler *et al*. 1990) (paediatric dose: 30 mg/kg a day), each for 21 days, are the

regimens of first choice for early disease. For penicillin-allergic individuals who should not take a tetracycline (pregnant or lactating women and children less than 9 years in age), cefuroxime axetil (500 mg twice daily) or erythromycin (250 mg four times a day) may be substituted (Nadelman *et al*. 1992). For individuals with a history of immediate hypersensitivity reactions to penicillin, cephalosporins should also be avoided. Outcomes with erythromycin or the newer macrolides, azithromycin and clarithromycin, have been less satisfactory than with the other choices (Steere *et al*. 1983b; Luft *et al*. 1996).

Maternal–fetal transmission of *B. burgdorferi* has been associated with stillbirth and neonatal death (Schlesinger *et al*. 1985; Weber *et al*. 1988). The direct role of *B. burgdorferi* in causing these adverse outcomes is unclear. Epidemiological surveys have demonstrated no congenital Lyme disease syndrome, however (Markowitz *et al*. 1986), and there is no evidence of fetal risk ascribable to maternal seropositivity alone (Strobino *et al*. 1993). There are no case reports of maternal fetal transmission associated with current recommended treatment. The available data favour aggressive treatment of Lyme disease in pregnancy.

About 10 per cent of patients experience a Jarisch–Herxheimer-like reaction shortly after the institution of therapy (Steere *et al*. 1983b). Regardless of which drug is selected and the duration of therapy, many patients have persistent symptoms particularly fatigue, for many weeks after its completion (Steere *et al*. 1983b; Dattwyler *et al*. 1990). These symptoms do not reflect continuing infection and do not respond to repeated courses of antibiotics, but eventually resolve in most patients. Persistent neurological complaints, however, should raise suspicion of inadequately treated nervous system infection which may progress over time even in seronegative individuals. The importance of careful neurological evaluation of all patients cannot be over emphasized.

Disseminated and chronic Lyme disease

Neurological disease

Optimal treatment of neurological involvement is still under investigation. Consensus has been slow to emerge because of the differences in the neurological features of European and North American Lyme disease and the slow, often incomplete, resolution of symptoms.

Acute neurological manifestations include meningitis, cranial neuropathy, and radiculoneuritis, and must be distinguished from the chronic manifestations described below. These respond to intravenous penicillin G, 20 million units a day for 10 days (Steere *et al*. 1983c); longer courses (14–21days) are recommended at present (Rahn and Malawista 1994). Ceftriaxone (Dattwyler *et al*. 1988) and cefotaxime (Pfister *et al*. 1989) have both been found effective when given intravenously for 14 days, and are superior to penicillin (Dattwyler *et al*. 1988). Headache usually begins to subside by the second day of therapy and disappears by 7 to 10 days. Radicular pain also resolves promptly but sensory and motor deficits frequently require 7 to 8 weeks for recovery (Steere *et al*. 1983c). Most experts currently recommend 4 weeks of therapy for all manifestations of central nervous Lyme disease.

Radiculopathy has been treated successfully in Europe with doxycycline, 100 mg twice daily for 14 days (Dotevall *et al*. 1988). This regimen has not been studied systematically in the United States. The recently recognized genetic differences between *B. burgdorferi* isolates (Baranton *et al*. 1992; Canica *et al*. 1993) suggest that future treatment should involve attempts to isolate and speciate the offending organism.

Chronic neurological manifestations, including encephalopathy and peripheral neuropathy, usually respond to antibiotic therapy, but response may be delayed or incomplete (Halperin *et al*. 1987; Halperin *et al*. 1989; Logigian *et al*. 1990). Some patients have persistent, non-progressive deficits after antibiotic therapy. Two to four weeks of treatment are generally recommended, with most authorities recommending 4 weeks.

Many individuals with ill-defined neurological complaints have been diagnosed to have neurological Lyme disease, without a history of exposure, serological confirmation, or even a clearly defined neurological lesion. The costs and risks attributable to misdiagnosis and inappropriate intravenous antibiotic therapy are considerable (Centers for Disease Control 1993; Lightfoot *et al*. 1993). It is best to withhold therapy until symptoms have been defined as precisely as possible. This may require neuropsychological testing for individuals with cognitive complaints, lumbar puncture for headache, and electrodiagnostic testing to evaluate sensorimotor complaints. Therapeutic trials should not be undertaken for treatment of vague symptoms without an objectively defined endpoint. The American College of Rheumatology has published a guideline recommending against intravenous therapy for seropositive individuals with non-specific symptoms within the spectrum of fibromyalgia (Britton *et al*. 1993).

Carditis

Carditis responds readily to intravenous antibiotic therapy with one of the above agents (Brown and Lane 1992) but it has also been treated successfully with oral antibiotics, salicylates or glucocorticoids (Steere *et al*. 1980). Carditis results from myocardial invasion by spirochaetes (de Koning *et al*. 1989), and should be treated with antibiotics. Prednisone, 40 to 60 mg a day in divided doses, rapidly reverses heart block (Steere *et al*. 1980) but may interfere with attempts to cure infection. For this reason, glucocorticoids should be given in short courses (fewer than 7 days) or avoided entirely. For patients with allergy to penicillin and cephalosporins, doxycycline (100 mg twice a day for 30 days) is a reasonable alternative choice. Cardiac pacing may be required temporarily (Steere *et al*. 1980; Clark *et al*. 1985).

Arthritis

Lyme arthritis responds to oral doxycycline (100 mg twice daily for 4 weeks) or amoxicillin/probenecid (500 mg each, four times daily). Either regimen cures most patients (Steere *et al*. 1994). Intravenous ceftriaxone is probably as effective when given for at least 2 weeks and offers the advantage of simultaneously treating neurological involvement but this must be balanced against the disadvantages of increased cost and increased risk of toxicity (antibiotic-associated colitis, line sepsis, acute cholecystitis, etc.) (Dattwyler *et al*. 1988; Steere *et al*. 1994). Parenteral penicillin is less effective for some patients than oral doxycycline or amoxicillin (Steere *et al*. 1985).

Treatment failures occur with all regimens, particularly in HLA-DR4 positive individuals (Steere *et al*. 1990). Longer courses of oral amoxicillin or tetracycline, or longer intravenous therapy with ceftriaxone or penicillin should be considered for those individuals who have not responded to an initial 4-week course, but this requires further study. In one study, there were no polymerase chain reaction-positive joint fluids after either 2 weeks of intravenous ceftriaxone or 8 weeks' oral therapy (Malawista *et al*. 1994). Antibiotic treatment failures may be treated successfully with synovectomy (Schoen *et al*. 1991). Recurrence after synovectomy has been rare.

Fibromyalgia

Fibromyalgia is the most common primary diagnosis for individuals coming to Lyme disease referral centres for treatment-resistant Lyme disease. In endemic areas in particular, fibromyalgia is often inappropriately treated with repeated, prolonged courses of antibiotics. Regardless of *B. burgdorferi* serological status, fibromyalgia should be treated in conventional ways.

Acrodermatitis chronica atrophicans

The infiltrative lesions of acrodermatitis chronica atrophicans are usually cured by 3 weeks of oral phenoxymethyl penicillin, 2 to 3 g daily in divided doses (Åsbrink *et al*. 1985).

Tick bite management

Ticks should be removed and the site observed for the appearance of erythema migrans. In one randomized, controlled trial in an endemic region in Connecticut, the risk of acquiring Lyme disease following a tick bite approximated the risk of adverse reaction to antibiotics administered prophylactically (Shapiro *et al*. 1992). In this trial, there were no instances of silent seroconversion and no individual developed disseminated disease, so the only risk associated with watchful waiting was a 1 per cent risk of developing erythema migrans. This low risk was much less than the tick infection rate, which was approx. 20 per cent. A cost-effectiveness analysis, based on the assumption that one-third of infected individuals would present with later disease and require intravenous therapy, concluded that prophylactic therapy was only indicated if the risk of acquiring Lyme disease following a tick bite exceeded 3 per cent (Magid *et al*. 1992). The current evidence favours an expectant, watchful, waiting approach, and not the prophylactic administration of antibiotics.

Vaccine development

Work in progress holds promise for the development of a Lyme vaccine. Vaccination with recombinant polypeptides from outer surface proteins (particularly OspA) have been shown to protect mice from infection and to provide at least limited cross-strain protection (Fikrig *et al*. 1992). Large-scale trials in humans are currently under way. The ability to induce immunity in humans, the range of cross-strain protection provided, and the durability of immunity must be determined (Lovrich *et al*. 1994). The trials currently in progress should define the potential for disease prevention through vaccination.

Summary of practical guidelines for management of Lyme disease

A pragmatic approach to the treatment of Lyme disease requires confirmation of the diagnosis, an assessment of the extent of disease (clinical staging), administration of antibiotic therapy appropriate to the extent of disease, and careful post-treatment follow-up. The choice of specific antibiotic, route of administration, and duration of therapy all are contingent upon an assessment of extent of disease. Generally, early disease (limited to primary and secondary skin sites) is treated with shorter-course antibiotic therapy administered orally, and more extensive disease requires longer oral and intravenous therapy.

My approach to management is summarized in Box 1.

Box 1 Management of Lyme disease

1. Early Lyme disease with clinical evidence limited to mild to moderate systemic symptoms, and one or multiple lesions of erythema migrans.

Always

(a) confirm diagnosis by observation of definite erythema migrans lesion(s) in individuals with a positive epidemiological history;
(b) search for evidence of neurological dissemination with a careful history and neurological examination;
(c) administer a 21-day course of oral antibiotic selected from the list in Table 3;
(d) follow up the patient at the end of the antibiotic course to ensure complete resolution of signs and symptoms of Lyme disease;
(e) counsel the patient about the possibility of relapse, especially neurological.

Often

(a) confirm the diagnosis serologically with enzyme immunoassay and Western blot, especially if the epidemiological history is questionable or skin lesions are atypical or absent;
(b) assess at the end of therapy for any persistent symptoms: delayed resolution may occur, especially of systemic symptoms such as fatigue and fibromyalgia-type complaints.

Sometimes

(a) perform a lumbar puncture if headache, paraesthesiae or facial nerve palsy are present at the time of diagnosis—abnormal spinal fluid should be considered evidence of neurological spread of infection and indicates the need for intravenous antibiotics;
(b) perform an electrocardiogram if the patient has had complaints possibly indicative of heart block;
(c) perform post-treatment serological tests if the initial diagnosis was questionable and initial serology was not positive.

2. Disseminated Lyme disease with clinical evidence of dissemination to joints, nervous system heart or other major organs.

Always

(a) confirm the diagnosis serologically (enzyme immunoassay and Western blot) and rule out alternative diagnoses to explain the clinical presentation;
(b) characterize the extent of disease before instituting antibiotics—perform a lumbar puncture if neurological signs or symptoms are present and aspirate joint fluid if joint swelling is present;
(c) administer antibiotic therapy as outlined in Table 3;
(d) follow up the patient serially until all symptoms are resolved.

Often

(a) observe the patient for continued resolution of signs and symptoms after completion of a course of antibiotic therapy if all signs of inflammation are not resolved;
(b) repeat initial studies to monitor for improvement and or progression after therapy (electromyography, nerve conduction studies, joint-fluid analysis, lumbar puncture).

Sometimes

(a) analyse spinal or joint fluid by polymerase chain reaction if the diagnosis is unclear by clinical and routine serological criteria;
(b) extend or repeat the course of antibiotic therapy if there is clinical evidence suggesting recurrence of inflammation after an initial response.

Chapter References

Ackerman, R., Rehse-Kupper, B., Gollmer, E., and Schmidt, R.(1988). Chronic neurologic manifestations of erythema migrans borreliosis. *Annals of the New York Academy of Science*, **539**, 16–23.

Afzelius, A. (1910). Verhandlungen der Dermatologischen Gesellschaft zu Stockholm on October 29, 1909. *Archives of Dermatology and Syphilis*, **101**, 404.

Aguero-Rosenfeld, M. E., Nowakowski, J., McKenna, D.F., Carbonato, C.A., and Wormser, G.P. (1993). Serodiagnosis in early Lyme disease. *Journal of Clinical Microbiology*, **31**, 3090–5.

Anderson, J. F., Johnson, R.C., Magnarelli, L.A., and Hyde, F.W. (1985). Identification of endemic foci of Lyme disease: isolation of *Borrelia burgdorferi* from feral rodents and ticks (*Dermacentor variabilis*).

Anderson, J. F., Johnson, R.C., Magnarelli, L.A., and Hyde, F.W. (1986). Involvement of birds in the epidemiology of the Lyme disease agent *Borrelia burgdorferi*. *Infection and Immunity*, **51**, 394–6.

Åsbrink, E. and Hovmark, A. (1985). Successful cultivation of spirochetes from skin lesions of patients with erythema chronicum migrans Afzelius and acrodermatitis chronica atrophicans. *Acta Pathologica Microbiologica et Immunologica Scandinavica* (Sect. B), **93**, 161–3.

Åsbrink, E. and Olsson, I. (1985). Clinical manifestations of erythema chronicum migrans Afzelius in 161 patients. A comparison with Lyme disease. *Acta Dermatologica et Venereologica* (Stockholm), **65**, 43–52.

Åsbrink, E., Hovmark, A., and Hederstedt, B. (1985). Serological studies of erythema chronicum migrans Afzelius and acrodermatitis chronica atrophicans with indirect immunofluorescence and enzyme-linked immunosorbent assays. *Acta Dermatologica et Venereologica* (Stockholm), **65**, 509–14.

Bannwarth, A. (1944). Zur Klinik und Pathogenese der chronischen lymphozytaren Meningitis. *Archiv für Psychiatrie und Nervenkrankheiten*, **117**, 161–85.

Baranton, G., Postic, D., Saint Girons, I., *et al*. (1992). Delineation of *Borrelia burgdorferi* sensu stricto, *Borrelia garinii* sp. nov., and group VS461 associated with Lyme borreliosis. *International Journal of Systematic Bacteriology*, **42**, 378–83.

Barbour, A. G., Heiland, R.A., and Howe, T.R. (1985). Heterogeneity of major proteins in Lyme disease *Borrelia*: a molecular analysis of North American and European isolates. *Journal of Infectious Diseases*, **152**, 478–84.

Barbour, A. G. (1988). Plasmid analysis of *Borrelia burgdorferi*, the Lyme disease agent. *Journal of Clinical Microbiology*, **26**, 475–8.

Benach, J. L., Bosler, E.M., Hanrahan, J.P., *et al*. (1983). Spirochetes isolated from the blood of two patients with Lyme disease. *New England Journal of Medicine*, **308**, 740–2.

Berger, B. W. (1984). Erythema chronicum migrans of Lyme disease. *Archives of Dermatology*, **120**, 1017–21.

Berger, B. W., Kaplan, M.H., Rothenberg, I.R., and Barbour, A.G. (1985). Isolation and characterization of the Lyme disease spirochete from the skin of patients with erythema chronicum migrans. *Journal of the American Academy of Dermatology*, **3**, 444–9.

Bradley, J. F., Johnson, R.C., and Goodman, J.L. (1994). The persistence of spirochetal nucleic acids in active Lyme arthritis. *Annals of Internal Medicine*, **120**, 487–9.

Britton, M. C., Gardner, P., Kaufman, R.L., *et al*. (1993). Appropriateness of parenteral antibiotic treatment for patients with presumed Lyme disease. A joint statement of the American College of Rheumatology and the Council of the Infectious Diseases Society of America. *Annals of Internal Medicine*, **119**, 518.

Brown, R. N. and Lane, R. S. (1992). Lyme disease in California: a novel enzootic transmission cycle of *Borrelia burgdorferi*. *Science*, **256**, 1439–42.

Burgdorfer, W. (1989). Vector/host relationships of the Lyme disease spirochete, *Borrelia burgdorferi*. *Rheumatic Disease Clinics of North America*, **15**, 775–87.

Burgdorfer, W., Barbour, A.G., Hayes, S.F., *et al*. (1982). Lyme disease—a tick-borne spirochetosis? *Science*, **216**, 1317–19.

Burgdorfer, W., Lane, R.S., Barbour, A.G., Gresbrink, R.A., and Anderson, J.R. (1985). The western black-legged tick, *Ixodes pacificus*: a vector of *Borrelia burgdorferi*. *American Journal of Tropical Medicine and Hygiene*, **34**, 925–30.

Canica, M. M., Nato, F., du Merle, L., *et al*. (1993). Monoclonal antibodies for identification of *Borrelia afzelii* sp. nov. associated with late cutaneous manifestations of Lyme borreliosis. *Scandinavian Journal of Infectious Diseases*, **25**, 441–8.

Centers for Disease Control (1993). Ceftriaxone-associated biliary complications of treatment of suspected disseminated Lyme disease—New Jersey, 1990–1992. *Journal of the American Medical Association*, **269**, 979–80.

Clark, J. R., Carlson, R.D., Casaki, C.T., Pachner, A.R., and Steere, A.C. (1985). Facial paralysis in Lyme disease. *Laryngoscope*, **95**, 1341–5.

Comstock, L. E. and Thomas, D. D. (1989). Penetration of endothelial cell monolayers by *Borrelia burgdorferi*. *Infection and Immunity*, 1626–8.

Craft, J. E., Grodzicki, R.L., and Steere, A.C. (1984). The antibody response in Lyme disease: evaluation of diagnostic tests. *Journal of Infectious Diseases*, **149**, 789–95.

Craven, R. and Dennis, D. (ed.) (1993). *Lyme disease surveillance summary*. U.S. Dept. of Health and Human Services, Fort Collins.

Dattwyler, R. J., Volkman, D.J., Halperin, J.J., and Luft, B.J. (1988). Treatment of late Lyme borreliosis—randomized comparison of ceftriaxone and penicillin. *Lancet*, **i**, 1191–4.

Dattwyler, R. J., Volkman, D.J., Luft, B.J., *et al*. (1989). Dissociation of specific T- and B-lymphocyte responses to *Borrelia burgdorferi*. *New England Journal of Medicine*, **319**, 1441–6.

Dattwyler, R. J., Volkman, D.J., Conaty, S.M, Platkin, S.P., and Luft, B.J. (1990). Amoxycillin plus probenecid versus doxycycline for treatment of erythema migrans borreliosis. *Lancet*, **336**, 1404–6.

Dekonenko, E. J., Steere, A.C., Berardi, V.P., and Kravchuk, L.N. (1988). Lyme borreliosis in the Soviet Union: a cooperative US–USSR report. *Journal of Infectious Diseases*, **158**, 748–53.

de Koning, J., Hoogkamp-Korstanje, J.A.A., van der Linde, M.R., and Crijns, H.J.G.M.. (1989). Demonstration of spirochetes in cardiac biopsies of patients with Lyme disease. *Journal of Infectious Diseases*, **160**, 150–2.

Dorward, D. W. and Garon, C. F. (1990). DNA is packaged within membrane-derived vesicles of Gram-negative but not Gram-positive bacteria. *Applied Environmental Microbiology*, **56**, 1960–2.

Dotevall, L., Alestig, K., Hanner, P., Norfrans, G., and Hagberg, L. (1988). The use of doxycycline in nervous system *Borrelia burgdorferi* infection. *Scandinavian Journal of Infectious Diseases*, **53**, 74–9.

Dressler, F., Whalen, J.A., Reinhardt, B.N., and Steere, A.C. (1993). Western blotting in the serodiagnosis of Lyme disease. *Journal of Infectious Diseases*, **167**, 392–400.

Fikrig, E., Barthold, S.W., Kantor, F.S., and Flavell, R.A. (1992). Long-term protection of mice from Lyme disease by vaccination with OspA. *Infection and Immunity*, **60**, 773–7.

Garcia-Monco, J. C., Villar, B.F., Alen, J.C., and Benach, J.L. (1990). *Borrelia burgdorferi* in the central nervous system: experimental and clinical evidence for early invasion. *Journal of Infectious Diseases*, **161**, 1187–93.

Garin-Bujadoux, C. (1922). Paralysie par les tiques. Journal *de Medicine*, Lyon, **71**, 765–7.

Ginsberg, H.S. (1995). Vector management to reduce the risk of Lyme disease. In *: Ecology and environmental management of Lyme disease* (ed. M.L. Wilson and R.D. Deblinger), pp. 126–56. Rutgers University Press, Newark, NJ.

Halperin, J. J., Krupp, L.B., Golightly, M.G., and Volkman, D.J. (1986). Lyme borreliosis-associated encephalopathy. *Neurology*, **40**, 1340.

Halperin, J. J., Little, B.W., Coyle, P.K., and Dattwyler, R.J. (1987). Lyme disease: cause of a treatable peripheral neuropathy. *Neurology*, **37**, 1700–6.

Halperin, J. J., Luft, B.J., Anand, A.K., *et al*. (1989). Lyme neuroborreliosis: central nervous system manifestations. *Neurology*, **39**, 753–9.

Hanrahan, J. P., Benach, J.L., Coleman, J.L., *et al*. (1984). Incidence and cumulative frequency of endemic Lyme disease in a community. *Journal of Infectious Diseases*, **150**, 489–96.

Hardin, J. A., Steere, A.C., and Malawista, S.E. (1979*a*). Immune complexes and the evolution of Lyme arthritis: dissemination and localization of abnormal Clq binding activity. *New England Journal of Medicine*, **301**, 1358–63.

Hardin, J. A., Walker, L.C., Steere, A.C., *et al*. (1979*b*). Circulating immune complexes in Lyme arthritis: detection of the ^{125}I-Clq binding, Clq solid phase and Raji cell assays. *Journal of Clinical Investigation*, **63**, 468–77.

Hardin, J. A., Steere, A.C., and Malawista, S.E. (1984). The pathogenesis of arthritis in Lyme disease: humoral immune responses and the role of intra-articular immune complexes. *Yale Journal of Biological Medicine*, **57**, 589–93.

Hsu, V. M., Patella, S.J., and Sigal, L.H. (1993). 'Chronic Lyme disease' as the incorrect diagnosis in patients with fibromyalgia. *Arthritis and Rheumatism*, **36**, 1493–500.

Johnston, Y. E., Duray, P.H., Steere, A.C., *et al*. (1985). Lyme arthritis: spirochetes found in synovial microangiopathic lesions. *American Journal of Pathology*, **118**, 26–34.

Kalish, R. A., Leong, J.M., and Steere, A.C. (1993). Association of treatment-resistant chronic Lyme arthritis with HLA-DR4 and antibody reactivity to OspA and OspB of *Borrelia burgdorferi*. *Infection and Immunity*, **61**, 2774–9.

Kaufman, L. D., Gruber, B.L., Phillips, M.E., and Benach, J.L. (1989). Late cutaneous Lyme disease: acrodermatitis chronica atrophicans. *American Journal of Medicine*, **86**, 828–30.

Keller, T. L., Halperin, J.J., and Whitman, M. (1992). PCR detection of *Borrelia burgdorferi* DNA in cerebrospinal fluid of Lyme neuroborreliosis patients. *Neurology*, **42**, 32–42.

Klempner, M. S., Noring, R., and Rogers, R.A.(1993). Invasion of human skin fibroblasts by the Lyme disease spirochete, *Borrelia burgdorferi*. *Journal of Infectious Diseases*, **167**, 1074–81.

Kuiper, H., van Dam, A.P., Spanjaard, L., *et al*. (1994). Isolation of *Borrelia burgdorferi* from biopsy specimens taken from healthy-looking skin of patients with Lyme borreliosis. *Journal of Clinical Microbiology*, **32**, 715–20.

Lightfoot, R. W., Benjamin, J.L., Rahn, D.W., *et al*. (1993). Empiric parenteral antibiotic treatment of patients with fibromyalgia and fatigue and a positive serologic result for Lyme disease. *Annals of Internal Medicine*, **119**, 503–9.

Logigian, E. L., Kaplan, R.F., and Steere, A.C. (1990). 'Chronic neurologic manifestations of Lyme disease. *New England Journal of Medicine*, **323**, 1438–44.

Lovrich, S. D., Callister, S.M., Lim, L.C.L., DuChateau, B.K., and Schell, R.F. (1994). Seroprotective groups of Lyme borreliosis spirochetes from North America and Europe. *Journal of Infectious Diseases*, **170**, 115–21.

Luft, B. J., Steinman, C.R., Neimark, H.C., *et al*. (1992). Invasion of the central nervous system by *Borrelia burgdorferi* in acute disseminated infection. *Journal of the American Medical Association*, **267**, 1364–7.

Luft, B.J., Dattwyler, R.J., Johnson, R.C., *et al*. (1996). Azithromycin compared with amoxicillin in the treatment of erythema migrans. *Annals of Internal Medicine*, **124**, 885–91.

Luger, S. W. and Krauss, E. (1990). Serologic tests for Lyme disease: interlaboratory variability. *Archives of Internal Medicine*, **150**, 761–3.

McAlister, H. F., Klementovicz, P., Andrews, C., *et al*. (1989). Lyme carditis: an important cause of reversible heart block. *Annals of Internal Medicine*, **110**, 339–45.

Magid, D., Schwartz, B., Craft, J., and Schwartz, J.S. (1992). Prevention of Lyme disease after tick bites. *New England Journal of Medicine*, **327**, 534–41.

Magnarelli, L. A. Miller, J.N., Anderson, J.F., and Riviere, G.R.. (1990). Cross-reactivity of nonspecific treponemal antibody in serologic tests for Lyme disease. *Journal of Clinical Microbiology*, **28**, 1276–9.

Malawista, S. E., Barthold, S.W., and Persing, D.H. (1994). Fate of *Borrelia burgdorferi* DNA in tissues of infected mice after antibiotic treatment. *Journal of Infectious Diseases*, **170**, 1312–16.

Marcus, L. C., Steere, A.C., Duray, P.H., Anderson, A.E., and Mahoney, E.B. (1985). Fatal pancarditis in a patient with coexistent Lyme disease and babesiosis. Demonstration of spirochetes in the myocardium. *Annals of Internal Medicine*, **103**, 374–6.

Markowitz, L. E., Steere, A.C., Benach, J.L., Slade, J.D., and Broome, C.V. (1986). Lyme disease during pregnancy. Journal of the American Medical Association, **255**, 3394–6.

Marshall, W. F., Telford, S.R., Rys, P.N., *et al*. (1994). Detection of *Borrelia burgdorferi* DNA in museum specimens of *Peromyscus leucopus*. *Journal of Infectious Diseases*, **170**, 1027–32.

Mather, T. N., Wilson, M.L., Moore, S.I., Ribeiro, J.M.C., and Spielman, A. (1989). Comparing the relative potential of rodents as reservoirs of the Lyme disease spirochete (*Borrelia burgdorferi*). *American Journal of Epidemiology*, **130**, 143–50.

Maupin, G. O., Gage, K.L., Piesman, J., *et al*. (1994). Discovery of an enzootic cycle of *Borrelia burgdorferi* in *Neotoma mexicana and Ixodes spinipalpis* from Northern Colorado; an area where Lyme disease is nonendemic. *Journal of Infectious Diseases*, **170**, 636–43.

Montgomery, R. R., Nathanson, M.H., and Malawista, S.E. (1994). Fc and non-Fc mediated phagocytosis of *Borrelia burgdorferi* by macrophages. *Journal of Infectious Diseases*, **170**, 890–3.

Nadelman, R., Luger, S., Frank, E., and Wisniewski, M. (1992). Comparison of cefuroxime axetil and doxycycline in the treatment of early Lyme disease. *Annals of Internal Medicine*, **117**, 273–80.

Nocton, J. J., Dressler, F., Rutledge, B.J., *et al*. (1994). Detection of *Borrelia burgdorferi* DNA by polymerase chain reaction in synovial fluid from patients with Lyme arthritis. *New England Journal of Medicine*, **330**, 229–34.

Oliver, J. H., Owsley, M.R., Hutcheson, H.J., *et al*. (1993). Conspecificity of the ticks *Ixodes scapular* and *I. dammini* (Acari: Ixodidae). *Journal of Medical Entomology*, 30, 54–63.

Pachner, A. R. and Steere, A. C. (1985). The triad of neurologic manifestations of Lyme disease: meningitis, cranial neuritis, and radiculoneuritis. *Neurology*, **35**, 47–53.

Pachner, A. R., Steere, A.C., Sigal, L.H., and Johnson, C.J. (1985). Antigen-specific proliferation of CSF lymphocytes in Lyme disease. *Neurology*, **35**, 1642–4.

Pachner, A. R., Duray, P., and Steere, A.C. (1989). Central nervous system manifestations of Lyme disease. *Archives of Neurology*, **46**, 790–5.

Persing, D. H., Telford, S.R., Spielman, A., and Barthold, S.W. (1990). Detection of *Borrelia burgdorferi* infection in *Ixodes dammini* ticks by using the polymerase chain reaction. *Journal of Clinical Microbiology*, **28**, 566–72.

Persing, D. H., Rutledge, B.J., Rys, P.N., *et al*. (1994). Target imbalance: disparity of *Borrelia burgdorferi* genetic material in synovial fluid from Lyme arthritis patients. *Journal of Infectious Diseases*, **169**, 664–8.

Pfister, H-W., Preac-Mursic, V., Wilske, B., and Einhaup, K.M. (1989). Cefotazime versus penicillin G for acute neurologic manifestations in Lyme borreliosis: a prospective randomized study. *Archives of Neurology*, **46**, 1190–4.

Pfister, H-W., Preac-Mursic, V., Wilske, B., Einhaupl, K.-M., and Weinberger, K. (1990). Latent Lyme neuroborreliosis: presence of *Borrelia burgdorferi* in the central nervous system: experimental and clinical evidence for early invasion. *Journal of Infectious Diseases*, **161**, 1187–93.

Rahn, D. W. and Malawista, S. E. (1994). *Treatment of Lyme disease*. Mosby–Year Book Inc., St Louis.

Reik, L., Steere, A.C., Bartenhagen, N.H., Shope, R.E., and Malawista, S.E (1979). Neurologic abnormalities of Lyme disease. *Medicine*, **58**, 281–94.

Reik, L. J., Smith, L., Khan, A., and Nelson, W. (1985). Demyelinating encephalopathy in Lyme disease. *Neurology*, **35**, 267–9.

Schlesinger, P. A., Duray, P.H., Burke, B.A., and Steere, A.C. (1985). Maternal–fetal transmission of the Lyme disease spirochete, *Borrelia burgdorferi*. *Annals of Internal Medicine*, **103**, 67–8.

Schoen, R. T., Aversa, J.M., Rahn, D.W., and Steere, A.C. (1991). Treatment of refractory chronic Lyme arthritis with arthroscopic synovectomy. *Arthritis and Rheumatism*, **34**, 1056–9.

Schwartz, B. S., Goldstein, M.D., Ribeiro, J.M.C., Schultz, T.L., and Shahied, S.I. (1989). Antibody testing in Lyme disease. *Journal of the American Medical Association*, **262**, 3431–4.

Shapiro, E. D., Gerber, M.A., Holabird, N.B., *et al*. (1992). A controlled trial of antimicrobial prophylaxis for Lyme disease after deer-tick bites. *New England Journal of Medicine*, **327**, 1769–73.

Shrestha, M., Grodzicki, R.L., and Steere, A.C. (1985). Diagnosing early Lyme disease. *American Journal of Medicine*, **78**, 235–40.

Sigal, L. H. (1990). Summary of the first 100 patients seen at a Lyme disease referral center. *American Journal of Medicine*, **88**, 577–81.

Snydman, D. R., Schenkein, D.P., Berardi, V.P., Lastavica, C.C., and Pariser, K.M. (1986). *Borrelia burgdorferi* in joint fluid in chronic Lyme arthritis. *Annals of Internal Medicine*, **104**, 798–800.

Spielman, A., Wilson, M.L., Levine, J.F., and Piesman, J. (1985). Ecology of *Ixodes dammini*-borne human babesiosis and Lyme disease. *Annual Reviews of Entomology*, **30**, 439–60.

Stanek, G., Klein, J., Bittner, R., and Glogar, D. (1990). Isolation of *Borrelia burgdorferi* from the myocardium of a patient with long-standing cardiomyopathy. *New England Journal of Medicine*, **322**, 249–52.

Steere, A. C. and Malawista, S. E. (1979). Cases of Lyme disease in the United States: locations correlated with distribution of *Ixodes dammini*. *Annals of Internal Medicine*, **91**, 730–3.

Steere, A. C., Hardin, J.A., and Malawista, S.E., *et al*. (1977*a*). Erythema chronicum migrans and Lyme arthritis. Cryoimmunoglobulins and clinical activity of skin and joints. *Science*, **196**, 1121–2.

Steere, A. C., Malawista, S.E., Snydman, D.R., *et al*. (1977*b*). Lyme arthritis: an epidemic of oligoarticular arthritis in children and adults in three Connecticut communities. *Arthritis and Rheumatism*, **20**, 7–17.

Steere, A. C., Malawista, S.E., Hardin, J.A, *et al*. (1977*c*). Erythema chronicum migrans and Lyme arthritis: the enlarging clinical spectrum. *Annals of Internal Medicine*, **86**, 685–98.

Steere, A. C., Broderick, T.E., and Malawista, S.E. (1978). Erythema chronicum migrans and Lyme arthritis: epidemiologic evidence for a tick vector. *American Journal of Epidemiology*, **108**, 312–21.

Steere, A. C., Gibofsky, A., Pattarroyo, M.E., *et al*. (1979). Chronic Lyme arthritis: clinical and immunogenetic differentiation from rheumatoid arthritis. *Annals of Internal Medicine*, **90**, 286–91.

Steere, A. C., Batsford, W.P., Weinberg, M., *et al*. (1980). Lyme carditis: cardiac abnormalities of Lyme disease. *Annals of Internal Medicine*, **93**, 8–16.

Steere, A. C., Bartenhagen, N.H., Craft, J.E., *et al*. (1983*a*). The early clinical manifestations of Lyme disease. *Annals of Internal Medicine*, **99**, 76–82.

Steere, A. C., Hutchinson, G.J., Rahn, D.W., *et al*. (1983*b*). Treatment of the early manifestations of Lyme disease. *Annals of Internal Medicine*, **99**, 22–6.

Steere, A. C., Pachner, A.R., and Malawista, S.E. (1983*c*). Neurologic abnormalities of Lyme disease: successful treatment with high-dose intravenous penicillin. *Annals of Internal Medicine*, **99**, 767–72.

Steere, A. C., Grodzicki, R.L., Kornblatt, A.N., *et al*. (1983*d*). The spirochetal etiology of Lyme disease. *New England Journal of Medicine*, **308**, 733–40.

Steere, A. C., Green, J., Schoen, R.T., *et al*. (1985). Successful parenteral antibiotic therapy of established Lyme arthritis. *New England Journal of Medicine*, **312**, 869–74.

Steere, A. C., Taylor, E., Wilson, M.W., Levine, J.R., and Spielman, A.(1986). Longitudinal assessment of the clinical and epidemiological features of Lyme disease in a defined population. *Journal of Infectious Diseases*, **154**, 295–300.

Steere, A. C., *et al*. (1987). The clinical evolution of Lyme arthritis. *Annals of Internal Medicine*, **107**, 725–31.

Steere, A. C., Dwyer, E., and Winchester, R. (1990). Association of chronic Lyme arthritis with HLA-DR4 and HLA-DR2 alleles. *New England Journal of Medicine*, **323**, 219–23.

Steere, A. C., Taylor, E., McHugh, G.L., and Logigian, E.L. (1993). The overdiagnosis of Lyme Disease. Journal of the American Medical Association, **269**, 1812–16.

Steere, A. C., Levin, R.E., Molloy, P.J., *et al*. (1994). Treatment of Lyme arthritis. *Arthritis and Rheumatism*, **37**, 878–88.

Strobino, B. A. Williams, C.L., Abid, S., Chalson, R., and Spierling, P. (1993). 'Lyme disease and pregnancy outcome: a prospective study of two thousand prenatal patients. *American Journal of Obstetrics and Gynecology*, **169**, 367–74.

Thone, A. W. (1968). *Ixodes ricinus* and erythema chronicum migrans (Afzelius). *Dermatologica*, **137**, 57–60.

van Dam, A. P., Kuiper, H., Vos, K., *et al*. (1993). Different genospecies of *Borrelia burgdorferi* are associated with distinct clinical manifestations of Lyme borreliosis. *Clinical Infectious Diseases*, **17**, 708–17.

Wallis, R. C., Brown, S.E., Kloter, K.O., and Main, J. (1978). Erythema chronicum migrans and Lyme arthritis: field study of ticks. *American Journal of Epidemiology*, **108**, 322–7.

Weber, K., Schierz, G., Wilske, B., and Preac-Mursic, V. (1984). European erythema migrans disease and related disorders. *Yale Journal of Biological Medicine*, **57**, 463–71.

Weber, K., Bratzke, H.J., and Neubert, U. (1988). *Borrelia burgdorferi* in a newborn despite oral penicillin for Lyme borreliosis during pregnancy. *Pediatric Infectious Diseases* Journal, **7**, 286–9.

Weber, K., Burgdorfer, W., and Schierz, G. (1993). *Aspects of Lyme disease borreliosis*, Vol. 3, *Ultrastructure of* Borrelia burgdorferi, pp. 20–43. Springer, Berlin.

Wienecke, R., Zochling, N., Neubert, U., *et al*. (1994). Molecular subtyping of *Borrelia burgdorferi* in erythema migrans and acrodermatitis chronica atrophicans. *Journal of Investigative Dermatology*, **103**, 19–22.

Wilske, B., Preac-Mursic, V., Gobel, U.B., *et al*. (1993). An OspA serotyping system for *Borrelia burgdorferi* based on reactivity with monoclonal antibodies and OspA sequence analysis. *Journal of Clinical Microbiology*, 340–50.

Zoschke, D. C., Skemp, A.A., and Defosse, D.L. (1991). Lymphoproliferative responses to *Borrelia burgdorferi* in Lyme disease. *Annals of Internal Medicine*, **114**, 285–9.

5.3.5 Viral arthritis

Stanley J. Naides

Viruses may affect the joints by a number of mechanisms. The mechanisms employed vary with the infecting virus and are based on mode of tissue entry, tissue tropism, mechanisms of replication, direct viral effects on cellular functions, the ability to establish persistent infection, local immune response, expression of host-like antigens, ability to alter host antigens, host age and genetic makeup, and the infection history of the host. Several viruses directly infect the cells of the synovium. The mechanism of injury may be through lysis of target cells. The target cells may die by one of three mechanisms:

1. Viral infection may result in classic cell necrosis with karyorrhexis.
2. The virus may initiate the cellular machinery for programmed cell death or apoptosis.
3. The virus may express virally encoded antigens on the cell surface which elicit an immune response which targets the killing of virally-infected cells.

Direct infection may also result in non-lytic mechanisms of viral arthritis pathogenesis. Immune activation may occur by transactivation of host genes by viral gene products. The infected cell elicits an immune response with itself as a target or recruits other cytokine-responsive cells. Viral infection may lead to expression of viral antigens on the cell surface. Such antigens may be seen as foreign and elicit an immune response. Alternatively, molecular mimicry of host autoantigens may break immune tolerance resulting in generation of an autoimmune response. Immune complex disease may result when the humoral response generates sufficient antibody to cause deposition of immune complexes either locally, at the site of viral infection, or systemically with deposition of circulating immune complex in synovium.

Parvovirus B19

Human parvovirus B19 was first discovered serendipitously in 1975 (Pattison 1988). It is a member of the family Parvoviridae, consisting of the smallest known DNA viruses, and the genus erythrovirus, autonomously replicating in erythroid precursors. Numerous autonomous parvoviruses are known to infect mammalian animal species. However, these viruses are extremely species specific and are not known to cross species barriers. B19 is a non-enveloped, single-stranded DNA virus measuring approximately 23 nm in diameter. Although infection of other tissue types may occur, reproduction is usually not as brisk in cells other than erythroid progenitors.

Epidemiology

B19 infection is common and geographically widespread. Seroepidemiological studies of community outbreaks of B19 infection demonstrate that a large proportion of B19 infections remain asymptomatic (Mosley 1994) or present as undiagnosed, non-specific viral illnesses. Up to 60 per cent of the general population has serological evidence for past B19 infection (Anderson *et al*. 1986). Outbreaks of B19 infection occur in late winter and spring, although epidemics have also been reported in summer and autumn. Within a community, B19 outbreaks tend to cycle every 3 to 5 years, representing the period of time for a fresh cohort of susceptible children to enter the school system. Since the seroprevalence of anti-B19 IgG antibodies is only approximately 50 per cent in adults, these periodic outbreaks often involve susceptible adults as well. The risk of infection in adults may be as high as 50 per cent with multiple exposures. Workers in occupations with increased exposure to children, such as school teachers, day-care workers, and hospital personnel, have increased risk of infection (Bell *et al*. 1989; Gillespie *et al*. 1990). Sporadic cases occur between outbreaks. Transmission is presumed to be via respiratory tract secretions.

The incubation period between infection and onset of symptoms is 7 to 18 days. In human volunteer studies, introduction of B19 nasally was followed in 7 days by a flu-like illness associated with viraemia, viral shedding in nasal secretions, and a reticulocytosis. At approximately 11 days postinfection, an incipient anti-B19 IgM antibody response was associated with clearing of viraemia, cessation of nasal shedding of virus, and a second phase of clinical illness with rash, arthralgia, and arthritis. Onset of the anti-B19 IgG antibody response occurred almost concurrently with the IgM response (Anderson *et al*. 1985). In natural infections, the temporal distinction between the two phases of clinical illness is often blurred.

Clinical features

Since 1981, well-defined, clinical syndromes have been attributed to B19 infection. B19 is the cause of transient aplastic crisis in the setting of chronic haemolytic anaemia, such as sickle cell disease, hereditary spherocytosis, a- and b-thalassaemias, pyruvate kinase deficiency, glucose-6-phosphate dehydrogenase deficiency, pyrimidine 5'-nucleotidase deficiency, hereditary stomatocytosis, autoimmune haemolytic anaemia, and HEMPAS (hereditary erythrocytic multinuclearity associated with a positive acidified—HAMS—test) (Naides 1992). B19 is the aetiological agent of erythema infectiosum, or fifth disease, a common rash illness of children characterized by bright red 'slapped cheeks' and a macular, maculopapular, and occasionally vesicular or haemorrhagic, eruption on the torso and extremities (Fig. 1).

Infection in children may be asymptomatic, and when symptoms do occur they tend to be mild and include sore throat, headache, fever, cough, anorexia, vomiting, diarrhoea, and arthralgia. Erythema infectiosum may also be seen in adults not previously infected. In adults, the rash tends to be more subtle and the bright red 'slapped cheeks' absent. A number of uncommon dermatological manifestations of B19 infection have been reported including a vesiculopustular eruption, purpura with or without thrombocytopenia, Henoch–Schönlein purpura, and a gloves and socks erythema (Mortimer *et al*. 1985; Feldmann *et al*. 1994).

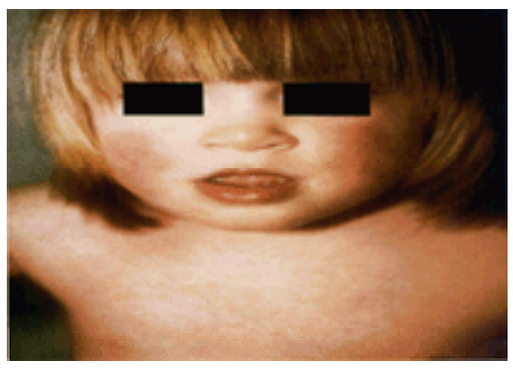


Fig. 1 Classic 'slapped cheeks' of a child with erythema infectiosum, or fifth disease, caused by parvovirus B19. A lacy, macular erythematous eruption is also present on the trunk but is not in focus. (Reproduced from Feder (1994), with permission.)

B19 infection may be associated with paraesthesias in the fingers. Rarely, progressive arm weakness may occur as may numbness of the toes. In such instances, nerve conduction studies may show mild slowing of nerve conduction velocities and decreased amplitudes of motor and sensory potentials (Faden *et al*. 1990).

B19 may cross the placenta to infect the fetus. Clinically affected fetuses develop hydrops fetalis on the basis of a B19-induced aplastic crisis, resulting in a high output cardiac failure, or viral cardiomyopathy, both resulting in hydrops fetalis. B19 has been reported to cause less commonly pancytopenia, isolated anaemia, thrombocytopenia, leucopenia, myocarditis, neuropathy, or hepatitis (Naides 1992; Luban 1994). Recent reports have suggested that B19 may be associated with vasculitis in some cases (Corman and Dolson 1992). Patients with congenital or acquired immune deficiencies, including prior chemotherapy for lymphoproliferative disorders, immunosuppressive therapy for transplantation, or human acquired immune deficiency syndrome (**AIDS**), may fail to clear B19 infection. Such individuals may have chronic or recurrent anaemia, thrombocytopenia, or leucopenia. B19 infection is the leading cause of pure red cell aplasia in patients with AIDS (Frickhofen *et al*. 1990).

Among immune competent children with B19 infection, arthralgia may occur in about 5 per cent and joint swelling in only approximately 3 per cent of children under 10 years of age. In adolescents, joint pain and swelling occurrs in about 12 per cent and 5 per cent, respectively. However, joint pain occurs in about 77 per cent and joint swelling in 60 per cent of adults 20 years of age or older (Ager *et al*. 1966). In adults, B19 infection may be associated with a severe, flu-like illness in which polyarthralgia and joint swelling are prominent. The distribution of involved joints is rheumatoid like with prominent symmetrical involvement of the metacarpophalangeal, proximal interphalangeal, wrist, knee, and ankle joints. Patients usually experience sudden onset polyarthralgia or polyarthritis. Onset of joint symptoms may or may not be preceded by a viral prodrome consisting of fever, malaise, chills, and myalgias. Most present with acute, moderately severe, symmetrical polyarthritis that usually starts in the hands or knees and within 24 to 48 h spreads to include the wrists, ankles, feet, elbows, and shoulders. Spinal involvement is uncommon. Joint symptoms in adults are usually self limited but a minority of adults may have symptoms for prolonged periods of time. Chronic symptoms fall into one of two patterns. Approximately two-thirds of patients have continuous symptoms of morning stiffness and arthralgia with intermittent flares. The remaining one-third of patients will be symptom free between flares. Morning stiffness is prominent. About one-half of the patients meet diagnostic criteria for rheumatoid arthritis. Rheumatoid factor may be present in low to moderate titre during the acute phase of infection but usually resolves. Anti-DNA, antilymphocyte, antinuclear, and antiphospholipid antibodies may also be found acutely. Joint erosions and rheumatoid nodules have not been recorded. Chronic B19 arthropathy may last for up to 8 years, the longest follow-up to date. Several weeks after the initial infection, symptoms of acute synovitis tend to resolve. Pain remains a prominent feature in patients who continue to report morning stiffness. Approximately 12 per cent of patients presenting with 'early synovitis' have B19-induced, rheumatoid-like arthropathy, the majority of whom are women (Naides *et al*. 1990). Adults usually lack the classic 'slapped-cheek' rash seen in children.

The distribution of joint involvement in B19 arthropathy and its symmetry may suggest a diagnosis of a rheumatoid arthritis. About half of all patients with chronic B19 arthropathy meet the criteria of the American Rheumatism Association for a diagnosis of rheumatoid arthritis—morning stiffness which may last for more than an hour, symmetrical involvement, involvement of at least three joints, and involvement of the hand joints. Joint erosions and rheumatoid nodules are absent (Silman 1988). While an initial report suggested that chronic B19 arthropathy may be associated with HLA-DR4, as is seen in classic erosive rheumatoid arthritis, subsequent studies by the same group have demonstrated no increased association with DR4. The absence of rheumatoid nodules or joint destruction aids in the differential diagnosis of B19 arthropathy from classic, erosive rheumatoid arthritis (Naides *et al*. 1990).

Diagnosis

Diagnosis is based on laboratory confirmation in the appropriate clinical setting. A number of approaches and methodologies have been used in the laboratory to confirm B19 infection. Immune electron microscopy, detection of B19 DNA during viraemia, and detection of anti-B19 IgM antibody may be used. However, the most useful modality in the rheumatology clinic is the IgM serology because patients usually have anti-B19 IgM antibodies and have begun to clear viraemia at the time of presentation with polyarthralgia/polyarthritis. Both radioimmunoassays (RIA) and enzyme-linked immunoabsorbent assays (ELISA) have been used to detect B19 antigen and specific antibody to B19 capsid (Cohen *et al*. 1983; Anderson *et al*. 1986; Bell *et al*. 1989; Naides *et al*. 1990). A number of laboratories are developing recombinant B19 antigens for B19 diagnosis in response to the difficulty in obtaining B19 viraemic serum to use as an antigen source.

The anti-B19 IgM antibody response is usually positive for at least 2 months following onset of joint symptoms, but may wane shortly thereafter. However, the IgM antibody may be detectable in occasional patients for 6 months or longer. Because of the high seroprevalence of anti-B19 IgG in the adult population, detection of anti-B19 IgG antibody shortly after presentation of acute-onset joint symptoms in a patient in the absence of anti-B19 IgM suggests past B19 infection and other diagnoses should be pursued. Failure to obtain B19 serological testing at presentation may leave the diagnostic IgM antibody response undetected and result in failure to diagnose B19 arthropathy in those patients in whom joint symptoms persist.

Pathogenesis

Anti-B19 IgM antibody and acute phase IgG antibody (less than 1 week postinoculation) recognize determinants on the major capsid protein, VP2. In convalescent serum, anti-B19 IgG antibody recognizes determinants on the minor capsid protein, VP1 structural protein (Kurtzman *et al*. 1989). B19 VP1 and VP2 are products of alternate transcription of the same open reading frame and VP1 contains an additional 227 N-terminal amino acids not present in VP2. VP1 therefore contains unique determinants not present in the truncated form represented by VP2; these determinants may be in the unique non-overlapping N-terminal region or, alternatively, represent conformational differences in the sequences shared between the two proteins. Western blot analysis of serum from individuals with congenital immune deficiency, prior chemotherapy, or AIDS demonstrated the absence of convalescent anti-B19 IgG antibodies directed against VP1. These sera were unable to neutralize B19 virus in experimental bone marrow culture systems (Kurtzman *et al*. 1989; Sato *et al*. 1991). In the absence of neutralizing antibodies to B19, B19 persists in the bone marrow and may cause chronic or intermittent suppression of one or more haematopoietic lineages.

Management

There is no specific vaccine or treatment for B19 infection at this time. Neutralizing activity to B19 is found in commercially available pooled immunoglobulin since seroprevalence of anti-B19 IgG antibodies in the adult population is approximately 50 per cent (Anderson *et al*. 1986). Intravenous immunoglobulin has been successful in the treatment of bone marrow suppression and B19 persistence in immunocompromised patients. However, this may not be applicable to chronic arthropathy patients. Treatment is symptomatic with non-steroidal anti-inflammatory agents (Naides *et al*. 1990).

Rubella virus

Rubella virus is the sole member of the genus rubivirus in the Togaviridae family of enveloped RNA viruses. The spherical rubella virion measures 50 to 70 nm in diameter with a 30 nm dense core. An envelope is acquired by budding at vesicles or the cell surface. Spike-like projections on the envelope measuring 5 to 6 nm contain haemagglutinin activity that is detected by agglutination of erythrocytes from a variety of animal species (Frey 1994).

Epidemiology

Rubella host range is restricted to humans. Like B19 infection, transmission is by nasopharyngeal secretions with peak incidence in late winter and spring. Widespread rubella vaccination altered the epidemiology of rubella infection, which had occurred in 6 to 9 year cycles prior to vaccination. Most cases were in children. Now the age profile has shifted toward young adults whose risk of infection of 10 to 20 per cent is comparable to that during the prevaccine era. Recent rubella outbreaks in college students and in adults underscores the public health need for maintaining vaccination programmes.

Incubation time from infection to onset of the rash is 14 to 21 days. Viraemia occurs 6 to 7 days before eruption, peaks immediately prior to eruption, and clears within 48 h of the rash. Virus shedding in nasopharyngeal secretions may be detected from 7 days before and until 14 days after eruption, but is maximal just before onset of the rash until 5 to 6 days posteruption (Wolinsky 1990).

Clinical features

The spectrum of clinical disease in children and adults ranges from asymptomatic infection to a classic syndrome of low-grade fever, rash, coryza, malaise, and prominent posterior cervical, postauricular, and occipital lymphadenopathy. Constitutional symptoms may precede the skin eruption by 5 days. The eruption may vary during a brief 2 to 3 day period, starting as a morbilliform facial eruption before spreading to the torso and upper, then lower, extremities. The eruption may coalesce on the face and clear as the extremities become involved. Alternatively, the eruption may be limited to a transient blush.

Joint complaints are common in adult infection, especially in women. Joint symptoms may occur 1 week before or after onset of the rash. Joint involvement is usually symmetrical and may be migratory, resolving over a few days to 2 weeks. Arthralgias are more common than frank arthritis. Stiffness is prominent. The metacarpophalangeal and proximal interphalangeal joints of the hands, the knees, wrists, ankles, and elbows are most frequently involved. Periarthritis, tenosynovitis, and carpal tunnel syndrome may be seen. In some patients, symptoms may persist for several months or years (Smith *et al*. 1987; Ueno 1994).

Live attenuated vaccines have been employed in rubella vaccination with a high frequency of postvaccination arthralgia, myalgia, arthritis, and paraesthesias. The HPV77/DK12 strain is the most arthritogenic of the vaccine strains that have been available. The pattern of joint involvement is similar to natural infection. Arthritis usually occurs 2 weeks postinoculation and lasts less than a week. However, symptoms may persist in some patients for more than a year. The currently used vaccine RA27/3 may cause postvaccination joint symptoms in as many as 15 per cent or more of recipients (Howson and Fineberg 1992; Mitchell *et al*. 1993).

In children, two syndromes of rheumatological interest may occur. In the 'arm syndrome,' a brachial radiculoneuropathy causes arm and hand pain and dysaesthesias that are worse at night. The 'catcher's crouch' syndrome is a lumbar radiculoneuropathy characterized by popliteal fossa pain on arising in the morning. Those affected assume a 'catcher's crouch' position. The pain gradually decreases through the day. Both syndromes occur 1 to 2 months postvaccination. The initial episode may last up to 2 months but recurrences are usually shorter in duration. Episodes of 'arm syndrome' and 'catcher's crouch syndrome' may recur for up to 1 year but there is no permanent damage (Schaffner *et al*. 1974).

Diagnosis

Rubella is readily cultured from tissues and body fluids including throat swabs. Virus is detected in either direct assays of cytopathic effects in tissue culture or in an indirect assay of interference of enterovirus growth in primary African green monkey kidney cell culture. Detection of antirubella IgM antibody or anti-IgG antibody seroconversion is diagnostic of rubella infection. Antirubella IgM and IgG are usually present at the onset of joint symptoms. IgM antibody peaks 8 to 21 days after symptoms then decreases over the next 4 to 5 weeks to undetectable levels in most patients. Therefore, detection of antirubella IgM indicates recent infection, usually in the last 1 to 2 months. Since antirubella IgG rises rapidly over a period of 7 to 21 days after the onset of symptoms, a diagnosis of rubella infection based on IgG serology can only be made with paired acute and convalescent sera. The presence of IgG in a single serum sample only documents immunity (Meurman 1978).

Pathogenesis

Failure to mount an adequate immune response to specific epitopes may allow rubella virus to persist in patients with rubella arthritis. Virus may be detected in synovial fluid during arthritis flares and in lymphocytes years after symptom resolution. Onset of rash and arthritis is coincident with the appearance of antibodies, including neutralizing antibodies to whole virus suggesting a role for antibody or immune complexes in the synovitis (Wolinsky 1990).

Management

Non-steroidal anti-inflammatory agents may be used for symptom control. Low to moderate doses of steroids may be needed to control symptoms and viraemia (Mitchell *et al*. 1993).

Hepatitis B virus

Hepatitis B virus (**HBV**) is a member of the family Hepadnaviridae, genus orthohepadnavirus. HBV is an enveloped, double-stranded DNA, icosahedral virus measuring 42 nm in diameter (Hollinger 1990; Seeger 1994).

Epidemiology

HBV is transmitted by the parenteral and sexual routes. HBV infection occurs worldwide, but prevalence of hepatitis B surface antigen (Australian antigen) is higher in Asia, the Middle East, and sub-Sahara Africa. The prevalence in China may be as high as 10 per cent compared to 0.01 per cent in the United States. There is no known seasonality to primary HBV infections. Most acute infections in endemic regions occur at an early age with many acquired perinatally from infected mothers and it is usually asymptomatic. Incidence of infection in children may be as high as 5 per cent annually, with gradual decline of carriage rates and specific antibody with advanced age. In the west, most infections are acquired during adulthood during sexual or needle exposures. Adult infection is more often associated with acute hepatitis and 5 to 10 per cent of those with hepatitis develop persistent infection. In endemic regions, HBV is a common cause of chronic liver disease and a leading cause of hepatocellular carcinoma (Robinson 1994).

Clinical features

The incubation period from infection to hepatitis is usually 45 to 120 days. A preicteric prodromal period, lasting several days to a month, may be associated with fever, myalgia, malaise, anorexia, nausea, and vomiting. HBV infection may cause an immune-complex-mediated arthritis during this period. Significant viraemia occurs early in infection; soluble immune complexes with circulating hepatitis B virus surface antigen (**HBsAg**) are formed as antihepatitis B surface antigen antibodies (**HBsAb**) are produced. Arthritis onset is usually sudden and often severe. Joint involvement is usually symmetrical with simultaneous involvement of several joints at onset, but arthritis may be migratory or additive. The joints of the hands and knees are most often affected, but wrists, ankles, elbows, shoulders, and other large joints may be involved as well. Fusiform swelling may be seen in the small joints of the hand. Morning stiffness is common. Arthritis and urticaria may precede jaundice by days to weeks and may persist several weeks after jaundice. However, arthritis and rash usually subside soon after the onset of clinical jaundice. While arthritis is usually limited to the preicteric prodrome, those patients who develop chronic active hepatitis or chronic HBV viraemia may have recurrent arthralgias or arthritis. Polyarteritis nodosa is frequently associated with chronic hepatitis B viraemia (Guillevin *et al*. 1995).

Diagnosis

Urticaria in the presence of polyarthritis should raise the possibility of HBV infection. Acute hepatitis may be asymptomatic but elevated bilirubin and transaminases are usually present when the arthritis appears. Examination of joint fluid is not diagnostic. At the time of arthritis onset, peak levels of serum HBsAg are detectable. Virions, viral DNA, polymerase, and hepatitis Be antigen may be detectable in serum. Antihepatitis B core antigen IgM antibodies are present and indicate acute HBV infection as opposed to past or chronic infection (Hoofnagle 1981).

Pathogenesis

HBV arthritis is thought to be mediated by immune complex deposition in the synovium. Immune complexes containing HBsAg, antibody, and complement components may be detected.

Management

Management is limited to supportive measures including non-steroidal anti-inflammatory agents.

Hepatitis C virus

Hepatitis C virus (**HCV**) is a member of the family Flaviviridae. HCV is an enveloped, single-stranded RNA, spherical virus measuring 38 to 50 nm in diameter (Purcell 1994).

Epidemiology

HCV is distributed worldwide. Using current diagnostic tools, seroprevalence is less than 1 per cent in developed western countries but is higher in Africa and Asia where it may cause a quarter of acute and chronic hepatitis. In Japan, this figure may reach 50 per cent. HCV is transmitted by the parenteral and sexual routes, although the latter is uncommon. HCV is responsible for 95 per cent of post-transfusion hepatitis in countries routinely screening donated blood for HBV. More than half of all cases of non-A, non-B hepatitis are attributable to HCV infection (Bhandari and Wright 1995). HCV genotypic variants have been described and these differ in their pathogenicity, including severity of disease and response to a interferon. To date, 11 HCV subtypes have been delineated (Bhandari and Wright 1995).

Clinical features

Acute HCV infection is usually benign. Up to 80 per cent of post-transfusion infections are anicteric and asymptomatic. Liver enzyme elevations are usually minimal, when present. Community-acquired cases present because of more symptomatic illness in which significant enzyme elevations occur. However, fulminant HCV hepatitis is rare. HCV is strongly associated with HBV-negative hepatocellular carcinoma, especially in Africa and Japan.

Acute onset polyarthritis in a rheumatoid distribution, including the small joints of the hand, wrists, shoulders, knees, and hips, may occur in acute HCV infection (Siegel *et al*. 1993). Hepatitis C virus is often associated with type II cryoglobulinaemia. It may present as essential mixed cryoglobulinaemia—a triad of arthritis, palpable purpura, and cryoglobulinaemia. Indeed, a majority of patients with essential mixed cryoglobulinaemia have HCV infection. HCV infection is also seen in non-essential secondary cryoglobulinaemia, although less commonly. The presence of anti-HCV antibodies in essential mixed cryoglobulinaemia is associated with more severe cutaneous involvement, for example Raynaud's phenomena, purpura, livedo, distal ulcers, and gangrene. HCV RNA may be found in 75 per cent of cryoprecipitates from patients with essential mixed cryoglobulinaemia and anti-HCV antibodies (Munoz-Fernandez *et al*. 1994).

Diagnosis

Serological tests utilize an array of antigens in an enzyme immunoassay while a recombinant strip immunoblot assay (**RIBA**) is confirmatory. Second generation RIBA-2 tests for reactivity to four viral antigens; c33c, c22–3, c100–3, and 5–1–1. A positive RIBA-2, especially to c33c and c22–3, is a sensitive assay of HCV infection (Van der Poel 1994). C33c positivity is associated with viraemia. A minority of patients may have HCV RNA detectable by polymerase chain reaction amplification methods in the absence of a positive serology.

Pathogenesis

Chronic HCV infection leads to cirrhosis, end stage liver failure, and hepatocellular carcinoma but the frequency of these sequelae and the mechanisms by which they occur are not known. HCV infection persists despite vigorous antibody response to an array of viral epitopes. A high rate of mutation in the envelope protein is responsible for emergence of neutralization escape mutants and quasispecies (Shimizu *et al*. 1994). Why HCV elicits cryoglobulins remains to be determined.

Management

Interferon a has been shown to be efficacious in the treatment of chronic HCV hepatitis and HCV-associated cryoglobulinaemia. Interferon a2b at a dose of three million units thrice weekly for 6 months suppresses viral titres and ameliorates clinical disease in about half of patients (Jenkins *et al*. 1996). Those with cryoglobulinaemia failing interferon therapy require immunosuppressive therapy. Relapse after completion of the initial course of therapy is common. There is controversy as to whether interferon therapy precipitates autoimmune disease such as autoimmune thyroiditis.

Retroviruses

Human immunodeficiency virus (see Chapter 5.3.6)

Several musculoskeletal syndromes have been described in human immunodeficiency virus (**HIV**) infected patients (Calabrese 1993). Whether reactive arthritis, Reiter's syndrome, and psoriatic arthritis are more prevalent in HIV-infected populations remains somewhat controversial. The incidence and prevalence of these rheumatic diseases may vary between populations studied and may depend on geography, mode of HIV transmission, exposure to different infectious agents, racial and ethnic makeup, risk behaviours, and patient ascertainment (Berman *et al*. 1991). Reiter's syndrome may have a prevalence as high as 11 per cent in some HIV-infected populations. These patients differ from 'idiopathic' Reiter's syndrome patients in that they do not have sacroiliitis or anterior uveitis, nor do they present with the classic triad of arthritis, urethritis, and uveitis. The prevalence of HLA-B27 positivity appears to be lower in the HIV-infected patients as compared to non-HIV associated Reiter's syndrome. In Zimbabwe, where the route of HIV transmission is predominantly heterosexual, approximately 40 per cent of HIV patients with joint symptoms have Reiter's syndrome, and another 40 per cent have a pauciarticular presentation without extra-articular features characteristic of Reiter's syndrome (Davis and Stein 1991). In the United States, psoriatic arthritis limited to a pattern of asymmetric oligoarthritis may be seen in as many as a third of HIV-infected patients with psoriasis, but the overall incidence of psoriasis does not appear to be significantly increased. Whether the different patterns of rheumatic disease expression are attributable to HIV infection itself or coinfection with other agents remains controversial. The caprine arthritis–encephalitis virus, a goat retrovirus, causes an inflammatory destructive arthritis and lends support to the notion that HIV infection alone may have musculoskeletal manifestations.

Initial HIV infection may be associated with a transient flu-like illness with arthralgias. Later, three pain syndromes not associated with synovitis may be seen. The concurrence of rheumatoid arthritis and HIV is thought to be very rare. An acute symmetrical polyarthritis involving the small joints of the hands and the wrists has been described in four patients but three had periosteal new bone formation about the involved joints, a feature not seen in rheumatoid arthritis. A subacute oligoarticular arthritis, primarily of the knees and ankles, may cause severe arthralgia and disability but is transient, peaks in intensity within 1 to 6 weeks, and responds to non-steroidal anti-inflammatory agents. The synovial fluid is non-inflammatory. Mononuclear cell infiltrates may be seen in the synovium of the involved joints. As many as 10 per cent of HIV-infected patients may experience 'painful articular syndrome' characterized by intermittent severe joint pain predominantly of the shoulders, elbows, and knees which lasts about a day. The pain may be incapacitating and require short-term narcotic analgesics. Fibromyalgia has been reported in HIV-infected patients with a prevalence as high as 29 per cent in one series. The role of HIV and other potential agents in these pain syndromes remains to be clarified (Calabrese 1993).

Human T lymphocyte leukaemia virus 1

Human T lymphocyte leukaemia virus 1 (**HTLV**) is endemic in Japan where it has been observed to be associated with oligoarthritis and a nodular rash. The patients have positive serology for anti-HTLV antibodies. Type C viral particles are seen in skin lesions. The presence of atypical synovial cells with lobulated nuclei and T cell synovial infiltrates suggests direct involvement of the synovial tissue by the leukaemic process (Nishioka *et al*. 1993).

Alphaviruses

Chikungunya virus

Chikungunya virus was originally isolated during an epidemic of febrile arthritis in Tanzania in 1952 to 1953. The local tribal word, Chikungunya, 'that which twists or bends up', was applied to the virus and the disease. Retrospectively, it is likely that similar epidemics occurred in Indonesia, Africa, India, Asia, and possibly the southern United States from 1779 to 1828 (Ross 1956; Peters and Dalrymple 1990). Humans are the major reservoir for Chikungunya virus which is transmitted by *Aedes* mosquitoes. The reinfestation of *Aedes aegypti* and the introduction of *Aedes albopictus* into the western hemisphere raises the spectre of an expanded geographic distribution.

Epidemiology

It is transmitted from its reservoir hosts (baboons, monkeys, and, in Senegal, *Scotophilus* bat species) to man by *Aedes* mosquitoes in south and west central Africa, Thailand, Vietnam, and India. *Mansonia africana* and mosquitoes from other genera may also act as a vector (Jupp and McIntosh 1990). In a 1964 epidemic in Bangkok, Thailand, an estimated 40 000 patients out of an urban population of two million were infected (Halstead *et al*. 1969a). Thirty-one per cent of the prospectively studied cohort seroconverted to Chikungunya virus antibody positivity. Communities, particularly urban centres, that have not seen Chikungunya fever, either endemically or epidemically, in a long period of time and that have a number of school age children who have not experienced the virus are at risk for significant outbreaks.

Clinical features

Chikungunya fever has an explosive onset associated with fever and severe arthralgia (Tesh 1982). Constitutional symptoms and rash follow an illness that lasts from 1 to 7 days. The incubation period is usually 2 to 3 days but ranges from 1 to 12 days. Fever elevations occur quickly, reaching 39 to 40°C, and are accompanied by rigors. The acute illness may last 2 to 3 days with a range of 1 to 7 days. Following the acute illness, the fever may resolve for 1 to 2 days before recrudescence. Polyarthralgia is migratory and predominantly affects the small joints of the hands, wrists, feet, and ankles with less prominent involvement of the large joints. Previously injured joints may be more severely affected. Stiffness and swelling may occur but large effusions are uncommon. Severe cases may have persistence of symptoms for months before resolution. Approximately 10 per cent of patients will have joint symptoms at 1 year post infection. Generalized myalgia and back and shoulder pain are common. Skin eruption is characterized by facial and neck flushing followed by macular or maculopapular eruption beginning 1 to 10 days after illness onset. Typically, a rash occurs after 2 to 5 days and is associated with defervescence. The rash may last 1 to 5 days and may recur with fever. It is located on the torso, extremities, and occasionally the face, palms, and soles. It may be pruritic. In some patients, the affected skin desquamates (Halstead *et al*. 1969d; Peters and Dalrymple 1990).

Significant haemorrhage usually does not occur but isolated petechias and mucosal bleeding may occur. Suffusion of the conjunctiva is prominent. Sore throat, pharyngitis, headache, photophobia, retro-orbital pain, anorexia, nausea, vomiting, and abdominal pain may accompany the acute illness. Lymphadenopathy may be tender but is usually not massive. Symptoms in children tend to be milder (Halstead *et al*. 1969b; Halstead *et al*. 1969c). In symptomatic children, nausea and vomiting, pharyngitis, and facial flushing are prominent features but arthralgia, arthritis, and rash are uncommon. Children may present with a mild dengue-like haemorrhagic fever, headache, pharyngeal injection, vomiting, abdominal pain, constipation, diarrhoea, cough, or lymphadenopathy. Arthralgia and arthritis in children are milder and briefer in duration. A destructive arthropathy may occur in a few adult patients with chronic symptoms (Brighton and Simson 1984). Low titre rheumatoid factor may be found in those with long-standing symptoms.

This infection has not been studied intensively enough to allow conclusions regarding pathogenesis to be made. As noted above, few patients may go on to have chronic symptoms of arthralgia. Case reports would suggest that a few patients go on to have destructive lesions resulting from chronic disease manifestations.

Diagnosis

Chikungunya fever should be considered in any febrile patient resident in or returning from endemic areas. A history of epidemic occurrence should be sought. Mayaro virus, Ross River virus, rubella virus, parvovirus B19, and hepatitis B virus infections may present similarly. Synovial fluid shows decreased viscosity with poor mucin clot and 2000 to 5000 white cells/mm^3. Therefore the diagnosis depends on laboratory confirmation.

Virus may be isolated during days two to four. In some patients, viral antigen may be detected in acute sera by haemagglutination assay due to the intensity of the viraemia. Haemagglutination inhibition antibodies develop as viraemia is cleared. Complement fixation antibodies are positive by the third week and slowly decrease over the subsequent year. Neutralizing antibody production parallels haemagglutination inhibition activity. Chikungunya virus-specific IgM antibodies may be found for 6 months or longer (Nakitare *et al*. 1983).

Pathogenesis

Following mosquito bite, intense viraemia occurs within 48 h. Viraemia begins to wane around day three. The appearance of haemagglutination inhibition activity and neutralizing antibody is associated with clearing of the viraemia. Affected skin shows erythrocyte extravasation from superficial capillaries and perivascular cuffing. The virus absorbs to human platelets causing aggregation, suggesting a mechanism for bleeding. Synovitis in Chikungunya fever probably results from direct viral infection of synovium.

Management

Management for the patient is supportive. During the acute attack, range of motion exercises ameliorate stiffness. Non-steroidal anti-inflammatory agents are useful. However, chloroquine phosphate (250 mg/day) has been used when non-steroidal anti-inflammatory agents have failed (Brighton 1984).

O'nyong-nyong virus

O'nyong-nyong fever is closely related to Chikungunya virus. O'nyong-nyong virus was first described in the Acholi province of north-western Uganda in February, 1959. Within 2 years, it had spread through Uganda and the surrounding region, affecting two million people. Serologically determined attack rates ranged from 50 to 60 per cent with 9 to 78 per cent of infected individuals becoming symptomatic (Williams *et al*. 1962; Williams *et al*. 1965a; Williams *et al*. 1965b). Disease spread at a rate of 2 to 3 km daily. After the epidemic, the virus was not detected again until it was isolated from *Anopheles funestus* mosquitoes in Kenya in 1978. *Anopheles gambiae* also serves as a vector. Serological surveys indicate that O'nyong-nyong virus is endogenous. The non-human vertebrate reservoir for O'nyong-nyong virus is not known.

Clinical features

O'nyong-nyong fever is clinically similar to Chikungunya infections (Shore 1961). The name derives from the Acholi word meaning 'joint breaker'. The incubation period lasts at least 8 days and is followed by sudden-onset polyarthralgia/polyarthritis. Four days later, the appearance of skin eruptions is typically associated with improvement in joint symptoms. The eruption is uniform in nature and lasts 4 to 7 days before fading. The fever is less prominent but postcervical lymphadenopathy may be marked. Although residual joint pain often persists, there appears to be no long-term sequelae.

Diagnosis

Viral isolation by intracerebral injection into suckling mice produces runting, rash, and alopecia (Williams *et al*. 1962). Haemagglutination inhibition or complement

fixation tests identify the virus (Williams *et al*. 1962). The differential diagnosis is similar to that of Chikungunya fever. Mouse antisera raised against Chikungunya virus or O'nyong-nyong virus react equally well with O'nyong-nyong virus, but O'nyong-nyong antisera does not react well with Chikungunya virus. The mechanisms of O'nyong-nyong virus pathogenesis are unknown.

Management

Management is symptomatic. Patients recover without sequelae.

Igbo ora virus

Igbo ora virus is serologically similar to Chikungunya and O'nyong-nyong viruses (Moore *et al*. 1975). Infection was first observed in a single patient with fever, sore throat, and arthritis. In 1984, an epidemic of fever, myalgias, arthralgias, and skin eruption occurred in four villages in the Ivory Coast. Igbo ora was coined as 'the disease that breaks your wings'. The virus was isolated from *Anopheles funestus* and *Anopheles gambiae* mosquitoes and from affected individuals.

Ross River virus (epidemic polyarthritis)

Epidemics of fever and rash have been observed in Australia since 1928. Epidemics occurred among soldiers stationed in Australia during World War II. Isolation of Ross River virus from mosquitoes, its serological association with epidemic polyarthritis, and the isolation of the virus from epidemic polyarthritis patients in Australia confirmed Ross River virus is the aetiological agent of epidemic polyarthritis (Aaskov *et al*. 1985). Epidemic polyarthritis from Ross River virus has been seen in New South Wales. Antibodies to Ross River virus have been observed in the sera of endogenous populations in Papua New Guinea, West New Guinea, the Bismarck Archipelago, Rossel Island, and the Solomon Islands (Scrimgeour *et al*. 1987a; Scrimgeour *et al*. 1987b). From 1979 to 1980, a major epidemic of febrile polyarthritis occurred in the Fiji Islands, affecting over 40 000 individuals (Bennett *et al*. 1980). Serological surveys suggested that a low level of Ross River virus infection was present throughout the Fiji Islands before 1979 but that following the epidemic up to 90 per cent of the residents of some communities had antibody. A similar epidemic occurred in the Cook Islands early in 1980. Weber's line is a hypothetical line separating the Australian geographic zone from the Asiatic zone; west of Weber's line, antibodies to Ross River virus are not found (Peters and Dalrymple 1990).

In Australia, both endemic cases and epidemics occur in tropical and temperate regions. Significant numbers of cases are reported in Queensland and New South Wales, although cases and outbreaks are described in other regions as well. Seroprevalence may reach only 6 to 15 per cent in temperate coastal zones but it is 27 to 39 per cent in the plains of the Murray Valley river system (Boughton *et al*. 1984). High rain fall usually precedes epidemic periods, as this causes increased mosquito populations. Cases occur from the spring to the autumn.

Aedes vigilax is the major vector on the eastern coast of Australia where the mosquito breeds in salt marshes. *Aedes camptorhynchus* similarly breeds in salt marshes of southern Australia. *Culex annulirostris* is a fresh-water breeding vector. Other Australian *Aedes* species and *Mansonia uniformus* may also serve as vectors. Several mammalian species may serve as intermediate hosts, including domestic animals, rodents, and marsupials. In the Pacific islands outbreaks, *Aedes polynesiensis, Aedes aegypti, Aedes vigilax*, and *Culex annulirostris* may have contributed as well (Peters and Dalrymple 1990).

In Australia, infection rates range from 0.2 to 3.5 per cent per year. During epidemics in Fiji and New South Wales, the majority of those infected were symptomatic (Hawkes *et al*. 1985). While male and female infection rates were similar, there was a predominance of women in presenting cases. Children have a case to infection ratio lower than adults.

A newly described alphavirus in Australia, Barmah forest virus, may present in a fashion similar to epidemic febrile polyarthritis (Lindsay *et al*. 1995).

Clinical features

Arthralgias occur abruptly after a 7 to 11 day incubation period (Fraser 1986). A macular, papular, or maculopapular skin eruption, which may be pruritic, typically follows the onset of arthralgia by 1 to 2 days but in some patients may precede or follow joint symptoms by 11 days or 15 days, respectively. Occasionally, vesicles, papules, or petechias are seen. The trunk and extremities are typically involved although involvement of the palms, soles, and face may occur. The rash resolves by fading to a brownish discoloration or by desquamation. Despite its name, half of patients have no fever, and in those who do, modest fevers may last only 1 to 3 days. Headache, nausea, and myalgia are common. Mild photophobia, respiratory symptoms, and lymphadenopathy may occur.

A majority of patients have severe, incapacitating arthralgia. Joint distribution is often asymmetrical and migratory, with metacarpophalangeal and finger interphalangeal joints, wrists, knees, and ankles commonly involved. Shoulders, elbows, and toes may also be involved. Axial, hip, and temporomandibular involvement occasionally occurs. Arthralgias are worse in the morning and after periods of inactivity. Mild exercise tends to improve joint symptoms. One-third will have frank synovitis. Polyarticular swelling and tenosynovitis are common. As many as one-third have paraesthesias, palm, or sole pain. Some patients have classic carpal tunnel syndrome. Half of all patients are able to resume their activities of daily living within 4 weeks although residual polyarthralgia may be present. Joint symptoms recur but episodes of relapse gradually resolve. A few patients continue to have joint symptoms for up to 3 years (Fraser 1986; Peters and Dalrymple 1990).

Diagnosis

The diagnosis of Ross River virus infection should be considered in anyone with a febrile arthritis in the appropriate geographic setting. Acute rubella arthritis may present in a similar fashion although the signs and symptoms of an upper respiratory infection are more prominent in rubella. Patients may present without a rash. The differential diagnosis then would include early seronegative rheumatoid arthritis, systemic lupus erythematosus, parvovirus B19 infection, hepatitis B virus infection, hepatitis C virus infection, other alphavirus infections, Henoch–Schönlein purpura, and drug hypersensitivities. In those individuals who develop vesicles, differential from varicella or parvovirus B19 infection would need to be considered.

Synovial fluid cell counts range from 1500 to 13 800 cells/mm ^{3}. Monocytes and vacuolated macrophages dominate with few neutrophils. Virus has been isolated only from antibody negative sera. In the Australian epidemics prior to 1979, patients were antibody positive at the time of presentation. However, in the Pacific Island epidemics of 1979 to 1980, patients remained viraemic and serologically negative for up to a week following onset of the symptoms (Aaskov *et al*. 1981). Ross River virus antigen is detectable by fluorescent antibody staining of C6/36 cells inoculated with acute patient serum. Virus in serum is stable for up to a month at 0 to -10°C (Tesh *et al*. 1981).

Pathogenesis

Ross River viral antigen may be detected by specific immunofluorescence in monocytes and macrophages early but intact virus is not identifiable by electron microscopy or cell culture (Fraser *et al*. 1981). The dermis shows mild perivascular mononuclear cell, mostly T lymphocytic, infiltrate in both erythematous and purpuric eruptions. The purpuric form of eruption also shows extravasation of erythrocytes. Ross River virus antigen may be found in epithelial cells in the erythematous and purpuric lesions, and in the perivascular zone in the erythematous lesion. However, viral antigen was not found in the perivascular zone in the purpuric lesions (Fraser *et al*. 1983).

Management

Management of the acute infection is symptomatic. Aspirin or non-steroidal anti-inflammatory drugs provide relief for joint pain. Occasional patients may develop more persistent joint symptoms but full recovery is usual.

Sindbis virus

Sindbis virus is the prototype alphavirus used for molecular virology studies. It was isolated from *Culex* mosquitoes in the Egyptian village of the same name in 1952.

Epidemiology

Sindbis virus infection occurs in Sweden, Finland, and the neighbouring Karelian isthmus of Russia where it is known locally as Okelbo disease, Pogosta disease, or Karelian fever (Tesh 1982). *Aedes, Culex*, and *Culiseta* species transmit the virus to humans. Birds are an intermediate host (Niklasson 1988). Cases are confined to predominately forested areas and individuals involved in outdoor activities or occupations are at risk. Sindbis virus infection has also been reported from Uganda, South Africa, Zimbabwe, central Africa, and Australia, where sporadic cases and small outbreaks occur (Peters and Dalrymple 1990).

Clinical features

Skin eruption and arthralgia are the initial symptoms although one may precede the other by a few days. A fever may be present, although not high. Constitutional symptoms including headache, fatigue, malaise, nausea, vomiting, pharyngitis, and paraesthesias may be present but are usually not severe. The macular skin eruption typically begins on the torso, spreading to the arms and legs, palms, soles, and occasionally head. The eruption evolves to form papules which have a tendency to vesiculate. Vesiculation is particularly prominent on pressure points including the palms and soles. As the eruption fades, a brownish discoloration is left. Vesicles on the palms and soles may become haemorrhagic. The rash may recur during convalescence (Tesh 1982).

Arthralgia and arthritis may involve the small joints of the hands and feet, wrists, elbows, ankles, and knees. Occasionally, the axial skeleton becomes involved. Tendonitis is common, often involving the extensor tendons of the hand and the Achilles tendon. Non-erosive chronic arthropathy is common in both Swedish and Finnish reports, with up to one-third of patients having arthropathy two or more years after onset. A smaller number had symptoms for as long as 5 to 6 years (Niklasson *et al*. 1988).

Diagnosis

The diagnosis may be established by haemagglutination inhibition and complement fixation tests. Antibodies appear during the first week of illness.

Pathogenesis

Little is known about the pathogenesis of Sindbis virus disease. Virus has been isolated from a skin vesicle in the absence of viraemia. Skin lesions show perivascular oedema, haemorrhage, lymphocytic infiltrates, and areas of necrosis. Antiviral IgM may persist for years, raising the possibility that Sindbis virus arthritis is associated with viral persistence and direct viral effect on the synovium (Niklasson *et al*. 1988).

Management

Management is supportive.

Mayaro virus

Mayaro virus was first recognized in Trinidad in 1954 and has caused epidemics in Bolivia and Brazil. Mayaro virus has a monkey reservoir and is transmitted to man by *Haemogogus* mosquitoes feeding in the south American tropical rain forest. Mayaro virus was responsible for an outbreak in Belterra, Brazil in 1988, in which 800 out of 4000 exposed latex gatherers were infected with a clinical attack rate of 80 per cent. Illness was characterized by sudden onset of fever, headache, dizziness, chills, and arthralgias in the wrists, fingers, ankles, and toes. About 20 per cent had joint swelling. Unilateral, inguinal lymphadenopathy was seen in some patients. Leucopenia was common. Viraemia was present during the first 1 to 2 days of illness,. After 2 to 5 days, fever resolved but a maculopapular rash on the trunk and extremities appeared. The rash lasted about 3 days. Recovery was complete, although some patients had persistent arthralgias at 2 month follow-up (Hoch *et al*. 1981; LeDuc *et al*. 1981; Pinheiro *et al*. 1981). It is of interest that Mayaro virus has been isolated from a bird in Louisiana (Calisher *et al*. 1974).

Other viruses

Apart from specific viral infections noted above in which arthralgia and/or arthritis is typically a prominent feature, there are a host of commonly encountered viral syndromes in which joint involvement is occasionally seen. Children with varicella have been reported rarely to develop brief monoarticular or pauciatricular arthritis that is thought to be viral in origin. Adults who develop mumps occasionally have small or large joint synovitis lasting up to several weeks. Arthritis may precede or follow parotitis by up to 4 weeks.

Infection with adenovirus and coxsackieviruses A9, B2, B3, B4, and B6 have been associated with recurrent episodes of polyarthritis, pleuritis, myalgia, rash, pharyngitis, myocarditis, and leucocytosis. Epstein–Barr virus associated mononucleosis is frequently accompanied by polyarthralgia but occasional monoarticular knee arthritis occurs. Polyarthritis, fever, and myalgias due to echovirus 9 infection has been reported in a few cases. Arthritis associated with herpes simplex virus or cytomegalovirus infections are likewise rare. Herpes hominis occasionally causes arthritis of the knee known as herpes gladiatorum because it is seen in wrestlers. Vaccinia virus has been associated with postvaccination knee arthritis in only two reported cases.

Chapter References

Aaskov, J.G., Mataika, J.U., Lawrence, G.W., Rabukawaqa, V., Tucker, M.M., Miles, J.A., and Dalglish, D.A. (1981). An epidemic of Ross River virus infection in Fiji, 1979. *American Journal of Tropical Medicine and Hygine*, **30**, 1053–9.

Aaskov, J.G., Ross, P.V., Harper, J.J., and Donaldson, M.D. (1985). Isolation of Ross River virus from epidemic polyarthritis patients in Australia. *Australian Journal of Experimantal Biology and Medical Science*, **63**, 587–97.

Ager, E.A., Chin, T.D.Y., and Poland, J.D. (1966). Epidemic erythema infectiosum. *New England Journal of Medicine*, **275**, 1326–31.

Anderson, M.J., Higgins, P.G., Davis, L.R., Willman, J.S., Jones, S.E., Kidd, I.M., Pattison, J.R., and Tyrrell, D.A. (1985). Experimental parvoviral infection in humans. *Journal of Infectious Diseases*, **152**, 257–65.

Anderson, L.J., Tsou, R.A., Chorba, T.L., Wulff, H., Tattersall, P., and Mortimer, P.P. (1986). Detection of antibodies and antigens of human parvovirus B19 by enzyme-linked immunosorbant assay. *Journal of Clinical Microbiology*, **24**, 522–6.

Bell, L.M., Naides, S.J., Stoffman, P., Hodinka, R.L., and Plotkin, S.A. (1989). Human parvovirus B19 infection among hospital staff members after contact with infected patients. *New England Journal of Medicine*, **321**, 485–91.

Bennett, N.M., Cunningham, A.L., Fraser, J.R., and Speed, B.R. (1980). Epidemic polyarthritis acquired in Fiji. *Medical Journal of Australia*, **1**, 316–17.

Berman, A., Reboredo, G., Spindler, A., Lasala, M.E., Lopez, H., and Espinoza, L.R. (1991). Rheumatic manifestations in populations at risk for HIV infection: the added effect of HIV. *Journal of Rheumatology*, **18**, 1564–7.

Bhandari, B.N. and Wright, T.L. (1995). Hepatitis C: An overview. *Annual Review of Medicine*, **46**, 309–17.

Boughton, C.R., Hawkes, R.A., Naim, H.M., Wild, J., and Chapman, B. (1984). Arbovirus infections in humans in New South Wales. Seroepidemiology of the alphavirus group of togaviruses. *Medical Journal of Australia*, **141**, 700–4.

Brighton, S.W. (1984). Chloroquine phosphate treatment of chronic Chikungunya arthritis. An open pilot study. *South African Medical Journal*, **66**, 217–8.

Brighton, S.W. and Simson, I.W. (1984). A destructive arthropathy following Chikungunya virus arthritis–a possible association. *Clinical Rheumatology*, **3**, 253–8.

Calabrese, L.H. (1993). Human immunodeficiency virus (HIV) infection and arthritis. *Infection and Arthritis*, **19**, 477–88.

Calisher, C.H., Gutierrez, E., Maness, K.S., and Lord, R.D. (1974). Isolation of Mayaro virus from a migrating bird captured in Louisiana in 1967. *Bulletin of the Pan American Health Organization*, **8**,

243–8.

Chantler, J.K., Tingle, A.J., and Petty, R.E. (1985). Persistent rubella virus infection associated with chronic arthritis in children. *New England Journal of Medicine*, **313**, 1117–23.

Cohen, B.J., Mortimer, P.P., and Pereira, M.S. (1983). Diagnostic assays with monoclonal antibodies for the human serum parvovirus-like virus (SPLV). *Journal of Hygiene*, **91**, 113–30.

Corman, L.C. and Dolson, D.J. (1992). Polyarteritis nodosa and parvovirus B19 infection. *Lancet*, **339**, 491

Davis, P. and Stein, M. (1991). Human immunodeficiency virus-related connective tissue diseases: a Zimbabwean perspective. (Review). *Rheumatic Disease Clinics of North America*, **17**, 89–97.

Faden, H., Gary, G.W., Jr., and Korman, M. (1990). Numbness and tingling of fingers associated with parvovirus B19 infection. *Journal of Infectious Diseases*, **161**, 354–5.

Feder, H.M. Jr (1994). Fifth disease. *New England Journal of Medicine*, **331**, 1062.

Feldmann, R., Harms, M., and Saurat, J.-H. (1994). Papular-purpuric 'gloves and socks' syndrome: Not only parvovirus B19. *Dermatology*, **188**, 85–7.

Fraser, J.R.E. (1986). Epidemic polyarthritis and Ross River virus disease. *Clinics in Rheumatic Diseases*, **12**, 369–88.

Fraser, J.R., Cunningham, A.L., Clarris, B.J., Aaskov, J.G., and Leach, R. (1981). Cytology of synovial effusions in epidemic polyarthritis. *Australian and New Zealand Journal of Medicine*, **11**, 168–73.

Fraser, J.R., Ratnamohan, V.M., Dowling, J.P., Becker, G.J., and Varigos, G.A. (1983). The exanthem of Ross River virus infection: histology, location of virus antigen and nature of inflammatory infiltrate. *Journal of Clinical Pathology*, **36**, 1256–63.

Frey, T.K. (1994). Molecular biology of rubella virus. *Advances in Virus Research*, **44**, 69–160.

Frickhofen, N., Abkowitz, J.L., Safford, M., *et al*. (1990). Persistent B19 parvovirus infection in patients infected with human immunodeficiency virus type 1 (HIV-1): a treatable cause of anemia in AIDS. *Annals of Internal Medicine*, **113**, 926–33.

Gillespie, S.M., Cartter, M.L., Asch, S., Rokos, J.B., Gary, G.W., Tsou, C.J., Hall, D.B., Anderson, L.J., and Hurwitz, E.S. (1990). Occupational risk of human parvovirus B19 infection for school and day-care personnel during an outbreak of erythema infectiosum. *Journal of the American Medical Association*, **263**, 2061–5.

Guillevin, L., Lhote, F., Cohen, P., Sauvaget, F., Jarrousse, B., Lortholary, O., Noël, L.H., and Trépo, C. (1995). Polyarteritis nodosa related to hepatitis B virus—a prospective study with long-term observation of 41 patients. *Medicine (Baltimore)*, **74**, 238–53.

Halstead, S.B., Nimmannitya, S., and Margiotta, M.R. (1969 *a*). Dengue and chikungunya virus infection in man in Thailand, 1962–1964. II. Observations on disease in outpatients. *American Journal of Tropical Medicine and Hygiene* **18**, 972–83.

Halstead, S.B., Scanlon, J.E., Umpaivit, P., and Udomsakdi, S. (1969 *b*). Dengue and chikungunya virus infection in man in Thailand, 1962–1964. IV. Epidemiologic studies in the Bangkok metropolitan area. *American Journal of Tropical Medicine and Hygiene* **18**, 997–1021.

Halstead, S.B., Udomsakdi, S., Scanlon, J.E., and Rohitayodhin, S. (1969 *c*). Dengue and chikungunya virus infection in man in Thailand, 1962–1964. V. Epidemiologic observations outside Bangkok. *American Journal of Tropical Medicine and Hygiene*, **18**, 1022–33.

Halstead, S.B., Udomsakdi, S., Singharaj, P., and Nisalak, A. (1969 *d*). Dengue and chikungunya virus infection in man in Thailand, 1962–1964. III. Clinical, epidemiologic, and virologic observations on disease in non-indigenous white persons. *American Journal of Tropical Medicine and Hygiene*, **18**, 984–96.

Hoch, A.L., Peterson, N.E., LeDuc, J.W., and Pinheiro, F.P. (1981). An outbreak of Mayaro virus disease in Belterra, Brazil. III. Entomological and ecological studies. *American Journal of Tropical Medicine and Hygiene*, **30**, 689–98.

Hollinger, F.B. (1990). Hepatitis B virus. In *Fields virology* (2nd edn) (eds B.N. Fields, D.M. Knipe, R.M. Chanock, M.S. Hirsch, J.L. Melnick, T.P. Monath, and B. Roizman), pp. 2171–236. Raven Press, New York.

Hoofnagle, J.H. (1981). Serologic markers of hepatitis B virus infection. *Annual Review of Medicine*, **32**, 1–11.

Howson, C.P. and Fineberg, H.V. (1992). Adverse events following pertussis and rubella vaccines. Summary of a report to the Institute of Medicine. *Journal of the American Medical Association*, **267**, 392–6.

Jenkins, P.J., Cromie, S.L., Bowden, D.S., Finch, C.F., and Dudley, F.J. (1996). Chronic hepatitis C and interferon alfa therapy: predictors of long term response. *Medical Journal of Australia*, **164**, 150–2.

Jupp, P.G. and McIntosh, B.M. (1990). *Aedes furcifer* and other mosquitoes as vectors of chikungunya virus at Mica, northeastern Transvaal, South Africa. *Journal of the American Mosquito Control Association* **6**, 415–20.

Kurtzman, G.J., Cohen, B.J., Field, A.M., Oseas, R., Blaese, R.M., and Young, N.S. (1989). Immune response to B19 parvovirus and an antibody defect in persistent viral infection. *Journal of Clinical Investigation*, **84**, 1114–23.

LeDuc, J.W., Pinheiro, F.P., and Travassos da Rosa, A.P. (1981). An outbreak of Mayaro virus disease in Belterra, Brazil. II. Epidemiology. *American Journal of Tropical Medicine and Hygiene*, **30**, 682–8.

Lindsay, M.D.A., Johansen, C.A., Broom, A.K., Smith, D.W., and Mackenzie, J.S. (1995). Emergence of Barmah Forest virus in western Australia. *Emerging Infectious Diseases (online)*, **1**, 1–6.

Luban, N.L.C. (1994). Human parvoviruses: implications for transfusion medicine. *Transfusion*, **34**, 821–7.

Meurman, O.H. (1978). Persistence of immunoglobulin G and immunoglobulin M antibodies after postnatal rubella infection determined by solid-phase radioimmunoassay. *Journal of Clinical Microbiology*, **7**, 34–8.

Mitchell, L.A., Tingle, A.J., Shukin, R., Sangeorzan, J.A., McCune, J., and Braun, D.K. (1993). Chronic rubella vaccine-associated arthropathy. *Archives of Internal Medicine*, **153**, 2268–74.

Moore, D.L., Causey, O.R., Carey, D.E., Reddy, S., Cooke, A.R., Akinkugbe, F.M., David-West, T.S., and Kemp, G.E. (1975). Arthropod-borne viral infections of man in Nigeria, 1964–1970. *Annals of Tropical Medicine and Parasitology*, **69**, 49–64.

Mortimer, P.P., Cohen, B.J., Rossiter, M.A., Fairhead, S.M., and Rahman, A.F.M.S. (1985). Human parvovirus and purpura. *Lancet*, **1**, 730–1.

Mosley, J.W. (1994). Should measures be taken to reduce the risk of human parvovirus (B19) infection by transfusion of blood components and clotting factor concentrates? *Transfusion*, **34**, 744–6.

Munoz-Fernandez, S., Barbado, F.J., Martin Mola, E., Gijon-Banos, J., Martinez Zapico, R., Quevedo, E., Arribas, J.R., Gonzalez Anglada, I., and Vazquez, J.J. (1994). Evidence of hepatitis C virus antibodies in the cryoprecipitate of patients with mixed cryoglobulinemia. (Review). *Journal of Rheumatology*, **21**, 229–33.

Naides, S.J. (1992). *Clinical virology manua*l, (2nd edn) (eds S. Specter and G. Lancz), pp. 547–69. Elsevier Science Publishers, Essex.

Naides, S.J., Scharosch, L.L., Foto, F., and Howard, E.J. (1990). Rheumatologic manifestations of human parvovirus B19 infection in adults. Initial two-year clinical experience. *Arthritis Rheumatism*, **33**, 1297–309.

Nakitare, G.W., Bundo, K., and Igarashi, A. (1983). Enzyme-linked immunosorbent assay (ELISA) for antibody titers against chikungunya virus of human serum from Kenya. *Tropical Medicine*, **25**, 119–28.

Niklasson, B. (1988). *The arboviruses: epidemiology and ecology* (ed. T.P. Monath), pp. 167–76. CRC Press, Boca Raton.

Niklasson, B., Espmark, A., and Lundstrom, J. (1988). Occurrence of arthralgia and specific IgM antibodies three to four years after Ockelbo disease. *Journal of Infectious Diseases*, **157**, 832–5.

Nishioka, K., Nakajima, T., Hasunuma, T., and Sato, K. (1993). Rheumatic manifestation of human leukemia virus infection. *Rheumatic Disease Clinics of North America*, **19**, 489–503.

Pattison, J.R. (1988). *Parvoviruses and human disease* (ed. J.R. Pattison), pp. 1–4. CRC Press, Boca Raton.

Peters, C.J. and Dalrymple, J.M. (1990). Alphaviruses. In *Fields virology*, (2nd edn) (eds B.N. Fields, D.M. Knipe, R.M. Chanock, M.S. Hirsch, J.L. Melnick, T.P. Monath, and B. Roizman), pp. 713–61.

Raven Press, New York.

Pinheiro, F.P., Freitas, R.B., Travassos da Rosa, J.F., Gabbay, Y.B., Mello, W.A., and LeDuc, J.W. (1981). An outbreak of Mayaro virus disease in Belterra, Brazil. I. Clinical and virological findings. *American Journal of Tropical Medicine and Hygiene*, **30**, 674–81.

Purcell, R.H. (1994). Hepatitis C virus. In *Encyclopedia of virology* (eds R.G. Webster and A. Granoff), pp. 569–74. Academic Press, San Diego.

Robinson, W.S. (1994). Hepatitis B viruses. In *Encyclopedia of virology* (eds R.G. Webster and A. Granoff), pp. 554–9. Academic Press, San Diego.

Ross, R.W. (1956). The Newala epidemic. III. The virus: isolation, pathogenic properties and relationship to the epidemic. *Journal of Hygiene*, **54**, 177–191.

Sato, H., Hirata, J., Kuroda, N., Shiraki, H., Maeda, Y., and Okochi, K. (1991). Identification and mapping of neutralizing epitopes of human parvovirus B19 by using human antibodies. *Journal of Virology*, **65**, 5485–90.

Schaffner, W., Fleet, W.F., Kilroy, A.W., Lefkowitz, L.B., Herrmann, K.L., Thompson, J., and Karzon, D.T. (1974). Polyneuropathy following rubella immunization: a follow-up study and review of the problem. *American Journal of Diseases of Children*, **127**, 684–8.

Scrimgeour, E.M., Aaskov, J.G., and Matz, L.R. (1987 *a*). Ross River virus arthritis in Papua New Guinea. *Transactions of the Royal Society for Tropical Medicine and Hygiene*, **81**, 833–4.

Scrimgeour, E.M., Matz, L.R., and Aaskov, J.G. (1987 *b*). A study of arthritis in Papua New Guinea. *Australia and New Zealand Journal of Medicine*, **17**, 51–4.

Seeger, C. (1994). Hepatitis B viruses: molecular biology (human). *Encyclopedia of virology* (eds R.G. Webster and A. Granoff), pp. 560–4. Academic Press, San Diego.

Shimizu, Y.K., Hijikata, M., Iwamoto, A., Alter, H.J., Purcell, R.H., and Yoshikura, H. (1994). Neutralizing antibodies against hepatitis C virus and the emergence of neutralization escape mutant viruses. *Journal of Virology*, **68**, 1494–500.

Shore, H. (1961). O'Nyong-nyong fever: an epidemic virus disease in east Africa. III. Some clinical and epidemiological observations in the northern province. *Transactions of the Royal Society for Tropical Medicine and Hygiene*, **55**, 361–73.

Siegel, L.B., Cohn, L., and Nashel, D. (1993). Rheumatic manifestations of hepatitis C infection. (Review). *Seminars in Arthritis and Rheumatism*, **23**, 149–54.

Silman, A.J. (1988). The 1987 revised American Rheumatism Association criteria for rheumatoid arthritis (editorial). *British Journal of Rheumatology*, **27**, 341–3.

Smith, C.A., Petty, R.E., and Tingle, A.J. (1987). Rubella virus and arthritis. (Review). *Rheumatic Disease Clinics of North America*, **13**, 265–74.

Tesh, R.B. (1982). Arthritides caused by mosquito-borne viruses. *Annual Review of Medicine*, **33**, 31–40.

Tesh, R.B., McLean, R.G., Shroyer, D.A., Calisher, C.H., and Rosen, L. (1981). Ross River virus (Togaviridae: Alphavirus) infection (epidemic polyarthritis) in American Samoa. *Transactions of the Royal Society for Tropical Medicine and Hygiene*, **75**, 426–31.

Ueno, Y. (1994). Rubella arthritis. An outbreak in Kyoto. *Journal of Rheumatology*, **21**, 874–6.

Van der Poel, C.L. (1994). Hepatitis C virus: Into the fourth generation. *Vox Sanguis*, **67** (Suppl. 3), 95–8.

Williams, M.C., Woodall, J.P., and Porterfield, J.S. (1962). O'Nyong-nyong fever: an epidemic virus disease in east Africa. V. Human antibody studies by plaque inhibition and other serological tests. *Transactions of the Royal Society for Tropical Medicine and Hygiene*, **56**, 166–72.

Williams, M.C., Woodall, J.P., Corbet, P.S., and Gillett, J.D. (1965 *a*). O'nyong-nyong fever: an epidemic in east Africa. VIII. Virus isolations from Anopheles mosquitoes. *Transactions of the Royal Society for Tropical Medicine and Hygiene*, **59**, 300–6.

Williams, M.C., Woodall, J.P., and Gillett, J.D. (1965 *b*). O'nyong-nyong fever: an epidemic in East Africa. VII. Virus isolations from man and serological studies up to July 1961. *Transactions of the Royal Society for Tropical Medicine and Hygiene*, **59**, 186–97.

Wolinsky, J.S. (1990). Rubella. In *Fields virology* (2nd edn) (eds B.N. Fields, D.M. Knipe, R.M. Chanock, M.S. Hirsch, J.L. Melnick, T.P. Monath, and B. Roizman), pp. 815–38. Raven Press, New York.

5.3.6 Rheumatological aspects of HIV infection

Robert Winchester and Silviu Itescu

Introduction

The epidemic of human immunodeficiency virus (**HIV**) infection is giving rise to several novel rheumatic conditions, directing renewed attention to the problem of musculoskeletal infection and providing new approaches to the study of the disease mechanisms of certain classic rheumatic illnesses that occur with undiminished intensity and frequency in stages of advanced acquired immune deficiency. The rheumatological aspects of HIV infection vary according to the stage of HIV infection (Table 1). Infection with HIV-1 initiates a complex pattern of injury and host responses that mark distinct stages in the evolution of the disease. For a variable number of years the continuing infection appears to be confined largely to cells of the monocyte lineage. It appears reasonably controlled by a specific immune response, during which time there is a striking degree of viral mutation and consequent diversification. The clinical symptomatology of this phase of the retroviral infection largely reflects the ongoing immune response to what have now become endogenous antigens with the potential for cytotoxic cellular and immune complex mechanisms of injury. In some, this phase may be without apparent symptoms. During this phase of clinical—but not viral—latency, the virus evolves toward a form that can more efficiently infect and injure the CD4 T cell. Ultimately, in most individuals and through mechanisms that are poorly understood, progressive deterioration of the immune system develops largely as the result of impairment in the function and numbers of the CD4 lineage of T cells leading to development of an acquired immune deficiency syndrome (**AIDS**).

Events in HIV infection	Consequences on spectrum of rheumatic disease
Chronic immune response to HIV antigens: humoral and cell-mediated	B-cell hyperactivity, autoantibody production and non-specific symptoms of chronic immune stimulation
	Lymphocytic infiltrative syndromes, e.g. DILS and inflammatory myopathy
	Nephropathy
	Vasculitides
	Mikulicz syndrome from B-cell lymphoma
Activation of intact or disregulated residual immune and inflammatory response mechanisms	Highly prevalent Reiter's syndrome and psoriatic arthritis
Selective immune deficiency affecting CD4+'helper' T cells	Opportunistic infections of musculoskeletal system
	Amelioration of CD4-dependent rheumatic diseases, e.g. rheumatoid arthritis

DILS, diffuse infiltrative lymphocytosis syndrome.

Table 1 Rheumatological consequences of HIV infection

The rheumatological aspects of HIV infection may be placed in three broad categories (Table 1). The first includes the host response to the HIV infection and specific conditions that appear to arise as a direct result of this response. Because of the endogenous location of the stimulating antigen there is a potential for a sustained immune response characterized by cellular infiltration, the production of certain autoantibodies, and vasculopathic and nephropathic responses that are mediated by immune complexes. The diseases in this group include polymyositis, various vasculitides, and the diffuse infiltrative lymphocytosis syndrome that superficially resembles Sjögren's disease. This mimicking of the features of Sjögren's syndrome by the host response to HIV as well as the rheumatic disorders following infection with HTLV-1 have re-energized interest in the possibility of relationships between autoimmune disorders and retroviruses (Kalden and Gay 1994). The constitutional manifestations associated with chronic viral infection may also simulate non-specific features of conventional inflammatory rheumatic diseases, especially in view of the production of some autoantibodies. The second category of rheumatological diseases comprises joint infections and other musculoskeletal manifestations that arise as a direct consequence of a deficient helper arm of the immune response associated with CD4 T-cell depletion. This group includes infectious arthritis and osteomyelitis resulting from conventional and opportunistic pathogens. The third category of rheumatic illnesses consists of certain conventional rheumatic diseases with an immune pathogenesis, such as Reiter's disease, psoriatic arthritis, and various undifferentiated spondylarthropathy syndromes, which otherwise might not be expected to occur in an immunosuppressed individual, since they are effectively treated by therapeutic immunosuppression. The existence of these diseases in this setting presumably reflects pathogenic mechanisms that operate through residual components of the immune system that have not been impaired significantly by the acquired immune deficiency (Table 2).

Disease example	Reiter's syndrome	Rheumatoid arthritis
Polymorphic MHC element associated with susceptibility	Class I (e.g. HLA B27)	Class II (e.g. DR4)
Autoantibodies present?	No	Yes
Response to selective immunosuppression of advanced HIV infection	Unchanged to worsened	Ameliorated
T-cell lineage presumably recognizing antigen in context of MHC molecule	CD8 (cytotoxic)	CD4 (help)
Implication for immunopathogenesis	CD8-MHC Class I	CD4-MHC Class II
	T-cell drive	T- and B-cell drive

Table 2 A classification of autoimmune disease according to hypothetical critical immune recognition events

Biology of HIV-1 infection

Epidemiology

Approximately 10 million people are infected with HIV-1 (Quinn 1990). HIV is transmitted either sexually or by exchange of blood and other bodily fluids. Accordingly, promiscuity or needle sharing confer a risk for acquisition of HIV. The virus may also be transmitted vertically *in utero*, perinatally at the time of delivery, or postnatally through breast feeding, with approximately one in four children born to infected mothers becoming persistently infected (Friedland *et al*. 1986; Fischl *et al*. 1987). The predominant mode of transmission differs in various regions of the world and this influences the segment of the population at risk for the infection. In African and Asian countries, heterosexual transmission is the predominant mode of spread and the male:female ratio is approximately 1:1. There is now particularly rapid heterosexual spread of HIV in certain parts of India and South-East Asia. In North America and Europe where the primary modes of transmission have been intravenous drug use and male homosexual relationships, the predominance of cases occur in homosexual or bisexual men, approximately one-third in intravenous drug users, and the rapidly growing balance occurs through heterosexual contact. Although early in the epidemic contaminated blood products represented a major problem that resulted in the infection of a large proportion of haemophiliacs, only 2 per cent of the prevalent cases have resulted from administration of contaminated blood products (Quinn 1990).

Retroviral taxonomy

HIV-1 is a member of the family of retroviridae, defined as RNA viruses replicating via a DNA intermediate that is integrated in the genome of the host cell (Table 3). The family of retroviridae consists of three subfamilies: spumaviridae, oncornaviridae, and lentiviridae. Spumaviruses cause largely inapparent infections. Oncornaviruses, including the human T-cell leukaemia virus type 1 (**HTLV-1**), are transforming viruses that cause neoplasms in the infected host. HIV is a member of the non-transforming lentivirus subfamily that causes a slowly evolving infection characterized by virus–host interactions that include prolonged periods of clinical stability, weak neutralizing antibody responses, and ongoing, extensive, viral genetic mutation and antigenic drift. In tissue culture the lentiviruses cause cytopathic effects including fusion of selected haematopoietic cell lineages. Lentiviruses also include ovine visna maedi virus and caprine arthritis/encephalomyelitis virus that primarily infect monocyte lineage cells and result in illness in sheep and goats with virus–host interactions similar to those occurring in HIV infection.

Oncornaviridae
HTLV-1
Lentiviridae
Human, e.g. HIV-1, HIV-2
Animal, e.g. simian immunodeficiency virus, caprine arthritis/encephalomyelitis virus, visna maedi virus
Spumaviridae

Table 3 The retroviridae

Structure and organization

The lentiviruses have a distinctive oblong viral capsid seen on electron microscopy, with a protein core shell consisting of p24 molecules. The HIV virion consists of the retroviral genome, which is a double-stranded 7 to10-kb molecule of RNA associated with lower molecular weight RNA species, nucleoproteins, and several enzymes, including reverse transcriptase, ribonuclease, DNA endonuclease, and a protease (Fig. 1). The virion is encapsulated in an internal protein core shell, which is covered by an outer envelope composed of virally encoded glycoprotein and host cell lipid bilayer membrane containing host-derived proteins such as b_2-microglobulin, HLA molecules, actin, and ubiquitin (Arthur *et al*. 1992).

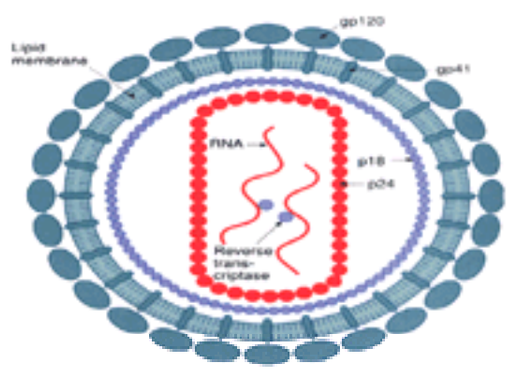


Fig. 1 Schematic representation of HIV-1 structure demonstrating location of various retroviral proteins.

The genomic RNA of all retroviruses contains three principal genes, *gag, pol*, and *env* (Fauci 1988; Greene 1991) that respectively encode structural proteins, reverse transcriptase, and envelope proteins. Each end of the genome contains the long terminal repeat sequence containing the viral promoter for RNA transcription, transcriptional stop signals, and polyadenylation sites (Fig. 2). The *gag* gene (group-associated antigen) encodes the virion structural proteins, including the protein core shell of p24 molecules. The *pol* gene encodes a precursor protein that is sequentially cleaved to produce the RNA-dependent DNA polymerase (reverse transcriptase) of the virus, a protease responsible for the cleavage, and an integrase. The *env* gene codes for two envelope-associated glycoproteins, a smaller transmembrane nchoring protein (gp41 in HIV) and a larger outer membrane protein (gp120 in HIV) which recognizes the target cell viral receptor and against which the host neutralizing antibodies are usually directed. In addition to the *gag, pol*, and *env* structural genes, the HIV-1 genome has at least six other accessory genes involved in regulation of viral replication (*tat, rev*, and *nef*), virion maturation (*vpr*) and morphogenesis (*vpu* and *vii*) (Fauci 1988; Greene 1991). The rev protein is of interest because it acts rather like the La/SSB molecule and is necessary for transportation of the large pool of unspliced or partially spliced viral mRNAs from the cell nucleus to the cytoplasm, where they are translated (Felber *et al*. 1989).

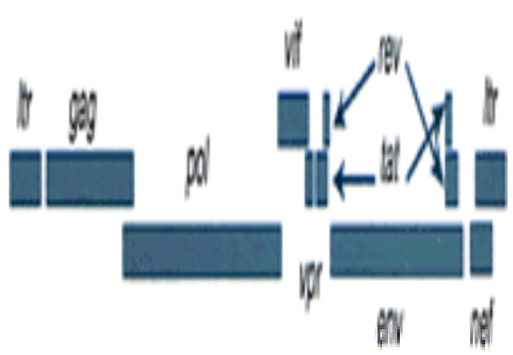


Fig. 2 Genomic organization of HIV-1. The genes encoding *tat* and *rev* exist as two non-contiguous elements, while the structural (*gag, pol, env*) and accessory genes (*vpr, vif, nef*) are shown as single elements.

Lifecycle and replication of HIV

Although the most striking clinical finding of HIV infection, AIDS, results from the infection of the CD4 T-cell lineage, infection of the monocyte plays a particularly central role in the biology of HIV-1 (Meltzer and Gendelman 1992). The monocyte/macrophage is usually the first cell that is infected. While the HIV infection progressively evolves towards a virus that has the ability to infect and injure T cells, this is not apparently a critical feature in the infectious cycle of the virus since often the virus is transmitted to another individual by an HIV-laden monocyte (Meltzer and Gendelman 1992). Macrophages are the major tissue reservoir of the virus at all stages of the infection (Meltzer and Gendelman 1992) with the virus evolving separately in different monocyte lineage compartments (Itescu *et al*. 1994a). As with the visna maedi virus, there is little or no cytopathic effect evident in the HIV-infected monocyte. To gain entry into a host cell (Fig. 3), the viral envelope gp120 molecule binds to the CD4 molecule on the surface of T-helper cells (Klatzmann *et al*. 1984), monocyte lineage cells (Koenig *et al*. 1986; Salahuddin *et al*. 1986), or skin Langerhans cells (Tschachler *et al*. 1987). Subsequently, fusion of cellular and viral membranes occurs, mediated by the HIV-1 gp41 protein. The change in HIV-1 tropism for macrophages or T cells correlates, in part, with structural differences within the HIV-1 gp120 envelope, most notably with the presence of particular residues in a 35-amino acid region within the highly variable principal neutralizing V3 domain (O'Brien *et al*. 1990; Cheng Mayer *et al*. 1991; Hwang *et al*. 1991; Shioda *et al*. 1991; Westervelt *et al*. 1992). Following internalization and uncoating of the virus particle, the viral RNA is reverse transcribed into DNA in the cytoplasm, transported to the nucleus, and integrated as provirus in the host genome.

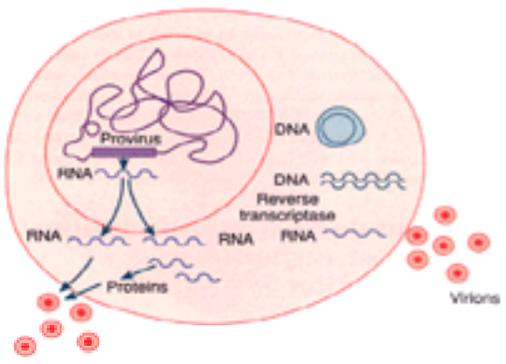


Fig. 3 HIV-1 lifecycle. Events depicted include cellular binding of HIV-1, virion internalization, reverse transcription of RNA genome, integration into host DNA as provirus, proviral transcription, virion transport and assembly, and budding of infectious particles.

The central features of HIV replication are depicted in Fig. 3. Regulation of HIV replication involves both host-regulated transcription initiation and virus-regulated post-transcriptional control. In T cells, HIV growth rates are closely linked to the growth state of the host's own T cell, with little or no virus production detected in non-proliferating T cells (Margolick *et al*. 1987). However, activation of these cells, which may occur during certain inflammatory disease states or following stimulation with antigens, mitogens such as phorbol myristate acetate or phytohaemagglutinin, and ytokines such as interleukin 2 (**IL-2**) or tumour necrosis factor-a (TNF-a), dramatically increases viral yields (Siekevitz *et al*. 1987). Coinfection with HTLV-1 also increases HIV-1 replication by a similar action of the transactivating protein of HTLV-1, tax (Siekevitz *et al*. 1987). All of these T-cell activators induce the intracellular production of various nuclear binding factors that mediate the programmes of T-cell activation, such as the NF_KB DNA transcription factor (Nabel *et al*. 1988) (Fig. 4). The presence of NF_KB, EBP-1, and Sp-1 binding sites in the long terminal repeat, promoter region of the virus (Nabel *et al*. 1988), accounts for the observed parallel activation of HIV replication and emphasizes how the virus has evolved to exploit normal immunological regulatory circuits of the host. For these reasons, the inflammation of rheumatic diseases arising in the HIV-infected individual may have an adverse impact on the host–viral balance, and it would appear reasonable to consider suppressing the inflammatory aspect of the rheumatic disease in an attempt to diminish the viral transactivation process that might initiate a period of accelerated immune decompensation. Similarly, the initiation of antiretroviral therapy should be considered in all HIV-infected persons developing inflammatory diseases. Indeed, the development of Reiter's syndrome has been rapidly followed by evidence of overt immune deficiency (Winchester *et al*. 1987).

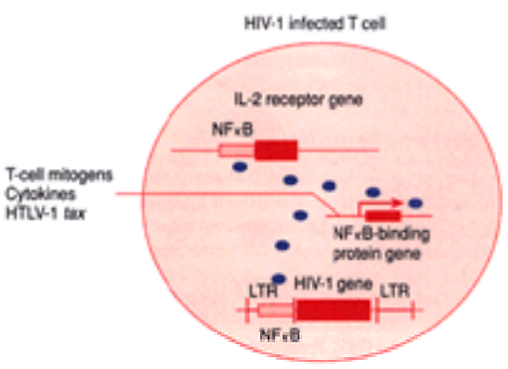


Fig. 4 HIV-1 transactivation. Schematic representation of NF_kB-binding protein induction by T-cell mitogens, cytokines, and HTLV-1, and initiation of both viral and host T-cell activation via NF_kB elements in the HIV-1 long terminal repeat (LTR) and the interleukin 2R (IL-2R) promoter.

The virus-specific post-transcriptional control of replication involves products of three viral genes: *tat, rev*, and *nef*. The tat protein is a powerful transactivator of the HIV genome. It requires specific transactivation responsive (TAR) regions in the viral mRNA transcripts (Muesing *et al*. 1987), leading to as much as a 1000-fold increase in viral mRNA expression. In addition to transactivation of HIV provirus, the tat protein is secreted extracellularly and taken up by uninfected cells (Viscidi *et al*. 1989). The nef protein is associated with cytoplasmic membranes by myristylation, and is a member of the G protein family, both binding GTP and displaying GTPase activity (Guy *et al*. 1987). It is thought to be critical for HIV pathogenesis since HIV strains with defective *nef* genes have been isolated from individuals with non-progressive disease (Kirchhoff *et al*. 1995).

Rheumatological manifestations of the host response to HIV-1 infection

General features

Following infection with HIV-1, specific antibodies develop that are directed against the major structural products of the viral genes *gag*, including p24, *pol* and *env*, including gp120 and gp41. IgM antibodies against gag or env proteins usually appear within 2 weeks of the acute clinical illness associated with primary HIV infection and by 3 months (Tindall *et al*. 1990) are replaced by IgG antibodies. The development of antibodies is occasionally delayed for several months or more. Viral infection can also be directly detected by p24 antigen capture assays, HIV culture, and viral RNA/DNA amplification using the polymerase chain reaction (PCR) (Piatak *et al*. 1993).

B-cell abnormalities

Involvement of the humoral (B-cell) arm of the immune system is reflected by the presence of polyclonal hypergammaglobulinaemia, circulating immune complexes, increased spontaneous B-cell proliferation and development of B-cell lymphomas (Solinger and Hess 1991). These features occur despite an inability to mount antigen-specific B-cell responses (Lane *et al*. 1983) and poor B-cell responsiveness to T-cell factors involved in proliferation and differentiation (Lane and Fauci 1985). Peripheral blood B cells from HIV-infected individuals are polyclonally activated and spontaneously secrete high levels of immunoglobulins (Lane *et al*. 1983). This may be a result of coinfection with known polyclonal B-cell activating viruses such as Epstein–Barr virus or cytomegalovirus (Lane and Fauci 1985), a direct consequence of effects on B cells by HIV-encoded proteins (Pahwa *et al*. 1985), or of B-cell stimulation by IL-6 which is induced in monocytes by HIV infection (Nakajima *et al*. 1989).

As anticipated because of the B-cell hyperreactivity in HIV infection there are a variety of autoantibodies; however, the well-characterized autoantibodies associated with classic rheumatic syndromes are not commonly observed. Rheumatoid factors are rarely detected (Solinger *et al*. 1988; Solinger and Hess 1991). Low titres of antinuclear antibodies are seen in a few patients, without being associated with analogous clinical syndromes in HIV-negative individuals (Solinger *et al*. 1988; Rynes 1990). Circulating immune complexes, measured by the Clq binding assay, are elevated in many HIV-infected individuals at all stages of disease, reflecting at least in part viral–antiviral complexes. However, C3 and C4 levels are not significantly decreased (Mayer-Siuta *et al*. 1988) nor is immune complex disease found in parallel. High titres of IgG anticardiolipin antibodies occur in 20 to 30 per cent of HIV-infected individuals (Bernard *et al*. 1990). While apparently not associated with the development of thrombotic events, these antibodies have been reported to be associated with thrombocytopenia (Canoso *et al*. 1987) and with cerebral perfusion defects (Rubbert *et al*. 1994). Autoantibodies against circulating lymphocytes have been reported in HIV-positive individuals at all stages of infection (Williams *et al*. 1984; Dorsett *et al*. 1990). These may be lymphocytotoxic, recognizing an extracellular domain of the CD4 molecule that is distinct from the HIV-1 gp120-binding region (Kowalski *et al*. 1989), or directed against the gp41 component of the HIV-1 envelope (Golding *et al*. 1988). In addition to these reactivities, antibodies directed to antineutrophil cytoplasmic antigens (**ANCA**) have been reported in up to 42 per cent of infected individuals (Klaassen *et al*. 1992; Savige *et al*. 1993). In these persons, ANCA antibodies are not associated with cutaneous or systemic vasculitis.

T-cell alterations

The host response to HIV includes an increase in the numbers of CD8+ T cells (Zolla-Pazner *et al*. 1987) that are either HIV-specific MHC class I restricted or unrestricted CD8 cytotoxic cells (Riviere *et al*. 1989). These are particularly prominent early in the course of HIV infection and tend to diminish with disease progression. Shortly after infection, the plasma viral load rises sharply and then decreases within weeks concomitantly with the development of host immune responses including HIV-specific cytotoxic T cells and neutralizing antibodies. At this initial stage the illness may assume a clinical form resembling infectious mononucleosis (De Wolf *et al*. 1988; Connor *et al*. 1993). Both neutralizing antibodies (Looney *et al*. 1988; Rusche *et al*. 1988; Javaherian *et al*. 1989; Koup *et al*. 1994) and cytotoxic T cells (Takahashi *et al*. 1990; Koup *et al*. 1994) elicited by natural infection recognize the immunodominant V3 loop of the envelope gp120 protein, and the high rates of sequence change in this region are thought to confer adaptive value to HIV-1 by allowing it to evade these host immune responses (Zwart *et al*. 1991). In addition, cytotoxic responses to various conserved viral proteins can be detected in most infected individuals throughout the disease course (Walker *et al*. 1987; Koenig *et al*. 1990; Phillips and McMichael 1993).

During this period the HIV envelope evolves to a T-cell tropic form associated with declining numbers of CD4 lymphocytes and increases in viral load. These events are predictive of progression from asymptomatic status to AIDS (Goedert *et al*. 1987; De Wolf *et al*. 1988; Moss *et al*. 1988; Connor *et al*. 1993; Weiss 1993). The median duration from initial infection to the development of frank immune deficiency is 7 to 10 years (Lui *et al*. 1988; Bacchetti and Moss 1989). The HLA genetics of the host appear to influence the course of the infection. Increased rates of progression to CD4 T-cell depletion and opportunistic infections have been reported in individuals with HLA B35 (Scorza-Smeraldi *et al*. 1986; Sheehy *et al*. 1989) and the HLA A1, B8, DR3 haplotype (Steel *et al*. 1988; Kaslow *et al*. 1990). In contrast, slower progression to immune incompetence occurs in infected individuals with the MHC class II alleles DRB1*1102 and DRB1*1301 (Itescu *et al*. 1994). These observations suggest that immunogenetic regulation of the host response to HIV infection may be critical in controlling HIV replication and limiting the progressive increase in viral burden.

Along with the progressive depletion of CD4+ T cells, HIV-infected individuals develop a qualitative helper cell defect, first evident as a deficient T-cell proliferative response to recall soluble antigens, and then to T-cell mitogens (Fahey 1986). These abnormalities are accompanied by diminished production of and response to IL-2 (Prince and John 1987), that results in a predominant T_{H1} pattern of immune deficit. These functional defects are simply demonstrated by cutaneous anergy to routine recall test antigens. They frequently precede the precipitous decline in CD4 cell numbers in late stages of the infection, may occur in the asymptomatic seropositive person with a CD4 count in the range of 400 to 500/mm^3, and lead to increased susceptibility to conventional and opportunistic infections. Because of the role of CD4 T cells in inducing CD8 T-cell maturation, the progressive loss of CD4-derived inductive signals ultimately leads to a functional deficiency of CD8 cell cytotoxicity against opportunistic pathogens, such as cytomegalovirus or perhaps HIV itself.

Rheumatic disorders related to T-cell proliferation

In a subset of individuals, the level of CD8 T cells remain high and may reach levels of greater than 2500/mm^3. Several distinctive syndromes occur in the setting of these elevated levels of CD8 T cells (Couderc *et al*. 1987; Itescu *et al*. 1989; Malbec *et al*. 1994) including the diffuse infiltrative lymphocytosis syndrome (**DILS**), generalized lymphadenopathy, and pseudotumoural splenomegaly. Each of these syndromes may be seen early in the course of HIV infection, while DILS is also found in long-term survivors.

Diffuse infiltrative lymphocytosis syndrome: a Sjögren's-like disease

The number of entities responsible for keratoconjunctivitis sicca and xerostomia have expanded over the past decade to include chronic graft-versus-host disease following bone marrow transplantation (Fox *et al*. 1986), infection with HTLV-1 (Vernant *et al*. 1988), and HIV-1 infection (Couderc *et al*. 1987; Itescu *et al*. 1989). The sicca syndrome occurring in HIV infection is a part of a syndrome designated 'diffuse infiltrative lymphocytosis', DILS, that differs from classic Sjögren's syndrome in a number of features. It appears to reflect a specific and seemingly beneficial host immune response to HIV (Itescu *et al*. 1994). DILS is encountered in a bimodal age distribution, being seen in children up to age 14 and in adults ranging in age from 22 to 62 years. DILS is not restricted to a particular risk group or ethnicity.

Clinical features

At presentation most patients with DILS meet criteria for AIDS-related complex, or stage 3B (Haverkos *et al*. 1985), with fevers, lymphadenopathy, and weight loss. In part reflecting ascertainment by a rheumatology service, bilateral parotid or submandibular gland enlargement is present in over 90 per cent of patients with DILS, being massive in two-thirds. Xerostomia occurs in 85 per cent, while xerophthalmia and keratoconjunctivitis sicca, diagnosed by Schirmer's test or rose bengal staining of the cornea, occur less frequently, in 40 per cent. Sicca symptoms are less evident in children with DILS. The decreased glandular secretions and associated lymphoid tissue enlargement predisposes to recurrent sinus and middle ear infections. Gallium scanning demonstrates bilateral isotope uptake in the involved tissue (Fig. 5). Computed tomography (CT) scans and magnetic resonance imaging (MRI) techniques show massive symmetrical and often cystic enlargement of major salivary glands (Itescu *et al*. 1993) (Fig. 6). The minor salivary gland tissues have focal lymphocytic infiltration indistinguishable from that of classic Sjögren's syndrome by conventional microscopy. There is a spectrum from complete preservation of glandular architecture to varying degrees of atrophic duct epithelium, canal dilatation, and interstitial fibrosis. Involved ductules usually express class II MHC molecules, suggesting local effects of cytokine release by infiltrating T cells (Fig. 7).

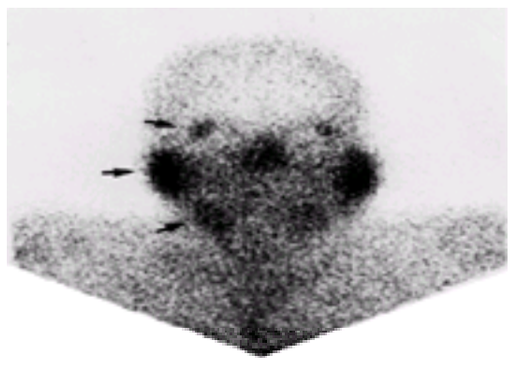


Fig. 5 Gallium scan demonstrating increased uptake bilaterally in the lacrimal, parotid, and submandibular glands (arrows) in an HIV-infected individual with diffuse infiltrative lymphocytosis syndrome.

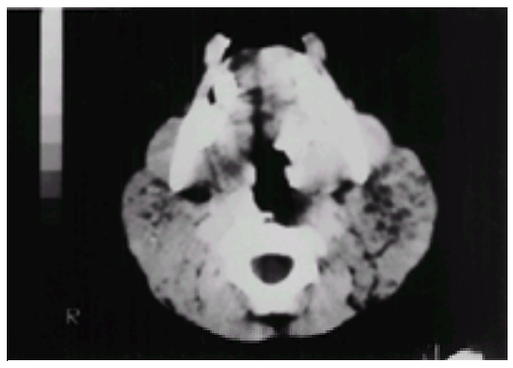


Fig. 6 CT scan of parotid glands in an HIV-infected child with DILS, demonstrating bilateral glandular enlargement with cystic changes.

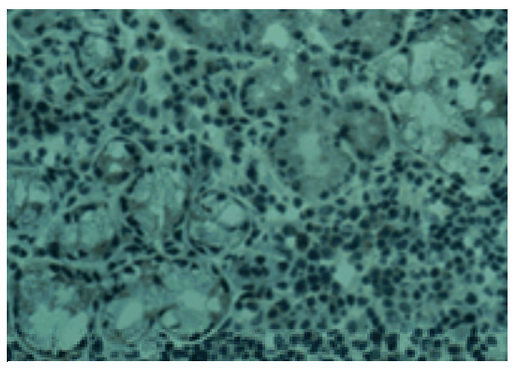


Fig. 7 Immunohistological study of minor salivary gland specimen from a patient with DILS demonstrating staining for HIV gp120 antigens expressed in the cytoplasm of several cells of the monocyte lineage. The positive cells are characterized by abundant cytoplasm and express CD14 lineage markers. They are in a periacinar location and adjacent to a focus of infiltrating lymphocytes.

Extraglandular involvement is particularly prominent. The development and extent of glandular and visceral lymphocytic infiltration in DILS is loosely correlated with the absolute numbers of circulating CD8 T cells, suggesting that lymphocytic infiltration is a direct consequence of the expanded population of circulating CD8 cells in these patients. Pulmonary involvement as a result of lymphocytic interstitial pneumonitis (**LIP**), which occurs in over 50 per cent of patients, appears to be the most serious complication of DILS and may be the initial manifestation of the syndrome. Affected patients may progress to frank respiratory insufficiency and endstage lung disease. Chest radiographs reveal bilateral interstitial infiltrates and gallium scanning often shows diffuse uptake throughout the lung fields. Diagnosis requires histological confirmation, and pulmonary infections, particularly *Pneumocystis pneumonia* or *Mycobacterium tuberculosis*, must be excluded. Superimposed bacterial pneumonias can complicate LIP. Enlargement of the tonsils, adenoids, and associated lymphoid tissue may be massive as can be seen in Fig. 6. Gastrointestinal manifestations include lymphocytic hepatitis, causing hepatomegaly and moderate elevations in transaminases and alkaline phosphatase, and gastric lymphocytic infiltration causing disorders in food intake resembling a linitis plastica. Lymphocytic interstitial nephritis, without glomerular involvement, may cause aseptic progressive renal insufficiency and a type IV renal tubular acidosis. Neurological involvement may cause lymphocytic meningitis, unilateral or bilateral palsy of the VIIth cranial nerve, and symmetric sensorimotor neuropathies. Other manifestations have included uveitis, lymphocytic thymoma, and of particular interest, lymphocytic mastitis which is illustrated in Fig. 8. Biopsy of the involved breast tissue revealed a massive CD8 T-cell infiltration, and the presence of HIV viral mRNA and viral proteins p24 and gp120 exclusively in infiltrating monocytes. The appearance of mastitis and LIP are the major features of visna maedi virus disease, suggesting that it may be a model of DILS.

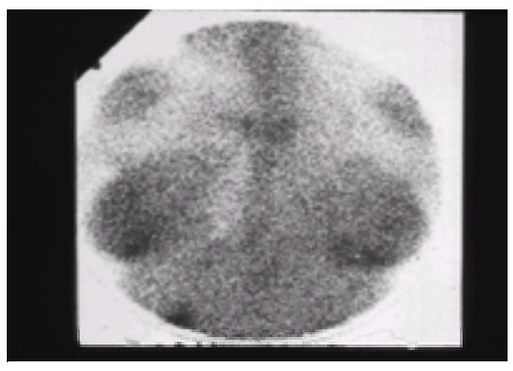


Fig. 8 Gallium scan of a female with DILS illustrating extensive bilateral uptake of the radionuclide in breast tissue. Biopsy of this tissue demonstrated CD8-predominant lymphocytic mastitis with the presence of HIV antigens evident in cells of the monocyte lineage. Analogous findings occur in ovine infection with the visna maedi agent.

HIV-infected individuals with DILS have a relatively low rate of progression to frank immune deficiency (Itescu *et al.* 1990), for durations of up to 10 years, CD4 T-cell levels decline by less than 10 per cent annually. Opportunistic infections and disseminated fungal or viral infections are rarely seen early in the course of DILS, but there is a considerably increased risk of developing high-grade B-cell salivary gland lymphomas. This complication should be suspected when there is a sudden, generalized increase in the size of the parotid glands, or when there is marked asymmetry. Biopsy is of great importance in making this diagnosis. Circulating cryoglobulins or monoclonal light chains may appear with lymphomatous transformation (Itescu 1991).

Pathogenesis

Immunophenotypic studies have demonstrated that the expanded circulating and infiltrative CD8 lymphocyte population in DILS has a memory phenotype (Itescu *et al*. 1993), expressing CD29 and CD11a/CD18 (LFA-1) molecules involved in effector functions including cell adhesion and cytotoxicity (Martz 1986; Springer *et al*. 1987; Springer 1990). The specific trafficking of CD8 lymphocytes to salivary gland and other tissues in DILS is probably regulated by the interaction between homing receptors on the lymphocytes and their ligands on postcapillary venule endothelial cells, including intercell adhesion molecule 1 (**ICAM-1**), a ligand for CD11a/CD18 (LFA-1) (Springer *et al*. 1987; Springer 1990), perhaps being induced by cytokines such as tumour necrosis factor-a (Springer *et al*. 1987; Springer 1990), produced in high quantities by HIV-infected monocytes (Roux Lombard *et al*. 1989). That the CD8 infiltration occurs in response to local replication of HIV is suggested by the marked diminution in salivary gland size frequently observed following treatment with antiretroviral agents, and by the demonstration of HIV-1 envelope proteins in salivary gland monocyte/macrophage lineage cells located in periacinar and perivascular areas adjacent to lymphoid aggregates and to postcapillary venule endothelial cells that selectively express ICAM-1 (Itescu *et al*. 1993). The T cells infiltrating the minor salivary glands in DILS are relatively oligoclonal, being drawn from a limited number of *V* and *J* gene segment subgroups but exhibiting extensive diversity in VDJ joining (Dwyer *et al*. 1993). The restricted repertoire of both the a- and b-chains of the T-cell receptor exemplified by common structural motifs, both germline and somatically encoded, constitutes compelling evidence that the infiltration of T cells into salivary tissue is the result of a subpopulation of lymphocytes that is responding in a very specific manner to structurally unique antigenic stimuli.

Predisposition to DILS is associated in black subjects with HLA DR5 (Itescu *et al*. 1989) and DR13 (Itescu *et al*. 1994). DNA sequence analysis demonstrates preferential association with the DRB1*1102 (JVM) allele of HLA DR5 (Itescu *et al*. 1994), which is extremely rare in white subjects (Johnson *et al*. 1989; Fernandez-Vina *et al*. 1990), but not with the HLA DR5 allele, DRB1*1101, which is predominant in white subjects. The DRB1*1301 subtype of HLA DRw6, found at the greatest increase in frequency, shares a b_1-chain diversity region motif with DRB1*1102, suggesting that this region may be of importance in predisposing a pattern of response to HIV.

HIV-1 is difficult to isolate and the predominant strains are tropic for monocyte lineage cells. However, viral isolates from individuals with longer disease duration are highly cytopathic and T-cell tropic. Analysis of the HIV-1 V3 domain in the T cells of persons with DILS revealed significantly lower nucleotide evolutionary divergence and amino acid heterogeneity than was observed in controls (Itescu *et al*. 1994).

Diagnosis

HIV infection should be considered in high-risk individuals presenting with the sicca syndrome, or in all individuals with the sicca syndrome that have atypical features, such as paediatric onset, male gender, prominent extraglandular manifestations, low or absent titres of autoantibodies, and reversed circulating CD4/CD8 T-cell ratio. Tentative criteria for the diagnosis of DILS are: (i) HIV-seropositivity shown by enzyme-linked immunosorbent assay (**ELISA**) and Western blot analysis, (ii) the presence of bilateral salivary gland enlargement or xerostomia persisting for more than 6 months, and (iii) histological confirmation of salivary or lacrimal gland lymphocytic infiltration, in the absence of granulomatous or neoplastic involvement. Typically, two or more foci of at least 50 lymphocytes/4 mm^2 of minor salivary gland tissue are seen, or grade 4 according to established criteria (Chisholm and Mason 1968). Rheumatoid factors are found in only 17 per cent of individuals with DILS and antinuclear antibodies with speckled pattern in 13 per cent, all at titres of less than 1:640. Antibodies against SSA/Ro and SSB/La, determined by ELISA using bovine and rabbit substrates as antigen sources (Buyon *et al*. 1989), are present in only 8 per cent of patients. LIP is an important element of DILS to delineate. After tissue confirmation of lymphocytic interstitial pneumonitis in adults, either by transbronchial or open lung biopsy, gallium scanning and pulmonary function studies are performed to assess the degree of pulmonary involvement. Appropriate biopsies are performed to measure the lymphocytic infiltration of other tissues because of the risk of other complications of HIV infection such as lymphoma or *Mycobacterium avium* infection that could masquerade as a feature of DILS.

Treatment

Treatment with zidovudine or other newer antiretroviral agents should be the first line of therapy. Zidovudine has often, but not uniformly, resulted in diminution of all of the manifestations of DILS including salivary gland enlargement. Discontinuation of zidovudine may result in striking re-enlargement of parotid glands. Circulating T-cell subsets are monitored regularly. Patients responding to zidovudine have demonstrated progressive diminution in CD8 cell numbers and either no change or a concomitant increase in CD4 cells. Symptomatic and progressive visceral involvement, not responsive to zidovudine, is very cautiously treated with 40 to 60 mg of prednisone daily, or other immunosuppressive agents such as chlorambucil, for a period not exceeding 8 to 12 weeks. Prior to commencing this therapy, the degree of circulating HIV antigen load is assessed usually by evaluating for the presence of p24 antigenaemia. Circulating p24 antigen is uncommon in DILS, but if elevated would constitute a relative contraindication to immunosuppressive therapy. More prolonged periods of steroid use and immunosuppression have resulted in the appearance of the development of opportunistic infections. Evidence of enhanced viral replication, ascertained by methods such as p24 antigen assays and cocultivation HIV reverse transcriptase levels, in patients who are being treated with prednisone would be an indication for immediate discontinuation of therapy. Treatment of the salivary gland lymphomas must be aggressive as these are high grade and associated with poor outcome.

Myopathy

Clinical features

HIV-infected individuals may develop a myopathy as a result of polymyositis-like inflammatory muscle disease, zidovudine therapy, or opportunistic infections (e.g. toxoplasmosis) (Snider *et al*. 1983; Dalakas *et al*. 1986; Dalakas *et al*. 1987; Baguley *et al*. 1988; Dalakas and Pezeshkpour 1988; Till and MacDonell 1990). Most patients with HIV-associated polymyositis present with an insidious onset of proximal muscle weakness and atrophy, as well as muscle pain and tenderness. Systemic features such as fever and weight loss may be present and typical skin lesions of dermatomyositis may be seen, such as heliotrope discoloration, periungual erythema, and erythematous plaques over the wrists and knuckles. In approximately 50 per cent of patients with myopathy, it is the initial manifestation of HIV infection, while in the remainder myopathy develops after the occurrence of opportunistic infections. Differential diagnosis includes vasculitic syndromes, which are usually associated with neuropathies and systemic organ involvement, and pyomyositis, which is usually a discrete unilateral process. Most patients with either inflammatory myopathy associated with HIV-1 or zidovudine-induced myopathy have elevations of muscle enzymes, as well as similar abnormalities on electromyography. Indeed, in one study, 16 per cent of all HIV-infected individuals taking zidovudine for more than 6 months developed abnormally elevated muscle enzymes, though only a few became symptomatic (Till and MacDonell 1990).

Diagnosis and pathogenesis

Muscle biopsy is the procedure of choice for determining the aetiology of HIV-associated myopathy. The entities are heterogeneous. In those with a polymyositis picture, classic myopathic changes are seen, including variation in fibre size, vacuolar change, and fibre destruction. In addition, type II atrophy and nemaline rod myopathy have been described (Dalakas *et al*. 1986; Dalakas *et al*. 1987; Panegyres *et al*. 1988). Most prominent, however, is the inflammatory perivascular and interstitial mononuclear cell infiltrate (Dalakas *et al*. 1986; Dalakas *et al*. 1987). In most, the infiltration consists predominantly of CD8+ lymphocytes, although CD4+ T cells and macrophages are also present (Espinoza *et al*. 1991), raising the possibility that the entity is related to the CD8 T-cell host response. HIV antigens have been demonstrated by immunofluorescence techniques in both muscle tissue and infiltrating mononuclear cells (Dalakas *et al*. 1986; Espinoza *et al*. 1991), however HIV-1 has not been cultured from muscle fibres in patients with polymyositis. The possible role of coinfection with HTLV-1, which is tropic for myocytes (Wiley *et al*. 1989) and may be associated clinically with a myositis (Mora *et al*. 1988), remains to be determined. In contrast, muscle biopsy in zidovudine-associated myopathy demonstrates similar myopathic changes but much less of an inflammatory infiltrate (Dalakas *et al*. 1990; Till and MacDonell 1990). In addition, ragged-red fibres suggestive of abnormal mitochondria are consistently observed. Electron microscopic studies confirm mitochondrial damage, with wide variation in size, swelling, degeneration, and laminar bodies present in these organelles (Dalakas *et al*. 1990; Till and MacDonell 1990). These mitochondrial abnormalities may be a result of inhibition by zidovudine of g-DNA polymerase (Mitsuya and Broder 1986), an enzyme required for mitochondrial DNA replication (Zimmerman *et al*. 1980). High blood lactate:pyruvate ratios, when determined repeatedly, may be a sensitive test for detecting the mitochondrial toxicity of zidovudine (Chariot *et al*. 1994).

Treatment

Therapy of zidovudine-induced myopathy requires discontinuation of the drug. In most cases, clinical improvement and decrease in creatine kinase values occur within 1 to 2 weeks. Careful reinstitution of lower-dose zidovudine may then be attempted. If no improvement occurs after zidovudine withdrawal, or if the affected individual was not taking the drug, muscle biopsy should be performed. Significant inflammatory infiltrates without evidence of mitochondrial changes or opportunistic pathogens indicate an immune-mediated myositis which usually responds to corticosteroid therapy. While these individuals appear to tolerate 40 to 60 mg of

prednisone daily, the fact that many patients with polymyositis are at advanced stages of HIV disease, as well as reports of Kaposi's sarcoma developing after treatment with prednisone and methotrexate (Espinoza *et al*. 1991), emphasize the need to taper steroids to the lowest dose effective in symptomatic control and to remain vigilant for possible infectious or malignant complications. Zidovudine therapy should be initiated or continued in polymyositis as it appears to be of benefit in the diminution of myopathic symptoms (Simpson 1988; Dalakas *et al*. 1990).

Vasculitis

Epidemiology

A number of vasculitic syndromes have been reported in HIV-infected individuals (Calabrese 1991; Gherardi *et al*. 1993). This is an area of considerable interest with the strong likelihood that some of these entities will be shown to be directly related to the host response to HIV infection. However, a direct causal relationship between infection with HIV and development of these syndromes remains to be established. Such determinations are complicated by the similarity in clinical manifestations between vasculitides and specific neurological, renal, pulmonary, or cardiac manifestations of HIV infection, and by the coexistence of pathogens such as Epstein–Barr virus, cytomegalovirus, and hepatitis B, which have also been causally related to various vasculitic syndromes (Sergent 1980; Marcellin *et al*. 1991; Louthrenoo 1993; Angulo *et al*. 1994). As with each of the other manifestations in the HIV-infected person, the rheumatologist is challenged by difficult and subtle differential diagnoses. The problems of distinguishing between the host response to hepatitis B and HIV infection has been informatively discussed (Angulo *et al*. 1994).

Clinical features

The reported vasculitic syndromes in HIV infection include involvement of small vessels in the hypersensitivity vasculitis group of either the neutrophilic or mononuclear cell type (Farthing *et al*. 1985; Chren *et al*. 1989; Potashner *et al*. 1990; Gherardi *et al*. 1993), lesions of medium-sized vessels in the polyarteritis nodosa group (Said *et al*. 1988; Gherardi *et al*. 1989; Valeriano *et al*. 1989; Angulo *et al*. 1994; Marks and Kuskov 1995), systemic granulomatous processes (Anders *et al*. 1989; Marks and Kuskov 1995), primary angiitis of the central nervous system (Yanker *et al*. 1986; Scaravalli *et al*. 1989), and involvement of the heart and great vessels (Marks and Kuskov 1995), with a fibroproliferative or aneurysmal process.

The most frequently encountered syndrome, hypersensitivity vasculitis, has been reported to occur either limited to the skin, presenting as palpable purpura, or in association with gut and renal involvement as part of Henoch–Schönlein purpura (Thompson *et al*. 1989; Gherardi *et al*. 1993). In some instances this could be attributed to a drug reaction. Polyarteritis nodosa-like forms of arteritides primarily involve muscles and nerves, and cause symmetric sensorimotor neuropathies, mononeuritis multiplex, muscle pain, and digital ischaemia and gangrene (Calabrese 1991; Jurgensen *et al*. 1995; Libman *et al*. 1995). Skin, gastrointestinal, and renal involvement are less common. Electromyographic studies help differentiate this condition from myopathy, demonstrating a pattern of axonal loss.

Churg–Strauss syndrome, characterized by purpuric skin lesions, bronchospasm, and eosinophilia, has been reported (Cooper and Patterson 1989). Other granulomatous processes, including lymphomatoid granulomatosis, have been reported in HIV-infected individuals, presenting chiefly with pulmonary and central nervous disease (Anders *et al*. 1989). Primary angiitis of the central nervous system may present either as a progressive loss of neurological function or as a fulminating encephalitic illness. Diagnosis is made by angiography or tissue biopsy. Positive serological findings or cultures for HIV in these cases may not be present in the periphery, and may be limited to the cerebrospinal fluid (Yanker *et al*. 1986).

Vasculitis resulting from cytomegalovirus, a major cause of morbidity among all immunocompromised patients because of infection of the endothelial cells by the virus, has been carefully reviewed (Golden *et al*. 1994). Virally induced proliferation and secondary ischaemia and host-mediated immune response to the virus contribute to the consequences of the vasculitis. Cutaneous and gastrointestinal involvement are prominent and must be diagnosed by biopsy.

Pathogenesis

Biopsy of involved skin in hypersensitivity angiitis demonstrates typical small vessel leucocytoclastic vasculitis with IgM and complement deposition within dermal capillaries. In Henoch–Schönlein purpura, the immune deposits in vessel walls contain IgA (Gherardi *et al*. 1993). Hypersensitivity to various drugs, particularly penicillins and sulphonamides, accounts for approximately 20 per cent of small vessel vasculitides (Gherardi *et al*. 1993). Productive infection with cytomegalovirus can be documented by the demonstration of viral inclusions in vascular endothelial cells (Gherardi *et al*. 1993; Golden *et al*. 1994). Polyarteritis nodosa-like disorders are associated with necrotizing vasculitic lesions in medium-sized vessels within muscle or epineurium. ANCA have been detected (Klaassen *et al*. 1992; Savige *et al*. 1993), however these antibodies are found at high prevalence in asymptomatic HIV-infected individuals. Although anticardiolipin antibodies occur in 20 to 30 per cent of HIV-infected individuals (Bernard *et al*. 1990), these antibodies have been reported to be associated with cerebral perfusion defects (Rubbert *et al*. 1994). In contrast, HIV p24 antigen has been reported within the vascular lesions (Bardin *et al*. 1987; Gherardi *et al*. 1993), and HIV RNA has been identified in perivascular cells of monocyte lineage (Gherardi *et al*. 1989; Gherardi *et al*. 1993), suggesting a possible direct role for HIV in the pathogenesis of the polyarteritis nodosa-like syndromes. HIV p24 antigen has also been detected within vascular endothelial cells in the midst of granulomatous inflammation associated with lymphomatoid granulomatosis or with primary central nervous angiitis (Anders *et al*. 1989; Calabrese 1991). The latter is further characterized by multinucleated giant cells within the internal elastic lamina on the surface of the cortex, brainstem, and associated leptomeninges. These findings may also be found in central nervous angiitis associated with herpes zoster infection (Eidelberg *et al*. 1986). The host response to bacterial infections, such as *Neisseria gonorrhoea*, have also been characterized by vasculitis (Ostlere *et al*. 1993).

Treatment

Any drugs possibly contributing to hypersensitivity vasculitis should be withdrawn. Life-threatening vasculitic complications involving the lungs, kidneys, or central nervous system require treatment with corticosteroids or other immunosuppressive agents. Additional treatment required may include antiretroviral agents and prophylactic therapy for herpes zoster and pneumocystis pneumonia. If cytomegalovirus inclusions are demonstrated within involved vascular endothelium, antiviral agents such as ganciclovir or foscarnet should be initiated promptly and immunosuppressive therapy concomitantly reduced (Golden *et al*. 1994).

HIV-associated nephropathy

HIV-associated nephropathy is another entity that is a component of the host response to HIV and which causes differential diagnostic problems for the rheumatologist, especially since, in concert with some of the autoantibodies that characterize the serum of patients with HIV, it may suggest the diagnosis of systemic lupus erythematosus. HIV-associated nephropathy is the most common finding among adult HIV-positive patients who present with nephrotic range proteinuria (D'Agati *et al*. 1989; Bourgoignie 1990). It is a clinicopathological complex characterized by varying degrees of focal segmental and global glomerulosclerosis with collapse of capillary tufts, mesangial cell proliferation, visceral epithelial cell hypertrophy and hyperplasia, microcystic tubular dilatation with tubular atrophy and degeneration, interstitial infiltration with mononuclear cells including a predominance of CD8 T cells, interstitial fibrosis, and endothelial tubuloreticular inclusions. Clinically, HIV-associated nephropathy differs from most idiopathic forms of focal segmental glomerulosclerosis in the greater severity of the renal pathology, the accelerated course to renal failure, and to a lesser degree, tubular damage.

Prevalence

The reported frequency of specific renal involvement in HIV varies to a certain degree according to whether ascertainment is by biopsy of those with symptomatic renal disease or at autopsy of all infected individuals (Bourgoignie 1990). In particular, diffuse mesangial cell hyperplasia and interstitial nephritis appear more frequent at autopsy, 31 and 12.5 per cent, respectively, compared with 5.7 and 1.7 per cent in patients undergoing biopsy, implying they did not give rise to findings that prompted biopsy (Bourgoignie 1990). In contrast, focal segmental glomerulosclerosis usually results in overt renal disease, accounting for 83 per cent of biopsy diagnoses of renal disease and from 1 to 15 per cent of diagnoses of renal pathology at autopsy according to population and criteria (Bourgoignie 1990). An overall prevalence of at least 7 per cent among HIV-infected adults is a reasonable estimate. The findings in children differ from those of adult series, with diffuse mesangial hyperplasia and focal segmental glomerulosclerosis found in approximately equal frequency (Strauss *et al*. 1989).

Epidemiology

The large majority of HIV-associated nephropathy in adults occurs in the stage of asymptomatic carrier or ARC and only 20 per cent occurs at the time when fully developed AIDS ensues (Haddoum *et al*. 1987; Pardo *et al*. 1987; Bourgoignie *et al*. 1988a; Carbone *et al*. 1989). In children, the average age at diagnosis of

HIV-associated nephropathy is approximately 3 years (range 6 months to 9 years) (Pardo *et al*. 1987; Connor *et al*. 1988; Strauss *et al*. 1989). Black subjects are far more likely to develop HIV-associated nephropathy than white subjects; this is apparent in all risk groups and in both adults and children, with a ratio of black to white patients approaching 11 to 1 (Bourgoignie *et al*. 1989; Strauss *et al*. 1989; Bourgoignie 1990; Ingulli *et al*. 1991). Because HIV-associated nephropathy develops at a stage of the infection where there is still a reasonably intact immune system which is responding to the virus with a specific immune response, it seems reasonable to hypothesize that this immune response may be the basis of the development of the nephropathy. In this sense HIV-associated nephropathy could be analogous to the development of diffuse infiltrative lymphocytosis syndrome, which is another, and possibly related, host response to HIV infection in which isolated renal tubular acidosis occurs.

Clinical features

Presenting manifestations of HIV-associated nephropathy, other than proteinuria and azotaemia, include nephrotic syndrome in 9 to 60 per cent (Pardo *et al*. 1987; Bourgoignie *et al*. 1988b), and hypertension in 39 per cent of individuals in one study (Carbone *et al*. 1989; D'Agati *et al*. 1989), although elsewhere this is less common (Chandler and Treser 1987; Rao *et al*. 1987; Bourgoignie *et al*. 1988b; Connor *et al*. 1988; Rao and Friedman 1989). Among adults, endstage renal disease ensues within 4 to 16 weeks (Rao *et al*. 1987; Carbone *et al*. 1989; D'Agati *et al*. 1989). A case of HIV-associated nephropathy with focal segmental membranous glomerulopathy that evolved to central nervous vasculitis and infarction, emphasizes that there may be a spectrum from HIV-associated nephropathy to some of the vasculitides (Bass *et al*. 1994). Cyclosporine and prednisone have been used in small numbers of patients and reported to control HIV-associated nephropathy in some, suggesting a role for participation of the immune response and cytokines in the mechanism of disease (Strauss *et al*. 1989; Ingulli *et al*. 1991; Ifudu *et al*. 1994). Other studies suggest that the progression to renal failure in HIV-associated nephropathy is slowed by zidovudine therapy (Ifudu *et al*. 1994), also directing attention to the importance of HIV replication in the pathogenesis of the syndrome.

Disorders occurring as a direct result of CD4 helper T-cell dysfunction

Musculoskeletal infections, including septic arthritis, osteomyelitis, and pyomyositis

Because of the central role played by the CD4 T cell in regulating both cellular and humoral immunity, the profound cell-mediated and humoral immunodeficiency in the later stages of HIV-1 infection predisposes individuals to a variety of infectious diseases. Infections may primarily develop in the musculoskeletal system as septic arthritis, osteomyelitis, and pyomyositis, or they may be systemic in nature and secondarily involve the musculoskeletal system, with either direct infection of bone joint and muscle or indirectly involving these structures with myalgias and arthralgias. In addition to these diagnostic challenges, the therapeutic approach to treating infection in the immunocompromised host must be altered to reflect the fact that there is an enhanced tendency for the microbial infections to persist after completion of conventional antibiotic therapy regimens, reflecting the failure of effective immune elimination of the residual organisms.

Septic arthritis and osteomyelitis are the two most common secondary infectious complications reported in HIV-infected individuals in case series studied in all parts of the world (Berman *et al*. 1991; Hughes *et al*. 1992; Ho 1993; Adebajo and Davis 1994; Goldenberg 1994). Septic arthritis may result from opportunistic or conventional infections that involve various joints, including those commonly affected by septic arthritis such as the hip, knee, wrist, metacarpophalangeal or interphalangeal joints, or those rarely involved such as the sacroiliac and sternomanubrial or sternoclavicular joints. The septic arthritis may be the first manifestation of AIDS, presenting the rheumatologist with the difficult diagnosis of both the septic process and the underlying HIV infection, or the arthritis may occur when an individual with known immune deficiency progresses into an advanced stage.

Organisms cultured from various joints have included, in addition to the anticipated organisms such as *Streptococcus pneumoniae* and *Staphylococcus aureus, Sporothrix schenckii* (Lipstein-Kresch *et al*. 1985), *Cryptococcus neoformans* (Ricciardi *et al*. 1986), *Histoplasma capsulatum* (Calabrese 1989), and *Mycobacterium avium intracellulare* species (Blumenthal *et al*. 1990). Other unusual pathogens cultured from septic joints in HIV-infected individuals include *Salmonella* spp. (Winchester *et al*. 1987; Gutierrez *et al*. 1993), *Haemophilus influenzae* type B (Lawrence *et al*. 1991), and *Campylobacter fetus* (Winchester *et al*. 1987).

Certain patterns of extra-articular involvement are the signatures of particular organisms such as the multifocal cutaneous cellulitis of *Campylobacter (Helicobacter) cinaedi* (Burman *et al*. 1995) associated with monoarticular or oligoarticular arthritis and an indolent febrile illness. Disseminated sporotrichosis infection may similarly present with diffuse skin lesions and oligo- or polyarthritis (Heller and Fuhrer 1991). Staphylococcal or streptococcal septic arthritis usually presents as an acutely swollen, painful, erythematous process in one or several joints, accompanied by systemic symptoms of bacteraemia. Staphylococcal infections may be widespread, causing septic bursitis (Jacobson *et al*. 1988; Buskila and Tenenbaum 1989), juxta-articular osteomyelitis (Masters and Lentino 1984), extensive periarticular soft-tissue involvement, and pyomyositis (Gaut *et al*. 1988).

Risk factor for HIV acquisition may influence the predominant infectious complications

In addition to the problem attributed to the effect of immune deficiency, the various high-risk behaviour patterns or underlying conditions associated with HIV infection may themselves independently predispose to infection with particular pathogens. *Staph. aureus, Pseudomonas aeruginosa*, or *Candida albicans* are usually the most frequently encountered organisms in the arthritis found in intravenous drug users regardless of whether HIV infection is present or absent (Rivera *et al*. 1992). If these infections, particularly *Pseudomonas*, occur in an individual, they should direct attention to the possibility of parenteral drug abuse (Munoz Fernandez *et al*. 1991), but some regions experience little Gram-negative disease among addicts (Munoz Fernandez *et al*. 1991). Among intravenous drug users infected with HIV, *Staph. aureus* infection, most commonly of the hip, accounts for about two-thirds of cases of non-gonococcal septic arthritis (Goldenberg 1991; Munoz Fernandez *et al*. 1991; Covelli *et al*. 1993). Pyogenic sacroiliitis or isolated sternoclavicular septic arthritis occurs predominantly among intravenous drug users and is usually caused by infection with *Staph. aureus* or Gram-negative organisms, including *Fusobacterium* as well as *P. aeruginosa* (Guyot *et al*. 1987).

Conversely, in Africa, where HIV is usually acquired by sexual promiscuity, gonococcal arthritis is the most common musculoskeletal infection and infection by *Staph. aureus, P. aeruginosa*, or *C. albicans* is uncommon (Blanche *et al*. 1993). Among HIV-infected haemophiliacs, the predominant cultured organisms are *Staph. aureus* and *Strep. pneumonia* (Pappo *et al*. 1989).

There is clear evidence that advanced infection of HIV by itself predisposes to a greatly increased risk of a variety of musculoskeletal infections (Blanche *et al*. 1993; Adebajo and Davis 1994). Similarly, an increased incidence of septic arthritis among HIV-infected haemophiliacs compared with those uninfected underlines the fact that immune deficiency of HIV infection itself predisposes to septic arthritis (Pappo *et al*. 1989). It has been argued that infections of the musculoskeletal system complicating HIV infection are rather more prevalent than the literature of reported cases might suggest (Ho 1993).

Osteomyelitis and pyomyositis

In addition to arthritis, osteomyelitis is a common problem. Juxta-articular osteomyelitis may be difficult to distinguish from arthritis, whereas other bony sites may be clinically more distinctive. Organisms reported to have caused osteomyelitis in HIV-infected individuals include *C. albicans* (Boix *et al*. 1990), *Mycobacterium kansasii* (Crawford and Baird 1987), and *Nocardia asteroides* (Masters and Lentino 1984). Staphylococcal juxta-articular osteomyelitis involving the olecranon process (Masters and Lentino 1984) or distal clavicle (Zimmermann *et al*. 1989) may present with subacute or chronic joint pain and swelling. Extensive periarticular infections, usually by *Staph. aureus*, can simulate multidigit dactylitis or arthritis.

Pyomyositis is a particular complication of staphylococcal infection and usually occurs in the more advanced stage of AIDS. Acute onset of unilateral thigh pain associated with soft-tissue swelling, erythema and woody induration of the distal thigh is characteristic of pyomyositis (Watts *et al*. 1987; Gaut *et al*. 1988; Goldenberg 1991). In the majority of cases, the cause is a single staphylococcal abscess; however, multiple collections may occur within the quadriceps muscle. Diagnosis is facilitated by imaging techniques, including ultrasound, CT, and MRI scans.

Clinical features and differential diagnosis

Several reviews deal with the changing pattern of musculoskeletal infection during this era and the general principles of diagnosis and therapy (Hughes *et al*. 1992; Ho 1993; Goldenberg 1994). Gram stain and culture of synovial fluid specimens in staphylococcal or streptococcal septic arthritis are diagnostic, and are frequently accompanied by positive blood cultures. Arthritis associated with *M. tuberculosis* may present its usual challenges in diagnosis. Septic arthritis caused by opportunistic organisms tends to be a more indolent and subtle process, often with minimal inflammation. In these patients, the underlying disorder may be suggested by the presence of associated extra-articular manifestations, such as necrotic, crusted skin lesions in sporotrichosis (Lipstein-Kresch *et al*. 1985). Synovial fluid aspirate often reveals low numbers of leucocytes with a relative increase in the proportion of monocytes. Culture of some of these organisms may require special

attention from the laboratory and the use of molecular biological methods. The radiological manifestations of musculoskeletal complications of HIV infection have been informatively summarized (Steinbach *et al*. 1993) and emphasize the extensive diagnostic evaluations that may be required to document the extent of structures involved with the infection.

A major problem in the diagnosis of infectious processes of the joints is distinguishing them from the inflammatory non-infectious entities that occur in HIV infection, in particular those of the spondylarthropathy group of illnesses. In addition, other entities that are part of HIV infection but not commonly encountered in rheumatology, such as B-cell lymphoma, Kaposi's sarcoma, and bacillary angiomatosis (an infection by an agent resembling that of trench fever, *Rochalimnaea quintana*) must be considered in the differential diagnosis (Steinbach *et al*. 1993). For example, in pyogenic sacroiliac joint infection, the pronounced pain, exquisite localized tenderness, asymmetry, and lack of involvement of other joints helps differentiate it from inflammatory sacroiliitis associated with the spondylarthropathies. Scintigraphy demonstrating unilateral involvement is also more suggestive of infection, especially when found in the absence of enthesopathy and other arthritis, however, diagnostic aspiration may be required (Guyot *et al*. 1987).

The fact that many patients present with atypical features of the infection because of the attenuated inflammatory response has been emphasized (Hughes *et al*. 1992). Systemic infections such as acute or subacute endocarditis occur with increased frequency in HIV and may present with arthralgias, arthritis, and back pain (Ho 1993). An instructive example of the diagnostic challenges that secondary infections present in the immunologically compromised host is the initial presentation of HIV infection as chronic syphilitic polyarthritis with an efflorescence of autoantibodies that suggested the presence of a rheumatic disease such as rheumatoid arthritis or systemic lupus erythematosus (Burgoyne *et al*. 1992). The rheumatologist has a particular problem in the approach to a patient who is not previously known to be infected with HIV. The general rule that is emerging in endemic areas of HIV infection is that the appearance of acute arthritis should prompt HIV testing (Blanche *et al*. 1993) and, by extension, in Europe and North America where HIV infection is more stratified, at least a strong suspicion of HIV infection should be entertained. Frequently the atypical manner of presentation directs attention to the probable presence of HIV infection, as for example the presentation of gonococcal arthritis involving a single hip and a sternoclavicular joint (Strongin *et al*. 1991). Atypical, invasive or extensive infection of the musculoskeletal system in a young individual should direct attention to the possibility of an underlying immune deficiency. For example, a report of sternoclavicular arthritis associated with *Strep. pneumoniae* as the presenting manifestation of AIDS illustrates that septic arthritis of a joint that is otherwise seldom involved by infectious processes should prompt a thorough search for predisposing conditions such as HIV infection (Leon *et al*. 1994).

Treatment

Septic arthritis or bursitis is initially treated with broad-spectrum coverage until culture results are available. Some organisms such as streptococcal and staphylococcal joint infections in HIV-positive individuals, or infection with organisms such as *Campylobacter (Helicobacter) cinaedi* (Burman *et al*. 1995), often respond as well to conventional antibiotic treatment as in HIV-negative intravenous drug users (Goldenberg 1991), yet streptococcal and staphylococcal bursitis is difficult to eradicate and requires prolonged therapy (Buskila and Tenenbaum 1989). Opportunistic infections are treated with appropriate antibiotic therapy, including intravenous amphotericin B for fungal infections. Sporotrichosis and allied infections will probably not be eradicated and require chronic maintenance therapy (Heller and Fuhrer 1991).

Some form of joint drainage is initiated soon after the diagnosis is reasonably established, although the precise method is still controversial (Ho 1993; Goldenberg 1994). Surgical drainage is required for appropriate treatment of osteomyelitis and pyomyositis in HIV infection in addition to antibiotics. Several less invasive procedures, such as 'tidal irrigation' or arthroscopic drainage appear to offer the promise of more rapidly controlling the sometimes strikingly rapid destructive consequences of septic arthritis in this group and have the advantage of avoiding the considerable morbidity of an arthrotomy (Ho 1993; Goldenberg 1994). The duration of antibiotic therapy must be given individual consideration in light of an assessment of the integrity of the residual immune function and with the knowledge that most infections in HIV-positive individuals require prolonged therapy and sometimes chronic suppressive therapy. Because most of these individuals are in an advanced stage of HIV infection, they require a comprehensive immunological staging evaluation and most probably appropriate retroviral therapy. Osteomyelitis may require permanent antimicrobial maintenance.

Immune-mediated arthritis occurring with the same or increased intensity and frequency in individuals with selective depletion of CD4 lineage T cells

The epidemiology of rheumatic disease is altered in the setting of HIV infection. The development of certain specific rheumatic disorders, and the absence of others, in individuals infected with HIV is consistent with a classification of immune-mediated rheumatic illnesses according to a hypothetical underlying immune recognition event (Table 2). Classic rheumatic disorders associated with prominent autoantibody and cellular responses, such as rheumatoid arthritis and systemic lupus erythematosus, are rarely found among HIV-infected individuals and, when seen, are often (Bijlsma *et al*. 1988; Amor 1989; Calabrese *et al*. 1989; Molina *et al*. 1995) but not always (Kerr and Spiera 1991), reported to improve with progression of HIV infection. This decrease emphasizes the probable importance of CD4 T cells in their immunopathogenesis.

Reiter's disease, reactive arthritis, psoriatic arthritis, and undifferentiated spondylarthropathy syndromes

Because advancing HIV infection is characterized by progressive immune deficiency, it came as a considerable surprise that certain of the seronegative spondylarthropathies occurred with undiminished, if not increased, frequency in the setting of frank AIDS, despite the fact that these rheumatic diseases respond to iatrogenic immunosuppression (Winchester *et al*. 1987). Both the musculoskeletal system, in the form of the spondylarthropathies (Winchester *et al*. 1987), and the skin, in the form of psoriasiform disease (Duvic 1995), are distinctive clinical markers for advanced HIV infection. The distinctions between the various seronegative spondylarthropathies are blurred in the HIV-infected individual. While HIV-infected patients presenting with classic Reiter's syndrome are at one end of the spectrum and those with classical psoriasis and psoriatic arthritis at the other, many develop features of an undifferentiated spondylarthropathy with or without cutaneous manifestations. Moreover, in the setting of HIV infection, a subset of individuals with Reiter's syndrome have a more distinctively fulminant and extensive disorder simulating psoriatic arthritis with unusually severe enthesopathy, upper limb joint manifestations, and cutaneous involvement which often becomes indistinguishable from pustular psoriasis (Winchester *et al*. 1987). The development of these disorders emphasizes their distinct mechanisms of immunological drive independent from the CD4 lineage as well as the broad spectrum of host–virus relationships that can occur in HIV infection, although psoriasis in isolated form is not associated with decreased survival (Obuch *et al*. 1992). Behçet's disease may fall into this category since its development has been attributed to antecedent HIV infection (Buskila *et al*. 1991). Curiously, ankylosing spondylitis has not been noted to occur at the expected frequency in this group.

Epidemiology

The precise frequency with which Reiter's syndrome, psoriatic arthritis, and the related spondylarthropathies occur in HIV-positive individuals continues to be the subject of inquiry, although there is universal agreement that their prevalence is not decreased and in a number of studies is variably increased (Berman *et al*. 1991; Calabrese *et al*. 1991; Monteagudo *et al*. 1991; Cuellar *et al*. 1994). Prevalence rates of spondylopathic disorders among HIV-infected individuals in the United States ranging up to nearly 5 per cent have been reported when ascertainment was performed by rheumatologists (Berman *et al*. 1988; Winchester *et al*. 1988). Calabrese and colleagues in a longitudinal study found psoriatic arthritis and Reiter's syndrome each at a frequency of 1.7 per cent and undifferentiated forms of oligo- or monoarticular rheumatism at a frequency of 11.1 per cent (Calabrese *et al*. 1991). For some of these patients, the standardized criteria for Reiter's syndrome or psoriatic arthritis are not proving adequate for classification. The uncommon clinical features of the spondylarthropathies that occur in these patients have been pointed out by a number of authors (Winchester *et al*. 1987; Altman *et al*. 1994; Kellner *et al*. 1994). Psoriasiform skin disease, sometimes occurring with spondylarthropathic disease, occurs at increased frequency among those infected with HIV (Cockerell 1991; Duvic 1991; Duvic 1995). It appears that both forms of Reiter's syndrome, the sexually transmitted venereal or endemic form preceded and apparently initiated by urethritis, and the epidemic or postdysenteric form following infection of the gastrointestinal tract, are seen in association with advanced HIV infection. Using strict criteria, the prevalence rates of 'complete' Reiter's syndrome among HIV-infected individuals in the United States were not found to be higher than those for HIV-negative groups matched demographically (Solinger and Hess 1993) and by high-risk behaviour (Clark *et al*. 1989; Hochberg *et al*. 1990). However, in these same studies, the frequency of psoriatic arthritis was over fivefold higher than expected in the HIV-positive cohort when compared with a demographically matched population (Solinger and Hess 1993). Studies from West and sub-Saharan Africa have reported between four- and sixfold higher HIV prevalence rates among patients with various spondylarthropathies than among the local populations (Blanche *et al*. 1993; Mijiyawa 1993). Together, these observations strongly suggest that there is a causal relationship between HIV infection and the development of various spondylarthropathy syndromes; however, their clinical expression may differ in comparison with analogous syndromes in the HIV-negative host.

Psoriasiform lesions, with or without arthritis, may be the first clinical manifestations of HIV infection. Alternatively, in individuals with pre-existent psoriasis, infection with HIV may significantly exacerbate the psoriatic condition, including the joint manifestations. In general, psoriasiform lesions in HIV infection are of greater severity

than in HIV-negative individuals and sometimes distinctive in distribution (Duvic *et al*. 1987; Duvic 1995). There is some divergence in studies on the prevalence of the psoriasiform disorders. Certain groups have found an increased prevalence of both psoriasiform lesions and musculoskeletal involvement plus psoriasiform skin lesions in HIV infection (Berman *et al*. 1988; Winchester *et al*. 1988; Solinger and Hess 1990; Calabrese *et al*. 1991). These studies have suggested that the prevalence of arthritis that resembles psoriatic arthritis in HIV infection is 1 to 3 per cent, compared with 0.05 to 0.14 per cent in HIV-negative individuals, and of psoriasiform lesions is 3 to 6 per cent, compared with 1 to 3 per cent in the HIV-negative population. Other studies have found similar prevalence rates of psoriasiform lesions but lower rates of arthritis, approximating those in the general population (Duvic *et al*. 1987; Kaplan *et al*. 1989). Possible explanations for these discrepancies may be differences in categorization of patients between centres, ethnic differences, variable use of diagnostic procedures such as confirmatory skin biopsies, and the high frequency of undifferentiated spondylarthropathy syndromes in HIV infection which do not meet criteria fully for either psoriatic arthritis or Reiter's syndrome.

Clinical features

The predominant finding in HIV-infected individuals with Reiter's syndrome is often relatively severe arthritis and enthesopathy, frequently accompanied by skin and nail disease (Fig. 9) (Winchester *et al*. 1987; Altman *et al*. 1994; Kellner *et al*. 1994). The course of the arthritis in HIV-associated Reiter's syndrome can take two general forms: an accumulative pattern evolving to full intensity over several weeks to months, or, more commonly, a milder intermittent pattern with recrudescences and remissions. The accumulative form is often associated with widespread polyarticular but asymmetric arthritis and is characterized by synovial thickening, erosions, and juxta-articular osteoporosis. The degree of upper extremity involvement and the accumulative pattern suggest features seen in psoriatic arthritis. The intermittent form usually has oligoarticular knee or ankle joint involvement and more closely resembles the clinical evolution of Reiter's syndrome in HIV-negative individuals. Enthesopathy is often unusually severe and profoundly disabling, while ocular involvement may be rather mild or absent. Cutaneous changes are usually striking in both intensity and extent. The distribution of cutaneous lesions is sometimes more like that found in pustular psoriasis.

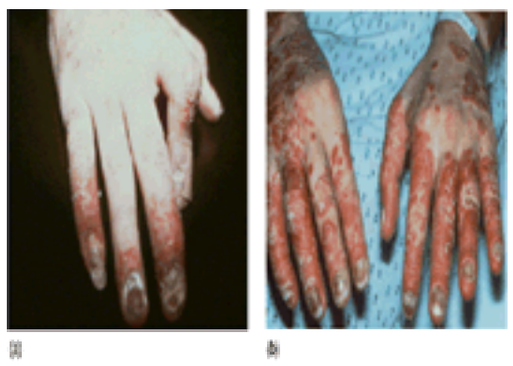


Fig. 9 Two patterns of HIV-associated spondylarthropathic hand involvement in HLA B27-positive persons; (a) illustrates findings typical of Reiter's syndrome, with onychodystrophy, subungual and acral hyperkeratosis, pseudoparonychia, fusiform swelling, and flexion contracture of the digits; (b) illustrates more extensive psoriasiform involvement of the skin of the hand and other parts of the body. Acral intensification and onychodystrophy are present.

The foot and ankle are the most commonly involved sites. Severe enthesopathy of the Achilles tendon, plantar fascia, and anterior or posterior tibial tendons may cause some patients to exhibit a characteristic broad-based 'AIDS' gait, walking with the feet in inversion and extension in an attempt to diminish pain by distributing weight on the lateral margins. New bone formation at the insertion of the Achilles tendon and/or plantar fascia may often be seen radiographically as typical fluffy periostitis. Multidigit dactylitis frequently occurs and, in combination with plantar fasciitis and extensor tenosynovitis, may simulate cellulitis or pedal oedema. While synovitis of the knee is prominent, hip disease and shoulder-girdle involvement are uncommon. The prevalence of axial involvement appears to be significantly less common than in the HIV-negative forms of the disease, with sacroiliitis only occasionally being seen, and spinal ankylosis very rarely. Synovitis at the elbow and wrist may result in early flexion contractures and fusion, while asymmetric involvement of the distal interphalangeal joints may cause progressive hand deformities (Fig. 9).

The cutaneous manifestations of HIV-associated Reiter's syndrome vary considerably from individual to individual, but are often very conspicuous and sustained, in contrast to HIV-negative Reiter's syndrome. The most prominent is keratoderma blenorrhagicum, a papulosquamous and pustular eruption that usually occurs on the palms and soles. In some instances (Fig. 10), the sole is involved with a uniform dyskeratosis. The rash may spread progressively over the body in a pattern indistinguishable from pustular psoriasis, except that there is a greater tendency for involvement of the groin and intertriginous regions (inverse or sebopsoriasis). A progressive intensification of changes in the distal digits is often very prominent. Acrokeratosis is common, often associated with erythema and periungual pseudoparonychia formation. Severe alterations in the nails of the hands and feet often accompanies involvement of the distal interphalangeal joints, and is manifested clinically as onychodystrophy with or without subungual hyperkeratoses and yellow discoloration of the nails (Fig. 9). Milder degrees of onychodystrophy are present without involvement of the distal interphalangeal joints. Conjunctivitis and iritis appear to be much less prominent than in HIV-negative Reiter's syndrome.

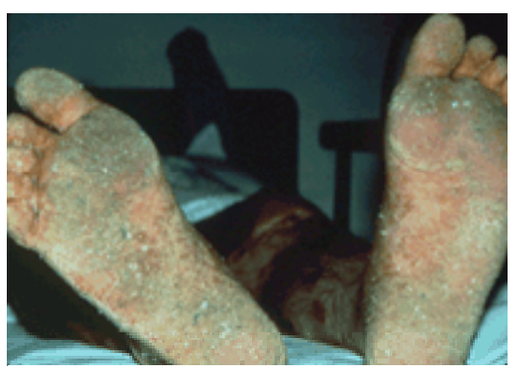


Fig. 10 Severe keratoderma blenorrhagica on the soles of an HIV-infected individual with lower limb arthritis and extensive psoriasiform skin lesions.

The cutaneous psoriasiform manifestations include lesions of psoriasis vulgaris, guttate psoriasis, keratoderma or pustular psoriasis, sebopsoriasis of the groin and axilla, and erythroderma (Fig. 11). Among the spectrum of psoriasiform skin diseases in HIV-positive patients, atypical features are present that are not seen in classic psoriasis, suggesting the existence of distinct disease mechanisms (Kaplan *et al*. 1987). The spondylarthropathy-like peripheral musculoskeletal involvement is equivalent to that described above in the section on Reiter's disease and oligoarthritis syndromes, but also includes individuals with preponderant distal interphalangeal joint disease including pencil-in-cup deformities. The severity of psoriatic arthritis in HIV-positive individuals has been emphasized (Bulbul *et al*. 1995). Enthesopathy and dactylitis, especially of the foot, are particularly prominent. Onychodystrophy is a common presenting symptom and is highly correlated with arthritis, especially in distal interphalangeal joints of the hands or feet. A significant number of patients with psoriatic skin manifestations or onychodystrophy only have limited musculoskeletal findings, such as dactylitis or enthesopathy, and do not meet the criteria for psoriatic arthritis (Berman *et al*. 1988; Espinoza *et al*. 1988; Winchester *et al*. 1988).

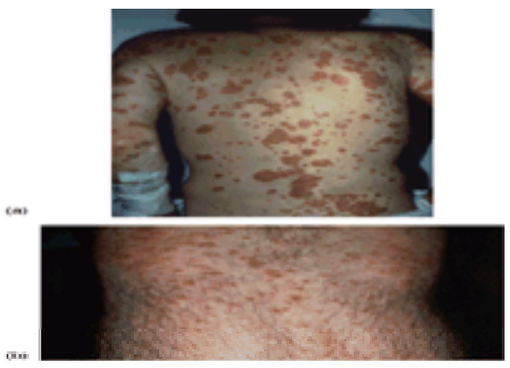


Fig. 11 Two patterns of psoriasiform skin involvement; (a) shows extensive psoriasiform involvement of the trunk in the HIV-infected patient shown in Fig. 9(a), initially presenting with features of Reiter's syndrome and progressing to severe multijoint arthritis with deformities. Extensive involvement of intertriginous areas (sebopsoriasis) was present; (b) illustrates psoriasis vulgaris in an HIV-positive person without significant arthritis.

Pathogenesis

The frequency of HLA B27 in HIV-positive Caucasian individuals with Reiter's syndrome approaches 80 per cent (Winchester *et al*. 1987; Berman *et al*. 1988; Forster *et al*. 1988), the same frequency observed in conventional Reiter's syndrome (Tiwari and Terasaki 1985). Over 30 per cent of cases are preceded by gastrointestinal infection with *Shigella flexneri, Salmonella* spp., *Yersinia enterocolitica* and *pseudotuberculosis*, and *Campylobacter jejuni* (Winchester *et al*. 1987; Forster *et al*. 1988). Temporal associations of Reiter's syndrome with infection by other organisms, including *Giardia lamblia* and atypical *Mycobacteria* spp., have also been observed but are of unknown significance (Winchester *et al*. 1987). High titres of antichlamydial antibodies have been reported in 33 per cent of HIV-positive individuals not necessarily afflicted with Reiter's syndrome, compared with 1.7 per cent in normal subjects (Gutierrez *et al*. 1990).

Whereas psoriasis vulgaris and psoriatic arthritis in the general population are associated with increased frequencies of HLA class I alleles Cw6, B13, B17, and B38 (White *et al*. 1972; Arnett 1985), no HLA associations have been detected with these disorders in HIV-infected individuals in two independent studies (Duvic *et al*. 1987; Winchester *et al*. 1988), suggesting differences in underlying pathogenesis and further arguing against a necessity for these conditions to occur at similar frequencies. In contrast, the presence of pustular psoriasis, and the arthritis that may accompany it, is associated with an increased frequency of HLA B27 in both HIV-negative and HIV-positive individuals (Arnett 1985; Winchester *et al*. 1988). It is this subgroup that most resembles Reiter's syndrome.

Progressive depletion of CD4 cells in HIV infection may be a permissive factor for greater severity, and possibly increased prevalence, of Reiter's syndrome by allowing the establishment of persistent infection with, or greater invasiveness of, gut micro-organisms such as *C. jejuni* (Perlman *et al*. 1988), or by diminishing help for B-cell dependent bacterial clearance mechanisms, which have been shown to be important in attenuating experimental chlamydial arthritis (Rank *et al*. 1988). In addition to quantitative depletion of CD4 cells, infection with HIV also leads to qualitative defects reflected by diminished interleukin 2 production following antigenic challenge (Antonen and Krohn 1986). Such defects have been observed in HLA B27-positive individuals developing Reiter's syndrome after a salmonella epidemic (Inman *et al*. 1989), and may contribute to selective microbial persistence and the development of Reiter's syndrome in HIV-infected individuals.

Biopsy of involved tissue in HIV-associated psoriasis is superficially similar to that in the idiopathic variety, demonstrating epidermal proliferation, dermal inflammatory cell infiltrate, tortuous dermal capillaries, and decreased numbers of epidermal Langerhans cells. Similar blood vessel changes, inflammatory cell infiltrate, and proliferative tissue are also present in psoriatic synovium (Espinoza *et al*. 1982). The depletion of epidermal Langerhans cells, which are CD4+ and readily infected by HIV *in vivo* (Belsito *et al*. 1984; Grelen *et al*. 1987), may be a permissive factor for the development of psoriatic lesions. In addition, products of the HIV proviral DNA may directly cause epidermal proliferation (Ensoli *et al*. 1990). In this regard, mice transgenic for the HIV *tat* gene preferentially express the tat protein in the skin and develop epidermal hyperkeratosis and acanthosis (Vogel *et al*. 1988). There is growing evidence for the participation of CD8 T cells in the pathogenesis of psoriasis and in the effect of cytokines (Duvic 1991; Winchester 1994).

The occurrence of these entities in the setting of HIV-induced immunosuppression and CD4 T-cell depletion suggests that the critical cells involved in disease pathogenesis may be residual components of the immune system, such as CD8 T lymphocytes or cells of the monocyte lineage. Indeed, immunopathological studies of synovium from HIV-infected patients with Reiter's syndrome demonstrate a lymphocytic infiltrate that is predominantly CD8+ (Espinoza *et al*. 1990). As the natural ligand for the CD8 structure on the surface of cytotoxic/suppressor cells is an MHC class I molecule such as HLA B27, we have postulated that, in Reiter's syndrome, cells of the CD8 lineage may be critically involved in an immune recognition event interacting with a particular antigen presented by HLA B27 molecules on the surface of cells of the monocyte/macrophage lineage (Meiser *et al*. 1982; Swain 1983) (Table 2). Two mechanistic hypotheses appear likely. In one, the pathogenesis of the spondylarthropathy is the same as when it occurs in a person not infected with HIV, while in the other, HIV-encoded peptides play an aetiological role in the rheumatic disease. HIV has been cultured from synovial fluid (Withington *et al*. 1987), abundant p24 antigen can be demonstrated in synovial tissue (Forster *et al*. 1988; Espinoza *et al*. 1990), and HIV DNA has been detected in synovial dendritic cells (Hughes *et al*. 1990), suggesting that the cellular infiltrate may, at least in part, be reactive to retroviral peptides. Parallel transactivation of HIV in this inflammatory milieu (Siekevitz *et al*. 1987) may act to increase HIV replication within these activated monocytes and T cells and lead to more rapid progression to AIDS. In support of this possibility is the fact that the appearance of Reiter's syndrome is an unfavourable prognostic sign, with many patients developing their first opportunistic infection within several months after the initial manifestations of Reiter's syndrome (Winchester *et al*. 1988).

Infectious agents thought to trigger psoriasis in the general population include streptococci in guttate psoriasis (Whyte and Baughman 1964) and staphylococci in pustular psoriasis (McFayden and Lyell 1971). Both organisms have also been implicated in psoriatic arthropathy (Mustakellio and Lassus 1964; Vasey *et al*. 1982). In HIV infection, the presence of psoriasis or psoriatic arthritis has been reported to be exacerbated by staphylococcal infections in almost 50 per cent of individuals (Duvic *et al*. 1987).

Treatment

The management of these entities is a difficult challenge (Duvic *et al*. 1987; Kaplan *et al*. 1989). Joint erosions, ankylosis, and osteolysis, together with chronic or recurrent enthesopathy, can rapidly lead to fibrosis, deformity, and functional disability. These are frequently compounded by generalized weakness, resulting from progressive muscle loss, and cachexia. Physical and rehabilitative therapy to maintain joint range of motion, prevent contractures, and strengthen muscle function is a central component in the comprehensive care of these patients. Optimal management involves a team-oriented approach including rheumatologists, physical and occupational therapists, and mental health care experts for management of the reactive depression that frequently accompanies the severe physical disability.

While the musculoskeletal pain and inflammation in mild cases may respond to conventional non-steroidal anti-inflammatory agents, severe manifestations of the spondylarthropathy syndromes in HIV infection are more effectively treated with phenylbutazone, given in 100 mg doses twice to three times daily. Monitoring of haematological parameters is recommended during this therapy, although no untoward cytopenias have been observed. Sulphasalazine, at doses of 0.5 to 1.5 g twice daily, may be administered together with phenylbutazone in severe cases. While controlled studies have not been performed, at least one-third of patients respond to this slow-acting drug. Maintenance sulphasalazine therapy is continued while phenylbutazone is gradually withdrawn. Intra-articular corticosteroid injection may sometimes be beneficial, and has not been associated with deleterious effects, in contrast to systemic corticosteroids which may cause extensive candidiasis and opportunistic infections in these patients (Duvic *et al*. 1987; Winchester *et al*. 1987). Prolonged and aggressive antibiotic therapy may, in theory, be of benefit in diminishing microbial persistence. Etretinate has been reported to be particularly efficacious in the treatment of both the joint and cutaneous manifestations of severe Reiter's syndrome previously unresponsive to non-steroidal anti-inflammatory drugs and topical corticosteroids, and in psoriasis (Belz *et al*. 1989; Louthrenoo 1993).

Zidovudine has been reported to improve skin disease (Ruzicka *et al*. 1987; Kaplan *et al*. 1989). In an unblinded study, 90 per cent of HIV-infected persons had either partial or complete improvement of their psoriatic skin disease following zidovudine treatment at a dosage of 1200 mg/day (Duvic *et al*. 1994). The associated arthritis in these individuals, however, did not improve with zidovudine. While zidovudine has no documented beneficial effect on the arthritic symptoms *per se*, therapy with this agent should be considered in all patients with Reiter's syndrome, whether presenting at advanced stages of immunosuppression or as the initial manifestation of HIV infection, in order to prevent the enhanced HIV replication that may secondarily occur as a result of CD4 cell activation. Zidovudine-induced myopathy may

become a confounding variable.

Methotrexate and other immunosuppressive agents are capable of strikingly ameliorating the skin and joint disease (Winchester *et al*. 1987; Maurer *et al*. 1994), but they should be used only with great caution because their use has been followed by the development of frank AIDS and death from opportunistic infection (Winchester *et al*. 1987). However, methotrexate administered to three individuals with HIV-associated psoriatic arthritis did not worsen the underlying immune deficiency in two of the three (Maurer *et al*. 1994), suggesting that it may have a limited place when used cautiously. Phototherapy has been asserted to be helpful and without evident problems (Ranki *et al*. 1991; Meola *et al*. 1993), but has also been associated with the appearance or exacerbation of Kaposi's sarcoma (Duvic *et al*. 1987) and the possibility of HIV activation has been emphasized (Duvic 1995). Studies in mice transgenic for HIV have shown that HIV may be activated by ultraviolet radiation, raising concerns for the use of phototherapy in humans (Morrey *et al*. 1991; Vogel *et al*. 1992). These mainstays of therapy in the HIV-negative patient with psoriasis or psoriatic arthritis are probably best reserved for very difficult situations that have failed to respond to all other approaches, because their mode of action parallels the consequences of HIV infection, although continued experimental experience with these therapies may define a risk–benefit balance that is acceptable in certain situations.

Chapter References

Adebajo, A. and Davis, P. (1994). Rheumatic diseases in African blacks. *Seminars in Arthritis and. Rheumatism*, **24**, 139–53.

Altman, E.M., Centeno, L.V., Mahal, M., and Bielory, L. (1994). AIDS-associated Reiter's syndrome. *Annals of Allergy*, **72**, 307–16.

Amor, B. (1989). Rheumatoid arthritis and AIDS. *Journal of Rheumatology*, **16**, 845.

Anders, K.H., Latta, H., Chang, B.S., Tomiyasu, U., Quddusi, A.S., and Vinters, H.V. (1989). Lymphoid granulomatosis and malignant lymphoma of the central nervous system in acquired immunodeficiency syndrome. *Human Pathology*, **20**,326–34.

Angulo, J.C., Lopez, J.I., Garcia, M.E., Peiro, J., and Flores, N. (1994). HIV infection presenting as renal polyarteritis nodosa. *International Urology and Nephrology*, **26**, 637–41.

Antonen, J. and Krohn, K. (1986). Interleukin 2 production in HTLV III/LAV infection. Evidence of defective antigen-induced, but normal mitogen-induced, IL-2 production. *Clinical and Experimental Immunology*, **65**, 489–96.

Arnett, F.C. (1985). Psoriatic arthritis: relationship to other spondyloarthropathies. In *Psoriatic arthritis* (ed. L.H. Gerber and L.R. Espinoza), p. 95. Grune and Stratton, Orlando, FL.

Arthur, L.O. *et al*. (1992). Cellular proteins bound to immunodeficiency viruses: implications for pathogenesis and vaccines. *Science*, **258**, 1935–8.

Bacchetti, P. and Moss, A.R. (1989). Incubation time of AIDS in San Francisco. *Nature*, **338**, 251–3.

Baguley, E., Wolfe, C., and Hughes, G.R. (1988). Dermatomyositis in HIV infection. *British Journal of Rheumatology*, **27**, 493–500.

Bardin, T. *et al*. (1987). Necrotizing vasculitis in human immunodeficiency virus (HIV) infection. (Abstract.) *Arthritis and Rheumatism*, **30**, S105.

Bass, P.S. *et al*. (1994). AIDS presenting as focal segmental membranous glomerulopathy. *Journal of Clinical Pathology*, **47**, 179–81.

Belsito, D.V., Sanchez, M.R., Baer, R.L., Valentine, F., and Thorbecke, G.J. (1984). Reduced Langerhans cell IA antigen and ATPase activity in patients with AIDS. *New England Journal of Medicine*, **310**, 1279–82.

Belz, J., Breneman, D.L., Nordlund, J.J., and Solinger, A. (1989). Successful treatment of a patient with Reiter's syndrome and acquired immunodeficiency syndrome using etretinate. *Journal of the American Academy of Dermatology*, **20**, 898–903.

Berman, A. *et al*. (1988). Rheumatic manifestations of human immunodeficiency virus infection. *American Journal of Medicine*, **85**, 59–64.

Berman, A., Reboredo, G., Spindler, A., Lasala, M.E., Lopez, H., and Espinoza, L.R. (1991). Rheumatic manifestations in populations at risk for HIV infection: the added effect of HIV. *Journal of Rheumatology*, **18**, 1564–7.

Bernard, C., Exquis, B., Reber, A., and de Moerloose, P. (1990). Determination of anticardiolipin and other antibodies in HIV-1 infected patients. *Journal of Acquired Immune Deficiency Syndromes*, **3**, 536–9.

Bijlsma, J.W., Derksen, R.W., Huber-Bruning, O., and Borleffs, J.C. (1988). Does AIDS 'cure' rheumatoid arthritis? *Annals of the Rheumatic Diseases*, **47**, 350–1.

Blanche, P. *et al*. (1993). Acute arthritis and human immunodeficiency virus infection in Rwanda. *Journal of Rheumatology*, **20**, 2123–7.

Blumenthal, D.R., Zucker, J.R., and Hawkins, C.C. (1990). *Mycobacterium avium* complex-induced septic arthritis and osteomyelitis in a patient with the acquired immunodeficiency syndrome. *Arthritis and Rheumatism*, **33**,757.

Boix, V., Tovar, J., and Martin-Hidalgo, A. (1990). Candida spondylodiscitis, chronic illness due to heroin analgesia in an HIV positive person. *Journal of Rheumatology*, **17**, 563.

Bourgoignie, J.J. (1990). Renal complications of human immunodeficiency virus type 1. *Kidney International*, **37**, 1571–84.

Bourgoignie, J.J., Meneses, R., and Pardo, V. (1988*a*). The nephropathy related to acquired immune deficiency syndrome. *Advances in Nephrology from the Necker Hospital*, **17**, 113–26.

Bourgoignie, J.J., Meneses, R., Ortiz, C., Jaffe, D., and Pardo, V. (1988 *b*). The clinical spectrum of renal disease associated with human immunodeficiency virus. *American Journal of Kidney Disease*, **12**, 131–7.

Bourgoignie, J.J., Ortiz Interian, C., Green, D.F., and Roth, D. (1989). Race, a cofactor in HIV-1-associated nephropathy. *Transplantation Proceedings*, **21**, 3899–901.

Bulbul, R., Williams, W.V., and Schumacher, H.R., Jr. (1995). Psoriatic arthritis. Diverse and sometimes highly destructive. *Postgraduate Medicine*, **97**, 97–9, 103–6.

Burgoyne, M., Agudelo, C., and Pisko, E. (1992). Chronic syphilitic polyarthritis mimicking systemic lupus erythematosus/rheumatoid arthritis as the initial presentation of human immunodeficiency virus infection. *Journal of Rheumatology*, **19**, 313–15.

Burman, W.J., Cohn, D.L., Reves, R.R., and Wilson, M.L. (1995). Multifocal cellulitis and monoarticular arthritis as manifestations of *Helicobacter cinaedi* bacteremia. *Clinical Infectious Diseases*, **20**, 564–70.

Buskila, D. and Tenenbaum, J. (1989). Septic bursitis in human immunodeficiency virus infection. *Journal of Rheumatology*, **16**, 1374.

Buskila, D., Gladman, D.D., Gilmore, J., and Salit, I.E. (1991). Behçet's disease in a patient with immunodeficiency virus infection. *Annals of the Rheumatic Diseases*, **50**, 115–16.

Buyon, J.P. *et al*. (1989). Acquired congenital heart block. *Journal of Clinical Investigation*, **84**, 627–34.

Calabrese, L.H. (1989). The rheumatic manifestations of infection with the human immunodeficiency virus. *Seminars in Arthritis and Rheumatism*, **18**, 225.

Calabrese, L.H. (1991). Vasculitis and infection with the human immunodeficiency virus. *Rheumatic Diseases Clinics of North America*, **17**, 131–47.

Calabrese, L.H., Wilke, W.S., Perkins, A.D., and Tubbs, R.R. (1989). Rheumatoid arthritis complicated by infection with the human immunodeficiency virus and the development of Sjögren's syndrome. *Arthritis and Rheumatism*, **32**, 1453–7.

Calabrese, L.H., Kelley, D.M., Myers, A., O'Connell, M., and Easley, K. (1991). Rheumatic symptoms and human immunodeficiency virus infection. The influence of clinical and laboratory variables in a longitudinal cohort study. *Arthritis and Rheumatism*, **34**, 257–63.

Canoso, R.T., Zon, L.I., and Groopman, J.E. (1987). Anticardiolipin antibodies associated with HTLV-III infection. *British Journal of Haematology*, **65**, 495–8.

Carbone, L., D'Agati, V., Cheng, J.T., and Appel, G.B. (1989). Course and prognosis of human immunodeficiency virus-associated nephropathy. *American Journal of Medicine*, **87**, 389–95.

Chander, P. and Treser, G. (1987). Ultrastructural markers of AIDS nephropathy. (Abstract.) *Kidney International*, **31**, 335.

Chariot, P. *et al*. (1994). Determination of the blood lactate:pyruvate ratio as a noninvasive test for the diagnosis of zidovudine myopathy. *Arthritis and Rheumatism*, **37**, 583–6.

Cheng Mayer, C., Shioda, T., and Levy, J.A. (1991). Host range, replicative, and cytopathic properties of human immunodeficiency virus type 1 are determined by very few amino acid changes in tat and gp120. *Journal of Virology*, **65**, 6931–41.

Chisholm, D.M. and Mason, D.K. (1968). Labial salivary gland biopsy in Sjögren's disease. *Journal of Clinical Pathology*, **21**, 656.

Chren, M.M., Silverman, R.A., Sorsensen, R.U., and Elmets, C.A. (1989). Leukocytoclastic vasculitis in a patient infected with human immunodeficiency virus. *Journal of the American Academy of Dermatology*, **21**, 1161–4.

Clark, M., Kinsolving, M., and Chernoff, D. (1989). The prevalence of arthritis in two HIV-infected cohorts. (Abstract.) *Arthritis and Rheumatism*, **32**, S585.

Cockerell, C.J. (1991). Human immunodeficiency virus infection and the skin. A crucial interface. *Archives of Internal Medicine*, **151**, 1295–303.

Connor, E. *et al*. (1988). Acquired immunodeficiency syndrome-associated renal disease in children. *Journal of Pediatrics*, **113**, 39–44.

Connor, R.I., Mohri, H., Cao, Y., and Ho, D.D. (1993). Increased viral burden and cytopathicity correlate temporally with CD4+ T-lymphocyte decline and clinical progression in human immunodeficiency virus type 1-infected individuals. *Journal of Virology*, **67**, 1772–7.

Cooper, L.M. and Patterson, J.A.K. (1989). Allergic granulomatosis and angiitis of Churg–Strauss: case report in a patient with antibodies to human immunodeficiency virus and hepatitis B virus. *International Journal of Dermatology*, **28**, 597.

Couderc, L.J., D'Agay, M.F., Danon, F., Harzic, M., Brocheriou, C., and Clauvel, J.P. (1987). Sicca complex and infection with human immunodeficiency virus. *Archives of Internal Medicine*, **147**, 898–901.

Covelli, M., Lapadula, G., Pipitone, N., Numo, R., and Pipitone, V. (1993). Isolated sternoclavicular joint arthritis in heroin addicts and/or HIV positive patients: three cases. *Clinical Rheumatology*, **12**, 422–5.

Crawford, E.J.P. and Baird, P.R.E. (1987). An orthopedic presentation of AIDS: brief report. *Journal of Bone and Joint Surgery*, **69B**, 672.

Cuellar, M.L., Silveira, L.H., and Espinoza, L.R. (1994). Recent developments in psoriatic arthritis. *Current Opinion in Rheumatology*, **6**, 378–84.

D'Agati, V., Suh, J.I., Carbone, L., Cheng, J.T., and Appel, G. (1989). Pathology of HIV-associated nephropathy: a detailed morphologic and comparative study. *Kidney International*, **35**, 1358–70.

Dalakas, M.C. and Pezeshkpour, G.H. (1988). Neuromuscular diseases associated with human immunodeficiency virus infection. *Annals of Neurology*, **23**, S38–S48.

Dalakas, M.C., Pezeshkpour, G.H., Gravell, M., and Sever, J.L. (1986). Polymyositis associated with AIDS retrovirus. *Journal of the American Medical Association*, **256**, 2381–3.

Dalakas, M.C., Pezeshkpour, G.H., and Flaherty, M. (1987). Progressive nemaline (rod) myopathy associated with HIV infection. *New England Journal of Medicine*, **317**, 1602–3.

Dalakas, M.C., Illa, I., Pezeshkpour, G.H., Laukaitis, J.P., Cohen, B., and Griffin, J.L. (1990). Mitochondrial myopathy caused by long-term zidovudine therapy. *New England Journal of Medicine*, **322**, 1098–105.

De Wolf, F. *et al*. (1988). Numbers of CD4+ cells and the levels of core antigens of and antibodies to the human immunodeficiency virus as predictors of AIDS among seropositive homosexual men. *Journal of Infectious Diseases*, **158**, 615–22.

Dorsett, B.H., Cronin, W., and Joachim, H.L. (1990). Presence and prognostic significance of antilymphocyte antibodies in symptomatic and asymptomatic human immunodeficiency virus infection. *Archives of Internal Medicine*, **150**, 1025–8.

Duvic, M. (1991). Papulosquamous disorders associated with human immunodeficiency virus infection. *Dermatologic Clinics*, **9**, 523–30.

Duvic, M. (1995). Human immunodeficiency virus and the skin: selected controversies. *Journal of Investigative Dermatology*, **105**, 117S–121S.

Duvic, M., Johnson, T.M., Rapini, R.P., Freeze, T., Brewton, G., and Rios, A. (1987). Acquired immunodeficiency syndrome associated psoriasis and Reiter's syndrome. *Archives of Dermatology*, **123**, 1622–32.

Duvic, M., Crane, M.M., Conant, M., Mahoney, S.E., Reveille, J.D., and Lehrman, S.N. (1994). Zidovudine improves psoriasis in human immunodeficiency virus-positive males. *Archives of Dermatology*, **130**, 447–51.

Dwyer, E., Itescu, S., and Winchester, R. (1993). Characterization of the primary structure of T-cell receptor b-chains in cells infiltrating the salivary gland in the sicca syndrome of HIV-1 infection. Evidence of antigen-driven clonal selection suggested by restricted combinations of VbJb gene segment usage and shared somatically encoded amino acid residues. *Journal of Clinical Investigation*, **92**, 495–502.

Eidelberg, D., Sotrel, A., Horoupian, D.S., Neumann, P.E., Pumarola-Sune, T., and Price, R.W. (1986). Thrombotic cerebral vasculopathy associated with herpes zoster. *Annals of Neurology*, **19**, 7–14.

Ensoli, B., Barillari, G., Salahuddin, S.Z., Gallo, R.C., and Wong-Staal, F. (1990). Tat protein of HIV-1 stimulates growth of cells derived from Kaposi's sarcoma lesions of AIDS patients. *Nature*, **345**, 84–6.

Espinoza, L.R., Vasey, F.B., Espinoza, T.S., and Germain, B.F. (1982). Vascular changes in psoriatic synovium. A light and electron microscopic study. *Arthritis and Rheumatism*, **25**, 677–84.

Espinoza, L.R., Berman, A., Vasey, F.B., Cahalin, C., Nelson, R., and Germain, B.F. (1988). Psoriatic arthritis and acquired immunodeficiency syndrome. *Arthritis and Rheumatism*, **31**,1034–40.

Espinoza, L.R. *et al*. (1990). HIV-associated arthropathy: HIV antigen demonstration in the synovial membrane. *Journal of Rheumatology*, **17**, 1195–201.

Espinoza, L.R. *et al*. (1991). Characteristics and pathogenesis of myositis in human immunodeficiency virus infection—distinction from azidothymidine-induced myopathy. *Rheumatic Diseases Clinics of North America*, **17**, 117–29.

Fahey, J.L. (1986). Immunologic aspects of human immunodeficiency virus infection and AIDS. *Clinical Aspects of Autoimmunity*, **1**, 12–33.

Farthing, C.F., Staughton, R.C.D., and Rowland Payne, C.M.E. (1985). Skin disease in homosexual patients with acquired immune deficiency syndrome (AIDS) and lesser forms of human T cell leukaemia virus (HTLV-III) disease. *Clinical and Experimental Dermatology*, **10**, 3.

Fauci, A.S. (1988). The human immunodeficiency virus: infectivity and mechanisms of pathogenesis. *Science*, **239**, 617–22.

Felber, B.K., Hadzopoulou-Cladaras, M., Cladaras, C., Copeland, T., and Pavlakis, G.N. (1989). Rev protein of human immunodeficiency virus type 1 affects the stability and transport of the viral mRNA. *Proceedings of the National Academy of Sciences* (*USA*), **86**, 1495–9.

Fernandez-Vina, M., Shumway, JW., and Stastny, P. (1990). DNA typing for class II HLA antigens with allele-specific or group specific amplification. II. Typing for alleles of the DRw52-associated group. *Human Immunology*, **28**, 51–64.

Fischl, M.A., Dickinson, G.M., Scott, G.B., Klunias, N., Fletcher, M.A., and Parks, W. (1987). Evaluation of heterosexual partners, children and household contacts of adults with AIDS. *Journal of the American Medical Association*, **257**, 640.

Forster, S.M. *et al*. (1988). Inflammatory joint disease and human immunodeficiency virus infection. *British Medical Journal*, **296**, 1625–7.

Fox, R.I., Robinson, C.A., Curd, J.G., Kozin, F., and Howell, F.V. (1986). Sjögren's syndrome: criteria for classification. *Arthritis and Rheumatism*, **29**, 577.

Friedland, G.H. *et al*. (1986). Lack of transmission of HTLV-III/LAV infection to household contacts of patients with AIDS or ARC with oral candidiasis. *New England Journal of Medicine*, **314**, 344.

Gaut, P., Wong, P.K., and Meyer, R.D. (1988). Pyomyositis in a patient with the acquired immunodeficiency syndrome. *Archives of Internal Medicine*, **148**, 1608–10.

Gherardi, R., Lebargy, F., Gaulard, P., Mhiri, C., Bernaudin, J.F., and Gray, F. (1989). Necrotizing vasculitis and HIV replication in peripheral nerves. *New England Journal of Medicine*, **321**, 685–6.

Gherardi, R. *et al*. (1993). The spectrum of vasculitis in human immunodeficiency virus-infected patients. A clinicopathologic evaluation. *Arthritis and Rheumatism*, **36**, 1164–74.

Goedert, J.J. *et al*. (1987). Effect of T4 count and cofactors on the incidence of AIDS in homosexual men infected with human immunodeficiency virus. *Journal of the American Medical Association*, **257**, 331–4.

Golden, M.P., Hammer, S.M., Wanke, C.A., and Albrecht, M.A. (1994). Cytomegalovirus vasculitis. Case reports and review of the literature. *Medicine*, **73**, 246–55.

Goldenberg, D.L. (1991). Septic arthritis and other infections of rheumatologic significance. *Rheumatic Diseases Clinics of North America*, **17**,149–56.

Goldenberg, D.L. (1994). Bacterial arthritis. *Current Opinion in Rheumatology*, **6**, 394–400.

Golding, H. *et al*. (1988). Identification of homologous regions in HIV-Igp 41 and human MHC class II b $_1$ domain. I. Monoclonal antibodies against the gp41-derived peptide and patients' sera react with native HLA class II antigens, suggesting a role for autoimmunity in the pathogenesis of AIDS. *Journal of Experimental Medicine*, **167**, 914–23.

Greene, W.C. (1991). The molecular biology of HIV type 1 infection. *New England Journal of Medicine*, **324**, 308–17.

Grelen, V., Schnitt, D., Degritter-Dambuyant, C., Nicholas, J.F., and Thivolet, J. (1987). AIDS and Langerhans cells: CD1, CD4 and HLA class II antigen expression. *Journal of Investigative Dermatology*, **89**, 324A.

Gutierrez, F. *et al*. (1990). Serologic evidence for chlamydia infection in human immunodeficiency virus-infected patients. (Abstract.) *Revista Mexicana de Rheumatologia*, **5**, 62.

Gutierrez, C., Cruz, L., Olive, A., Tena, X., Romeu, J., and Raventos, A. (1993). Salmonella septic arthritis in HIV patients. *British Journal of Rheumatology*, **32**, 88.

Guy, B. *et al*. (1987). HIV F/3_ orf encodes a phosphorylated GTP-binding protein resembling an oncogene product. *Nature*, **330**, 266–9.

Guyot, D.R., Manoli, A.II, and Kling, G.A. (1987). Pyogenic sacroiliitis in IV drug abusers. *American Journal of Roentgenology*, **149**, 1209–11.

Haddoum, F., Dosquet, P., Mougenot, B., Bourrat, E., Viron, B., and Mignon, F. (1987). Nephrotic syndrome disclosing AIDS. *Presse Medicale*, **16**, 1373.

Haverkos, H.W., Gotlieb, M.S., Killen, J.Y., and Edelman, R. (1985). Classification of HTLV-III/LAV-related diseases. *Journal of Infectious Diseases*, **152**, 1095.

Heller, H.M. and Fuhrer, J. (1991). Disseminated sporotrichosis in patients with AIDS: case report and review of the literature. *AIDS*, **5**, 1243–6.

Ho, G., Jr. (1993). Bacterial arthritis. *Current Opinion in Rheumatology*, **5**, 449–53.

Hochberg, M.C., Fox, R., Nelson, K.E., and Saah, A. (1990). HIV infection is not associated with Reiter's syndrome: data from the Johns Hopkins Multicenter AIDS cohort study. *AIDS*, **4**, 1149–51.

Hughes, R.A., Macatonia, S.E., Rowe, I.F., Keat, A.C., and Knight, S.C. (1990). The detection of HIV DNA in dendritic cells from the joints of patients with aseptic arthritis. *British Journal of Rheumatology*, **29**, 166–70.

Hughes, R.A., Rowe, I.F., Shanson, D., and Keat, A.C. (1992). Septic bone, joint and muscle lesions associated with human immunodeficiency virus infection. *British Journal of Rheumatology*, **31**, 381–8.

Hwang, S.S., Boyle, T.J., Lyerly, H.K., and Cullen, B.R. (1991). Identification of the envelope V3 loop as the primary determinant of cell tropism in HIV-1. *Science*, **253**, 71–3.

Ifudu, O., Rao, T.K.S., Tan, C.C., Fleischman, H., Chirgwin, K., and Friedman, E.A. (1994). Zidovudine improves prognosis in HIV-associated nephropathy. *Journal of the American Society of Nephrology*, **4**, 277.

Ingulli, E., Tejani, A., Fikrig, S., Nicastri, A., Chen, C.K., and Pomrantz, A. (1991). Nephrotic syndrome associated with acquired immunodeficiency syndrome in children. *Journal of Pediatrics*, **119**, 710–16.

Inman, R.D., Chiu, B., Johnson, M.E., Vas, S., and Falk, J. (1989). HLA class I-related impairment in IL-2 production and lymphocyte response to microbial antigens in reactive arthritis. *Journal of Immunology*, **142**, 4256–360.

Itescu, S. (1991). Diffuse infiltrative lymphocytosis syndrome in human immunodeficiency virus infection—a Sjögren's-like disease. In *Rheumatic disease clinics of North America*, Vol. 17, No. 1: *AIDS and rheumatic disease* (ed. R. Winchester), pp. 99–115. W.B. Saunders, Philadelphia.

Itescu, S., Brancato, L.J., and Winchester, R. (1989). A sicca syndrome in HIV infection: association with HLA-DR5 and CD8 lymphocytosis. *Lancet*, **ii**, 466–8.

Itescu, S. *et al*. (1990). A diffuse infiltrative CD8 lymphocytosis syndrome in human immunodeficiency virus (HIV) infection: a host immune response associated with HLA-DR5. *Annals of Internal Medicine*, **112**, 3–10.

Itescu, S., Dalton, J., Zhang, H., and Winchester, R. (1993). Tissue infiltration in a CD8 lymphocytosis syndrome associated with human immunodeficiency virus-1 infection has the phenotypic appearence of an antigenically driven response. *Journal of Clinical Investigation*, **91**, 2216–25.

Itescu, S., Simonelli, P.F., Winchester, R.J., and Ginsberg, H.S. (1994 *a*). Human immunodeficiency virus type 1 strains in the lungs of infected individuals evolve independently from those in peripheral blood and are highly conserved in the C-terminal region of the envelope V3 loop. *Proceedings of the National Academy of Sciences* (*USA*), **91**, 11378–82.

Itescu, S., Rose, S., Dwyer, E., and Winchester, R. (1994 *b*). Certain HLA-DR5 and -DR6 major histocompatibility complex class II alleles are associated with a CD8 lymphocytic host response to human immunodeficiency virus type 1 characterized by low lymphocyte viral strain heterogeneity and slow disease progression. *Proceedings of the National Academy of Sciences* (*USA*),**91**, 11472–6.

Jacobson, M.A., Geller, H., and Chambers, H. (1988). *Staphylococcus aureus* bacteremia and recurrent staphylococcal infection in patients with acquired immunodeficiency syndrome and AIDS-related complex. *American Journal of Medicine*, **85**, 172.

Javaherian, K. *et al*. (1989). Principal neutralizing domain of the human immunodeficiency virus type 1 envelope protein. *Proceedings of the National Academy of Sciences* (*USA*), **86**, 6768–72.

Johnson, A.H., Rosen-Bronson, S., and Hurley, C.K. (1989). Heterogeneity of the HLA-D region in American Blacks. *Transplantation Proceedings*, **21**, 3872–3.

Jurgensen, O., Altermatt, H.J., von Overbeck, J., and Berthold, H. (1995). Necrotizing vasculitis of the tongue. A contribution to the differential diagnosis of ulcerative mucosal changes in HIV-infected patients. *Schweizer Monatsschrift für Zahnmedizin*, **105**, 54–62.

Kalden, J.R. and Gay, S. (1994). Retroviruses and autoimmune rheumatic diseases. *Clinical and. Experimental Immunology*, **98**,1–5.

Kaplan, M.H., Sadick, N., McNutt, S., Meltzer, M., Sarngadharan, M.G., and Pahwa, S. (1987). Dermatologic findings and manifestations of acquired immunodeficiency syndrome (AIDS). *Journal of the American Academy of Dermatology*, **16**, 485–506.

Kaplan, M.H., Sadick, N.S., Weider, J., Farber, B.F., and Neidt, G.W. (1989). Anti-psoriatic effects of zidovudine in HIV-associated psoriasis. *Journal of the American Academy of Dermatology*, **20**, 76–82.

Kaslow, R.A., Duquesnoy, R., VanRaden, M., Kingsley, L., and Mann, D. (1990). A1, Cw7, B8, DR3 HLA antigen combination associated with rapid decline of T-helper lymphocytes in HIV-1 infection. *Lancet*, **335**, 927–30.

Kellner, H., Fuessl, H.S., and Herzer, P. (1994). Seronegative spondylarthropathies in HIV-infected patients: further evidence of uncommon clinical features. *Rheumatology International*, **13**, 211–13.

Kerr, L.D. and Spiera, H. (1991). The coexistence of active classic rheumatoid arthritis and AIDS. *Journal of Rheumatology*, **18**, 1739–40.

Kirchhoff, F., Greenough, T.C., Brettler, D.B., Sullivan, J.L., and Desrosiers, R.C. (1995). Brief report: absence of intact nef sequences in a long-term survivor with nonprogressive HIV-1 infection. *New England Journal of Medicine*, **332**, 228–32.

Klaassen, R.J. *et al*. (1992). Anti-neutrophil cytoplasmic autoantibodies in patients with symptomatic HIV infection. *Clinical and Experimental Immunology*, **87**, 24–30.

Klatzmann, D. *et al*. (1984). Selective tropism of lymphadenopathy-associated virus (LAV) for helper-inducer T lymphocytes. *Science*, **225**, 59–63.

Koenig, S. *et al*. (1986). Detection of AIDS virus in macrophages in brain tissue from AIDS patients with encephalopathy. *Science*, **233**, 1089–93.

Koenig, S. *et al*. (1990). Mapping the fine specificity of a cytolytic T cell response to HIV-1 nef protein. *Journal of Immunology*, **145**, 127–35.

Koup, R.A. *et al*. (1994). Temporal association of cellular immune responses with the initial control of viremia in primary HIV-1 syndrome. *Journal of Virology*, **68**, 4650–5.

Kowalski, M. *et al*. (1989). Antibodies to CD4 in individuals infected with human immunodeficiency virus type 1. *Proceedings of the National Academy of Sciences* (*USA*), **86**, 3346–50.

Lane, H.C. and Fauci, A.S. (1985). Immunologic abnormalities in the acquired immune deficiency syndrome. *Annual Review of Immunology*, **3**, 477.

Lane, H.C. *et al*. (1983). Abnormalities of B cell activation and immunoregulation in patients with the acquired immunodeficiency syndrome. *New England Journal of Medicine*, **309**, 453–8.

Lawrence, J.M., Osborn, T.G., Paro, R., Eaton, C., Hyers, T.M., and Moore, T.L. (1991). Septic arthritis caused by hemophilus influenza type B in a patient with HIV-1 infection. *Journal of Rheumatology*, **18**, 1772–3.

Leon, M., Ramos, M., Saavedra, J., Dominguez, A., Ferrer, T., and Pujol, E. (1994). *S. pneumoniae* sternoclavicular arthritis in a patient with HIV infection. *Anales de Medicina Interna*, **11**, 395–7.

Libman, B.S., Quismorio, F.P., Jr., and Stimmler, M.M. (1995). Polyarteritis nodosa-like vasculitis in human immunodeficiency virus infection. *Journal of Rheumatology*, **22**, 351–5.

Lipstein-Kresch, E., Isenberg, H.D., Singer, C., Cooke, O., and Greenwald, R.A. (1985). Disseminated sporothrix schenkii infection with arthritis in a patient with acquired immunodeficiency syndrome. *Journal of Rheumatology*, **12**, 805–8.

Looney, D.J. *et al*. (1988). Type-restricted neutralization of molecular clones of human immunodeficiency virus. *Science*, **241**, 357–9.

Louthrenoo, W. (1993). Successful treatment of severe Reiter's syndrome associated with human immunodeficiency virus infection with etretinate. Report of 2 cases. *Journal of Rheumatology*, **20**, 1243–6.

Lui, K.-J., Darrow, W.W., and Rutherford, G.W., III (1988). A model-based estimate of the mean incubation period for AIDS in homosexual men. *Science*, **240**, 1333–5.

Malbec, D., Pines, E., Boudon, P., Lusina, D., and Choudat, L. (1994). CD8 hyperlymphocytosis syndrome and human immunodeficiency virus infection: 5 cases. *Revue de Medecine Interne*, **15**, 630–3.

Marcellin, P. *et al*. (1991). Latent hepatitis B virus (HBV) infection in systemic necrotizing vasculitis. *Clinical and Experimental Rheumatology*, **9**, 23–8.

Margolick, J.B., Volkman, D.J., Folks, T.M., and Fauci, A.S. (1987). Amplification of HTLV-III/LAV infection by antigen-induced activation of T cells and direct suppression by virus of lymphocyte blastogenic responses. *Journal of Immunology*, **138**, 1719–23.

Marks, C. and Kuskov, S. (1995). Pattern of arterial aneurysms in acquired immunodeficiency disease. *World Journal of Surgery*, **19**, 127–32.

Martz, E. (1986). LFA-1 and other accessory molecules functioning in adhesions of T and B lymphocytes. *Human Immunology*, **18**, 3–37.

Masters, D.L. and Lentino, J.R. (1984). Cervical osteomyelitis related to *Nocardia asteroides*. *Journal of Infectious Diseases*, **149**, 824.

Maurer, T.A., Zackheim, H.S., Tuffanelli, L., and Berger, T.G. (1994). The use of methotrexate for treatment of psoriasis in patients with HIV infection. *Journal of the American Academy of Dermatology*, **31**, 372–5.

Mayer-Siuta, R., Keil, L.B., and De Bari, V.A. (1988). Autoantibodies and circulating immune complexes in subjects infected with human immunodeficiency virus. *Medical Microbiology and Immunology*, **177**, 189–94.

McFayden, T. and Lyell, A. (1971). Coagulase positive staphylococci in pustular psoriasis: evidence for bacteremia and good response to treatment. In *International symposium on psoriasis* (ed. E. Farber and A. Cox), pp. 79–85. Stanford University Press.

Meiser, S.C., Schlossman, S.F., and Reinhertz, E.L. (1982). Clonal analysis of human cytotoxic T lymphocytes: T4+ and T8+ effector cells recognize products of different major histocompatibility complex regions. *Proceedings of the National Academy of Sciences* (*USA*), **79**, 4395.

Meltzer, M.S. and Gendelman, H.E. (1992). Mononuclear phagocytes as targets, tissue reservoirs, and immunoregulatory cells in human immunodeficiency virus disease. *Current Topics in Microbiology and Immunology*, **181**, 239–63.

Meola, T., Soter, N.A., Ostreicher, R., Sanchez, M., and Moy, J.A. (1993). The safety of UVB phototherapy in patients with HIV infection. *Journal of the American Academy of Dermatology*, **29**, 216–20.

Mijiyawa, M. (1993). Spondyloarthropathies in patients attending the rheumatology unit of Lomé Hospital. *Journal of Rheumatology*, **20**, 1167–9.

Mitsuya, H. and Broder, S. (1986). Inhibition of the in vitro infectivity and cytopathic effect of human T-lymphotrophic virus type III/lymphadenopathy-associated virus (HTLV-III/LAV) by 2,3-dideoxynucleosides. *Proceedings of the National Academy of Sciences* (*USA*), **83**, 1911–15.

Molina, J.F. *et al*. (1995). Coexistence of human immunodeficiency virus infection and systemic lupus erythematosus. *Journal of Rheumatology*, **22**, 347–50.

Monteagudo, I., Rivera, J., Lopez-Longo, J., Cosin, J., Garcia-Monforte, A., and Carreno, L. (1991). AIDS and rheumatic manifestations in patients addicted to drugs. An analysis of 106 cases. *Journal of Rheumatology*, **18**, 1038–41.

Mora, C.A. *et al*. (1988). Seroprevalence of antibodies to HTLV-1 in patients with chronic neurological disorders other than tropical spastic paraparesis. *Annals of Neurology*, **23**, S192–S195.

Morrey, J.D. *et al*. (1991). *In vivo* activation of human immunodeficiency virus type 1 long terminal repeat by UV type A (UV-A) light plus psoralen and UV-B light in the skin of transgenic mice. *Journal of Virology*, **65**, 5045–51.

Moss, A.R. *et al*. (1988). Seropositivity for HIV and the development of AIDS or AIDS-related conditions: three year follow-up of the San Francisco General Hospital cohort. *British Medical Journal*, **296**, 745–50.

Muesing, M., Smith, D.H., and Capon, D.J. (1987). Regulation of mRNA accumulation by a human immunodeficiency virus transactivator protein. *Cell*, **48**, 691–701.

Munoz Fernandez, S. *et al*. (1991). Osteoarticular infection associated with the human immunodeficiency virus. *Clinical and Experimental Rheumatology*, **9**, 489–93.

Mustakellio, K.K. and Lassus, A. (1964). Staphylococcal a-antitoxin in psoriatic arthropathy. *British Journal of Dermatology*, **76**, 544.

Nabel, G.J., Rice, S.A., Knipe, D.M., and Baltimore, D. (1988). Alternative mechanisms for activation of human immunodeficiency virus enhancer in T cells. *Science*, **239**, 1299–302.

Nakajima, K. *et al*. (1989). Induction of IL-6 (B cell stimulatory factor 2/IFN-b2) production by HIV. *Journal of Immunology*, **142**, 531–6.

O'Brien, W.A. *et al*. (1990). HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp120 outside the CD4-binding domain. *Nature*, **348**, 69–73.

Obuch, M.L., Maurer, T.A., Becker, B., and Berger, T.G. (1992). Psoriasis and human immunodeficiency virus infection. *Journal of the American Academy of Dermatology*, **27**, 667–73.

Ostlere, L.S., Harris, D., Johnson, M., and Rustin, M.H. (1993). Gastrointestinal and cutaneous vasculitis associated with gonococcal infection in an HIV-seropositive patient. *Journal of the American Academy of Dermatology*, **29**, 276–8.

Pahwa, S. *et al*. (1985). Influence of the human T-lymphotropic virus/lymphadenopathy-associated virus and functions of human lymphocytes: evidence of immunosuppressive effects and polyclonal B-cell activation by banded viral and lymphocyte preparations. *Proceedings of the National Academy of Sciences* (*USA*), **82**, 8198–202.

Panegyres, P.K., Tan, N., Kakulas, B.A., Armstrong, J.A., and Hollingsworth, P. (1988). Necrotising myopathy and zidovudine. *Lancet*, **i**, 1050–1.

Pappo, A.S., Buchanan, G.R., and Johnson, A. (1989). Septic arthritis in children with hemophilia. *American Journal of Diseases of Children*, **143**, 1226.

Pardo, V. *et al*. (1987). AIDS-related glomerulopathy: occurrence in specific risk groups. *Kidney International*, **31**, 1167–73.

Perlman, D.M. *et al*. (1988). Persistent *Campylobacter jejuni* infections in patients infected with human immunodeficiency virus (HIV). *Annals of Internal Medicine*, **108**, 540–6.

Phillips, R.E. and McMichael, A.J. (1993). How does the HIV escape cytotoxic T cell immunity? *Chemical Immunology*, **56**, 150–64.

Piatak, M., Jr. *et al*. (1993). High levels of HIV-1 in plasma during all stages of infection determined by competitive PCR. *Science*, **259**, 1749–54.

Potashner, W., Buskila, D., Patterson, B., Karasik, A., and Keystone, E.C. (1990). Leukocytoclastic vasculitis in association with human immunodeficiency virus infection. *Journal of Rheumatology*, **17**, 1104–7.

Prince, H.E. and John, J.K. (1987). Abnormalities of interleukin 2 receptor expression associated with decreased antigen-induced lymphocyte proliferation in patients with AIDS and related disorders. *Clinical and Experimental Immunology*, **67**, 236–44.

Quinn, T.C. (1990). Global epidemiology of HIV infections. In *The medical management of AIDS* (ed. M.A. Sande and P.A. Volberding), pp. 3–22. W.B. Saunders, Philadelphia.

Rank, R.G., Ramsey, K.H., and Hough, A.J. (1988). Antibody-mediated modulation of arthritis induced by *Chlamydia*. *American Journal of Pathology*, **132**, 372–81.

Ranki, A., Puska, P., Mattinen, S., Lagerstedt, A., and Krohn, K. (1991). Effect of PUVA on immunologic and virologic findings in HIV-infected patients. *Journal of the American Academy of Dermatology*, **24**, 404–10.

Rao, T. and Friedman, E. (1989). AIDS (HIV)-associated nephropathy: does it exist? *American Journal of Nephrology*, **9**, 441–53.

Rao, T.K.S., Friedman, E.A., and Nicastri, A.D. (1987). The types of renal disease in the acquired immunodeficiency syndrome. *New England Journal of Medicine*, **316**, 1062–8.

Ricciardi, D.D., Sepkowitz, D.V., Berkowitz, L.B., Bienenstock, H., and Maslow, M. (1986). Cryptococcal arthritis in AIDS patients. *Journal of Rheumatology*, **13**, 455–9.

Rivera, J., Monteagudo, I., Lopez-Longo, J., and Sanchez-Atrio, A. (1992). Septic arthritis in patients with acquired immunodeficiency syndrome with human immunodeficiency virus infection. *Journal of Rheumatology*, **19**, 1960–2.

Riviere, Y. *et al*. (1989). HIV-specific cytotoxic responses of seropositive individuals; distinct types of effector cells mediate killing of targets expressing gag and env proteins. *Journal of Virology*, **63**, 2270–7.

Roux Lombard, P., Modoux, C., Cruchaud, A., and Dayer, J.M. (1989). Purified blood monocytes from HIV 1-infected patients produce high levels of TNF-a and IL-1. *Clinical Immunology and Immunopathology*, **50**, 374–84.

Rubbert, A. *et al*. (1994). Anticardiolipin antibodies in HIV infection: association with cerebral perfusion defects as detected by 99mTc-HMPAO SPECT. *Clinical and Experimental Immunology*, **98**, 361–8.

Rusche, J.R. *et al*. (1988). Antibodies that inhibit fusion of human immunodeficiency virus-infected cells bind a 24-amino acid sequence of the viral envelope, gp120. *Proceedings of the National Academy of Sciences* (*USA*),**85**, 3198–202.

Ruzicka, T. *et al*. (1987). Treatment of HIV-induced retinoid-resistant psoriasis with zidovudine. *Lancet*, **ii**, 1469.

Rynes, R.I. (1990). HIV and rheumatologic autoimmune phenomena: imitator or illuminator. *Clinical and Experimental Rheumatology*, **8**, 103–6.

Said, G. *et al*. (1988). The peripheral neuropathy of necrotizing arteritis: a clinicopathological study. *Annals of Neurology*, **23**, 461.

Salahuddin, S.Z., Rose, R.M., Groopman, J.E., Markham, P.D., and Gallo, R.C. (1986). Human T lymphotropic virus type III infection of human alveolar macrophages. *Blood*, **68**, 281–4.

Savige, J.A., Chang, L., and Crowe, S.M. (1993). Anti-neutrophil cytoplasm antibodies in HIV infection. *Advances in Experimental Medicine and Biology*, **336**, 349–52.

Scaravalli, F. *et al*. (1989). Chronic basal meningitis and vasculitis in acquired immunodeficiency syndrome. *Archives of Pathology and Laboratory Medicine*, **113**, 192.

Scorza-Smeraldi, R., Fabio, G., Lazzarin, A., Eisera, N.B., Moroni, M., and Zanussi, C. (1986). HLA-associated susceptibility to acquired immunodeficiency syndrome in Italian patients with human-immunodeficiency-virus infection. *Lancet*, **ii**, 1187–9.

Sergent, J. (1980). Vasculitis associated with viral infection. *Clinics in the Rheumatic Diseases*, **6**, 339.

Sheehy, M.J. *et al*. (1989). A diabetes-susceptible HLA haplotype is best defined by a combination of HLA-DR and -DQ alleles. *Journal of Clinical Investigation*, **83**, 830–5.

Shioda, T., Levy, J.A., and Cheng Mayer, C. (1991). Macrophage and T cell-line tropisms of HIV-1 are determined by specific regions of the envelope gp120 gene. *Nature*, **349**, 167–9.

Siekevitz, M., Josephs, S.F., Dukovich, M., Peffer, N., Wong-Staal, F., and Green, W.C. (1987). Activation of the HIV-1 LTR by T-cell mitogens and the trans-activator protein of HTLV-1. *Science*, **238**, 1575–8.

Simpson, D.M. (1988). Myopathy associated with human immunodeficiency virus (HIV) but not with zidovudine. *Annals of Internal Medicine*, **109**, 842.

Snider, W.D., Simpson, D.M., Nielsen, S., Gold, J.W., Metroka, C.E., and Posner, J.B. (1983). Neurological complications of acquired immunodeficiency virus: analysis of 50 patients. *Annals of Neurology*, **14**, 403–18.

Solinger, A.M. and Hess, E.V. (1990). HIV and arthritis. *Arthritis and Rheumatism*, **17**, 562.

Solinger, A.M. and Hess, E.V. (1991). Induction of autoantibodies by HIV infection and their significance. *Rheumatic Diseases Clinics of North America*, **17**, 157–76.

Solinger, A.M. and Hess, E.V. (1993). Rheumatic diseases and AIDS—is the association real? *Journal of Rheumatology*, **20**, 678–83.

Solinger, A.M. *et al*. (1988). Acquired immune deficiency syndrome (AIDS) and autoimmunity—mutually exclusive entities? *Journal of Clinical Immunology*, **8**, 32–42.

Springer, T.A. (1990). Adhesion receptors of the immune system. *Nature*, **346**, 425–34.

Springer, T.A., Dustin, M.L., Kishimoto, T.K., and Marlin, S.D. (1987). The lymphocyte function-associated LFA-1, CD2, and LFA-3 molecules: cell adhesion receptors of the immune system. *Annual Review of Immunology*, **5**, 223–52.

Steel, C.M. *et al*. (1988). HLA haplotype A1, B8, DR3 as a risk factor for HIV-related disease. *Lancet*, **ii**, 1185–8.

Steinbach, L.S., Tehranzadeh, J., Fleckenstein, J.L., Vanarthos, W.J., and Pais, M.J. (1993). Human immunodeficiency virus infection: musculoskeletal manifestations. *Radiology*, **186**, 833–8.

Strauss, J. *et al*. (1989). Renal disease in children with the acquired immunodeficiency syndrome. *New England Journal of Medicine*, **321**, 625–30.

Strongin, I.S., Kale, S.A., Raymond, M.K., Luskin, R.L., Weisberg, G.W., and Jacobs, J.J. (1991). An unusual presentation of gonococcal arthritis in an HIV positive patient. *Annals of the Rheumatic Diseases*, **50**, 572–3.

Swain, S.L. (1983). T cell subsets and the recognition of MHC class. *Immunological Reviews*, **74**, 129.

Takahashi, H., Germain, R.N., Moss, B., and Berzofsky, J.A. (1990). An immunodominant class I-restricted cytotoxic T lymphocyte determinant of human immunodeficiency virus type 1 induces CD4 class II-restricted help for itself. *Journal of Experimental Medicine*, **171**, 571–6.

Thompson, I. *et al*. (1989). Henoch–Schönlein purpura and IgA glomerulonephritis associated with HIV infection. *Abstracts of the Fifth International Conference on AIDS: The Scientific and Social Challenge*, Quebec, MBP278.

Till, M. and MacDonell, K.B. (1990). Myopathy with HIV-1 infection: HIV-1 or zidovudine? *Annals of Internal Medicine*, **113**, 492–3.

Tindall, B., Imre, A., Donovan, B., Penny, R., and Cooper, D.A. (1990). Primary HIV infection: clinical, immunologic and serologic aspects. In *The medical management of AIDS* (ed. M.A. Sande and P.A. Volberding), pp. 68–84. W.B. Saunders, Philadelphia.

Tiwari, J.L. and Terasaki, P.I. (1985). *HLA and disease associations*. Springer Verlag, New York.

Tschachler, E. *et al*. (1987). Epidermal Langerhans cells—a target for HTLVIII/LAV infection. *Journal of Investigative Dermatology*, **88**, 233–7.

Valeriano, J., Lolita, B., and Kerry, L.D. (1989). HIV-associated polyarteritis nodosa (PAN) diagnosed by rectal biopsy; a case report. *Arthritis and Rheumatism*, **32**, S44.

Vasey, F.B. *et al*. (1982). Possible involvement of group A streptococci in the pathogenesis of psoriatic arthritis. *Journal of Rheumatology*, **9**, 719.

Vernant, J.C. *et al*. (1988). T-lymphocyte alveolitis, tropical spastic paresis, and Sjögren syndrome. *Lancet*, **i**, 177.

Viscidi, R.P., Mayur, K., Lederman, H.M., and Frankel, A.D. (1989). Inhibition of antigen-induced lymphocyte proliferation by Tat protein from HIV-1. *Science*, **246**, 1606–8.

Vogel, J. *et al*. (1988). The HIV tat gene induces dermal lesions resembling Kaposi's sarcoma in transgenic mice. *Nature*, **335**, 606.

Vogel, J., Cepeda, M., Tschachler, E., Napolitano, L.A., and Jay, G. (1992). UV activation of human immunodeficiency virus gene expression in transgenic mice. *Journal of Virology*, **66**, 1–5.

Walker, B.D. *et al*. (1987). HIV-specific cytotoxic T lymphocytes in seropositive individuals. *Nature*, **328**, 345–8.

Watts, R.A. *et al*. (1987). Pyomyositis associated with human immunodeficiency virus infection. *British Medical Journal*, **294**, 1524.

Weiss, R.A. (1993). How does HIV cause AIDS? *Science*, **260**, 1273–9.

Westervelt, P. *et al*. (1992). Macrophage tropism determinants of human immunodeficiency virus type 1 in vivo. *Journal of Virology*, **66**, 2577–82.

White, S.H. *et al*., (1972). Disturbance of HL-A antigen frequency in psoriasis. *New England Journal of Medicine*, **287**, 740.

Whyte, H.J. and Baughman, R.D. (1964). Acute guttate psoriasis and streptococcal infection. *Archives of Dermatology*, **89**, 350–6.

Wiley, C.A., Nerenberg, M., Cros, D., and Soto-Aguilar, M.C. (1989). HTLV-I polymyositis in a patient also infected with the human immunodeficiency virus. *New England Journal of Medicine*, **320**, 992–5.

Williams, R.C., Hasur, H., and Spera, T.J. (1984). Lymphocyte-reactive antibodies in acquired immunodeficiency syndrome. *Journal of Clinical Immunology*, **4**, 118–23.

Winchester, R. (1994). Psoriatic arthritis. In *Dermatology in general medicine* (ed. T.B. Fitzpatrick, A.Z. Eisen, K. Wolff, I.M. Freedberg, and K. Austen). McGraw-Hill, New York.

Winchester, R., Bernstein, D.H., Fischer, H.D., Enlow, R., and Solomon, G. (1987). The co-occurrence of Reiter's syndrome and acquired immunodeficiency. *Annals of Internal Medicine*, **106**, 19–26.

Winchester, R., Brancato, L., Itescu, S., Skovron, M.L., and Solomon, G. (1988). Implications from the occurrence of Reiter's syndrome and related disorders in association with advanced HIV infection. *Scandinavian Journal of Rheumatology*, **74**, 89–93.

Withington, R.H. *et al*. (1987). Isolation of HIV from synovial fluid of a patient with reactive arthritis. *British Medical Journal*, **294**, 484.

Yanker, B.A. *et al*. (1986). Cerebral granulomatous angiitis associated with isolation of human T-lymphotrophic virus III from the central nervous system. *Annals of Neurology*, **20**, 362.

Zimmermann, B., Erickson, A.D., and Mikolich, D.J. (1989). Septic acromioclavicular arthritis and osteomyelitis in a patient with acquired immunodeficiency syndrome. *Arthritis and Rheumatism*, **32**, 1175.

Zimmerman, W., Cher, S.M., Bolden, A., and Weissbach, A. (1980). Mitochondrial DNA replication does not involve DNA polymerase-a. *Journal of Biological Chemistry*, **255**, 11847–52.

Zolla-Pazner, S. *et al*. (1987). Nonrandom development of immunologic abnormalities after infection with human immunodeficiency virus: implications for immunologic classification of the disease. *Proceedings of the National Academy of Sciences* (*USA*), **84**, 5404–8.

Zwart, G. *et al*. (1991). Immunodominance and antigenic variation of the principal neutralization domain of HIV-1. *Virology*, **181**, 481–9.

5.3.7 Mycobacterial diseases

Sanjiv N. Amin

Tuberculosis and leprosy are infectious diseases characterized by chronic inflammation. Tuberculosis is caused chiefly by the organism *Mycobacterium tuberculosis* and much less commonly by *M. bovis* and atypical mycobacteria (*M. avium, intercellulare, scrofulaceum, gordonae, marium*, etc.). Leprosy is caused by *M. leprae*. Most of these infections are readily amenable to modern medical and surgical treatment, which almost always cures the infection.

Tuberculosis

Epidemiology and pathogenesis

Tuberculosis was rapidly declining in the United States, Europe, and Australia until 1985, after which the trend was interrupted by the appearance of acquired immune deficiency disease (Modilevsky *et al*. 1989). Tuberculosis in persons infected by human immunodeficiency virus is characterized by extrapulmonary disease in as many as 70 per cent. At present, one-third of the world's population, estimated at 1.7 billion persons, is infected with *M. tuberculosis*. This reservoir of infection affects around 8 million new patients and results in 19 million deaths annually (Arachi 1991). Fewer than one-tenth of all patients with tuberculosis residing in developed nations suffer from infection of the bones and joints (Davies *et al*. 1994), but the proportion in other populations is much higher.

In the majority of instances, *M. tuberculosis* is transmitted from person to person via the respiratory route. A patient with infectious pulmonary tuberculosis (sputum smear-positive) coughs and produces an aerosol of small droplets, 1 to 5 µm in size, containing bacilli. The small droplets evaporate within a short distance from the mouth, and the desiccated bacilli remain airborne for long periods. Infection of a host occurs when a few bacilli are inhaled. These are sufficiently small to reach the pulmonary alveoli and are phagocytosed by alveolar macrophages.

The *M. tuberculosis* grows slowly, dividing approximately every 10 to 24 h. It has no known endotoxins or exotoxins, and there is no immediate host response to infection. Growth and multiplication of the organism are essentially unimpeded, until a specific cell-mediated immune response develops after 4 to 8 weeks, and only if a threshold number of *M. tuberculosis* organisms is reached. The collection of activated T cells and macrophages forms granulomas that undergo central coagulative necrosis (caseation). Healing then occurs, often with calcification.

Tuberculosis as a clinically manifested disease develops in a minority of the patients who fail to contain the primary infection. In some individuals the disease may appear within a few weeks of primary infection, and in others the bacilli may remain dormant within the macrophages for many years before entering a phase of exponential multiplication to cause the disease. The bacilli reach the bloodstream either by being carried in the lymph to the draining lymph nodes and thence to the thoracic duct, or by erosion of blood vessels in the walls of developing tuberculous lesions in the lungs. The bacilli that enter the bloodstream disseminate throughout the body, and some are deposited in bones or joints.

Immunology

There appears to be variation between individuals in their immune response to mycobacteria, which perhaps is genetically determined. The ability of the host to control infection with *M. tuberculosis* resides in its ability to mount an effective cellular immune response. Cell-mediated immunity develops when T lymphocytes become sensitized after recognizing their specific antigen and then release mediators that modulate macrophage function.

Delayed cellular hypersensitivity is the associated immunological response in the majority of patients with tuberculosis (Stanford 1983; Lucas 1988). Cell-mediated immunity and delayed hypersensitivity are closely related phenomena that occur in the host as a result of T cells becoming specifically activated by *M. tuberculosis*. Both result from the same immunological mechanism and alter the response of the host to subsequent exposure to antigen. Delayed hypersensitivity is responsible for the tuberculin skin-test reaction. There is no single dominant antigen, and the infected (or artificially sensitized) host develops an immune response to an array of mycobacterial proteins. Many of the observed effects of tuberculosis are considered to be due to delayed hypersensitivity (Sifford and Bates 1991) and it has been implicated in caseation and cavitation. Caseation follows the early exudative lesions in the soft tissues and results when blood vessels adjacent to the areas of inflammation thrombose causing tissue necrosis. Delayed hypersensitivity is also involved in the liquefaction of caseous lesions, probably as a result of lymphokine production, which attracts and activates macrophages. Macrophages then release hydrolytic enzymes that digest the necrotic debris.

Mycobacterial infections also stimulate humoral antibody responses, and production of antibodies to polysaccharide and protein antigens has been demonstrated. However, there is no evidence that these antibodies play a part in immunity, hypersensitivity, or the pathogenesis of tuberculosis (Dutt 1989; Ellner *et al*. 1989). Attempts to link HLA phenotype and susceptibility to tuberculous infection have not yielded a significant correlation (Shoemaker 1986).

Clinical manifestations

Primary tuberculous infection is usually asymptomatic and typically occurs in lower or mid zones of the lung as a pneumonitis with enlargement of hilar lymph nodes (primary complex). It may either undergo remission spontaneously or progress at once to clinical disease. After remission, there may be reactivation many months or years later, resulting in a chronic wasting disease. The lungs and respiratory tracts are most frequently affected, followed by lymph nodes, skeleton, pericardium, brain and meninges, abdomen, and skin.

Tuberculosis of a bone or joint is usually a low-grade and slowly progressive infection with a variable degree of local and systemic manifestations (Wolfgang 1978; Butorac *et al*. 1987; Martini *et al*. 1988), depending on the virulence of the organism and the defensive response mounted by the host. The onset of symptoms is insidious; involvement is monoarticular or mono-osseous in the majority of patients. Simultaneous involvement of other viscera (lungs, lymph nodes) is common. There are associated constitutional symptoms (fever, fatigue, weight loss, poor appetite, night sweats). Pre-existing arthritis or old trauma, alcoholism, prolonged use of corticosteroids, and immunodeficiency diseases are significant predisposing factors.

Tuberculosis can affect any part of the spine, although up to about two decades ago the thoracolumbar junction was more commonly affected (Davidson and Horowitz 1970). The infection may begin in and remain confined to a vertebral body, eventually leading to collapse, or it may affect the end-plate of the vertebral body with early involvement of the adjacent disc and the next vertebral end-plate. The formation of a paravertebral abscess is usual, and progressive necrosis of bone and disc

with sequestration leads to kyphotic deformity or gibbus at that site. The abscess may remain localized at the same site or track along tissue planes and neurovascular bundles to cause symptoms at a remote site. In the lumbar spine, the abscess may track along the psoas muscle to present as a swelling in groin or thigh; from thoracic vertebrae it may follow the course of a rib and present anteriorly on the chest wall. The local symptoms are pain, muscle spasm, and limited movement. The pain is characteristically worse during sleep, which perhaps is due to relaxation of the protective muscle spasm. In the occasional patient, pain may not be the predominant symptom, and kyphosis or a cold abscess is the first manifestation.

Primary involvement of the posterior elements of vertebrae appears less common. Haematogenous spread of infection to the pedicle is usually the initial event. Destruction of the neural arch and vertebral body posteriorly characterize the disease progression. Later in the course of disease, the medial ends of ribs are eroded, followed by the transverse processes of affected vertebrae. In the majority of cases only a single vertebra is involved (Kumar 1985). Disc spaces remain preserved until late in the disease and paraspinal masses are prominent. Clinical presentation in these patients is often with early paraplegia.

Tuberculosis involving the first and second cervical vertebrae may begin in the retropharyngeal space with secondary involvement of the bone, or more rarely in the bone itself (Lifeso 1987) (Fig. 1). With progression there is increasing ligamentous involvement, with minimal osteolytic erosions, into the odontoid or the arch of the first cervical vertebra. This would allow anterior subluxation of C1 on C2, increasing rotatory subluxation, and proximal translocation of the odontoid. In a final stage of the disease, bone destruction increases, with complete loss of the C1 arch or fracture through the base of the odontoid, leading to a grossly unstable articulation between the occiput and C2. The diagnosis is suspected if an individual with prolonged neck pain has restriction in all ranges of movement. The most serious complication of spinal tuberculosis is involvement of the spinal cord causing a neurological deficit. This may occur incompletely and slowly, with the patient complaining of difficulty in walking, or it can appear more dramatically, with complete spastic paraplegia or quadriplegia and loss of sphincter control. The neurological deficit may be secondary to medullary and radicular inflammation or cord compression by an abscess, tuberculoma, or subluxation of a vertebral body.

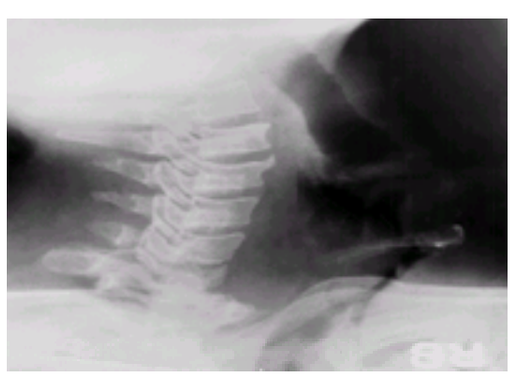


Fig. 1 Tuberculosis of C1 and C2 with a large retropharyngeal abscess.

Skeletal tuberculosis other than in the spine in adults more commonly occurs in the joints of the lower extremities (Fig. 2 and Fig. 3). It usually is monoarticular. Weight-bearing joints are more frequently affected, and microtrauma to cartilage is thought to predispose to the infection. The joints commonly involved are the hip, knee, ankle, sacroiliac, wrist, and shoulder in that order (Garrido *et al*. 1988; Pouchot *et al*. 1988). The infected individual usually has a long history of mild or moderate joint or bone pain along with a swelling or large effusion. On examination, there is synovial thickening and mild warmth. There is always significant muscle atrophy around the joint. If a sinus track has formed, there may be signs of superimposed pyogenic infection. Even in the early stages, there is limitation in the range of motion by effusion and synovial thickening. As infection progresses, flexion contractures and joint deformity develop. The end-result may be fibrous or bony ankylosis.

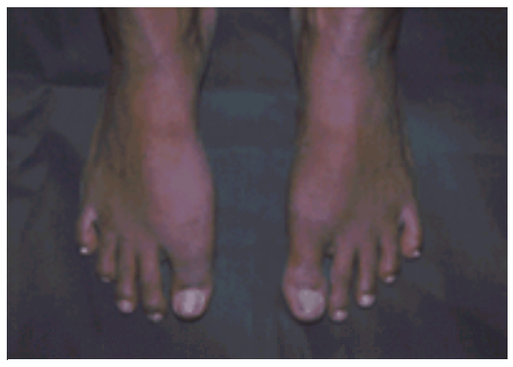


Fig. 2 The base of the right great toe was painful and swollen for 5 years in a 40-year-old premenopausal female; this was an indolent and slowly progressive lesion of tuberculosis.

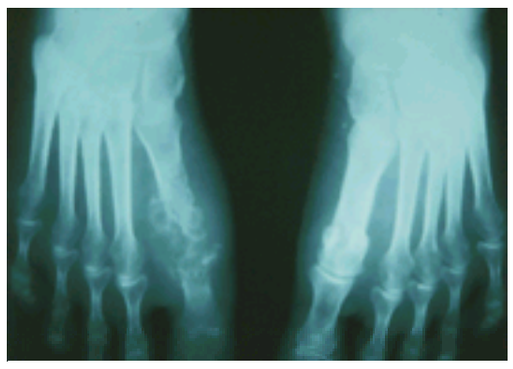


Fig. 3 Radiograph of the feet of the patient in Fig. 2 at 5 years after the onset of pain, revealing remarkable destruction of the bones and joints by tuberculous infection.

Tuberculous osteomyelitis may affect any long bone (Martini *et al*. 1986). The patient presents with local pain and sometimes a diffuse swelling may be evident, with or without a sinus opening. Involvement of the short bones of the hands or feet is termed tuberculous dactylitis. Soft tissue swelling is usually the first manifestation and the pain appears a few days or weeks thereafter. One or several fingers or toes may be affected.

Cystic tuberculosis of bone is another type of tuberculous osteomyelitis. Lesions are of variable size and well circumscribed. Multifocal cystic tuberculosis of bone was more common five decades ago, but it seems that solitary lesions are now predominant (Kumar *et al*. 1988). In the adult the common sites are the skull, axial skeleton, and shoulder and pelvic girdles. In children the metaphyses of long bones are often the sites of infection. This predilection is probably due to the vascular structure of the long bones in this region (Edeikin *et al*. 1963). Tubercle bacilli probably lodge in the small terminal branches of the arteries of the metaphyses and grow, caseate, and produce a cystic lesion. A soft tissue swelling is externally visible if the skull bone is affected (LeRoux *et al*. 1990). Most commonly, a circumscribed, punched out lesion, approx. 2 cm in diameter, results. At presentation, most patients are systemically well and manifest a painless swelling, which is firm if the periosteum is intact

or soft if not. Headache is uncommon, but if present, it is usually localized to the site of infection. Occasionally, there is a discharging sinus.

Poncet's disease is another pattern of joint involvement (Dall *et al*. 1989). Patients who have active visceral tuberculosis complain of polyarthritis or tenosynovitis affecting peripheral joints (Fig. 4), a form of reactive arthropathy. In some patients the symptoms and signs are suggestive of enthesopathy at the ankle, knee, hip, or elbow. There may be mild stiffness after inactivity, and swelling. These symptoms regress within a few weeks after starting treatment. Physicians from the Indian subcontinent do not doubt its existence as a definite entity.

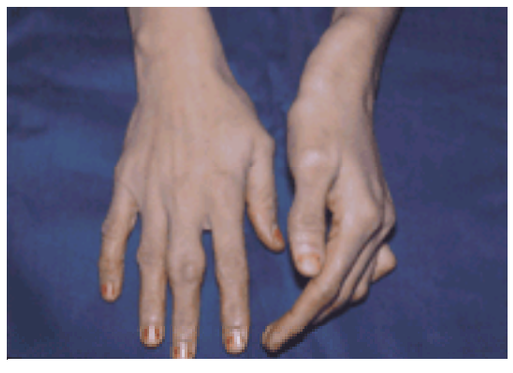


Fig. 4 Tuberculous inflammation of the right middle proximal interphalangeal joint and left abductor pollicis longus tendon sheath in a 36-year-old woman.

Other, less frequently seen clinical features of tuberculosis include necrotic skin ulcers of erythema nodosum, shoulder-hand syndrome, parotid gland swelling, and red eye caused by uveitis and chorioretinitis. Secondary amyloidosis may appear many years after the tuberculous infection.

Investigations and diagnosis

Tuberculosis is diagnosed if the clinical pattern is suggestive and acid-fast bacilli are demonstrated in the lesions. However, the latter is often not possible and other criteria have to be sought as additional evidence.

To culture tubercle bacilli on the Lowenstein–Jensen medium requires 4 to 8 weeks to detect growth. Nowadays, radiometric techniques using highly selective media allow cultivation in 1 or 2 weeks, but confirmation of the identity of an isolated organism may require more time. Advances in molecular genetics have resulted in impressive DNA hybridization probes that are non-radioactive and easily utilized in a clinical microbiology laboratory. However, these probes require growth of the organism in culture because they are unreliable in detecting the relatively small numbers of organisms in clinical specimens such as tuberculous synovitis.

A smear from the infected site such as a paraspinal abscess usually provides the required evidence because the staining characteristics of *M. tuberculosis* are typical. A slender, curved, often polychromatically beaded rod is seen singly, in pairs, or in clumps of a few organisms, lying side by side. If stained with fluorescent auramine-rhodamine, the bacilli are visualized under the usual high-dry (100 ×) magnification. A more specific stain consists of carbol fuchsin and this requires careful scanning with oil-immersion (1000 ×) microscopy.

A simple smear from clinical specimens such as synovial fluid or infected bone scrapings seldom demonstrates the organism. The technique of polymerase chain reaction holds remarkable promise for this purpose. The key technological element of this procedure is a DNA polymerase that is heat-stable. The function of a DNA polymerase is to generate a complementary strand of DNA using a single strand of DNA as a template. A short segment of complementary DNA annealed to the template (which equates with the a short segment of double-stranded DNA) serves as a primer region. The DNA polymerase can add nucleotides to the primer in a sequential fashion based on the template sequence and result in a longer sequence of double-stranded DNA. The nucleotides are always added in the same direction from any given primer site. If a short primer sequence is added to the double-stranded DNA (chromosomal, plasmid) and the temperature is raised above the melting point of the double strands and then lowered, a number of the primers will anneal to their complementary segments on the chromosome or plasmid. Once annealed, the primer region is elongated by the addition of an appropriate DNA polymerase and deoxyribonucleotide triphosphates. Two primers from opposite ends and strands can anneal and result in a duplication of the target segment. Once the elongation is complete, the temperature is raised so that both the old and newly synthesized strands fall apart. By lowering the temperature, the target region of DNA is logarithmically amplified. Although the reaction mixture goes above 90 °C during each cycle, the heat-stable DNA polymerase is not completely inactivated and can result in millionfold amplifications of the target DNA sequence after 30 to 40 cycles. The temperatures for denaturing, annealing, and elongation steps are easily programmed into automated heat blocks that can reproducibly change temperatures with great precision. Thus, the polymerase chain reaction can amplify minute quantities of DNA to levels that are easily seen on routine agarose gel electrophoresis. However, the major limitation is that minute quantity of contaminating DNA also gets amplified and is seen on the agarose gel as a false-positive result. Present research is seeking to overcome this limitation. Also, the polymerase chain reaction does not differentiate dead from living organisms because it simply amplifies DNA. Thus, identifying a portion of the genome of an organism in a clinical specimen provides the ability to state that the organism was present, and it cannot distinguish between active and treated or inactive tuberculosis (Schluger *et al*. 1994). The correlation of clinical information with the greater sensitivity of the polymerase chain reaction is currently under study.

The typical histological feature is a tuberculous granuloma with partial or complete caseation necrosis. The epithelioid cells are surrounded by a wall of mononuclear cells, which is a lymphocyte mantle comprised mainly of T-suppressor cells. Such typical granulomas are not always seen (particularly in the synovium) and only a tuberculoid infiltrate may be visible. In this infiltrate are seen irregular accumulations of epithelioid cells among mononuclear cells, with or without necrosis and giant cells (Levy *et al*. 1986).

Serological tests for the diagnosis of tuberculosis (Daniel 1989) remain experimental and are not satisfactorily established for routine clinical use. Advances in the definition of species-specific antigenic determinants of *M. tuberculosis* had created a firm basis for the production of standardized serological immunoassays (Brisson-Noel *et al*. 1989). However, in the majority of patients, musculoskeletal tuberculosis is a paucibacillary disease in which the desired sensitivity of these tests is not achievable (Ivanyi 1988). This is attributed to the lack of sufficient antibody formation rather than to limitations in the techniques of detection. A raised erythrocyte sedimentation rate is a non-specific serological measure of inflammation and occasionally is in the normal range though the patient has active tuberculosis.

The likelihood of infection with *M. tuberculosis* can be assessed by intradermal injection of 5 tuberculin units of purified protein derivative (American Thoracic Society 1990). The diameter of any induration is recorded 48 to 72 h later. Reactivity to tuberculin (³5 mm induration) is found if the person has been exposed to mycobacteria and has intact cell-mediated immunity. A history of vaccination with bacille Calmette-Guérin is ignored in interpreting the results of tuberculin skin testing in adults, because skin-test reactivity from bacille Calmette-Guérin usually declines by adulthood. If skin testing is done periodically, as in hospital employees, the boosting phenomenon must be considered. In some tuberculin reactors, although the sensitivity to tuberculin declines with time to become a negative skin test, administration of the skin test boosts immunological memory so that a second tuberculin skin test up to 2 years later will be positive. If the person is being tested annually, the positive second response and negative first response are erroneously thought to represent recent infection requiring treatment. Another shortcoming of the test is cross-reactivity between mycobacterial antigens, and a positive test may result from exposure to environmental non-tuberculous mycobacteria. Tuberculin test sensitivity falls with age, early treatment, and in all conditions that diminish delayed hypersensitivity. In addition, for as yet unknown reasons, about one-tenth of individuals with tuberculosis do not respond to the ordinary intermediate strength of tuberculin (anergy).

Plain radiographs of the skeleton will have features suggestive of tuberculous infection. Osteoporosis is the first sign of active infection. The reactive hyperaemia, which is intense in an exudative lesion, stimulates osteolysis. If long bones are affected (osteomyelitis), small zones of clearly defined radiolucency indicate granular foci. Diffuse demineralization surrounds these osteolytic areas. As caseation takes place, the osteolytic foci become more evident. When the healing process begins, the perifocal bone becomes sclerotic. If the central demineralized area is merely exudative, healing results in reossification with eventual return of the normal trabecular pattern. If central caseation occurs during the active phase of infection, with calcium deposition, then the dense image of a sequestrum is surrounded by an osteolytic ring representing the fibrous wall beyond which the bone is demineralized. In tuberculous dactylitis the soft tissue swelling is obvious and there may be mild or exuberant periostitis of phalanges, metacarpals, or metatarsals (Bush and Schneider 1984). Expansion of the bone accompanied by cystic change is termed spina

ventosa. In involvement of a peripheral joint the synovial shadow may be clearly visible in a slightly underexposed radiograph. If the arthritis is destructive, then the joint space narrows due to erosion of cartilage, and the subchondral cortex of bone becomes ragged and osteoporotic. Vague, irregular densities seen in the surrounding soft tissues may signify abscess formation.

The radiographic appearance of spinal tuberculosis depends on the extent of infection (Chang *et al*. 1989). Destructive changes occasionally confined to a vertebral body or part of two posterior elements of the vertebral complex such as the lamina or pars, but these are uncommon and easily confused with malignant disease. More typically seen is reduction of the disc space, with irregularity of adjacent end-plates, surrounded by a soft tissue swelling due to a paravertebral abscess. Later, more extensive destruction and increased abscess formation are evident. Kyphotic or scoliotic deformity develops as vertebral destruction progresses. In children, secondary changes, such as an increase in vertebral height, may develop in uninvolved adjacent vertebrae to produce a compensatory lordosis above and below the kyphotic deformity. A severe kyphotic deformity may increase with time, even after healing has occurred, because of the gravitational effect on the deformed spine.

Radiological evidence of healing after successful drug treatment is usually observed late on routine radiographs, both in limb bones or joints (Fig. 5) and the spine. Bone destruction or loss of vertebral height can progress for up to 14 months and recovery of vertebral height is often not seen earlier than 15 months after beginning chemotherapy (Boxer *et al*. 1992). Thus the progression of bone destruction whilst on treatment is not always an adverse feature. Sclerosis is a feature of healing, and appears variably from onset to within 5 months of starting drug treatment. The change from sclerosis to normal bone density takes 5 years or much longer. Paravertebral soft-tissue masses may also take as long as 15 months to resolve. Involvement of adjacent vertebrae is often associated with the reduction in disc space, fusion of the vertebrae, and the formation of syndesmophyte-like bone bridges (Fig. 6).

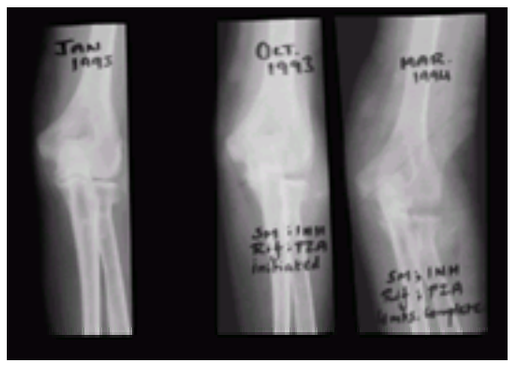


Fig. 5 Progressive radiological destruction in tuberculous infection of the left elbow during the first 5 months of adequate treatment.

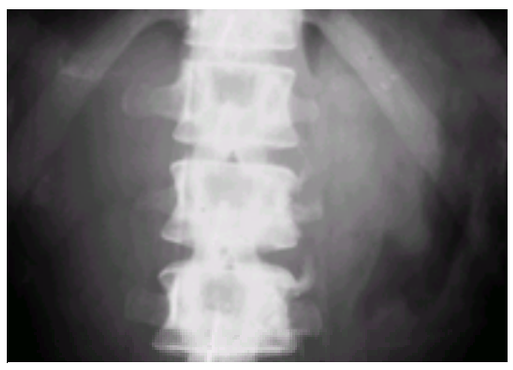


Fig. 6 Erosion of the left transverse process at the L2 vertebra with calcification along the track of caseous fluid to the above and below vertebrae.

Computed tomography (**CT**) is provenly superior to conventional radiographs in detecting and monitoring paravertebral abscesses, particularly those situated in the lumbar region (LaBerge and Brant-Zawadski 1984). The CT findings indicative of an abscess include an abnormal mass of low attenuation number, displacement of surrounding structures, obliteration of normal fascial planes, and a 'rind sign' consisting of a rim of increased tissue attenuation. However, none of these features is specific for tuberculosis, with the differential diagnoses being haematoma or neoplasia (Whelan *et al*. 1985), brucellosis, and other pyogenic infections. Tuberculous abscesses are often characterized by macroscopic calcifications. These 'rice bodies', occasionally seen on CT, are diagnostic of tuberculosis infection.

Magnetic resonance imaging (**MRI**) features of vertebral osteomyelitis are characteristic (Smith *et al*. 1989). The T_1-weighted images show a confluent decreased signal intensity at the involved vertebral bodies and intervening disc, whereas the T_2-weighted images demonstrate increased signal intensity. In addition, the normal biconcave configuration of the disc changes and the normal intranuclear cleft is lost. Abscesses can cause an abnormal area of low signal intensity on a T_1-weighted image and a relatively increased signal intensity on T_2-weighted images because of increased fluid content. The MRI can clearly delineate abscesses as distinct from the adjacent spinal cord, psoas muscle, and other surrounding paravertebral soft tissues. To differentiate tuberculous from other pyogenic infections, the clinical features, plain radiographs and CT have to be considered.

Treatment

The cornerstone of the treatment of musculoskeletal tuberculosis is a good regimen of antituberculous drugs, with surgical intervention in selected cases (Cooke 1985; Goldberger 1988; Davidson 1989). Because of increasing drug resistance, drug susceptibility testing is recommended on all *M. tuberculosis* isolates. The available drugs are listed in Table 1. In the treatment of newly diagnosed patients, the regimens should include, whenever possible, the main sterilizing drugs and those most effective in preventing the emergence of resistance. Patients in whom drug resistance is unlikely are prescribed isoniazid, rifampicin, and pyrazinamide. Some of the currently prescribed regimens for adults are:

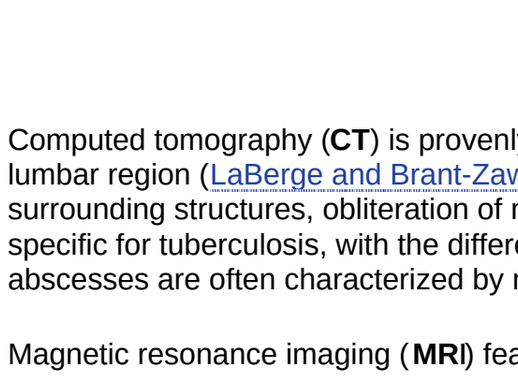


Table 1 Drugs for tuberculosis

1. rifampicin 450–600 mg, isoniazid 300 mg, pyrazinamide 1000–1500 mg, daily for 2 months, followed by rifampicin and isoniazid daily for 6 months;
2. rifampicin 450–600 mg, streptomycin 0.5–1.0 g, isoniazid 300 mg, pyrazinamide 1000–1500 mg or ethambutol 800 mg for the first 2 months, followed by

   a. rifampicin 450–600 mg and isoniazid 300 mg daily for 4 months,
   b. isoniazid 300 mg and ethambutol 800 mg daily for 6 months,
   c. isoniazid 300 mg, rifampicin 600 mg, and streptomycin 1 g twice weekly for 6 months;
3. rifampicin 450–600 mg, isoniazid 300 mg daily for 1 month, followed by rifampicin 600 mg and isoniazid 900 mg twice weekly for 8 months;
4. streptomycin 0.5 g, isoniazid 300 mg, ethambutol 800 mg daily for 3 months, followed by isoniazid and ethambutol for 12 to 18 months.

The duration of treatment is extended if there is no convincing evidence of improvement in clinical features. Surveillance is kept for 12 months after completion of treatment, as most relapses are likely to occur within this period. The introduction of short-course chemotherapy is a major step forward in the treatment of tuberculosis (Hannachi 1988). The twice-weekly drug regimens were developed with the intention to improve patient compliance.

Patients at increased risk for drug resistance should receive isoniazid, rifampicin, pyrazinamide, and ethambutol until susceptibility results are available. Ofloxacin and ciprofloxacin are bactericidal quinolones with low toxicity profiles and penetrate most tissues. They have been used for the treatment of drug-resistant tuberculosis, but reports of prospective controlled trials are not yet available. For adults, ofloxacin 600 mg/day or ciprofloxacin 1000 mg/day may be prescribed in combination with the other drugs.

The surgical treatment of skeletal tuberculosis depends on the tissue involved, whether it is bone, bursa or joint, and the severity of infection. Surgical treatment of bone tuberculosis without joint involvement consists of debridement of the abscess (if it is present) after starting the antituberculous drugs. Suction–irrigation systems are not necessary in tuberculous osteomyelitis. In an occasional patient, if structural instability is anticipated, grafting with cancellous bone chips is done at the time of initial debridement. The treatment of tuberculous arthritis depends on the extent of involvement at the time of detection. Surgery is not indicated if the infection is limited to the synovium (or bursa), with little or no radiographic involvement of the adjacent bone. If the synovium is affected and adjacent bone and cartilage are partially eroded but without gross instability of the joint, there is an argument for synovectomy together with antituberculous drugs to reduce the time needed for convalescence and reduce limitation in the range of joint movement. For advanced joint destruction (complete loss of cartilage and disorganization of the bones), the appropriate surgery is synovectomy, debridement, and fusion of the involved joint. Total arthroplasty is contraindicated in the presence of active tuberculosis. After the infection has been controlled, arthroplasty may be planned for weight-bearing joints (Kim 1988).

All adult patients with spinal tuberculosis are treated with the standardized drug regimen described above (ICMR/BMRC 1989; Medical Research Council 1993). Neurological complications are more frequent when the disease involves the upper and mid-thoracic spine (Omari *et al*. 1989). Surgical intervention is often necessary, especially to obtain adequate specimens for diagnostic purposes. In difficult cases, operative exploration of the lesion is often necessary for the management of complications such as abscesses and sinuses or myelopathy.

Summary of management

The diagnosis of tuberculosis is suspected from the clinical features and appropriate imaging. The final diagnosis is by smear of the aspirate for acid-fast bacilli or from histopathological examination of excised tissue. Appropriate specimens must be obtained before treatment is begun. The material is cultured whenever possible, chiefly to ascertain sensitivity to antituberculous drugs.

The initial drug treatment is with rifampicin, isoniazid, and pyrazinamide for the first 2 months, followed by rifampicin, isoniazid, and ethambutol for 7 or more months. Drug resistance is suspected if there is worsening of clinical signs and symptoms after the first 3 or 4 months of treatment and imaging reveals increased tissue destruction. If reports on culture and antituberculous drug sensitivity are available, then at least four of the appropriate drugs are selected, of which two or three must be bactericidal. In the absence of drug sensitivity reports the treatment must be changed to kanamycin, isoniazid, ethionamide, cycloserine, and ofloxacin or ciprofloxacin. The duration of treatment is then extended by 18 or 24 months.

Surgical intervention is most often necessary to obtain a specimen for diagnosis. In limb bones and vertebrae, necrotic tissue is debrided during the first operation and stabilization, if needed, is done at another operation after 6 or 8 months of drug treatment. Immobilization of the affected region is essential, and for spinal tuberculosis the plaster or acrylic cast must be worn for at least the first 3 months. If spinal instability is demonstrable clinically and on plain radiographs taken in the appropriate postures, then immobilization is prolonged to 6 months or more. Joint infection seldom necessitates synovectomy, except for the occasional patient in whom the diagnosis was delayed and there is a discharging sinus. Most patients not requiring synovectomy can remain ambulatory; if a leg joint is affected, partial weight bearing is recommended for the first 3 or 4 months. This is followed by physiotherapy to build lost muscle mass and tone. Surveillance for 1 year after the end of drug treatment is recommended.

Leprosy

Epidemiology and pathogenesis

The earliest written records describing leprosy come from China and India and date from 600 BC. The true scientific era began when Hensen published his tentative conclusion that the rod-shaped organisms consistently observed in material taken from patients with leprosy were probably responsible for the disease. The World Health Organization estimates that there are over 11 million cases in the world today; that number has not changed much in recent decades, with the majority of patients residing in the poorer nations (Shepard 1982). Leprosy can affect an individual at any age, but cases in infants of less than 1 year old are extremely rare.

Mycobacterium leprae is virtually non-toxic and may infest the human body in large numbers without causing symptoms. Most symptoms of the disease are due to immune reactions against the bacilli. The course after infection in individual patients is variable and determined by the ability of the host to mount an immune response that will limit the bacillary multiplication. There are patients with subclinical infection whose immunological reactivity after exposure to *M. leprae* changes and they do not develop clinical signs. Another group of patients develops skin lesions, which sometimes heal spontaneously, and this is termed the indeterminate form. The third group (estimated as around 10 per cent of all infected) are patients who have the chronic disease, and they chiefly develop dermal and neural involvement. The Ridley and Jopling (1966) classification is accepted widely and provides a good description of variations in clinical course as a basis of accurate clinical diagnosis. It comprises a continuous spectrum with the immunologically stable tuberculoid (**TT**), in which there are a few lesions containing few bacilli, at one pole, and the lepromatous (**LL**) at the other pole, which is a multibacillary disease with fulminant lesions. These poles are bridged by borderline borderline (**BB**), and borderline lepromatous (**BL**).

Leprosy is acquired by direct person-to-person transmission (Reich 1987). Spread within family members and others coming into physical contact with patients is facilitated by the indolence of clinical symptoms and the long incubation period, which varies from 6 months to several decades.

Intense bacillaemia is very common in patients with lepromatous leprosy and the organisms can be seen in stained smears of peripheral blood or buffy coats, but signs of toxaemia, including high fever, are absent. Even in the advanced stages of the disease, the lesions are restricted to the skin, peripheral nerves, anterior portion of eyes, upper respiratory tract, and testes. In lepromatous leprosy, collections of bacilli are also found in bone marrow liver, and spleen.

Although the usual course of leprosy is indolent, occasional interruption by the two types of 'reactional states' is observed in patients who are either untreated or already receiving antileprosy drugs. The type 1 lepra reaction can complicate all the three borderline conditions (BT, BB, and BL) and chiefly consists of a change in the course of the disease towards either the lepromatous or tuberculoid pole. The type 2 reaction (or erythema nodosum leprosum) occurs in patients with lepromatous or borderline leprosy, most frequently in the second half of the first year of treatment. Tender subcutaneous nodules develop and the associated features are low-grade fever, lymphadenopathy, and arthritis. Arthritis in this situation is a classical example of an immune-complex deposition disease.

Immunology and histopathology

Although both cell-mediated and humoral immunities occur simultaneously in leprosy, it is the cell-mediated component that assumes significance (Harboe 1985). Individuals capable of developing cell-mediated immunity localize the disease in the form of tuberculoid in the form of tuberculoid leprosy, while those with depleted cell-mediated immunity express it in the lepromatous form.

The ratio of CD4/CD8 lymphocytes is normal in patients with tuberculoid leprosy and it may be decreased or normal in patients with lepromatous leprosy, depending on the bacterial load. At the site of actual inflammation, CD4 cells constitute about 95 per cent of the lymphocyte population in tuberculoid granulomas and CD8 cells

dominate to the extent of 85 per cent in lepromatous lesions (Narayanan 1988).

There is speculation that macrophage function is defective in patients with lepromatous leprosy (Birdi *et al*. 1989). T cells from patients with lepromatous leprosy fail to produce interleukin 2 after exposure to *M. leprae*. The nature of the humoral immune response in leprosy is not clearly defined (Sehgal *et al*. 1989). As in other chronic infections, the concentrations of serum gammaglobulins increase in leprosy.

The earliest histopathological events in human leprosy are not known. Perhaps the majority of infections are overcome at the site of entry (possibly the mucosa of the respiratory tract) and the bacilli destroyed. The Schwann cells of the nerves in the upper respiratory tract may be the first to harbour the bacilli and haematogenous spread occurs thereafter. The skin is probably involved via the endothelial cells of small vessels. Bacilli are seen in superficial nerve plexuses and perivascular macrophages in the skin in early lesions, with mild local lymphocytic infiltration. In many cases (indeterminate leprosy) these lesions resolve spontaneously with eradication of bacilli. If the bacilli do persist and multiply in skin and/or nerves, the inflammatory reactions are amplified and lesions of leprosy appear. These lesions can be placed along an immunohistopathological spectrum that ranges from organized epithelioid-cell granulomas containing giant cells with few or no bacilli, through intermediate stages with less organized epithelioid cells containing more bacilli, to lepromatous leprosy with mainly macrophages and abundant bacilli. In the reactional state (type 1), there is increased inflammation in the lesions with activation of epithelioid cells and marked oedema. In type 2 states the classical lesions show oedema, polymorphonuclear leucocyte infiltration, and often a necrotizing vasculitis.

The intraosseous lesions of leprosy are characterized by granulomatous tissue reactions that lead to trabecular destruction. The lesions are evident in the epiphysis and metaphyses of tubular bones, and direct involvement of the medullary canal can also occur.

Clinical manifestations

The first signs of leprosy are usually cutaneous. Single or multiple, hypopigmented macules or plaques may appear. Sometimes an anaesthetic patch is first noticed by the patient, but often sensation remains preserved in early lesions, particularly on the face. In tuberculoid leprosy, the fully developed skin lesions are densely anaesthetic and have lost sweat glands and hair follicles. Their distribution is not symmetrical. Nerve involvement occurs early and the superficial nerve leading from lesion is enlarged. The supraorbital, facial, greater auricular, ulnar, median, radial cutaneous, common peroneal, sural, anterior, or posterior tibial nerves may be grossly enlarged and easily palpated. There may be severe paraesthesias initially, followed by muscle atrophy. If the disease progresses and the hands and feet are affected, then contractures develop (Paterson and Rad 1961). Repeated trauma in the absence of protective sensation results in ulcers that get secondarily infected (Fig. 7). In a more advanced stage, osteolysis of the terminal phalanges occurs. In lepromatous leprosy the skin lesions are macules, nodules, papules, or plaques with predilection for the face, wrists, elbows, buttocks, and knees. Involvement of major nerve trunks in common and leads to glove-and-stocking paraesthesias in the extremities.

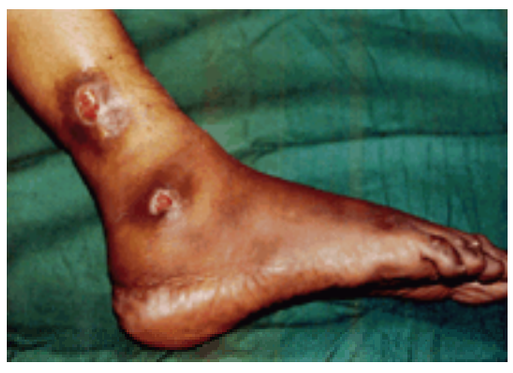


Fig. 7 Trophic ulcers in a patient with tuberculoid leprosy.

The incidence of direct involvement of the skeleton in leprosy is probably under-reported. Symmetrical, peripheral, inflammatory polyarthritis of insidious onset with a pattern of exacerbation and remission is observed in some patients (Atkin *et al*. 1989), which is different from the joint involvement of erythema nodosum leprosum or lepra type 1 reaction. Affected joints can include the wrist, knees, and hands and feet. Morning stiffness is variable from 30 min to 1 h. Symptomatic improvement is observed within a few weeks of starting multidrug treatment.

A different pattern of involvement is with changes that are usually confined to the small bones of the face, the hands, and feet (Atkin *et al*. 1987). At the fingers and toes it may appear as a dactylitis. The osseous involvement is probably due to extension of infection from overlying dermal or mucosal areas. The periosteum is contaminated initially (leprous periostitis) and subsequently the cortex and marrow are infected (leprous osteitis and osteomyelitis). In an occasional patient, haematogenous spread of infection can occur, leading to other intramedullary foci in the tubular long bones and the ribs. Overall, the progression of the lesions is very slow.

Neuroarthropathy is a progressive, degenerative joint change due to lesions of peripheral nerves and is observed more frequently in patients admitted to hospital. The skeletal changes may follow the involvement of nerves by two or three decades. The bones of the hands and feet are most susceptible. In the feet the changes usually start in the medial arch and later involve the lateral arch, talus, and calcaneus (Horibe *et al*. 1988). In extreme cases, dissolution of the mid-foot results in separation of the forefoot and the hind foot, and the tibia is driven downwards to become weight-bearing. Infection and bone injury, which can occur separately as sequelae to neuropathy and trophic ulcers, tend to accelerate the skeletal changes, leading to disintegration of the affected bones and joints.

The acute and chronic arthritis associated with the reactional state of erythema nodosum leprosum is common (Karat *et al*. 1966). The development of synovitis and/or dactylitis coincides closely with the appearance of fulminant skin lesions (Fig. 8), and the patient is febrile and toxic.

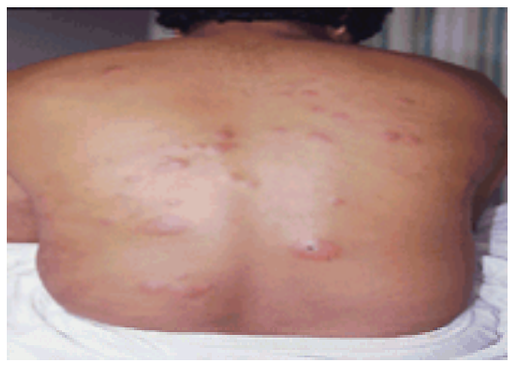


Fig. 8 Erythema nodosum lesions in the lepra reaction.

Investigations and diagnosis

The principal criteria for the diagnosis of leprosy are:

1. a hypopigmented patch of skin with sensory impairment;
2. thickening of the peripheral nerves;

3. the presence of *M. leprae* in slit-skin smears—graded on the Ridley Jopling bacteriological index (**BI**) as 1+ to 6+;
4. the characteristic histopathological changes in a biopsy specimen of a skin lesion or a peripheral nerve;
5. positive serological tests.

Patients with lepromatous leprosy frequently have mild anaemia, an elevated erythrocyte sedimentation rate, and hypergammaglobulinaemia; 10 to 20 per cent of patients show a false-positive reaction in serological tests for syphilis, and also antinuclear antibodies.

Myocobacterium leprae is an acid-fast bacillus morphologically and biochemically similar to *M. tuberculosis*. It does not grow in artificial media or tissue cultures but is propagated consistently in the footpads of mice and nine-banded armadillos for the purpose of epidemiological studies and drug evaluation. The bacillus is slow to multiply and the doubling time is around 12 days in optimal conditions in murine footpads. The technique is expensive and time-consuming, and it takes at least 10 months for the organism to grow.

An enzyme-linked immunosorbent assay to detect antibodies to the *M. leprae* specific surface antigen has been developed (Burgess *et al*. 1988). The antigen used in the assay is a phenolic glycolipid 1 and the antibodies detected are predominantly IgM. Titres increase from the tuberculoid to the lepromatous pole of the leprosy spectrum. The early estimates of the sensitivity of this test were 95 per cent for the lepromatous pole and 30 per cent for the tuberculoid pole. This test is not yet available for use in clinical practice. In research laboratories, the polymerase chain-reaction assay is used to detect *M. leprae* DNA in skin biopsy samples (Jamil *et al*. 1993). Owing to its ability to detect small numbers of organisms, the polymerase chain reaction may prove to be a useful tool in paucibacillary disease.

The radiographic features of leprous osteitis are enlarged nutrient foramina, osteoporosis, endosteal thinning, and cyst-like lesions that are best appreciated in the phalanges. Bone sclerosis, which signifies the healing process, may also be seen. In the face, nasal destruction is characteristic. Destruction of the alveolar process and the anterior nasal spine of the maxilla appear to be related to primary involvement of the bone as well as to secondary infection. In patients with neuroarthropathy the typical changes begin with erosion of the terminal tufts of the phalanges and in the more advanced cases the entire length of terminal group of phalanges may be resorbed (MacMoran *et al*. 1987). A rare but specific radiographic finding is calcification of a large peripheral nerve such as the radial or ulnar.

Treatment

Multidrug therapy is the mainstay of treatment for leprosy (Ellard 1988; Hasting *et al*. 1988; Ye 1988). Rifampicin is by far the most potent bactericidal antileprosy drug. Dapsone is well tolerated, and until two decades ago it was the only drug used (monotherapy) in treatment of all forms of leprosy. Clofazimine is the third drug; it has the additional property of contributing to the control of erythema nodosum leprosum through its anti-inflammatory action. The main drawback of clofazimine is that it causes a deep-brown pigmentation of skin, more evident in light-skinned persons, which is slow to regress after cessation of treatment.

The current multidrug therapy recommended by the World Health Organization for leprosy in adults is, for multibacillary disease (BI>2+), rifampicin, 600 mg orally once monthly, dapsone 100 mg orally daily, and clofazimine, 300 mg orally once monthly and additionally 50 mg daily. The duration of treatment is at least 2 years and until negative skin tests are obtained, which occasionally extends to beyond 5 years. For paucibacillary disease (BI, 2+) the regimen is rifampicin, 600 mg orally once monthly, and dapsone, 100 mg orally daily. The duration of treatment is 6 months. All drugs are prescribed in full doses from the beginning of treatment and continued without interruption, even during the reactional states.

Evidence of clinical improvement should appear after 8 to 12 weeks from the beginning of treatment. The clinical response to adequate therapy may be confused by reactional states. Mild lepra reactions are controlled by non-steroidal anti-inflammatory drugs such as aspirin or indomethacin. In severe cares, prednisolone (60-100 mg/daily) may be required. For the reactional state of erythema nodosum leprosum, clofazimine 100 mg orally three times a day for 2 or 3 weeks followed by tapering off to a lower dose may be helpful. An alternative, effective drug is thalidomide, especially for those patients who later take a chronic course with erythema nodosum leprosum. The usual dose of thalidomide is 200 mg orally twice daily, which is then gradually tapered off to a maintenance dose of 50 to 100 mg daily. Thalidomide is absolutely contraindicated in pregnancy.

Non-steroidal anti-inflammatory drugs usually fail to alleviate the symptoms of the polyarthritis, and they regress within a few weeks after starting antibiotic treatment. Those who present with neuroarthropathy are far more difficult to treat, particularly when the feet are involved (Warren 1973). Absolute rest for the affected limbs and control of superimposed infection are mainstays of treatment. Immobilization in plaster until all ulcerative lesions of the skin have healed may be necessary. Regular exercises are required to maintain the flexibility of the other joints until the ulcers heal. Thereafter, graded weight bearing is permitted. Amputation may be indicated if advanced disintegration of bones and joints has occurred, as the patient would be better off with a prosthesis. Advanced deformities of the hands, associated with tendon rupture, can be functionally improved by appropriate tendon-transfer surgery.

Much effort in research is being directed to control of leprosy. Trials with bacille Calmette–Guérin have so far yielded conflicting results, with some revealing only modest efficacy. Two new vaccines (Antia *et al*. 1988), one named the ICRC and the other containing viable bacille Calmette–Guérin with heat-killed *M. leprae*, are currently in field trials.

Summary of management

The diagnosis is usually evident from the clinical examination of the skin, and tender nerves are palpable at the elbows, wrists, and near the head of the fibula. Polyarthritis and enthesitis are sometimes observed in all stages of the disease. Multidrug treatment is the mainstay of therapy, and the duration largely depends on the extent of dermal involvement. Ideally the bacillary count on the slit-skin smear enables one to judge if the patient has pauci- or multibacillary disease. Dapsone and rifampicin are given for 6 or 12 months to patients with paucibacillary disease. Those with multibacillary disease are given dapsone, rifampicin, and clofazimine for 2 years, and if the slit-skin smears remain positive the treatment may be extended to 5 years. The symptoms of arthritis and enthesitis regress within 3 or 4 months of the start of treatment. In late stages, some patients develop a neuropathic foot and subluxation of joints that require surgical rehabilitation. Reactive states are occasionally observed in the first year of treatment and the patient is then given aspirin or low-dose prednisolone for 6 to 8 weeks to alleviate the symptoms, while the multidrug therapy is continued.

Chapter References

American Thoracic Society (1990). Diagnostic standards and classification of tuberculosis. *American Review of Respiratory Diseases*, **142**, 725–35.

Antia, N. H. *et al*. (1988). Leprosy vaccine—a reappraisal. *International Journal of Leprosy and Other Mycobacterial Diseases*, **56**, 231–7.

Arachi, A. (1991). The global tuberculosis situation and new control strategy of the World Health Organisation. *Tubercle*, **72**, 1–6.

Atkin, S. L. *et al*. (1987). Clinical and laboratory studies of inflammatory polyarthritis in patients with leprosy in Papua New Guinea. *Annals of the Rheumatic Diseases*, **46**, 688–90.

Atkin, S. L. *et al*. (1989). Clinical and laboratory studies of arthritis in leprosy. *British Medical Journal*, **298**, 1423–5.

Birdi, T. J. *et al*. (1989). The macrophage in leprosy: A review of the current status. *International Journal of Leprosy and Other Mycobacterial Diseases*, **57**, 511–25.

Boxer, D. *et al*. (1992). Radiological features during and following treatment of spinal tuberculosis. *British Journal of Radiology*, **65**, 476–9.

Brisson-Noel, A. *et al*. (1989). Rapid diagnosis of tuberculosis by amplification of mycobacterial DNA in clinical samples. *Lancet*, **ii**, 1069–71.

Burgess, P. J. *et al*. (1988). Serologic tests in leprosy. The sensitivity, specificity and predictive value of ELISA tests based on phenolic glycolipid antigens and the implications of their use in epidemiological studies. *Epidemiology of Infections*, **101**, 159–71.

Bush, D. C. and Schneider, L. H. (1984). Tuberculosis of the hand and wrist. *Journal of Hand Surgery*, **9A**, 391–8.

Butorac, R. *et al*. (1987). Mycobacterial disease in the musculoskeletal system. *Medical Journal of Australia*, **147**, 388–91.

Chang, K. H. *et al*. (1989). Tuberculous arachnoiditis of the spine: findings on myelography, CT and MR imaging. *American Journal of Neuroradiology*, **10**, 1255–62.

Cooke, N. J. (1985). Treatment of tuberculosis. *British Medical Journal*, **291**, 497–8.

Dall, L. *et al*. (1989). Poncet's disease: tuberculous rheumatism. *Review of Infectious Diseases*, **11**, 105–7.

Daniel, T. M. (1989). Rapid diagnosis of tuberculosis: laboratory techniques applicable in developing countries. *Review of Infectious Diseases*, **11** (Suppl. 2), 5471–8.

Davidson, P. T. (1989). The diagnosis and management of disease caused by *M. avium* complex, *M. kansasii* and other mycobacterias. *Clinics in Chest Medicine*, **10**, 431–43.

Davidson, P. T. and Horowitz, I. (1970). Skeletal tuberculosis: a review with patient presentations and discussion. *American Journal of Medicine*, **48**, 77–84.

Davies, P. D. *et al*. (1984). Bone and joint tuberculosis: a survey of notifications in England and Wales. *Journal of Bone and Joint Surgery*, **66B**, 326–30.

Dutt, A. K. (1989). Update in tuberculosis: an overview. *Seminars in Respiratory Infections*, **4**, 155–6.

Edeikin, J. *et al*. (1963). 'Cystic' tuberculosis of bone. *Clinical Orthopedics*, **28**, 163–8.

Ellard, G. A. (1988). Chemotherapy of leprosy. *British Medical Bulletin*, **44**, 775–90.

Ellner, J. J. *et al*. (1989). Immunologic aspects of mycobacterial infections. *Review of Infectious Diseases*, **11** (Suppl. 2), 455–9.

Garrido, G. *et al*. (1988). A review of peripheral tuberculous arthritis. *Seminars in Arthritis and Rheumatism*, **18**, 142–9.

Goldberger, M. J. (1988). Anti tuberculosis agents. *Medical Clinics of North America*, **72**, 661–8.

Hannachi, M. R. (1988). Comparison of 3 chemotherapeutic regimens of short duration (6 months) in osteoarticular tuberculosis—results after 5 years. *Bulletin of the International Union for Tubercular Lung Disease*, **63**, 14–15.

Harboe, M. (1985). Immunology of leprosy. In *Leprosy* (ed. R. C. Hastings), pp. 53–87. Churchill Livingstone, New York.

Hastings, R. C. *et al*. (1988). Chemotherapy of leprosy. *Annual Review of Pharmacology and Toxicology*, **28**, 231–45.

Horibe, S. *et al*. (1988). Neuroarthropathy of the foot in leprosy. *Journal of Bone and Joint Surgery*, **70**, 481–5.

ICMR/BMRC (1989). A controlled trial of short-course regimens of chemotherapy in patients receiving ambulatory treatment or undergoing radical surgery for tuberculosis of spine. *Indian Journal of Tuberculosis*, **36** (Suppl.), 1–21.

Ivanyi, J. (1988). Inmunodiagnostic assays for tuberculosis and leprosy. *British Medical Bulletin*, **44**, 635–49.

Jamil, S. *et al*. (1993). Use of polymerase chain reaction to assess efficacy of leprosy chemotherapy. *Lancet*, **342**, 264–8.

Karat, A. B. A. *et al*. (1966). Acute exudative arthritis in leprosy: rheumatoid arthritis like syndrome in association with erythema nodosum leprosum. *British Medical Journal*, **ii**, 770–3.

Kim, Y. H. (1988). Total knee, arthroplasty for tuberculous arthritis. *Journal of Bone and Joint Surgery*, **70A**, 1322–30.

Kumar, K. (1985). A clinical study and classification of posterior spinal tuberculosis. *International Orthopedics*, **9**, 147–52.

Kumar, K. *et al*. (1988). Multifocal osteoarticular tuberculosis. *International Orthopedics*, **12**, 135–8.

LaBerge, J. M. and Brant-Zawadzki, M. (1984). Evaluation of Pott's disease with computed tomography. *Neuroradiology*, **26**, 429–34.

LeRoux, P. *et al*. (1990). Tuberculosis of the skull—a rare condition: case report and review of the literature. *Neurosurgery*, **26**, 851–6.

Levy, A. *et al*. (1986). Early recognition of tuberculous arthritis assisted by CT scan and closed needle synovial biopsy. *Clinical Rheumatology*, **5**, 523–6.

Lifeso, R. M. (1987). Atlanto axial tuberculosis in adults. *Journal of Bone and Joint Surgery*, **69B**, 183–7.

Lucas, S. B. (1988). Histopathology of leprosy and tuberculosis-an overview. *British Medical Bulletin*, **44**, 584–99.

MacMoran, J. W. *et al*. (1987). Bone loss in limbs with decreased or absent sensation: ten year follow up of the hands in leprosy. *Skeletal Radiology*, **16**, 452–9.

Martini, M. *et al*. (1986). Tuberculous osteomyelitis. A review of 125 cases. *International Orthopedics*, **10**, 201–7.

Martini, M. *et al*. (1988). Bone and joint tuberculosis: a review of 652 cases. *Orthopedics*, **11**, 861–6.

Medical Research Council (1993). Controlled trial of short-course regimens of chemotherapy in the ambulatory treatment of spinal tuberculosis. *Journal of Bone and Joint Surgery*, **75B**, 240–8.

Modilevsky, T. *et al*. (1989). Mycobacterial disease in patients with human immunodeficiency virus infection. *Archives of Internal Medicine*, **149**, 2201–5.

Narayanan, R. B. (1988). Immunopathology of leprosy granuloma. Current status: a review. *Leprosy Review*, **59**, 75–82.

Omari, B. *et al*. (1989). Potts' Disease—a resurgent challenge to the thoracic surgeon. *Chest*, **95**, 145–50.

Paterson, D. E. and Rad, M. (1961). Bone changes in leprosy: their incidence, progress and prevention and arrest. *International Journal of Leprosy*, **29**, 393–422.

Pouchot, L. *et al*. (1988). Tuberculosis of the sacroiliac joint: clinical features, outcome, and evaluation of closed needle biopsy in 11 consecutive cases. *American Journal of Medicine*, **84**, 622–8.

Reich, C. V. (1987). Leprosy: cause, transmission and a new theory of pathogenesis. *Review of Infectious Diseases*, **9**, 590–4.

Ridley, D. S. and Jopling, W. H. (1966). Classification of leprosy according to immunity: a five group system. *International Journal of Leprosy*, **34**, 255–73.

Schluger, N. *et al*. (1994). Clinical utility of the polymerase chain reaction in the diagnosis of infections due to *Mycobacterium tuberculosis*. *Chest*, **105**, 1116–21.

Sehgal, V. N. *et al*. (1989). Immunology of leprosy: a comprehensive survey. *International Journal of Dermatology*, **28**, 571–84.

Shepard, C. C. (1982). Leprosy today. *New England Journal of Medicine*, **307**, 1640–1.

Shoemaker, S. A. (1986). Molecular biology and mycobacteriae. *Seminars in Respiratory Infections*, **1**, 265–9.

Sifford, M. and Bates J. H. (1991). Host determinants of susceptibility to mycobacterium tuberculosis. *Seminars in Respiratory Infections*, **6**, 44–50.

Smith, A. S. *et al*. (1989). MR imaging characteristics of tuberculous spondylitis vs. vertebral osteomyelitis, *American Journal of Roentgenology*, **153**, 399–405.

Stanford. J. I. (1983). Immunologically important constituents of mycobacteria. In *Biology of mycobacteria* (ed. C. Ratledge and J. L. Stanford), pp. 85–127. Academic Press, London.

Warren, G. (1973). The management of tarsal disintegration. *Leprosy Review*, **43**, 137–47.

Whelan, M. A. *et al*. (1985). Computed tomography of nontuberculous spinal infection. *Journal of Computer Assisted Tomography*, **9**, 280–7.

Wolfgang, G. I. (1978). Tuberculosis joint infection. *Clinical Orthopedics*, **136**, 257–63.

Ye, G. Y. (1988). Diagnosis of leprosy. *International Journal of Dermatology*, **27**, 263–4.

5.3.8 Brucellar arthritis

E. Pascual

Introduction

Infection of the joints is the most frequent, localized complication of brucellosis, and a common cause of infectious arthritis in the countries where the disease is endemic.

Epidemiology

Brucellosis occurs naturally in domesticated animals. It constitutes an important economic problem and a serious heath hazard in many countries, especially in the Mediterranean region, Arabian peninsula, Indian subcontinent, Mexico, and parts of Central and South America. Of the six recognized species of brucella, only four are known to be human pathogens (Young 1995); *Brucella melitensis*, the most virulent, causes disease in goats, sheep, and camels, *B. abortus* in cattle, *B. suis* in pigs, and *B. canis* in dogs. Multiple viobars have been identified in *B. melitensis*. Localization of brucellae in the male and female reproductive organs accounts for the major clinical manifestation, that is abortion. Human infection is contracted from infected animals and is closely linked to poor methods of animal husbandry, feeding habits, and hygiene standards. The disease is common in some professions due to handling of infected animals or viscera, where the organism is acquired either through breaches in the skin or by infectious aerosol reaching the conjunctiva or the airways. Ingestion of unpasteurized dairy products is the most common cause of infection due to *B. melitensis*. Because of their high infectiveness, brucellae are a common cause of accidental infection in laboratory workers (Miller *et al*. 1987; Young 1995).

Infection acquired during foreign travel, often through consumption of infected, illegally marketed dairy products, is the cause of imported disease (Arnow *et al*. 1984; Revak *et al*. 1989). The possibility of illegal import of such products should be borne in mind (Thapar and Young 1986).

General characteristics of the organism

Brucellae are small, aerobic, non-motile, gram-negative coccobacilli which grow well at 37°C in any high quality peptone-based medium enriched with blood or serum. Their growth requires a much longer incubation period than pyogenic organisms and many strains of *B. abortus* and *B. suis* require supplementary CO_2, as described below (Young 1995).

The host defence

Experiments in rats have shown that after entering the blood stream, brucellae are phagocytosed by polymorphonuclear leucocytes. Within hours, phagocytosed organisms can be seen in the mononuclear phagocytic cells of the liver sinusoids, lymph nodes, spleen, and bone marrow—and probably in other organs rich in mononuclear phagocytic cells— where they reproduce (Spink 1964; Smith and Ficht 1990). The frequent isolation of brucellae from bone marrow cultures in diseased humans (Gotuzzo *et al*. 1986) as well as from liver biopsies (Spink 1964) indicates a similar distribution in humans. The permanence and reproduction of the organisms inside the cells may be important in understanding some of the characteristics of the disease.

The tissue response in established disease is a non-specific granulomatous lesion, very similar to sarcoidosis (Hunt and Bothwell 1967). Brucellae cannot be identified in these tissues but can be cultured from them (Spink 1964). Localized areas of suppuration or caseation may occur, most frequently in *B. suis* (Spink 1964) but also in *B. melitensis*.

General characteristics of the disease

The manifestations of brucellosis are non-specific and no combination of signs or symptoms can be considered to be characteristic. *B. melitensis* causes more severe disease than *B. abortus* and *B. suis*, probably because of its greater ability to avoid the host's defences (Smith and Ficht 1990). *B. suis* has a higher tendency to suppurative complications. The large, published series have focused on patients infected by *B. melitensis*, and generalization of this data to disease caused by the other brucellae may be inaccurate. Reference to the other infections will be indicated in the text, as appropriate.

Brucellosis is more common in men than in women both in the Middle East and Mediterranean countries (Colmenero *et al*. 1985; Andonopoulos *et al*. 1986; Mousa *et al*. 1987; Lulu *et al*. 1988; Ariza 1988; Batlle *et al*. 1989). In the same areas, the disease occurs with equal frequency in children of both sexes (Feiz *et al*. 1978; Lubani *et al*. 1986; Gómez -Reino *et al*. 1986; Al-Eissa *et al*. 1990), as happens with adults in Peru (Gotuzzo *et al*. 1987). It may be that where brucellosis is contracted as a professional hazard, more males are affected. Ingestion of milk, which is the usual cause of the disease in children and Peruvian adults (Gotuzzo *et al*. 1987), would result in equal incidence in both sexes.

The incubation period may be as short as 1 week but usually lasts from 2 to 8 weeks (Young 1995). The disease usually presents with fever, often without an undulant pattern, sweats, which may be drenching, and a general feeling of malaise. Weakness, anorexia, myalgia, and arthralgia are common. Patients often recall contact with possibly infected animals or their unpasteurized products but this evidence may be absent. Physical examination may show lymphadenopathies and enlargement of liver or spleen. Localized infections, especially in the skeleton, are common. Other possible causes, such as neurobrucellosis, endocarditis, hepatitis, and epididimo-orchitis, should be considered.

Subclinical disease occurs and may heal spontaneously. Patients may present with relapses following previous subclinical disease or with late, localized complications, simulating a very long incubation period.

The concept of chronic brucellosis was coined before the antibiotic era, to refer to a group of patients who complained of ill health after having suffered from brucellosis, generally due to *B. abortus*. A careful study (Spink 1951) showed that some of these patients had a bacteriologically proven relapse, or localized disease; in others, no explanation was found for the symptoms and it was felt that they were related to an unstable emotional state. It is unclear whether chronic brucellosis in the absence of infection exists and such diagnosis must be handled with care.

Brucellar arthritis

Arthralgia is more common in brucellosis than in other febrile illnesses. It has been recorded in 88 per cent (Thapar and Young 1986), 65 per cent (Colmenero *et al*. 1986), 65 per cent (Ariza 1988), and 58 per cent (Lulu *et al*. 1988) of adult patients studied prospectively, and in 87 per cent (Feiz *et al*. 1978) and 74 per cent (Al-Eissa *et al*. 1990) of the children thus studied; its presence may be a clue to the disease. Bone scans of patients with brucellosis and musculoskeletal symptoms

are very frequently abnormal, often in multiple sites (El Desouki 1991). Perhaps some of the minor musculoskeletal symptoms of these patients are due to localized infection. Since the treatment of brucellar arthritis and that of uncomplicated brucellosis is the same, once the diagnosis of brucellosis has been reached, it is not worthwhile conducting a search for arthritis to explain minor or unclear musculoskeletal symptoms.

Infection in the joints is the most common form of localized disease in brucellosis. It appears in 25 per cent (Colmenero *et al*. 1991), 29 per cent (Ariza 1988), 26 per cent (Lulu *et al*. 1988), 24 per cent (Gotuzzo *et al*. 1987), and 22 per cent (Andonopoulos *et al*. 1986) of adult patients studied prospectively, and in 38 per cent (Al-Eissa *et al*. 1990), 28 per cent (Llorens-Terol and Busquets 1980), and 19 per cent (Feiz *et al*. 1978) of children. In areas where brucellosis is endemic, brucellar arthritis may outnumber tuberculous arthritis (Rajapakse 1987). All the above series refer to *B. melitensis*.

The general clinical characteristics of brucellar arthritis are similar to those of infectious arthritis due to other organisms. The disease is generally monoarticular, but in 18 per cent of our patients more than one joint was affected (Batlle *et al*. 1989). The large peripheral joints, sacroiliacs, and the spine are the usual sites of articular involvement. Bursitis and osteomyelitis may occur. Some patients have simultaneous infection in more than one joint. The possibility of reactive arthritis has been considered.

Clinical and radiological characteristics

Arthritis of peripheral joints

Large peripheral joints are a common site of localized infection, ranging from 20 to 73 per cent of all arthritides in different adult series (Serre *et al*. 1981; Andonopoulos *et al*. 1986; Gotuzzo *et al*. 1987; Ariza 1988; Al-Rawi *et al*. 1989; Batlle *et al*. 1989) (Table 1). It is of interest that children nearly always show peripheral joint involvement, either with *B. melitensis* (Lubani *et al*. 1986; Gómez-Reino *et al*. 1986; Gotuzzo and Carrillo 1988; Benjamin *et al*. 1992) or *B. abortus* infections (Adam *et al*. 1967). *B. canis* knee arthritis in a child has been reported (Young 1983). In all age groups, the hip followed by the knee are the most common locations. Ankles, shoulders, elbows, wrists, and sternoclavicular joints (Berrocal *et al*. 1993) contribute a small percentage of cases. The smaller hand and feet joints seem to be rarely affected. A 'rheumatoid like' distribution may occur (Gotuzzo *et al*. 1982).

	Peripheral joints*	Sacroiliacs	Spine
Adults	Frequent	Frequent	Frequent
Children	Frequent	Infrequent (older children)	Exceptional

*In order of frequency: hip, knee, ankle, shoulder, elbow, wrist, and sternoclavicular joint.

Table 1 Distribution of joint involvement

Peripheral joint arthritis often, but not always, predominates over other disease manifestations, and the patient's main complaint is centred on the joint. Usually it occurs during the acute phase of the disease or during a relapse (Gotuzzo *et al*. 1982); fever, or other general symptoms of the disease, frequently accompany the arthritis, which tends to be symptomatic from the start and so patients tend to seek prompt medical attention. The joints are swollen and painful and an effusion is generally seen. Joint inflammation is not as intense as it is in some septic arthritides due to pyogenic organisms, and there is seldom obvious redness or unusual warmth in the skin. Local complications, such as popliteal cyst rupture, may occur (Laajam 1985).

If there is delay in the diagnosis or treatment, the structure of the joint is damaged resulting in radiological abnormalities, such as joint space narrowing (Fig. 1) and damage to the joint surfaces, similar to those of other infectious arthritides. The interval between the initiation of the infection and the appearance of structural damage in the joint tends to be longer than in pyogenic arthritides, but probably shorter than in tuberculosis. The sequelae left after appropriate treatment depends on the damage present when treatment was started. If this is started early enough, no sequelae will remain. In children with arthritis of the hip, dislocation and aseptic necrosis have been observed (Benjamin and Khan 1994). Brucellae may infect prosthetic joints (Agarwal *et al*. 1991). Bursitis and tenosynovitis due to *B. melitensis* (Mousa *et al*. 1987), *B. abortus*, and *B. suis* also may occur (Kelly *et al*. 1960).

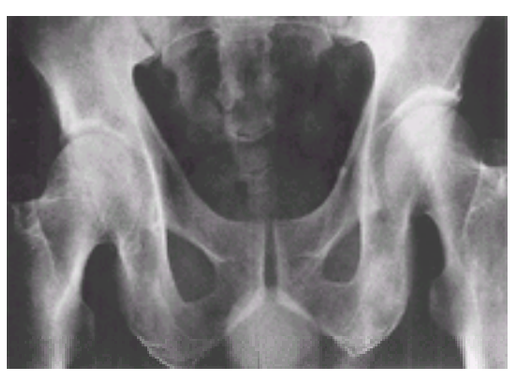


Fig. 1 Radiograph of the hips showing diminished radiological joint space in a patient with brucellar arthritis of his right hip.

Sacroiliitis

Involvement of the sacroiliac joint comprises a large percentage of brucellar arthritides. Higher frequencies have been reported in the west Mediterranean countries—43 per cent in men and 20 per cent in women (Rotés-Querol 1957); 34 per cent (Serre *et al*. 1981); 51 per cent (Ariza 1988); and 36 per cent (Batlle *et al*. 1989)—as compared to countries further east—11 per cent in Greece (Andonopoulos *et al*. 1986) and 13 per cent and 5 per cent in the Middle East (Mousa *et al*. 1987; Al-Rawi *et al*. 1989). An incidence of 50 per cent was found in Peru (Gotuzzo *et al*. 1982). The reasons for these differences are unclear. Sacroiliitis seems to be an unusual complication of *B. abortus* and *B. suis* (Kelly *et al*. 1960). Sacroiliitis is uncommon in children (Lubani *et al*. 1986; Gómez-Reino *et al*. 1986) but young adults have a clearly increased risk for arthritis in this particular joint (Rotés-Querol 1957; Gotuzzo *et al*. 1982).

Sacroiliitis tends to occur in the acute, febrile phase of the disease, is usually unilateral, and clearly symptomatic from its start. The pain in some of the patients becomes so intense in 2 or 3 days that they can hardly move from the bed (Mousa *et al*. 1987). Radiation of the pain to the buttock, posterior thigh, and even below, is not unusual. In these patients, any movement of the leg is very painful and tests for hip manoeuvres or straight leg elevation may seem positive. In this setting, clinical differentiation of sacroiliitis from hip disease or sciatica may be difficult (Mousa *et al*. 1987; Gotuzzo and Carrillo 1988); gently tapping on the lower surface of the heel while the patient keeps the leg extended, clearly localizes the pain to the sacroiliac area. Other sacroiliac manoeuvres may produce too much pain, and may be difficult to interpret in this group. Standard manoeuvres provide appropriate information in patients with less intense symptoms.

Due to the acute nature of the sacroiliitis, the presenting radiographs do not aid the diagnosis; some blurring of the articular margins and widening of the sacroiliac space (Fig. 2) may occur when the disease is of longer duration (Ariza *et al*. 1993). Computed tomography (**CT**) scan generally allows earlier demonstration of joint infection. A bone scan provides early evidence of sacroiliitis on most occasions (Mousa *et al*. 1987; Bahar *et al*. 1988; Madkour *et al*. 1988; Cordero Sánchez *et al*.

1990). Brucellar sacroiliitis is a mild disease associated with a good outcome similar to that observed for patients with uncomplicated brucellosis (Ariza *et al*. 1993).

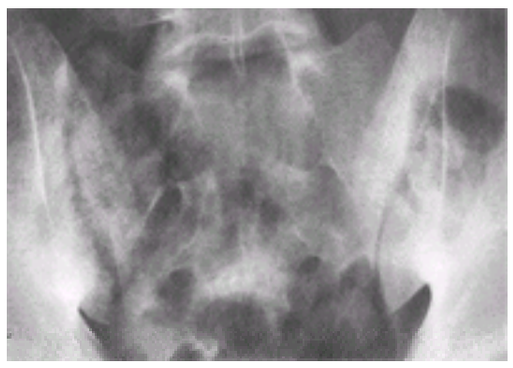


Fig. 2 Radiograph of the sacroiliac joints, showing widening and erosions of the right joint in a patient with long-standing right brucellar sacroiliitis.

Spondylitis

Infection in the spine is the cause of between 7 and 53 per cent of localized skeletal infection (Rotés-Querol 1957; Serre *et al*. 1981; Gotuzzo *et al*. 1982; Gotuzzo and Carrillo 1988; Ariza 1988; Al-Rawi *et al*. 1989; Batlle *et al*. 1989). Referral patterns to hospitals may account for some of the discrepancies. It is generally seen in older patients (Rotés-Querol 1957; Serre *et al*. 1981; Gotuzzo *et al*. 1982; Lifeso *et al*. 1985; Ariza *et al*. 1985a; Gotuzzo and Carrillo 1988 ; Colmenero *et al*. 1991) but not in children, neither with *B. melitensis* (Lubani *et al*. 1986; Gómez-Reino *et al*. 1986; Gotuzzo and Carrillo 1988) nor with *B. abortus* (Adam *et al*. 1967). The majority of patients from the United Kingdom or United States described as suffering from skeletal brucellosis are isolated cases or small series and are due to *B. abortus*, or *B. suis* (Kelly *et al*. 1960; Torres-Rojas *et al*. 1979; Manaster 1988); their clinical characteristics are similar to those caused by *B. melitensis*.

The lumbar spine is the most commonly affected segment, followed by the dorsal spine. The cervical spine accounts for a small percentage of the cases. Involvement of more than one disc space is not infrequent (Rotés-Querol 1957; Kelly *et al*. 1960; Serre *et al*. 1981; Ariza *et al*. 1985a; Lifeso *et al*. 1985; Gotuzzo and Carrillo 1988; Sharif *et al*. 1989). In contrast to peripheral arthritis or sacroiliitis, spondylitis is generally a late feature, often occurring in patients with only vague symptoms of a general infectious disorder, and not infrequently afebrile (Serre *et al*. 1981; Lifeso *et al*. 1985; Gotuzzo and Carrillo 1988), although careful register of the temperature may show a low grade fever and patients may complain of some malaise. In other series, 17 out of 20 patients had systemic symptoms, but these patients presented with the spine infection shortly after the start of the disease (Ariza *et al*. 1985a).

Onset of the disease is often insidious, and its hallmark is local pain of variable intensity, often moderate, frequently allowing the patient to maintain a fairly normal life. Pressure or percussion of the spinous processes of the vertebrae corresponding to the affected level often reproduces the pain, and some tenderness may also occur in the paravertebral muscles. Deformities are uncommon (Rotés-Querol 1957). Paravertebral abscesses appear in about 16 per cent of the patients (Ariza *et al*. 1985a) but are generally smaller than those of tuberculosis. In the cervical spine, abscesses may produce retropharyngeal swelling (Lifeso *et al*. 1985). Material obtained from the abscesses is usually sterile (Rotés-Querol 1957; Ariza *et al*. 1985a). Similar abscesses may occur with *B. abortus* (Torres-Rojas *et al*. 1979). Some patients may develop epidural abscesses from the infected disc space and develop a paraplegia. Compression of the medulla or nerve roots is found more frequently in cervical spondylitis, which should be considered a severe manifestation of the disease (Colmenero *et al*. 1992). In this case myelography shows extradural compression and these patients may need surgical decompression (Lifeso *et al*. 1985); despite surgery, severe sequelae may occur (Colmenero *et al*. 1991).

In the very early phases, radiographs of the affected segments are normal. After 2 months, nearly all patients show some radiographic alterations (Ariza *et al*. 1985a; Lifeso *et al*. 1985). The first sign seen is narrowing of the disc (Fig. 3), without bone abnormalities. Later, two different alterations are seen in the limiting vertebral end-plates: a limited form with an erosion, generally in the anterosuperior vertebral angle, with sclerotic base, which is considered as characteristic of brucellar spondylitis (Pedro Pons' sign) (Fig. 4); and a diffuse form, in which the corresponding vertebral end-plates show erosions, accompanied by disc-space narrowing (Fig. 5). Early signs of repair, with the appearance of osteophytes, occur (Rotés-Querol 1957; Serre *et al*. 1981; Gotuzzo *et al*. 1982; Ariza *et al*. 1985a; Lifeso *et al*. 1985; Gotuzzo and Carrillo 1988 ; Sharif *et al*. 1989). Both pathological (Rotés-Querol 1957) and radiological (Sharif *et al*. 1989) data suggest that the infection begins in the vertebra and then spreads to the disc. CT scans may show either localized epiphyseal changes (Sharif *et al*. 1989) or diffuse irregularities and erosions in the vertebral end-plates (Fig. 6). Paravertebral abscess formation (Lifeso *et al*. 1985; Manaster *et al*. 1988, Sharif *et al*. 1989) and epidural extension (Manaster *et al*. 1988) are also demonstrated. Experience with magnetic resonance imaging is limited, and has shown lesions of both localized epiphyseal changes and of diffuse disc-space infections; epidural extension was seen in all four patients studied in the absence of neurological damage (Sharif *et al*. 1990).

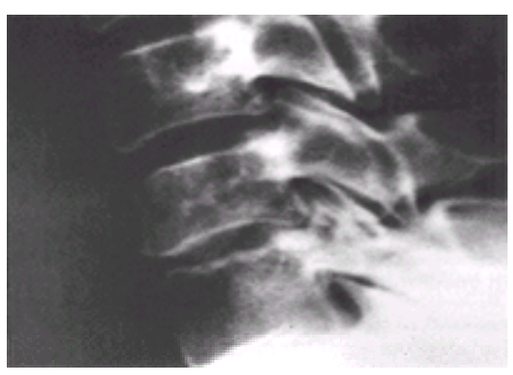


Fig. 3 Lateral radiograph of the lower cervical spine, showing diminished height of the disc space, in a patient with brucellar spondylitis

Fig. 4 Lateral radiograph of lumbar vertebrae, showing an upper corner lesion, with sclerotic base (Pedro Pons' sign), considered characteristic of brucellar spondylitis. Diminution in the height of the disc space is also seen.

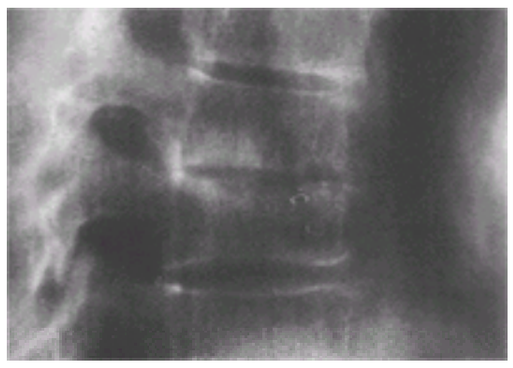


Fig. 5 Lateral tomography of the dorsal spine, showing diminished disc space height and vertebral end-plates erosions in a patient with brucellar spondylitis.

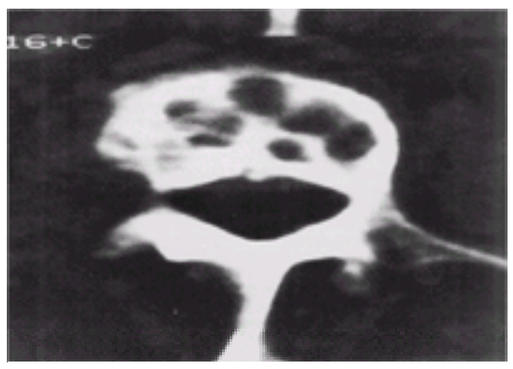


Fig. 6 CT scan through the vertebral end plate of a patient with brucellar spondylitis, showing erosive lesions.

Bone scans frequently show increased uptake at the level of the affected disc space, even when radiologically normal (Bahar *et al*. 1988; Madkour *et al*. 1988; Sharif *et al*. 1989) (Fig. 7). On the other hand, in a prospective study of brucellar spondylitis, bone scans were normal in some patients during the first 3 months of the disease, and later, at times, only slightly abnormal (Ariza *et al*. 1985a).

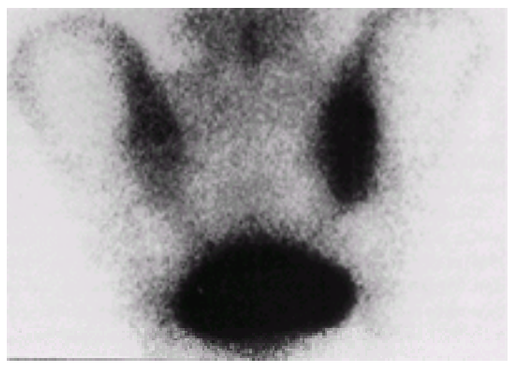


Fig. 7 Bone scan of the posterior pelvis of a patient with left brucellar sacroiliitis of short duration and normal radiograph.

Osteomyelitis

Osteomyelitis is an unusual feature of brucellosis, which may present as local pain and tenderness. *B. melitensis* has a predilection for the ribs and epiphyses of long bones, but may occur in other locations (Rotés-Querol 1957; Serre *et al*. 1981; Mousa *et al*. 1987). *B. suis* seems to prefer the long bones (Kelly *et al*. 1960; Keenan and Guttmann 1982; Bonfiglio *et al*. 1983).

Reactive arthritis

A report (Hodinka *et al*. 1978) of higher frequency of HLA-B27 in patients with brucellar spondylitis raised the possibility of brucellar reactive arthritis. Three HLA-B27-positive patients described as having reactive arthritis after brucellosis were subsequently reported (Dawes and Ghosh 1985). Other studies have failed to find such an association (Alarcón *et al*. 1981; Alarcón *et al*. 1985; Al-Rawi *et al*. 1987). My own experience with 86 prospectively followed patients with brucellar arthritis, is that none of them had any type of chronic arthritis remaining after antimicrobial treatment. As a rule, the symptoms improve clearly after appropriate treatment, both in peripheral and axial arthritides. It is of interest that brucella can be grown from erythema nodosum-like lesions (thought to be a hypersensitivity reaction) occurring in patients with brucellosis (Ariza *et al*. 1989).

The possibility of reactive brucellar arthritis is an open question; such diagnosis should be handled with care and only entertained in those patients in whom unexplained arthritis persists after successful antimicrobial treatment of proven brucellosis.

Other musculoskeletal manifestations

Acute brucellosis may present as a leucocytoclastic vasculitis (Vazquez Doval *et al*. 1991). Patients with brucellosis may also show lesions resembling panniculitis from which brucella can be grown (Ariza *et al*. 1989; Zuckerman *et al*. 1994). Soft tissue abscesses have been reported (Tovar *et al*. 1990)

General laboratory features

With the exception of serological and bacteriological data, the laboratory features of brucellosis are non-specific. The erythrocyte sedimentation rate generally shows some elevation—it was below 44 mm in 70 per cent of the patients with brucellar arthritis (Mousa *et al*. 1987), and normal in 16 per cent in another series (Rotés-Querol 1957). Frequent slight elevations or normal values have been found by others, both in adults (Colmenero *et al*. 1986; Al-Rawi *et al*. 1987; Lulu *et al*. 1988) and children (Gómez-Reino *et al*. 1986; Al-Eissa *et al*. 1990). A similar pattern is seen with *B. abortus* and *B. suis* (Kelly *et al*. 1960). Normal values are not unusual in late, localized forms, as in spondylitis (Serre *et al*. 1981). A normal or low leucocyte count, associated with lymphocytosis is usual (Crosby *et al*. 1984; Colmenero *et al*. 1986; Thapar and Young 1986; Mousa *et al*. 1987). Lymphopenia may be associated with more severe clinical manifestations (Crosby *et al*. 1984). Leucocytosis is unusual but may occur. Thrombocytopenia, anaemia, and pancytopenia may all occur (Crosby *et al*. 1984; Lulu *et al*. 1988).

Abnormality of liver function tests is a common feature of the disease (Colmenero *et al*. 1985; Thapar and Young 1986; Lulu *et al*. 1988), and occurs more frequently in the early phases.

Synovial fluid analysis

Published data of synovial fluid analysis in brucellar arthritis are scarce. Cell counts have been found to be below 50 000/mm ^{3}, (Gotuzzo and Carrillo 1988) or even lower—4460 to 8800/mm^3 (Andonopoulos *et al*. 1986); 6000 to 18 000/mm^3 (Mousa *et al*. 1987); and 8600 to 11 600/mm^3 (Mavridis *et al*. 1987). Glucose levels were within normal limits in the above series. Lactic acid levels have also been found to be normal (Mavridis *et al*. 1984). *B. melitensis* has been isolated from some of the synovial fluids in these studies. Our own unpublished data of the analysis of synovial fluids from 18 prospectively studied brucellar arthritides differ from the above results. The cell counts were 3200 to 90 000/mm^3; glucose was low in some of the fluids (0–100 mg/dl, mean 47 ± 33 SD), and lactic acid was also high in some samples (13–138 mg/dl, mean 74 ± 36 SD). *B. melitensis* was recovered from 73 per cent of these fluids, many of them with characteristics suggestive of non-infectious, inflammatory synovial fluid.

Diagnostic investigations

Importance of the clinical features

The clinical features of brucellosis are not specific. The manifestations of brucellar joint disease are similar to those of other infectious peripheral or axial arthritides, and may also resemble some inflammatory arthritides. Nevertheless, in the following circumstances, appropriate testing for brucellosis should always be done:

1. in all undiagnosed arthritides, with features fitting those of brucellosis, occurring in areas where the disease is endemic or in patients with a history of possible exposure to the disease;
2. in all cases of undiagnosed acute unilateral sacroiliitis;
3. in all cases of undiagnosed disc-space infection, especially those with radiological evidence of localized vertebral angle infection or infection of multiple levels.

About one-quarter of the patients do not recall exposure to possibly infected animals or unpasteurized diary products.

Bacteriological diagnosis

Isolation of brucellae from blood, synovial fluid, or other sources should always be attempted when the disease is suspected. It provides a definitive diagnosis and the possibility of differentiating *B. melitensis, B. abortus*, and *B. suis*, which cannot be done with the usual serological tests. When brucellosis is suspected, the laboratory should always be warned, since:

1. Brucellae are slow growing organisms and most cultures will be negative if not kept long enough—ideally up to 6 weeks (Rodríguez Torres 1988). In a series of 262 positive blood cultures, the growth was apparent in 8 per cent of the samples during the first week, 36 per cent during the second, 31 per cent during the third, 17 per cent during the fourth, 6 per cent during the fifth, and 2 per cent during the sixth week; the culture was not pursued further (Ariza 1988). More delayed growth was found in a smaller series, with a mean of 37 days (range 20 to 51 days) (Thapar and Young 1986).
2. Brucellae are a common cause of acquired infection in the laboratory and personnel should take special precautions (Miller *et al*. 1987).

Any high quality, peptone-based medium enriched with blood or serum is suitable for growing brucellae (Young 1995); a 10 per cent CO_2 atmosphere increases the yield and is necessary for most strains of *B. abortus*; culture systems with a solid phase are easier and less risky to handle in the laboratory (Rodríguez Torres 1988). Of practical importance, brucellae were isolated from seven blood cultures in a mean time of 2.1 days when processed by the lysis–centrifugation system; simultaneous standard blood cultures of the same samples were positive in only six cases, and the mean detection time was 20.6 days (Navas *et al*. 1993).

Blood cultures are often positive in disease due to *B. melitensis*. Table 2 shows the isolation rate of the organism from blood cultures obtained in different series of unselected patients and in patients with arthritis. Patients with relapsing disease have the same rate of positive blood cultures as new patients (Ariza 1988).

Unselected adults	
73 (Thapar and Young 1986)	
62 (Colmenero *et al*. 1986)	
78 (Ariza 1988)	
30 (Lulu *et al*. 1988)	
72 (Colmenero *et al*. 1991)	
Unselected children	
32 (Feiz *et al*. 1978)	
75 (Al-Eissa *et al*. 1990)	
Brucellar arthritis (adults)	
58 (Rotés-Querol 1957)	
41 (Batlle *et al*. 1969)	
41 (Al-Rawi *et al*. 1989)	
Brucellar arthritis (children)	
33 (Gómez-Reino *et al*. 1986)	
70 (Lubani *et al*. 1988)	

Table 2 Percentage of isolation of *Brucella melitensis* in blood cultures

Although blood cultures should be obtained when patients are febrile, in 30 per cent (Ariza 1988) and 31 per cent (Rodríguez Torres 1988) of afebrile patients the organism grew. Isolation of *B. melitensis* from blood cultures is less frequent in late, localized disease, and in spondylitis (Serre *et al*. 1981; Gotuzzo and Carrillo 1988; Colmenero *et al*. 1991). In these cases, bone marrow culture may be advantageous (Gotuzzo *et al*. 1986). Differences in disease characteristics, or laboratory procedures, probably account for the wide differences found between series; a 7 per cent positive rate was found in a large series in which blood cultures were frequently kept less than 10 days (Mousa *et al*. 1987). The isolation rate of *B. abortus* and *B. suis* is lower than that of *B. melitensis*.

Synovial fluid should always be cultured. *B. melitensis* grew in 73 per cent of the synovial fluids inoculated in blood culture flasks (Carro *et al*. 1988). Other series have obtained a 62 per cent (Gotuzzo and Carrillo 1988), 60 per cent (Andonopoulos *et al*. 1986), and 27 per cent (Gotuzzo *et al*. 1982) growth in the cultured samples. *B. abortus* has been cultured from synovial fluid (Al-Rawi *et al*. 1989) as well as *B. canis* (Young 1983). Material obtained by needle puncture from infected disc spaces (Seignon *et al*. 1980) or surgically from the sacroiliac joint (Porat and Shapiro 1984) may also grow brucellae. *B. abortus* or *B. suis* may grow in samples obtained from bone, bursa, tendons, and joints (Kelly *et al*. 1960).

Serological diagnosis

The attempts to isolate brucellae from blood or other sources are not always successful; moreover, a long incubation period is required. Under these circumstances, the possibility of serological diagnosis offers great advantages. Serological tests allow the detection of antibodies produced against the lipopolysaccharide of the bacterial cell wall (Diaz *et al*. 1968), which is common to *B. melitensis, B. abortus* and *B. suis*, but not *B. canis* which needs a specific antigenic suspension (Polt *et al*. 1982; Devi *et al*. 1987).

The standard tube agglutination test (**STA**) is the most widely used test. The test antigen is generally obtained from *B. abortus*; it reacts against *B. abortus, B. melitensis*, and *B. suis* and does not allow differentiation between them (Rodríguez Torres 1988; Young 1995). A positive STA is indicative of contact with brucellae. Although high titres are indicative of current infection, the presence of any positive titre, if the clinical features are compatible, must be investigated further. As in other serological investigations, the individual response, antigen preparations, and laboratory procedures influence the final titre (Rodríguez Torres 1988).

Blocking antibodies may result in a negative STA test at low dilutions, while a positive test is obtained with further serum dilution. This phenomenon occurs mainly in late, localized brucellosis (Rodríguez Torres 1988). The brucellar Coombs' test detects these blocking antibodies and allows the diagnosis to be made (Hall and Manion 1953). The brucellar Coombs' test always gives higher titres than the STA test (Rodríguez Torres 1988). A combination of the STA test and the brucellar Coombs' test allows detection of the large majority of infections, and appears adequate for routine clinical practice (Ariza 1988; Rodríguez Torres 1988). If simple

agglutination methods (Moyer *et al*. 1987) or the rose bengal slide agglutination test (Altwegg and Bohl 1985; Rodríguez Torres 1988) are used for screening, standard tests must be performed in the positive sera for definite serological diagnosis.

The pattern of the antigenic response may be measured with an enzyme-linked immunosorbent assay. IgM antibodies appear first and may disappear within a mean time of 9 months. IgG peaks at about 2 months but significant titres persist after 18 months or more (Ariza 1988; Gazapo *et al*. 1989). In a large group of serially followed patients, STA and IgM antibodies had a parallel decline, as did the Coombs' test and IgG antibodies (Ariza 1988; Ariza *et al*. 1992a). The measurement of specific antibodies allows the detection of occasional patients not discovered by other serological tests (Sippel *et al*. 1982; Gazapo *et al*. 1989; Ariza *et al*. 1992a). Detection of an elevation in the IgG antibody during the follow-up is a very useful serological sign for the diagnosis of relapses (Pellicer *et al*. 1988; Gazapo *et al*. 1989; Ariza I>et al. 1992a). The definitive diagnosis of a relapse requires either bacteriological or clinical evidence of the disease.

Occupationally exposed workers may have low, abnormal serological test results in the absence of disease. Due to similarities in the cell wall components, serological test for brucellosis may be positive in infections due to *F. tularensis, Yersinia enterocolitica* 0<9, vibriocholerae, and *Escherichia coli* 0<157 (Corbel *et al*. 1983); when a positive result is due to cross-reaction, titres tend to be lower against the cross-reacting antigenic suspension.

Treatment

The treatment of brucellar arthritis is similar to that of uncomplicated brucellosis (Box 1). The regimen recommended at present by the World Health Organization (Joint FAO/WHO expert committee on brucellosis 1986) includes a combination of rifampicin 900 mg combined with doxycycline at 200 mg, both given once a day for 6 weeks. The relapse rate of this regimen (5 per cent), is similar to that of doxycycline 200 mg/day for 45 days, combined with streptomycin 1 g/day for 14 days (4 per cent relapse rate) (Acocella *et al*. 1989). A study conducted after treatment for only 30 days, showed that the combination of rifampicin 15 mg/kg per day with doxycycline 100 mg every 12 h had a 38 per cent relapse rate, while a combination of tetracycline 0.5 g every 6 h and streptomycin 1 g/day (for only 21 days) had a 7.1 per cent relapse rate (Ariza *et al*. 1985b). In another trial, doxycycline 100 mg every 12 h and rifampicin 15 mg/kg per day in a single morning dose, for 45 days, showed a similar efficacy than a combination of doxycycline for 45 days plus streptomycin 1 g/day for 15 days, although the doxycycline–rifampicin combination was less effective in patients with spondylitis (Ariza *et al*. 1992b). A recent meta-analysis of six published randomized trials comparing the relative efficacy of rifampicin and doxycycline versus streptomycin and doxycycline, or another tetracycline, concluded that treatment with rifampicin and doxycycline presents a greater number of recurrences and lower number of cures than the classical treatment with streptomycin and tetracycline (Solera *et al*. 1994). Probably a regimen including streptomycin and a tetracycline is to be preferred for patients with severe complications, such as spondylitis of the cervical spine.

Box 1 Management of brucellar arthritis

Antimicrobial therapy:
Is the same as that used for uncomplicated brucellosis, i.e. a combination of rifampicin 900 mg/day and doxycycline 200 mg/day over 45 days (relapse rate 5%: relapsing patients : organisms have the same sensitivities as the pre-treatment isolates (relapse is not due to drug resistance). Patients should receive a second course of treatment.

For children under 8:
Use a combination of trimethoprim–sulphamethoxazole, 10–50 mg/kg/day, given twice daily, for 3 weeks and gentamicin, 5 mg/kg/day twice daily, intramuscularly given the first 5 days

Local management:
- Peripheral joints: no local manoeuvres (e.g. daily drainage) are required.
- Sacroileitis or spine infection: only bed rest and analgesia until pain eases.
- Epidural abscesses:
  Medical treatment may be enough, but the need for surgery is an open question.
  An antimicrobial regimen including streptomycin and a tetracycline may offer advantages.

Brucellae isolated from patients with relapses show similar drug sensitivities to the pretreatment isolates from the same patients (Ariza *et al*. 1986), indicating that the relapses are not due to drug resistance. A second course of treatment is generally effective in these patients.

A therapeutic study conducted on 1100 children with early disease, showed that very few relapses were seen after oral monotherapy combined with either streptomycin or gentamicin. These combinations fared better than other regimens. It was concluded that children of 8 years or younger should receive trimethoprim–sulphamethoxazole 10 to 50 mg/kg per day, twice daily for 3 weeks, with intramuscular injections of gentamicin 5 mg/kg per day, twice daily for the first 5 days. Children of 9 years or older should receive doxycycline 5 mg/kg per day, twice daily for 3 weeks, combined with intramuscular injections of gentamicin 5 mg/kg per day, twice daily for the first 5 days (Lubani *et al*. 1989). Monotherapy with trimethoprim–sulphamethoxazole (Gómez-Reino *et al*. 1986; Lubani *et al*. 1989) or with rifampicin (Llorens-Terol and Busquets 1980 ; Lubani *et al*. 1989) results in a higher relapse rate. Treatment of 113 children with 10 to 12 mg/kg trimethoprim, 50 to 60 mg/kg sulphamethoxazole, and rifampicin 15 to 20 mg/kg in two divided doses for 6 weeks resulted in only four relapses and offers a convenient oral therapy for children (Khuri Bulos *et al*. 1993).

Apart from antibiotics, patients with peripheral arthritis do not require repeated evacuation of the joint, as is necessary in pyogenic arthritides. Rarely, large paravertebral abscesses may require drainage. Epidural abscesses with cord compression may need surgery, although it is not clearly established if this more aggressive approach offers advantages over medical treatment alone.

Chapter References

Acocella, G., Bertrand, A., Beytout, J., Durrande, J.B., García Rodríguez, J.A., Kosmidis, J., Micoud, M., Rey, M., Rodríguez Zapata, M., Roux, J., and Stahl, J.P. (1989). Comparison of three different regimens in the treatment of acute brucellosis: a multicenter, multinational study. *Journal of Antimicrobial Chemotherapy*, **23**, 433–9.

Adam, A., MacDonald, A., and MacKenzie, I.G. (1967). Monoarticular brucellar arthritis in children. *The Journal of Bone and Joint Surgery*, **49B**, 652–7.

Agarwal, S., Kadhi, S.K., and Rooney, R.J. (1991). Brucellosis complicating bilateral total knee arthroplasty. *Clinical Orthopedics and Related Research*, **267**, 179–81.

Alarcón, G.S., Bocanegra, T.S., Gotuzzo, E., Hinostroza, S., Carrillo, C., Vasey, F.B., Germain, B.F., and Espinoza, L.R. (1981). Reactive arthritis associated with brucellosis: HLA studies. *Journal of Rheumatology*, **8**, 621–5.

Alarcón, G.S., Gotuzzo, E., Hinostroza, S.A., Carrillo, C., Bocanegra, T.S., and Espinoza, L.R. (1985). HLA studies in brucellar spondylitis. *Clinical Rheumatology*, **4**, 312–4.

Al-Eissa, Y.A., Kambal, A.M., al-Nasser, M.N., al-Habib, S.A., Al-Fawaz, I.M., and Al-Zamil, F.A. (1990). Childhood brucellosis: a study of 102 cases. *Pediatric Infectious Diseases Journal*, **9**, 74–9.

Al-Rawi, Z.S., Al-Khateeb, N., and Khalifa, S.J. (1987). Brucella arthritis among Iraqi patients. *British Journal of Rheumatology*, **26**, 24–7.

Al-Rawi, T.I., Thewaini, A.J., Shawket, A.R., and Ahmed, G.M. (1989). Skeletal brucellosis in Iraqi patients. *Annals of the Rheumatic Diseases*, **48**, 77–9.

Altwegg, M. and Bohl, E. (1985). Evaluation of a rapid, reliable, and inexpensive screening test for the serological diagnosis of human brucellosis. *Zentralblatt fur Bakteriologie, Mikrobiologie und Hygiene Series A. Medical Microbiology, Infectious Diseases, Virology and Bacteriology*, **260**, 65–70.

Andonopoulos, A.P., Asimakopoulos, G., Anastasiou, E., and Bassaris, H.P. (1986). Brucella arthritis. *Scandinavian Journal of Rheumatology*, **15**, 377–80.

Ariza, J. (1988). *Brucelosis: perspectiva actual de la enfermedad. Perfil de las inmunoglobulinas específicas en el curso de su evolución*. Doctoral thesis. University of Barcelona.

Ariza, J., Gudiol, F., Valverde, J., Pallarés, R., Fernández Viladrich, P., Rufí, G., Espadaler, L., and Fernández Nogués, F. (1985 *a*). Brucellar spondylitis: a detailed analysis based on current findings. *Reviews of Infectious Diseases*, **7**, 656–64.

Ariza, J., Gudiol, F,. Pallarés, G., Rufí, G., and Fernández Viladrich, P. (1985 *b*). Comparative trial of rifampin-doxicicline versus tetracycline-streptomycine in the therapy of human brucellosis. *Antimicrobial Agents and Chemotherapy*, **28**, 548–51.

Ariza, J., Bosch, J., Gudiol, F., Liñares, J., Fernández-Viladrich, P., and Martín, R. (1986). Relavance of in vitro antimicrobial susceptibility of brucella mellitensis to relapse rate in human brucellosis. *Antimicrobial Agents and Chemotherapy*, **30**, 958–60.

Ariza, J., Servitje, O., Pallarés, R., Fernández-Viladrich, P., Rufí, G., Peyrí, J., and Gudiol, F. (1989). Characteristic cutaneous lessions in patients with brucellosis. *Archives of Dermatology*, **125**, 380–3.

Ariza, J., Pellicer, T., Pallares, R., Foz, A., and Gudiol, F. (1992 *a*) Specific antibody profile in human brucellosis. *Clinics in Infectious Diseases*, **14**, 131–40.

Ariza, J., Gudiol, F., Pallares, R., Viladrich, P.F., Rufi, G., Corredoira, J., and Miravitlles, M.R. (1992 *b*) Treatment of human brucellosis with doxycycline plus rifampin or doxycycline plus streptomycin. A randomized, double-blind study. *Annals of Internal Medicine*, **117**, 25–30

Ariza, J., Valverde, J., Nolla, J.M., Rufi, G., Viladrich P.F., Corredoira. J.M., and Gudiol, F. (1993). Brucellar sacroiliitis: findings in 63 episodes and current relevance. *Clincs in Infectious Diseases*, **16**, 761–5.

Arnow, P.M., Smaron, M., and Ormiste, V. (1984). Brucellosis in a group of travelers to Spain. *Journal of the American Medical Association*, **251**, 505–7.

Bahar, R.H., Al-Suhaili, A.R., Mousa, A.M., Nawaz, M.K., Kaddah, N., and Abdel-Dayem, H.M. (1988). Brucellosis: appearance on skeletal imaging. *Clinical Nuclear Medicine*, **13**, 102–6.

Batlle, E., Pascual, E., Salas, E., Plazas, J., Román, J., and Vela, P. (1989). Brucellar arthritis: a study of 86 prospectively collected patients (abstract). *British Journal of Rheumatology*, **28** (suppl. 2), 25.

Benjamin, B. and Khan, M.R. (1994) Hip involvement in childhood brucellosis. *Journal of Bone and Joint Surgery* (British), **76**, 544–7.

Benjamin, B., Annobil, S.H., and Khan, M.R. (1992). Osteoarticular complications of childhood brucellosis: a study of 57 cases in Saudi Arabia. *Journal of Pediatric Orthopedics*, **12**, 801–5.

Berrocal, A., Gotuzzo, E., Calvo, A., Carrillo, C., Castañeda, O., and Alarcon, G.S. (1993). Sternoclavicular brucellar arthritis: a report of 7 cases and a review of the literature. *Journal of Rheumatology*, **20**, 1184–6.

Bonfiglio, M., Mickelson, M.R., and El-Khoury, G.Y. (1983). Case report 221. *Skeletal Radiology*, **9**, 208–11.

Carro, A., Batlle, E., Pascual, E., Castellano, J.A., and Plazas, J. (1988). Blood culture flasks for synovial fluid culture: a more sensitive method for the isolation of bacteria from synovial fluid, particularly brucella mellitensis (abstract). *British Journal of Rheumatology*, **27** (suppl. 2), 33.

Colmenero, D.J., Porras, J.J., Valdivieso, P., Porras, J.A., de Ramón, E., Cause, M., and Juárez, C. (1986). Brucelosis: estudio prospectivo de 100 casos. *Medicina Clínica* (Barcelona), **86**, 43–8.

Colmenero, J.D., Reguera, J.M., Fernandez-Nebro, A., and Cabrera-Franquelo, F. (1991). Osteoarticular complications of brucellosis. *Annals of the Rheumatic Diseases*, **50**, 23–6.

Colmenero, J.D., Cisneros, J.M., Orjuela, D.L., Pachon, J., Garcia-Portales, R., Rodriguez-Sampedro, F., and Juarez, C. (1992). Clinical course and prognosis of Brucella spondylitis. *Infection*, **20**, 38–42.

Corbel, M.J., Stuart, F.A., and Brewer, R.A. (1983). Observations on serological cross-reactions between smooth brucella species and organism of other genera. *Developement Biology Standard*, **56**, 341–8.

Cordero Sánchez, M., Alvarez Ruiz, S., López Ochoa, J., and García Talavera, J.R. (1990). Scintigraphic evaluation of lumbosacral pain in brucellosis. *Arthritis and Rheumatism*, **33**, 1052–5.

Crosby, E., Llosa, L., Miró Quesada, M., Carrillo, P., and Gotuzzo, E. (1984). Hematologic changes in brucellosis. *Journal of Infectious Diseases*, **150**, 419–24.

Dawes, P.T. and Ghosh, S.K. (1985). Tissue typing in brucellosis. *Annals of the Rheumatic Diseases*, **44**, 526–8.

Devi, S.J.N., Polt, S.S., Boctor, F.N., and Peter, J.B. (1987). Serological evaluation of brucellosis: importance of species in antigen preparation. *Journal of Infectious Diseases*, **156**, 658–61.

Diaz, R., Jones, L.M., Leong, D., and Wilson, J.B. (1968). Surface antigens in smooth brucellae. *Journal of Bacteriology*, **96**, 893–901.

El Desouki, M. (1991). Skeletal brucellosis: assessment with bone scintigraphy. *Radiology*, **181**, 415–8.

Feiz, J., Sabbaghian, H., and Miralai, M. (1978). Brucellosis due to *B. mellitensis* in children: clinical and epidemiologic observations on 95 patients studied in central Iran. *Clinical Pediatrics*, **17**, 904–7.

Gazapo, E., Gonzalez Lahoz, J., Subiza, J.L., Baquero, M., Gil, J., and Gómez de la Concha, E.G. (1989). Changes in IgM and IgG antibody concentrations in brucellosis over time: importance for diagnosis and follow-up. *Journal of Infectious Diseases*, **159**, 219–25.

Gómez-Reino, F.J., Mateo, I., Fuertes, A., and Gómez-Reino, J.J. (1986). Brucellar arthritis in children and its successful treatment with trimethoprim-sulphamethoxazole. *Annals of the Rheumatic Diseases*, **45**, 256–8.

Gotuzzo, E and Carrillo, C. (1988). Brucella arthritis. In *Infections in the rheumatic diseases*, (eds L. Espinoza, D. Goldenberg, F. Arnett, and G. Alarcón), pp. 31–41. Grune and Stratton, Orlando.

Gotuzzo, E., Alarcón, G.S., Bocanegra, T.S., Carrillo, C., Guerra J.C., Rolando, I., and Espinoza, L.R. (1982). Articular involvement in human brucellosis: a retrospective analysis of 304 cases. *Seminars in Arthritis and Rheumatism*, **12**, 245–55.

Gotuzzo, E,. Carrillo, C., and Guera, J. (1986). An evaluation of the diagnostic methods for brucellosis—the value of bone marrow culture. *Journal of Infectious Diseases*, **153**, 122–5.

Gotuzzo, E., Seas C., Guerra, J.G., Carrillo, C., Bocanegra, T.S., Calvo, A., Castañeda, O., and Alarcón, G.S. (1987). Brucellar arthritis: a study of 39 Peruvian families. *Annals of the Rheumatic Diseases*, **46**, 506–9.

Hall, W.H. and Manion, R.E. (1953). Comparison of the Coomb's test with other methods for brucella agglutinins in human serum. *Journal of Clinical Investigation*, **32**, 96–106.

Hodinka, L., Gomor, B., Meretey, K., Zahumenszky, Z., Geher, P., Telegdy, L., and Bozsoky, S. (1978). HLA-B27-associated spondylarthritis in chronic brucellosis (letter). *Lancet*, **i**, 499.

Hunt, A.C. and Bothwell, P.W. (1967). Histological findings in human brucellosis. *Journal of Clinical Pathology*, **20**, 267–72

Joint FAO/WHO expert committee on brucellosis. (1986). World Health Organization, Geneva.

Keenan, M.A. and Guttmann, G.G. (1982). Brucella osteomyelitis of the distal part of the femur. *Journal of Bone and Joint Surgery*, **64A**, 142–4.

Kelly, P.J., Martin, W.J., Shirger, A., and Weed, L.A. (1960). Brucellosis of the bones and joints. *Journal of the American Medical Association*, **174**, 347–53.

Khuri Bulos, N,A., Daoud, A,H., and Azab, S,M. (1993). Treatment of childhood brucellosis: results of a prospective trial on 113 children. *Pediatric Infectious Diseases Journal*, **12**, 377–81.

Laajam, M.A. (1985). Synovial rupture complicating brucella arthritis. *British Journal of Rheumatology*, **24**, 191–3

Lifeso, R.M., Harder, E., and McCorkell, S.J. (1985). Spinal brucellosis. *Journal of Bone and Joint Surgery*, **67B**, 345–51.

Llorens-Terol, J. and Busquets, R.M. (1980). Brucellosis treated with rifampin. *Archives of Diseases of Childhood*, **55**, 486–8.

Lubani, M.M., Sharda, D.C., and Helin, I. (1986). Brucella arthritis in children. *Infection*, **14**, 233–6.

Lubani, M.M, Dudin, K.I., Sharda, D.C., Mana Ndhar, D.S., Araj, G.F., Hafed, H.A., Al-Saleh, Q.A., Helin, I., and Salhi, M.M. (1989). A multicenter therapeutic study of 1100 children with brucellosis. *Pediatric Infectious Diseases Journal*, **8**, 75–8.

Lulu, A.R., Araj, G.F., Khateeb, M.I., Mustafa, M.Y., Yusuf A.R., and Fenech, F.F. (1988). Human brucellosis in Kuwait: a prospective study of 400 cases. *Quarterly Journal of Medicine*, **249**, 39–54.

Madkour, M.M., Sharif, H.S., Abed, M.Y., and Al-Fayez, M.A. (1988). Osteoarticular brucellosis: Results of bone scintigraphy in 140 patients. *American Journal of Radiology*, **150**, 1101–5

Manaster, B.J. (1988). Case report 469. Spondylitis (lumbar spine) due to *Brucella abortus*. *Skeletal Radiology*, **17**, 144–7.

Mavridis, A.K., Drosos, A.A., Tsolas, O., and Moutsopoulos, H.M. (1984). Lactate levels in brucella arthritis. *Rheumatology International*, **4**, 169–71.

Miller, C.D., Songer, J.R., and Sullivan, J.F. (1987). A twenty-five year review of laboratory-acquired human infections at the National Animal Disease Center. *American Industrial Hygiene Association Journal*, **48**, 271–5.

Mousa, A.R., Elhag, K.M., Khogali, M., and Marafie, A.A. (1987). The nature of human brucellosis in Kuwait: study of 379 cases. *Reviews of Infectious Diseases*, **10**, 211–7.

Moyer, N.P., Evins, G.M., Pigott, N.E., Hudson, J.D., Farshy, C.E., Feeley, J.C., and Hausler, W.J. (1987). Comparison of serological screening tests for brucellosis. *Journal of Clinical Microbiology*, **25**, 1969–72.

Navas, E., Guerrero, A., Cobo, J., and Loza, E. (1993). Faster isolation of Brucella spp. from blood by isolator compared with BACTEC NR. *Diagnostic Microbiology and Infectious Disease*, **16**, 79–81.

Pellicer, T., Ariza, J., and Foz, A. (1988). Specific antibodies detected during relapse of human brucellosis. *Journal of Infectious Diseases*, **157**, 918–24.

Polt, S.S., Dismukes, W.E., Flint, A., and Schaefer, J. (1982). Human brucellosis caused by *Brucella canis*: clinical features and immune response. *Annals of Internal Medicine*, **97**, 717–19.

Porat, S. and Shapiro, M. (1984). Brucella arthritis of the sacro-iliac joint. *Infection*, **12**, 205–7.

Rajapakse, C.N.A. (1987). The spectrum of rheumatic diseases in Saudi Arabia. *British Journal of Rheumatology*, **26**, 22–3.

Revak, D.M., Swain, R.A., Guthrie, R.M., and Lubbers, J.R. (1989). Brucellosis contracted during foreign travel. *Postgraduate Medicine*, **85**, 101–4.

Rodríguez Torres, A. (1988). Diagnóstico de la brucelosis humana. *Revista Española de Reumatología*, **15**, 204–14.

Rotés-Querol, J. (1957). Osteo-articular sites of brucellosis. *Annals of the Rheumatic Diseases*, **16**, 63–8.

Seignon, B., Weilbacher, H., Thorel, J.B., Bussières, J.L., Deshayes, P., and Gougeon, J. (1980). La ponction discale dans le diagnostic bactériologique des spondylodiscites bactériennes. *Revue du Rhumatisme et des Maladies Ostéoarticulaires*, **47**, 45–7.

Serre, H., Kalea, G., Brousson, A., Sany, J., Bertrand, A., and Simon, L. (1981) Manifestations ostéo-articulaires de la brucellose. *Revue du Rhumatisme et des Maladies Ostéoarticulaires*, **47**, 143–8.

Sharif, H.S., Aideyan, O.A., Clark, D.C., Madkour, M.M., Aabed, M.Y., Mattsson, T.A., Al-Deeb, S.M., and Moutaery, K.R. (1989). Brucellar and tuberculous spondylitis: comparative imaging features. *Radiology*, **171**, 419–25.

Sippel, J.E., El-Masry, N.A., and Farid, Z. (1982). Diagnosis of human brucellosis with ELISA. *Lancet*, **ii**, 19–21.

Smith, L.D. and Ficht, T.A. (1990). Pathogenesis of Brucella. *Critical Reviews in Microbiology*, **17**, 209–30.

Solera, J., Martinez-Alfaro, E., and Saez, L. (1994). Metaanalisis sobre la eficacia de la combinacion de rifampicina y doxiciclina en el tratamiento de la brucelosis humana. *Medicina Clinica* (Barcelona), **102**, 731–8.

Spink, W.W. (1951). What is chronic brucellosis? *Annals of Internal Medicine*, **35**, 358–74.

Spink, W.W. (1964). Host-parasite relationship in brucellosis. *Lancet*, **ii**, 161–4.

Thapar, M.K. and Young, E.J. (1986). Urban outbreak of goat cheese brucellosis. *Pediatric Infectious Diseases*, **5**, 640–3.

Torres-Rojas, J., Taddonio, R.F., and Sanders, C.V. (1979). Spondylitis caused by *Brucella abortus*. *Southern Medical Journal*, **72**, 1166–9.

Tovar, J.V., Moreno, R., Perez, A., Satorres, J., Navarro, F., and Royo, G. (1990). Abscesos gluteos por Brucella melitensis. Probable patogenia yatrogena. *Enfermedades Infecciosas y Microbiología Clínica*, **8**, 588.

Vazquez Doval, F.J., Ruiz de Erenchun Lasa, F., Sola Casas, M.A., Soto de Delas, J., and Quintanilla Gutierrez, E. (1991) Acute brucellosis presenting as leukocytoclastic vasculitis. *Journal of Investigative Allergology and Clinical Immunology*, **1**, 411–3

Young, E.J. (1983). Human brucellosis. *Reviews of Infectious Diseases*, **5**, 821–42.

Young, E.J. (1995). Brucella species. In *Principles and practice of infectious diseases*, (eds G.L. Mandell, J.E. Bennett, and R. Dolin), pp. 2053–60. Churchill Livingstone, New York.

Zuckerman, E., Naschitz, J.E., Yeshurun, D., Wellisch, G., Shajrawi, I., and Boss, J.H. (1994). Fasciitis-panniculitis in acute brucellosis. *International Journal of Dermatology*, **33**, 57–9.

5.3.9 Parasitic involvement

Tomás S. Bocanegra

Improvements in living and sanitary conditions, and aggressive eradication programmes, have reduced the prevalence of most parasitic diseases in Third World countries; however, these conditions continue to be endemic in many areas of the world (Cairncross 1995). In industrialized nations, parasites cause sporadic outbreaks of disease in urban communities (McAmilty *et al*. 1994; MacKenzie *et al*. 1994; Millard *et al*. 1994; Huang *et al*. 1995) and occasionally serious illness in immigrants and travellers (Encarnación *et al*. 1994). More importantly, in the last decade, there are reports of an increasing number of life-threatening infections with parasites in immunosuppressed patients and individuals infected with the human immunodeficiency virus (Pape *et al*. 1994; Ognibene *et al*. 1995; Weiss 1995).

The clinical presentation of parasitic infection varies from localized symptoms to a multisystem disease that may include musculoskeletal manifestations. The incidence of rheumatic manifestations in the different parasitic diseases is not known, but there are reports of a variety of musculoskeletal syndromes. Recently, the importance of considering parasitic disease in the differential diagnosis of patients presenting with rheumatic manifestations was recognized by the description of these syndromes in two major textbooks of rheumatology (Bocanegra 1993; Bocanegra 1994). The diagnosis of parasite-associated musculoskeletal syndrome is established on clinical grounds when manifestations develop in residents or travellers to endemic areas, in patients who have a documented infection with an intestinal or tissue-invasive parasite, and in those whose symptoms do not respond to conventional anti-inflammatory therapy but resolve after the eradication of the parasite (Bocanegra 1994; Doury 1994) (Table 1).

Residence in, or travelling to, endemic areas	
Documented parasitic infection	
Poor response to conventional treatment	
Resolution following eradication of parasite	

Table 1 Clinical characteristics of parasite-associated musculoskeletal syndromes

This chapter will review the rheumatic syndromes that develop in humans as a consequence of parasitic infestation or of the treatment of parasitic diseases.

Protozoal infections

There are reports of rheumatic syndromes with a variety of pathogenic, tissue-invasive, as well as oportunistic protozoa (see Box 1).

Box 1 Musculoskeletal syndrome induced by protozoa

	Articular	Muscular	Vascular	Treatment of choice
Toxoplasma	Polyarthritis	Polymyositis-like	Polyarteritis nodosa-like	Pyrimethamine and sulphadiazine
Amoeba	Arthralgias	None	Focal vasculitis	Metronidazole
Giardia	Oligoarthritis	None	Small vessel vasculitis	Metronidazole or tinidazole
Cryptosporidium	Oligoarthritis	None	None	Paromomycin
Isospora	Oligoarthritis	Focal and generalized myositis	Polyarteritis nodosa-like	Trimethoprim-sulphamethoxazole
Microsporidia	None	Generalized myopathy	None	Albendazole
Trypanosoma	None	Generalized myopathy (*T. rhodesience*), Polymyositis-like (*T. cruzi*)	None	Suramin (*T. rhodesience*) Nifurtimox (*T. cruzi*)
Plasmodium	None	Myalgias	Small vessel vasculitis	Chloroquine

Articular syndromes

A symmetrical polyarthritis of the small joints of the hands, wrists, and knees in a rheumatoid pattern may develop in patients infected with *Toxoplasma gondii*. Rheumatoid factor may or may not be present in serum, but all patients have serological evidence of acute toxoplasma infection (IgG toxoplasma antibodies) (Antezana 1979; Gemou *et al*. 1983; Balleari *et al*. 1991).

Arthralgia, back pain, and arthritis occurred in young adults with intestinal amoebiasis (Yonis 1943; Rappaport *et al*. 1951; Doury *et al*. 1977). Patients presented with a polyarthritis of recent onset affecting the small, medium, and large joints of the upper and lower limbs, preceded by mild gastrointestinal symptoms. One patient had concomitant urticaria. Joint symptoms resolved on treatment of the intestinal amoebiasis.

Giardiasis may cause joint symptoms, most commonly in children and young adults. The largest series of patients with articular symptoms associated with *Giardia lamblia* infection was reported by Goobar (1977). The most common pattern was a mild and self-limiting oligoarthritis of the large and medium-sized joints of the upper and lower limbs. The erythrocyte sedimentation rate was elevated but rheumatoid factor was negative with few exceptions. All patients had concomitant gastrointestinal manifestations and most had urticaria. Recently, there have been several other case reports of a similar pattern of joint involvement. However, in a few instances the small joints of the hands and feet were affected (Farthing *et al*. 1983; Woo and Panayi 1984; Barton *et al*. 1986; Shaw and Stevens 1987; Brougui and Richard 1990). Articular symptoms did not respond to non-steroidal anti-inflammatory agents but resolved after treatment with metronidazole.

Reactive arthritis may also develop after cryptosporidial gastroenteritis in young adults and children (Hay *et al*. 1987; Sheperd *et al*. 1989). Small, medium, and large joints are affected, but symptoms are usually self-limiting. In addition, there are reports of infectious and reactive arthritis in association with opportunistic infections by *Blastocystis hominis* and *Isospora belli* (Lee *et al*. 1990; Lakhanpal *et al*. 1991; Gonzalez-Dominguez *et al*. 1994; Krüger *et al*. 1994). In one case, *Blastocystis* was recovered from synovial fluid in the knee, supporting the possibility of an infectious cause; in the others, the oligoarthritis was preceded by diarrhoea and resolved

after antiparasitic treatment, suggesting a 'reactive' mechanism.

Muscular syndromes

Focal and diffuse myositis may occur in patients infected with protozoa, mostly due to invasion of muscle tissue by the parasite. However, in some cases, an immune-mediated mechanism may play a part.

Isospora hominis (Jeffrey 1974; Bonciou *et al*. 1981) may cause both a localized and a generalized myositis with proximal and distal muscle weakness. Toxoplasmosis may produce an acute and a chronic myositis with clinical manifestations and muscle enzyme abnormalities similar to those seen in idiopathic polymyositis (McNicholl and Underhill 1970; Greenlee *et al*. 1975; Topi *et al*. 1979; Roig-Quilis and Damjanov 1982). Patients usually have a concomitant febrile systemic illness and serological evidence of acute toxoplasma infection. Toxoplasma may or may not be found in the striated muscle. It is possible that toxoplasma is involved in idiopathic polymyositis, based on the higher than expected frequency of antitoxoplasma antibodies in patients with polymyositis and dermatomyositis (Magid and Kagen 1983), the improvement in dermatomyositis on treatment of associated toxoplasmosis (Harland *et al*. 1991), and the expression of major histocompatibility complex class I and II antigens in the inflammatory cells around the muscle fibres (Matsubara *et al*. 1990). However, there is no convincing proof for this hypothesis.

Granulomatous myositis due to microsporidial infection with pleistophora was reported in a patient with severe immunodeficiency (Ledford *et al*. 1985); clusters of parasitic spores were present in the striated muscle fibres. Clinically, the patient presented with muscle wasting, contractures, and cachexia. Muscle wasting is relatively common in American trypanosomiasis (Chagas' disease) caused by *Trypanosoma cruzi* and African trypanosomiasis caused by *T. rhodesiense* and *T. gambiense*, but it is unclear if it is due to direct invasion of the striated muscle by the parasite, denervation, or malnutrition. A symmetrical proximal myositis, similar to idiopathic polymyositis, occurred in a patient with Chagas' disease (Cossermelli *et al*. 1978). Encysted trypanosomes were identified in the muscle biopsy. The muscle weakness improved but did not resolve after treatment of the underlying trypanosomiasis.

Myalgia, muscle necrosis, and elevations of creatine kinase can occur in malaria (DeSilva *et al*. 1988; Swash and Schwartz 1993), and have been attributed to microvascular changes in the striated muscle.

Vascular syndromes

A polyarteritis nodosa-like syndrome is described in patients infected with *Toxoplasma gondii* and *Isospora hominis* (McGill 1957; Carmeni *et al*. 1991). These patients presented with fever, myalgias, weakness, paraesthesias, and weight loss. Muscle biopsy showed necrotizing vasculitis of small and medium arteries.

Localized arteritis due to parasitic invasion of the vessel wall may occur in toxoplasmosis of the central nervous system and in amoebic colitis, and may be the underlying cause of the tissue necrosis frequently observed in these conditions (Huang and Chou 1988; De La Torre and Gorraez 1989; Desphande *et al*. 1992). A systemic, immune complex-mediated vasculitis may occur in infections with *Plasmodium falciparum* and *Giardia lamblia*, resulting in uveitis, urticaria, erythema nodosum, and glomerulonephritis.

Helminthic infections

Musculoskeletal manifestations are more frequently reported in patients infected by helminths than in those with protozoal infections (see Box 2).

Box 2 Musculoskeletal syndrome induced by helminthies

	Articular	Muscular	Vascular	Treatment of choice
Taeniasis (intestinal)	Oligoarthritis and polyarthritis	None	None	Praziquantel
Cysticercus	Arthralgias	Myalgias, nodular myopathy	Localized around cysticercus	Albendazole
Hydatid cyst	Bone pain, mono- and polyarthritis	Asymptomatic muscular cysts	Polyarteritis nodosa-like	Albendazole, surgical excision of cyst
Coeneurosis	None	Asymptomatic nodules	None	Surgical excision
Sparganosis	None	Nodular, occasionally migratory	None	Surgical excision
Onathostoma	None	Nodular myositis, occasionally migratory	None	Albendazole, surgical excision
Strongyloides	Arthralgias, oligo- and polyarthritis	None	Leucocytoclastic vasculitis	Thiabendazole
Ascaris	Oligoarthritis	None	Churg-Strauss syndrome	Mebendazole
Trichuris	Oligoarthritis	None	None	Mebendazole
Ancylostoma	Oligoarthritis	None	None	Mebendazole
Toxocara	Arthralgias, monoarthritis	Localized myositis of the calf	None	Diethylcarbamazine
Dracunculus	Arthralgias, monoarthritis	Myalgias	None	Metronidazole
Filariae	Arthralgias, mono- and oligoarthritis	Modular myositis (onchocerca)	Leucocytoclastic vasculitis (loa-loa)	Diethylcarbamazine, ivermectin (onchocerca)
Schistosoma	Arthralgias, oligo- and polyarthritis, enthesitis	Diffuse myopathy	None	Praziquantel
Dirofilaria	Oligoarthritis	None	None	Self-limited
Trichina	None	Myalgias, polymyositis-like	Polyarteritis nodosa-like	Mebendazole

Articular syndromes

Arthritis occurs in humans infected with cestodes as definite host (taeniasis) and as intermediate hosts (cysticercosis and hydatid cyst). A symmetrical, 'rheumatoid-like' polyarthritis and an oligoarthritis of the knees have been reported in patients infected with *Taenia saginata* (Bocanegra *et al*. 1981; Bussiere *et al*. 1981). The patient with polyarthritis had elevated amounts of circulating immune complexes and IgE in serum and deposits of IgG and C3 in the synovium. Parasites were not found in the synovial fluid or tissues. Treatment with indomethacin was unsuccessful, but symptoms resolved after treatment of the parasitic infection.

Arthralgias and bone pain may occur in cysticercosis (Surianu *et al*. 1967) and hydatid disease (Fyfe *et al*. 1990). Bone involvement is seen in 1 to 2 per cent of patients with hydatid disease; about half of bone lesions are cysts in the thoracic or lumbar spine (Rao *et al*. 1991). Although the infection is probably acquired in childhood, clinical manifestations do not develop until adulthood, owing to the slow growth of the cyst. Radiographically, bone hydatic cysts present as a lytic lesion without sclerosis and are commonly confused with metastatic malignancy or plasmacytoma. Most commonly, bone hydatid disease affects the body of the vertebra but occasionally may affect the pedicles. Spinal compression with paraplegia may occur and carries a poor prognosis (Argenson *et al*. 1989). The pelvis and femur are affected less frequently, but their involvement may result in pathological fractures (Hooper and McLean 1977; Duran *et al*. 1978). A chronic granulomatous synovitis may occur when a cyst in the para-articular bone or muscle opens into the joint (Vigliani and Campailla 1977; Voutsinas *et al*. 1987). Patients may present with urticaria and eosinophilia in addition to monoarthritis. In one patient (Campoy *et al*. 1995), synovial fluid had increased numbers of eosinophils; synovial biopsy showed granulomas and the remains of hydatid cysts (Fig. 1).

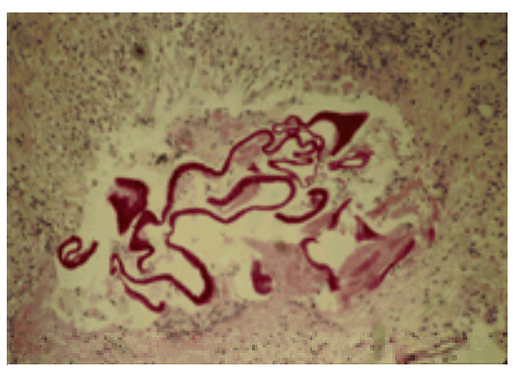


Fig. 1 Hydatid synovitis. Synovial biopsy of the knee showing inflammatory infiltrate, granulomas, and fragments of a hydatid cyst. Sections stained with periodic acid–Schiff. Original magnification ×10. (Reproduced from Campoy *et al*. 1995, with permission.).

There are reports of a seronegative polyarthritis of the small, medium, and large joints in a few patients with hydatid disease of the liver (Ballina-Garcia *et al*. 1987; Buskila *et al*. 1992). The arthritis did not respond to non-steroidal anti-inflammatory agents but resolved on removal of the hepatic hydatid cysts. The diagnosis of joint or bone involvement by hydatid disease is difficult and is often established accidentally by the finding of cysts during exploratory surgery. Needle biopsy of the cyst should not be performed, owing to the risk of disseminating the disease. The presence of eosinophilia in peripheral blood or synovial fluid, urticaria, and prior residence in an endemic area for hydatid disease in a patient presenting with arthritis or an unexplained cystic bone lesion should lead the physician to consider hydatid cyst in differential diagnosis. The treatment of bone and joint hydatid disease is of limited success because of the difficulty in removing the cyst and scoleces from the bone. In patients with arthritis associated with hepatic hydatid cysts, the symptoms resolve after the resection of the cyst. However, surgery carries a risk of anaphylactic shock due to the massive release of hydatid antigen in a patient already sensitized to it. Oral treatment with mebendazole may be beneficial in some patients but it has poor penetration of bone.

Joint pain and arthritis may develop in the course of invasive and intestinal infections with *Strongyloides stercoralis* (Bocanegra *et al*. 1981; Patey *et al*. 1990). Two articular syndromes have been reported: one characterized by symmetrical, polyarticular involvement similar to rheumatoid arthritis (Doury *et al*. 1975; Bocanegra *et al*. 1981; Amor *et al*. 1983; Forzy *et al*. 1988) and an oligoarthritis of the knees, ankles, hips, and sacroiliac joints (Amor *et al*. 1983; Akoglu *et al*. 1984; De Jonge-Bok *et al*. 1985; Menkes *et al*. 1987). Patients usually present with eosinophilia in peripheral blood and synovial fluid. Larvae of *Strongyloides* are found in stools and, in some patients, in the synovial fluid or in the synovium (Fig. 2). One patient with polyarthritis had immune complexes in serum and synovial fluid. HLA-B27 antigen was negative in all patients tested. Articular symptoms resolved in all patients on treatment with thiabendazole.

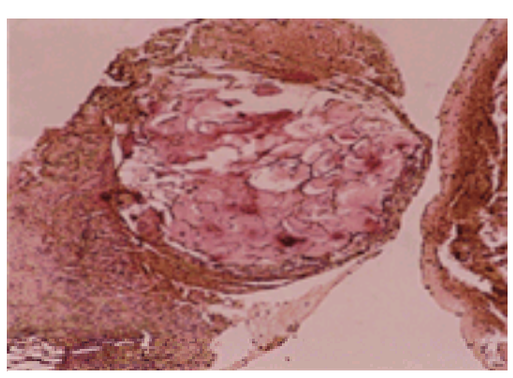


Fig. 2 *Strongyloides* synovitis: synovia from the ankle showing larvae of *Strongyloides stercoralis* surrounded by inflammatory infiltrate (reproduced from Akoglu *et al*. 1984, with permission,).

There are reports of arthralgia and monoarthritis of the knee in infections with *Toxocara canis* (canine ascarids) (William and Roy 1981; Le Luyer *et al*. 1990; Richardson de Corral *et al*. 1990; Van Linthoudt *et al*. 1990; Kraus *et al*. 1995). Some patients presented with panniculitis of the legs due to migration of the parasite through the subcutaneous tissues (Fig. 3). Diagnosis is confirmed by finding the parasite in the affected tissues or by a positive enzyme immunoassay for antitoxocara antibodies. The natural course of the disease is self-limiting, although albendazole has been used in some cases. Isolated cases of seronegative oligoarthritis of leg joints are reported in infections with *Ankylostoma duodenale, Ascaris lumbricoides*, and *Trichuris trichiura* (Treusch *et al*. 1981; Bissonnette and Beaudet 1983).

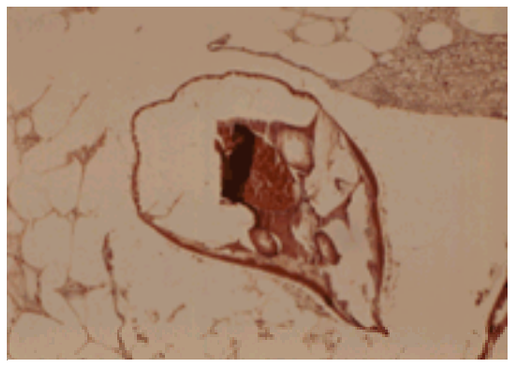


Fig. 3 *Toxocara* panniculitis: skin biopsy showing fragments of *Toxocara* and inflammatory cells in the subcutaneous tissue (reproduced from Kraus *et al*. 1995, with permission).

One-third to one-half of patients infected with *Dracunculus medinensis*, a tissue-dwelling nematode, develop arthralgias or myalgias in the legs (Garf 1985). Monoarthritis of the ankle or knee occurred in about 2 per cent of patients (Kothari *et al*. 1968) due to three possible mechanisms: (i) invasion of the joint by the parasite (Reddy and Sivaramappa 1968) with release of microfilariae into the joint—synovial fluid is inflammatory and contains microfilariae of the Guinea worm as well as increased eosinophils; (ii) a reactive arthritis due to the presence of adult worms in the neighbouring soft tissues (McLaughlin *et al*. 1984) —synovial fluid is inflammatory and may show eosinophilia but not microfilariae; (iii) infectious arthritis, most commonly due to staphylococci, which enter the joint through the skin ulcerations and sinus track created by the parasite (Greenwood 1968). Diagnosis is established by the presence of skin ulcerations and calcified parasites near the affected joints. Arthrocentesis and arthroscopy are helpful, particularly in cases of direct joint invasion and infectious arthritis. Non-steroidal anti-inflammatory agents are indicated in reactive arthritis. Antiparasitic treatment rarely results in parasite death and is not recommended. Surgical removal of Guinea worm from the soft tissue is difficult due to the length of the parasite and should not be attempted. However, arthroscopy or arthrotomy are indicated when Guinea worm is located in the joint. Intravenous antibiotics and arthrocentesis are the treatment of choice in infectious arthritis.

Joint pain and arthritis occur in filariasis caused by *Wuchereria bancrofti, Loa loa*, and *Onchocerca*. Intermittent polyarthralgias occur during the early phase of *W. bancrofti* filariasis. Arthritis occurs less often and then only in the late, obstructive phase of the infection (Alhadeff 1955). Among patients with arthritis, two-thirds have monoarticular involvement while the other one-third have oligoarthritis (less than four joints affected) (Ismail and Nagaratnam 1973). The joints most frequently affected are the knee, the ankle, and the hip (Salfield 1975). Involvement of the arms is rare, and when present is limited to large joints. The majority of patients do not have other manifestations of filariasis but may show eosinophilia, positive serology for filariasis, or microfilariae in peripheral blood or synovial fluid. Radiographs of the joints are unremarkable except for soft-tissue swelling. Synovial fluid is inflammatory with low cell counts. A few patients have a chylous effusion due to obstruction of the lymphatic vessels (Das and Sen 1968). Synovial biopsy may show mononuclear cell infiltrates and eosinophils or chronic fibrosis. The treatment of choice is ivermectin. Anti-inflammatory drugs alone are ineffective. Arthrocentesis is indicated only for diagnostic purposes but not for treatment. A similar articular syndrome has been noticed in loasis.

Polyarthralgias of large and medium joints of the limbs occurred in up to one-third of patients (Carme *et al*. 1989) and an oligoarthritis affecting the knees and ankles has been seen in a few cases (Bouvet *et al*. 1977). Radiographs of the joints show soft-tissue swelling and occasionally calcified filariae. Synovial fluid is inflammatory, with lymphomononuclear cells, eosinophilia, and occasionally polymorphonuclear cells and microfilariae (Doury *et al*. 1984; Roussel *et al*. 1989). Treatment is similar to that of *W. bancrofti* filariasis.

Infections with *Onchocerca volvulus* may occasionally produce back pain and arthritis of the hip, knee, or metatarsophalangeal joints (De Jou 1941; Commandre *et al*. 1976). Diagnosis is made by identification of microfilariae in skin-punch biopsy (skin-snip examination), in biopsies of subcutaneous nodules (onchocercomas), or in

the synovial fluid. The treatment of choice is ivermectin.

Dirofilaria immitis (dog heartworm) rarely infects humans. Most patients are asymptomatic, but a few develop respiratory symptoms including haemoptysis and pulmonary nodules. Arthritis of large leg joints was reported in patients infected with *D. immitis* and *D. tenuis* (Corman 1987; Langer *et al*. 1987). Synovial fluid has a low cell count with a predominance of mononuclear cells. Antiparasitic therapy is unnecessary since dirofilariae do not fully develop in man. Spontaneous death of the parasite leads to resolution of the symptoms.

Schistosomiasis in humans is caused by three species: *Schistosoma mansoni* and *S. japonicum* localized in the veins of the bowel, and *S. haematobium* in the veins of the genitourinary tract. Musculoskeletal manifestations may occur during the acute and chronic phases. Joint and muscle pain occur during the acute, serum sickness-like syndrome known as Katayama fever. Arthritis and enthesitis are seen in chronic schistosomiasis. Atkin *et al*. (1986) reported musculoskeletal manifestations in 80 per cent of patients. In general, these patients were older and had been infected with schistosomes for longer than those without rheumatic symptoms. The three syndromes identified were polyarthritis, in 60 per cent of patients, oligoarthritis in 30 per cent, and enthesitis in 10 per cent. However, enthesitis frequently occurred in patients with oligo- and polyarthritis and, all together, was the most common musculoskeletal syndrome. Polyarthritis affected the small joints of the hands and feet, knees and ankles in a symmetrical pattern, and was accompanied by morning stiffness. Oligoarthritis affected large joints of the legs. Enthesitis was not localized to any particular area. Synovial fluid was inflammatory, with cell counts of 7 000 to 30 000/ml and a clear predominance of polymorphonuclear leucocytes. An asymmetrical oligoarthritis of the large leg joints and unilateral sacroiliitis was the most common syndrome in another series of patients with schistosomiasis (Bassiouni and Kamel 1984). Many patients had enthesitis and/or calcaneal spurs. Synovia showed mononuclear cell infiltrates and, in some, ova of schistosomes (Bassiouni and Kamel 1984; Fachartz *et al*. 1993). Low-titre rheumatoid factor, higher levels of IgE and IgM antibodies, and circulating immune complexes were more common in patients with articular symptoms than in those without (Kamel *et al*. 1989; Bebars *et al*. 1992). The treatment of choice is a single oral dose of praziquantel.

Muscular syndromes

Cysticercosis develops in humans infected with *Taenia solium* as intermediate hosts. The most common locations of cysticerci are the central nervous system and the striated muscle. However, occasionally, the cysts may affect the extraocular muscles and the tongue (Stewart *et al*. 1993; Gupta *et al*. 1994). Among those with skeletal muscle involvement, a few develop symptoms consisting of muscle pain and nodules or generalized myopathy (Serre *et al*. 1970; Sawhney *et al*. 1976; Vilhena Lana-Peixoto *et al*. 1985). Deep-seated nodules can be palpated in the muscles of the pelvic and scapular girdles, mostly during muscle contraction. Occasionally, a diffuse, 'pseudohypertrophic' myopathy of the limb muscles is present. Despite an apparently impressive muscle development, many patients are weak. Commonly, patients have seizures or other symptoms indicative of involvement of the central nervous system. Muscle enzymes may be mildly elevated and the electromyogram may show a myopathic pattern. Diagnosis is made by demonstration of multiple muscle cysts by MRI or CT scan (Gupta *et al*. 1994), and muscle biopsy. Cysts contain non-viable larvae and are surrounded by an inflammatory infiltrate with mononuclear cells and eosinophils. Occasionally, cysts are calcified. The treatment of choice is albendazole. A short course of corticosteroids may be needed to prevent exacerbation of symptoms due to inflammation elicited by the dying larvae.

As in cysticercosis, man is an intermediate host of *Echinococcus granulosus* and *E. multilocularis*. Parasite larvae mainly encyst (hydatid cysts) in the liver and lungs. Muscular involvement is infrequent and may affect the chest wall, abdominal wall, pectoralis, sartorius, and proximal limb muscles (Schimrigk and Emser 1978; Menuier *et al*. 1983; Duncan and Tooke 1990; García-Picazo *et al*. 1995). The cysts are usually solitary and asymptomatic, presenting as a soft-tissue mass. CT scans or MRI show a well-defined cyst with a fluid-density signal. Diagnosis and treatment are achieved by complete excision of the cyst. Percutaneous needle biopsy is not recommended, as there is a risk of disseminating the scoleces along the needle tract and of precipitating anaphylactic shock from the release of large quantities of parasitic antigens. Intraoperative cyst aspiration and thorough irrigation of the cavity with hypertonic saline is recommended to inactivate the scoleces and reduce the risk of recurrence. A fibrous and inflammatory reaction with a mononuclear, neutrophilic, and eosinophilic cell infiltrate is seen on histological examination. Muscular manifestations may develop in infections caused by larvae of two other 'flat worms', *Taenia multiceps* (coenurosis) and *Spirometra* spp. (sparganosis). In coenurosis, larvae of *T. multiceps* encyst in the subcutaneous tissues and muscles of the neck, chest or abdomen causing painless nodules (Templeton 1968). In sparganosis acquired by the oral route, larvae of *Spirometra* are released in the intestine and migrate from there to different organs. During this process the larvae burrow through the subcutaneous tissue and muscles causing a slowly growing, occasionally migrating soft-tissue mass (Cho and Patel 1978; Nakamura *et al*. 1990). In both conditions, diagnosis and treatment are achieved by surgical excision of the mass.

Gnathostoma spinigerum (gnathostomiasis) may produce a localized soft-tissue abscess or a nodular, migratory, eosinophilic panniculitis due to migration of the larvae through the skin, subcutaneous tissues, and muscles (Rusnak and Lucey 1993; Stevens and Bryson 1994). Diagnosis is confirmed by positive serology. The treatment of choice is albendazole (Kraivichian *et al*. 1992).

An acute, transient myositis localized to the legs was reported in two children infected with *Toxocara canis* (visceral larva migrans) (Walsh *et al*. 1988). Both cases presented with eosinophilia and diffuse, non-tender swelling of the calf, which resolved spontaneously in 3 days. Diagnosis was established by positive serology (enzyme immunoassay) for *Toxocara*. *Onchocerca volvulus* may produce a nodular, localized, eosinophilic myositis of the abdominal wall (Neumann *et al*. 1985), and schistosomiasis has been associated with a diffuse myopathy of muscles of the shoulder and pelvic girdles (Mansour and Reese 1964).

Muscle involvement occurs during the invasive phase of trichinosis when larvae, released in the intestine by the adult female, migrate through the tissues and reach the striated muscle. The clinical presentation varies from asymptomatic infection to a severe, sometimes fatal, multisystem disease, depending on the immune status of the host and the parasite load; fewer than 10 larvae per gram of muscle rarely cause symptoms. Clinical manifestations characteristic of the invasive phase are: fever, chills, periorbital oedema, subconjunctival, retinal and subungual (splinter) haemorrhages, eosinophilia, myocarditis, and less frequently, encephalitis, nephritis and death. Myalgia and weakness affecting the proximal muscles of the arms and legs develop in two-thirds of patients infected by *T. nelsoni* (Ferraccioli *et al*. 1988). Occasionally, patients may present with manifestations indistinguishable from those of dermato- or polymyositis; that is, proximal muscle weakness, elevated creatine kinase, an electromyogram consistent with myositis, and extensive eosinophilic myositis on muscle biopsy (Herrera *et al*. 1985; MacLean *et al*. 1989; Durán-Ortiz *et al*. 1992; Louthrenoo *et al*. 1993). Larvae of trichinae are present in the muscle fibres (Fig. 4). Larvae of *Trichinella spiralis* are encapsulated while those of *T. pseudospiralis* are unencapsulated and mobile. The treatment of choice is with albendazole or mebendazole. Corticosteroids are indicated only in massive infections with severe myositis, myocarditis or central nervous involvement. Corticosteroids alone may aggravate muscle symptoms (Andrews *et al*. 1994).

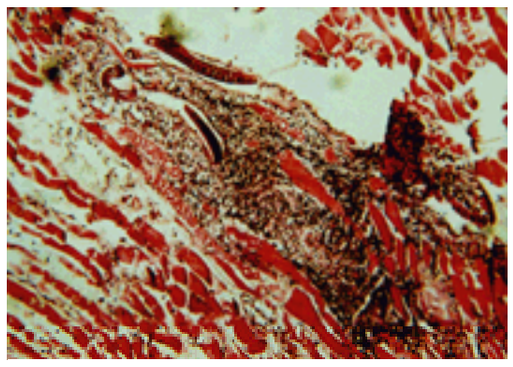


Fig. 4 Trichina myositis. Larvae of *Trichinella spiralis* and inflammatory cell infiltrate in a muscle biopsy of a child with trichinosis. Section stained with haematoxylin and eosin. Original magnification ×100. (Reproduced from Durán-Ortiz *et al*. 1992, with permission.).

Vascular syndromes

Vasculitis was induced in animals experimentally infected with larva migrans (Watzke *et al*. 1984) and is documented in a few patients infected with cestodes and nematodes. Vascular changes have been demonstrated in cysticercosis of the central nervous system and probably contribute to the basal arachnoiditis characteristic of racemous cysticercosis (Estañol *et al*. 1986). A case of histologically proven polyarteritis nodosa in a child with hepatic hydatidosis was reported recently (Bakkaloglu *et al*. 1994): treatment with prednisone and cyclophosphamide for 2 months was ineffective, but all symptoms resolved after treatment with mebendazole followed by surgical excision. Leucocytoclastic vasculitis characterized by palpable purpura, arthritis, eosinophilia, and immunoglobulin deposition in the vascular

walls has been reported in infections with *Strongyloides stercoralis* (Akoglu *et al*. 1984) and *Loa loa* (Portilla-Sogorb *et al*. 1991). As in other parasite-induced symptoms, all manifestations of vasculitis resolved after antiparasitic treatment. Nematode infections have been associated with more severe forms of necrotizing vasculitis. Reimann *et al*. (1943) and Frayha (1981) described cases of polyarteritis nodosa associated with trichinosis. More recently, Churg–Strauss vasculitis was reported in patients infected with ascarids (Chanham *et al*. 1990) and an *Angiostrongylus*-like nematode (Pirisi *et al*. 1995). In both cases the diagnosis was proven by histology. Grcevska (1993) reported recently a case of renal necrotizing vasculitis associated with glomerulonephritis in a patient infected with ascarids; the renal manifestations resolved completely after eradication of the parasite.

Musculoskeletal symptoms induced by antiparasitic treatment

Arthralgias and arthritis may develop during antiparasitic treatment, especially in patients infected with tissue-invasive parasites. In general, the frequency and severity of the symptoms vary according to the type of parasite, the parasite load, and the drug used. Mild joint and muscle pain occur in 2 to 3 per cent of patients with leishmaniasis treated with stibogluconate (Thakur and Kumar 1990), and in 10 to 15 per cent of patients with opisthochiasis treated with praziquantel (Viravan *et al*. 1986). More severe manifestations (Mazzotti reaction) develop in patients with filariasis treated with diethylcarbamazine or ivermectin (Ottesen 1987; Kumaraswami *et al*. 1988). Mazzotti reactions are seen more commonly in onchocerciasis and in lymphatic filariasis caused by *Bruggia* spp. The clinical manifestations of Mazzotti reaction vary from malaise and minor body aches to a severe, systemic, sometimes fatal reaction. Symptoms develop rapidly in the first few hours after the administration of diethylcarbamazine and consist of pruritus, rash, lymphadenopathy, fever, tachycardia, hypotension, and arthralgias. A second phase of the reaction is characterized by arthritis of the small and large joints, fever, and myalgia, developing in the first 3 to 5 days. Milder and less frequent symptoms have been reported with the use of ivermectin for the treatment of onchocerciasis and lymphatic filariasis.

The use of low initial doses of diethylcarbamazine often diminishes the severity of the early phase of the reaction. The treatment of the Mazzotti reaction is symptomatic; for patients with mild symptoms, analgesics and antipyretics are indicated. Corticosteroids, prednisone 60 mg/day, prevent the development of symptoms and are indicated in patients with heavy infections, although they may decrease the efficacy of diethylcarbamazine.

Chapter References

Akoglu, T., Tuncer, I., Erkon, E., Gürcay, A., Ozer, F. L., and Ozcan, K. (1984). Parasitic arthritis induced by *Strongyloides stercoralis*. *Annals of the Rheumatic Diseases*, **43**, 523–5.

Alhadeff, R. (1955). Clinical aspects of filariasis. *Journal of Tropical Medicine and Hygiene*, **58**, 173–9.

Amor, B., Benhamou, C. L., Dougados, M., and Grant, A. (1983). Arthrites a eosinophiles et revue generale de la signification de l'eosinophile articulaire. *Revue Rheumatisme Maladie Osteoarticulaire*, **50**, 659–64.

Andrews, J. R. H., Ainsworth, R., and Abernethy, D. (1994). *Trichinella pseudospiralis* in humans: description of a case and its treatment. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, **88**, 200–3.

Antezana, E. (1979). Polytenosynovitis caused by *Toxoplasma gondii*. *South African Medical Journal*, **56**, 746.

Argenson, C., Griffet, J., Lacour, C., Arcamone, H., Lovet, J., and de Peretti, F. (1989). Hydatid disease of the spine. Report of two cases. *Revue de Chirurgie Orthopedique*, **75**, 267–70.

Atkin, S. L., Kamel, M., El-Hady, A. M,. El-Badawy, S. A., El-Ghobary, A., and Dick, W. C. (1986). Schistosomiasis and inflammatory polyarthritis: a clinical, radiological and laboratory study of 96 patients infected by *S. mansoni* with particular reference to the diarthrodial joint. *Quarterly Journal of Medicine*, **59**, 479–87.

Bakkaloglu, A., Söylemezøglu, O., Tinaztepe, K., and Saatci, Ü. (1994). A possible relationship between polyarteritis nodosa and hydatid disease. *European Journal of Pediatrics*, **153**, 469.

Balleari, E., Cutolo, M., and Accardo, S. (1991). Adult onset Still's disease associated to *Toxoplasma gondii* infection. *Clinical Rheumatology*, 10, 326–7.

Ballina-Garcia, F. J., Rodriguez-Arboleya, L., Torre-Alonso, J. C., Mori-Garcia, A., and Rodriguez-Perez, A. (1987). Polyarthrite au cours d'une hydatidose hepatique. *Revue du Rheumatisme*, **54**, 779.

Barton, J. J., Burke, J. P., and Casey, E. B. (1986). Reactive arthritis: *Giardia lamblia*, another new pathogen. *Irish Medical Journal*, **79**, 223.

Bassiouni, M. and Kamel, M. (1984). Bilharzial arthropathy. *Annals of the Rheumatic Diseases*, **43**, 806–9.

Bebars, M. A., Soffar, S. A., Abbas, M. M., Ismail, M. S., and Tamara, F. A. (1992). Circulating immune complexes in bilharzial arthropathy. *Journal of the Egyptian Society of Parasitology*, **22**, 51–7.

Bissonnette, B. and Beaudet, F. (1983). Reactive arthritis with eosinophilic synovial infiltrations. *Annals of the Rheumatic Diseases*, **42**, 466–8.

Bocanegra, T. S. (1993). Musculoskeletal manifestations of parasitic diseases. In *Oxford textbook of rheumatology* (1st edn) (ed. P. I. Maddison, D. A. Isenberg, P. Woo, and D. N. Glass), pp. 589–98. Oxford University Press.

Bocanegra, T. S. (1994). Mycobacterial, fungal and parasitic arthritides. In *Rheumatology (ed*. J. H. Klippel and P. A. Dieppe), pp. 4.5.1–4.5.12. Mosby–Year Book Europe, London.

Bocanegra, T. S., Espinoza, L. R., Bridgeford, P. H., Vasey, F. B., and Germain, B. F. (1981). Reactive arthritis induced by parasitic infestation. *Annals of Internal Medicine*, **94**, 207–9.

Bonciou, C., Petronicu, M., and Panateiscu, D. (1981). Observations on a case of chronic myositis due to sarcosporidia. *Archives Romaines do Pathologie Experimentale et Microbiologie*, **40**, 361–3.

Bouvet, J. P., Thérizol, M., and Auquier, L. (1977) Microfilarial polyarthritis in a massive Loa-Loa infestation: a case report. *Acta Tropica*, **34**, 281–4.

Brougui, P. and Richard, P. (1990). Articular manifestations of *Giardia lamblia* infections. *Bulletin de la Societé de Pathologie Exotique et des ses Filiales*, **83**, 688–9.

Buskila, D., Sikenik, S., Klein, M., and Horowitz, J. (1992). Polyarthritis associated with hydatid disease (echinococcosis) of the liver. *Clinical Rheumatology*, **11**, 286–7.

Bussiere, J. L., Ristori, J. M., and Sauvezie, B. (1981). Inflammatory rheumatism associated with intestinal parasitosis. *Nouvelle Presse Médicale*, **10**, 2990.

Cairncross, S. (1995). Victory over guineaworm disease: partial or pyrrhic?, *Lancet*, **346**, 1440.

Campoy, E., Rodriguez-Morena, J., Del Blanco, J., Narvaez, J., Clavaguera, T., and Roig-Escofet, D. (1995). Hydatid disease. An unusual cause of monoarthritis. *Arthritis and Rheumatism*, **38**, 1338–9.

Carme, B., Mamboueni, J. P., Copin, N., and Noireau, F. (1989). Clinical and biological study of Loa-Loa filariasis in Congolese. *American Journal of Tropical Medicine and Hygiene*, **41**, 331–7.

Carmeni, G., Tato, A., Martusciello, S., Valiante, B., Magrino, A., and Meloni, F. (1991). Vascular involvement and toxoplasma infection. *British Journal of Dermatology*, **124**, 14.

Chanham, A., Scott, D. G. I., Neuberger, J., Gaston, J. S. H., and Bacon, P. A. (1990). Churg–Strauss vasculitis and Ascaris infection. *Annals of the Rheumatic Diseases*, **49**, 320–2.

Cho, C. and Patel, S. P. (1978). Human sparganosis in northern United States. *New York State Journal of Medicine*, **78**, 1456–8.

Commandre, F., Lapeyre, L., Viani, J. L., Revelli, G., Lefichoux, Y., and Jozan, S. (1976). Arthritis associated with filarioses. *Rheumatologie*, **18**, 27–33.

Corman, L. C. (1987). Acute arthritis occurring in association with subcutaneous *Dirofilaria tenuis* infection. *Arthritis and Rheumatism*, **30**, 1431–4.

Cossermelli, W. *et al*. (1978). Polymyositis in Chagas' disease. *Annals of the Rheumatic Diseases*, **37**, 277–80.

Das, G. C. and Sen, S. B. (1968). Chylous arthritis. *British Medical Journal*, **2**, 27–9.

De Jonge-Bok, J. M., Overbosch, D., and MacFarlane, J. D. (1985). Parasitic rheumatism presenting as oligoarthritis: a case report. *Tropical and Geographical Medicine*, **37**, 367–8.

De Jou, L. (1941). Surgical localizations of African filarioses: arthritis and suppurations of the soft tissue. *Medicine Tropicale*, **1**, 15–35.

De La Torre, F. E. and Gorraez, M. T. (1989). Toxoplasma-induced occlusive hypertrophic arteritis as the cause of discrete coagulative necrosis in the CNS. *Human Pathology*, **20**, 604.

DeSilva, H. J. *et al*. (1988). Skeletal muscle necrosis in severe falciparum malaria. *British Medical Journal*, **296**, 1039.

Desphande, R. B., Bharucha, M. A., Modhe, J. M., and Bhalerao, R. A. (1992). Necrotizing arteritis in amoebic colitis. *Journal of Postgraduate Medicine*, **38**, 151–2.

Doury, P. (1994). Parasitic arthritis and parasitic rheumatism. *Smaine des Hospitaux Paris*, **7**, 522–8.

Doury, P., Pattin, S., Durosoir, J. L., Voinesson, A., Bienot, B., and Duret, J. C. (1975). Rheumatism and strongyloidiasis. *Nouvelle Presse Médicale*, **4**, 805.

Doury, P., Pattin, S., Dienot, B., Roue, R., and Delayaye, R. P. (1977). Les rheumatismes parasitaires. *Smaine des Hospitaux Paris*, **53**, 1359–63.

Doury, P., Saliou, P., and Charmot, G. (1984). Eosinophilic joint effusions: *propos* of a case. *Revue du Rheumatisme*, **51**, 29–31.

Duncan, G. J. and Tooke, S. M. (1990). Echinococcus infestation of the biceps brachii. A case report. *Clinical Orthopaedics*, **26**, 247–50.

Duran, H. *et al*. (1978). Osseous hydatidosis. *Journal of Bone and Joint Surgery* (A), **60**, 685–90.

Durán-Ortiz, J. S. *et al*. (1992). Trichinosis with severe myopathic involvement mimicking polymyositis. Report of a family outbreak. *Journal of Rheumatology*, **19**, 310–12.

Encarnación, C. F., Giordano, M. F., and Murray, H. W. (1994). Onchocerciasis in New York City. The Moa–Manhattan Connection. *Archives of Internal Medicine*, **154**, 1749–51.

Estañol, B., Corona, T., and Abad, P. (1986). A prognostic classification of cerebral cysticercosis: therapeutic implications. *Journal of Neurology, Neurosurgery and Psychology*, **49**, 1131–4.

Fachartz, O. A., Kumer, V., and Al-Hilou, M. (1993). Synovial schistosomiasis of the hip. *Journal of Bone and Joint Surgery* (B), **75**, 602–3.

Farthing, M. J. G., Chong, S. K. F., and Walter-Smith, J. A. (1983). Acute allergic phenomenon in giardiasis. *Lancet*, **ii**, 1428.

Ferraccioli, G. F., Mercadanti, M., Salaffi, F., Bruschi, F., Melissari, M., and Pozio, E. (1988). Prospective rheumatological study of muscle and joint symptoms during *Trichinella nelsoni* infection. *Quarterly Journal of Medicine*, **69**, 973–84.

Forzy, G., Dhondt, J. L., Leloire, O., Shayeb, J., and Vincent, G. (1988). Reactive arthritis and Strongyloides. *Journal of the American Medical Association*, **259**, 2546–7.

Frayha, R. A. (1981). Trichinosis related polyarteritis nodosa. *American Journal of Medicine*, **71**, 307–82.

Fyfe, B., Amazon, K., Poppiti, R. J., and Razzetti, A. (1990). Intra-osseous echinococcosis: a rare manifestation of echinococcal disease. *Southern Medical Journal*, **83**, 66–8.

García-Picazo, D., Vasquez-Arragón, P., Palomares-Ortiz, G., Cascales-Sánchez, P., and Lopez-Fando de Castro, J. (1995). Multiple hydatid cysts of the liver, bones and muscles. *Revista Española de Enfermedades Digestivas*, **87**, 267–70.

Garf, A. E. (1985). Parasitic rheumatism: rheumatic manifestations associated with calcified Guinea worm: *Journal of Rheumatology*, **12**, 976–9.

Gemou, V., Messaritakis, J., Karpathios, T., and Kingo, A. (1983). Chronic polyarthritis of toxoplasmic etiology. *Helvetia Paediatrica Acta*, **38**, 295–6.

Gonzalez-Dominguez, J., Roldan, R., Vllanueva, J. L., Kindelám, J. M., Jurado, R., and Torre-Cisneros, J. (1994). *Isospora belli* reactive arthritis in a patient with AIDS. *Annals of the Rheumatic Diseases*, **53**, 618–19.

Goobar, J. P. (1977). Joint symptoms in giardiasis. *Lancet*, **i**, 1010–11.

Grcevska, L. and Polenakovic, M. (1993). Renal vasculitis associated with ascaridiasis with good prognosis. *Nephron*, **64**, 327–8.

Greenlee, J. E., Johnson, W. D., Campa, J. F., Adelman, L. S., and Sande, M. A. (1975). Adult toxoplasmosis presenting as polymyositis and cerebellar ataxia. *Annals of Internal Medicine*, **82**, 367–71.

Greenwood, B. M. (1968). Guinea-worm arthritis of knee joint. *British Medical Journal*, **1**, 314.

Gupta, P. K., Jaleel, M. A., Prasad, V. S., Naik, R. T., Sundaram, C., and Dinakar, I. (1994). Unusual clinical manifestations of cysticercosis. *Journal of the Association of Physicians of India*, **42**, 411–12.

Harland, C. C., Marsden, J. R., Vernon, S. A., and Allen, B. R. (1991). Dermatomyositis responding to treatment of associated toxoplasmosis. *British Journal of Dermatology*, **125**, 76–8.

Hay, E. M,. Windfield, J., and McKendrick, M. W. (1987). Reactive arthritis associated with crytosporidium enteritis. *British Medical Journal*, **295**, 248.

Herrera, R., Varela, E., Morales, G., Del Rio, A., and Gallardo, J. M. (1985). Dermatomyositis-like syndrome caused by Trichinae: report of two cases. *Journal of Rheumatology*, **12**, 782–4.

Hooper, J. and McLean, I. (1977). Hydatid disease of the femur. Report of a case. *Journal of Bone and Joint Surgery* (A), **59**, 974–6.

Huang, T. E. and Chou, S. M. (1988). Occlusive hypertrophic arteritis as the cause of discrete necrosis in CNS toxoplasmosis in the acquired immunodeficiency syndrome. *Human Pathology*, **19**, 1210–14.

Huang, P. *et al*. (1995). The first reported outbreak of diarrheal illness associated with *Cyclospora* in the United States. *Annals of Internal Medicine*, **123**, 409–14.

Ismail, M. M. and Nagaratnam, N. (1973). Arthritis possibly due to filariasis. *Transactions of the Royal Society for Tropical Medicine and Hygiene*, **67**, 405–9.

Jeffrey, H. C. (1974). Sarcosporidiosis in man. *Transactions of the Royal Society for Tropical Medicine and Hygiene*, **68**, 17–29.

Kamel, M., Sajwat, E., and El-Tayab, S. (1989). Bilharzial arthropathy. *Scandinavian Journal of Rheumatology*, **18**, 315–19.

Kothari, M. L., Pardnani, D. S., Mehta, L., and Amand, M. P. (1968) Guinea worm arthritis of the knee joint. *British Medical Journal*, **3**, 435–6.

Kraivichian, P., Kulkumthorn, M., Yingyourd, P., Akarabovan, P., and Paireepai, C. C. (1992). Albendazole for the treatment of human gnathostomiasis. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, **86**, 418–21.

Kraus, A., Valencia, X., Cabral, A. R., and De La Vega, G. (1995) Visceral larva migrans mimicking rheumatic diseases. *Journal of Rheumatology*, **22**, 497–500.

Krüger, K., Kamilli, I., and Schattenkirchner, M. (1994). *Blastocystis hominis* infection. A rare cause of arthritis. *Zeitschrift für Rheumatologie*, **53**, 83–5.

Kumaraswami, V. *et al*. (1988). Ivermectin for the treatment of *Wuchereria brancrofti* filariasis: efficacy and adverse reactions. *Journal of the American Medical Association*, **259**, 3150–3.

Lakhanpal, S., Cohen, S. B., and Fleischmann, R. M. (1991). Reactive arthritis from *Blastocystis hominis*. *Arthritis and Rheumatism*, **34**, 251–3.

Langer, H. E., Bialek, R., Mielke, H., and Klose, J. (1987). Human dirofilariasis with reactive arthritis—case report and review of the literature. *Klinische Wochenschrift*, **65**, 746–51.

Ledford, D. K., Overman, M. D,. Gonzalvo, A., Cali, A., Mester, S. W., and Lockey, R. F. (1985). Microsporidiosis myositis in a patient with the acquired immunodeficiency syndrome. *Annals of Internal Medicine*, **102**, 628–30.

Lee, M. G., Rawlins, S. C., Didier, M., and DeCeulaer, K. (1990). Infective arthritis due to *Blastocystis hominis*. *Annals of the Rheumatic Diseases*, **49**, 192–3.

Le Luyer, B., Menager, V., Audebert, C., Le Roux, P., Briquet, M. T., and Boulloche, J. (1990). Arthropathies inflammatoires revelatrices d'une larva migrans a *Toxocara canis*. *Annales Pediatria*, **7**, 445–8.

Louthrenoo, W., Mahanuphab, P., Sanguanmitra, P., and Thamprasert, K. (1993). Trichinosis mimicking polymyositis in a patient with human immunodeficiency virus infection. *British Journal of Rheumatology*, **32**, 1025–30.

McAmilty, J. M., Fleming, D. W., and Gonzalez, A. H. (1994). A community-wide outbreak of cryptosporidiosis associated with swimming at a wave pool. *Journal of the American Medical Association*,

272, 1597–600.

McGill, R. J. (1957). Sarcosporidiosis in a man with polyarteritis nodosa. *British Medical Journal*, **2**, 333–4.

MacKenzie, W. R. *et al*. (1994). A massive outbreak in Milwaukee of *Cryptosporidium* infection transmitted through the public water supply. *New England Journal of Medicine*, **331**, 161–7.

McLaughlin, G. E., Utsinger, P. D., Trakat, W. F., Resnick, D., and Moidel, R. A. (1984). Rheumatic syndromes secondary to Guinea worm infestation. *Arthritis and Rheumatism*, **27**, 694–7.

MacLean, J. D., Viallet, J., Law, C., and Staudt, M. (1989). Trichinosis in the Canadian Arctic: report of five outbreaks and a new clinical syndrome. *Journal of Infectious Diseases*, **160**, 573–20.

McNicholl, B. and Underhill, D. (1970). Toxoplasmic polymyositis. *Irish Journal of Medical Science*, **3**, 525–7.

Magid, S. K., and Kagen, L. J. (1983). Serologic evidence for acute toxoplasmosis in polymyositis-dermatomyositis: increased frequency of specific anti-toxoplasma IgM antibodies. *American Journal of Medicine*, **75**, 313–20.

Mansour, S. D. and Reese, H. H. (1964). A previously unexpected myopathy in patients with schistosomiasis. *Neurology*, **14**, 355–61.

Matsubara, S., Takamori, M., Adachi, H., and Kida, H. (1990). Acute toxoplasma myositis: an immunohistochemical and ultrastructural study. *Acta Neuropathologica*, **81**, 223–7.

Menkes, C. J., Papo, T., Carter, H., and Renoux, M. (1987). Reactive arthritis caused by strongyloides and ankylosing spondylitis. *Revue du Rheumatisme*, **54**, 504–6.

Menuier, Y., Danis, M., Nozais, J. P., and Gentilini, M. (1983). Muscular hydatid disease: two case reports. *Smaine des Hospitaux Paris*, **59**, 2785–6.

Millard, P. S. *et al*. (1994). An outbreak of cryptosporidiosis from fresh-pressed apple cider. *Journal of the American Medical Association*, **272**, 1592–6.

Nakamura, T., Hara, M., Matusuka, M., Kawabata, M., and Tsuji, M. (1990). Human proliferative sparganosis: a new Japanese case. *American Journal of Clinical Pathology*, **94**, 224–8.

Neumann, H., Herz, R., and Baum, C. (1985). Granulomatous and eosinophilic myositis due to *Onchocerca volvulus. Pathology*, **6**, 101–7.

Ognibene, F. P. *et al*. (1995). *Pneumocystis carinii* pneumonia—a major complication of immunosuppressive therapy in patients with Wegener's granulomatosis. *American Journal of Respiratory and Critical Care Medicine*, **151**, 795–9.

Ottesen, E. A. (1987). Description mechanisms and control of reactions to treatment in the human filariases. *Ciba Foundation Symposium*, **127**, 265–83.

Pape, J. W., Verdier, R. I., Boncy, M., Boncy, J., and Johnson, W. D. (1994). Cyclospora infection in adults infected with HIV. Clinical manifestations, treatment and prophylaxis. *Annals of Internal Medicine*, **121**, 654–7.

Patey, O., Bouhali, R., Breuil, J., Chapuis, L., Courillon-Mallet, A., and Lafaix, C. (1990). Arthritis associated with *Strongyloides stercoralis*. *Nouvelle Presse Médicale*, **4**, 805.

Pirisi, M. *et al*. (1995). Fatal human pulmonary infection caused by *Angiostrongylus*-like nematode. *Clinical Infectious Diseases*, **20**, 59–65.

Portilla-Sogorb, J., Sevilla-Linares, A., Carrion, A,. Cordoba, C., and Blesa, M. (1991). Vasculitis Leucocitoclástica por hipersensiblilidad a microfilarias. *Anales de Medicina Interna (Madrid)*, **8**, 30–2.

Rao, S., Parikh, S., and Kerr, R. (1991). Echinococcal Infestation of the spine in North America. *Clinical Orthopedics*, **271**, 164–9.

Rappaport, E. M., Rossien, A., and Roseblum, L. (1951). Arthritis due to intestinal amebiasis. *Annals of Internal Medicine*, **34**, 1224–31.

Reddy, C. R. and Sivaramappa, M. (1968). Guinea-worm arthritis of knee joint. *British Medical Journal*, **1**, 155–6.

Reimann, H. A., Price, A. H., and Herbut, P. A. (1943). Trichinosis and periarteritis nodosa: differential diagnosis, possible relationship. *Journal of the American Medical Association*, **12**, 274–9.

Richardson de Corral, V., Lozano-Garcia, J., and Ramos-Corona, L. E. (1990). Una presentación poco usual de toxocariasis sistémica. *Boletín Médico del Hospital Infantil de México*, **47**, 841–4.

Roig-Quilis, M. and Damjanov, I. (1982). Dermatomyositis as an immunologic complication of toxoplasmosis. *Acta Neuropathologica (Berlin)*, **58**, 183–6.

Roussel, F., Roussel, C., Brasseur, P., Gourmelen, O., and Le Loet, X. (1989). Aseptic knee effusion with Loa-Loa microfilariae in the articular fluid. *Acta Cytologica*, **33**, 281–3.

Rusnak, J. M. and Lucey, D. R. (1993). Clinical gnathostomiasis: a case report and review of the English language literature. *Clinics in Infectious Diseases*, **16**, 33–50.

Salfield, S. (1975). Filarial arthritis in the Sepik district of Papua New Guinea. *Medical Journal of Australia*, **1**, 264–7.

Sawhney, B. B., Chopra, J. S., Banerji, A. K., and Wahi, P. L. (1976). Pseudohypertrophic myopathy in cysticercosis. *Neurology*, **26**, 270–2.

Schimrigk, K. and Emser, W. (1978). Parasitic myositis by *Echinococcus alveolaris*: report of a family with myotonia congenita. *European Neurology*, **17**, 1–7.

Serre, H., Simon, L., Lamboley, C., and Regal, R. (1970). Muscular manifestations of human cysticercosis. *Rheumatologie*, **22**, 537–63.

Shaw, R. A. and Stevens, M. B. (1987). The reactive arthritis of *Giardia lamblia*. A case report. *Journal of the American Medical Association*, **258**, 2734–5.

Sheperd, R. C., Smail, P. J., and Sinha, G. P. (1989). Reactive arthritis complicating cryptosporidial infection. *Archives of Disease in Childhood*, **64**, 743–4.

Stevens, C. O. and Bryson, A. D. (1994). Gnathostomiasis as a cause of soft tissue swellings. *British Journal of Rheumatology*, **33** (Suppl. 65), 36.

Stewart, C. R., Salmon, J. F., Murray, A. D., and Sperryn, C. (1993). Cysticercosis as a cause of severe medial rectus muscle myositis. *American Journal of Ophthalmology*, **116**, 510–11.

Surianu, P., Medrea, B., Popa, M., and Karassi, A. (1967). Rheumatic manifestations of cutaneous cysticercosis. *Presse Médicale*, **65**, 1505.

Swash, M. and Schwartz, M. S. (1993). Malaria myositis. *Journal of Neurology, Neurosurgery and Psychiatry*, **56**, 1238.

Templeton, A. C. (1968). Human coeneurosis: a report of 14 cases from Uganda. *Transactions of the Royal Society for Tropical Medicine and Hygiene*, **62**, 251–5.

Thakur, C. P. and Kumar, K. (1990). Efficacy of prolonged therapy with stibogluconate in post kala-azar dermal leishmaniasis. *Indian Journal of Medical Research*, **91**, 144–8.

Topi, G. C., D'Alessandro, L., Catricala, C., and Zardi, O. (1979). Dermatomyositis-like syndrome due to toxoplasma. *British Journal of Dermatology*, **101**, 589–91.

Treusch, P. J., Swatnam, R. E., and Woelke, B. J. (1981). Eosinophilic joint effusion and intestinal nematodiasis. *Annals of Emergency Medicine*, **10**, 614–15.

Van Linthoudt, D., Mean, A. P., Favre, R., and Ott, H. (1990). Coxarthritis in a child associated with toxocariasis. *Schweizerische Rundschau fur Medizine Praxic (Bern)*, **79**, 1022–4.

Vigliani, F. and Campailla, E. (1977). Meta-echinococcal arthrosynovitis: a clinical, anatomical and pathological entity not previously recognized. *Italian Journal of Orthopaedics and Traumatology*, **3**, 103–10.

Vilhena Lana-Peixoto, M. I., Lana-Peixoto, M. A., and Campos, G. B. (1985). Pseudohypertrophic myopathy due to cysticercosis: a case report. *Arquivos Neuro-psiquiatria (Sao Paolo)*, **43**, 396–402.

Viravan, C., Bunnang, D., Harinasuta, T., Upatham, S., Kurathong, S., and Viyanant, V. (1986). Clinical field trial of praziquantel in opisthorchiasis in Nong Rangya Village, Khon Kaen Province, Thailand. *South East Asian Journal of Tropical Medicine and Public Health*, **17**, 63–6.

Voutsinas, S., Sayakos, J., and Smytnis, P. (1987). Echinococcus infestation complicating total hip replacement. A case report. *Journal of Bone and Joint Surgery* (A), **69**, 1456–8.

Walsh, S. S., Robson, W. J., and Hart, C. A. (1988) Acute transient myositis due to *Toxocara*. *Archives of Disease in Childhood*, **63**, 1087–8.

Watzke, R. C., Oaks, J. A., and Folk, J. C. (1984). *Toxocara canis* infection in the eye: correlation of clinical observations with developing pathology in the primate model. *Archives of Ophthalmology*, **102**, 282–91.

Weiss, L. M. (1995). And now microsporidiosis. *Annals of Internal Medicine*, **123**, 954–6.

William, D. and Roy, S. (1981). Arthritis and arthralgia associated with toxocaral infestation. *British Medical Journal*, **283**, 192.

Woo, P. and Panayi, G. S. (1984). Reactive arthritis due to infestation with *Giardia lamblia*. *Journal of Rheumatology*, **11**, 719.

Yonis, Z. (1943). Amebic arthritis. *Harefuah*, **24**, 135.

5.3.10 Fungal arthritis

Carol A. Kemper and Stanley C. Deresinski

Introduction

Fungal infection of joints is a challenging but uncommon clinical problem whose aetiology is often belatedly recognized. Fungal joint infection most often results from the haematogenous dissemination of the pathogen from a primary portal of infection (usually pulmonary) directly to the synovial tissue or may initially affect para-articular bone with subsequent rupture into a joint space. Less commonly, such infection occurs as the result of direct inoculation of the organism into the joint space or synovial tissue. An inflammatory aseptic arthritis may also occur in association with certain fungal infections (e.g. coccidioidomycosis or histoplasmosis) as a result of the immune response to the organism, rather than a result of infection of the joint itself.

Epidemiology

Only a handful of fungi, perhaps five or six species at most, are responsible for the majority of human mycotic musculoskeletal infections (Schwarz 1984; Bradsher 1988; Fader and McGinnis 1988; Silveira *et al*. 1991; Cuellar *et al*. 1992; Cueller *et al*. 1993) (Table 1), but virtually all of the approximately 100 fungi pathogenic in man have been reported to cause infection of bones and/or joints. The frequency with which arthritis occurs, its clinical presentation, and its outcome varies depending on the specific fungal agent as well as upon host variables. For example, fungal arthritis caused by the endemic dimorphic fungi, such as *Histoplasmosis capsulatum, Blastomycosis dermatiditis*, and *Coccidioides immitis*, is often seen in patients without overt immunodeficiency. In contrast, infection resulting from *Candida* spp. is usually found in association with intravascular infection in individuals with readily apparent host factors, such as those with indwelling central venous catheters (often in association with the administration of long-term antibiotic therapy and/or parenteral nutrition), or those undergoing haemodialysis, or intravenous drug users. Defects in cellular immunity are critical to the dissemination of certain fungi from their initial portal of infection and the secondary infection of joint spaces. Patients with haematological malignancy or the acquired immune deficiency syndrome (AIDS), organ transplants recipients receiving immunosuppressive agents, or those receiving chronic corticosteroids are especially at risk for fungal arthritis. HIV-infected individuals are especially vulnerable to disseminated infection with *Cryptococcus neoformans, C. immitis*, and *H. capsulatum*.

Table 1 Risk factors for infection and the clinical setting of fungal joint infection

Rarely, joint infection occurs secondary to direct inoculation of the organism into the joint during aspiration or injection, trauma, or surgical intervention. Human to human transmission of these mycoses does not, for all practical purposes, occur.

Exposure to *B. dermatiditis, C. immitis, H. capsulatum*, or *Paracoccidioides brasiliensis* ordinarily occurs within their respective endemic areas (see Table 1). *C. immitis* is limited to endemic zones in North, Central and South America, while *H. capsulatum* is found in areas of both hemispheres, often in association with avian and chiropteran habitats. *B. dermatiditis*, while most often acquired in the United States, has also been reported from Africa and the Middle East. *P. brasiliensis* is found in Central and South America, although rare cases have been described in North America. For each of these dimorphic fungi, inhalation of conidia or arthroconidia released by the mycelial phase of the organism results in a primary pulmonary infection which is either subacute or acute, and typically self-limited, or which, in some cases, becomes chronic. Secondary dissemination during the acute or chronic phase of pulmonary infection results in a varying incidence of clinical joint infection for each of these diseases. *C. neoformans* and *Sporothrix schenckii* have a worldwide distribution. In contrast to the other fungi, infection with candida is ordinarily the consequence of host invasion by endogenous colonizing organisms.

Clinical picture

The clinical presentation of joint infection is most often indolent, although the onset of some infections, such as those caused by *B. dermatiditis, Candida* spp., and, occasionally, other fungi, may be acute, with hot, erythematous and tender joints and accompanying fever. The presentation may thus resemble an acute bacterial septic arthritis. Most cases, however, present with the usual findings of arthritis with decreased range of motion, tenderness and swelling. There is often evidence of joint effusion, but in some cases of chronic infection resulting from *C. immitis*, joint swelling may be because of synovial proliferation rather than the accumulation of fluid. The initial list of differential diagnoses may therefore be quite broad, and includes septic arthritis, rheumatoid arthritis, mycobacterial infection, brucellosis, and pigmented villonodular synovitis. While fungal arthritis may present in the setting of widespread fungal infection, in many instances there is little clinical evidence of extra-articular infection. Large weight-bearing joints, particularly the knee, are the usual targets.

Radiographic examination generally reveals evidence of joint effusion. Other findings which may be seen with varied frequency, depending upon the aetiology, host factors, and the chronicity of the infection, include erosion of juxta-articular cortex, osteoporosis, and associated para-articular osteomyelitis. These radiographic findings are also common to those found in tuberculosis, rheumatoid arthritis, sarcoidosis, metastatic neoplasm, eosinophilic granuloma, and pigmented villonodular synovitis (MacKenzie *et al*. 1988).

Clinically important information about joint integrity and the presence of otherwise unapparent para-articular osteomyelitis may be provided by magnetic resonance imaging which has greater sensitivity and resolution than other conventional techniques (MacKenzie *et al*. 1988; Brown *et al*. 1990). However, the role of this

procedure in the clinical evaluation and management of fungal arthritis has not been critically evaluated. Nuclear medicine techniques, such as scanning after injection of technetium pyrophosphate, may serve to confirm clinical evidence of joint inflammation (Fig. 1).

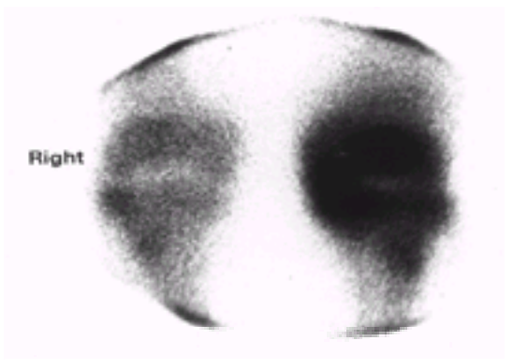


Fig. 1 Bone scintigraphy, using technetium-99m, demonstrating intense uptake of the radionuclide in the left knee of a patient with synovitis caused by *Sporothrix schenckii* (by courtesy of Jesse Hofflin MD).

Synovial fluid examination reveals an elevated white blood cell count. Although candidal and blastomycotic joint infections typically present with frankly purulent synovial fluid with a predominance of polymorphonuclear leucocytes, the other fungi often cause lesser degrees of inflammation with lower cell counts and variable predominance of either polymorphonuclear leucocytes or lymphocytes. The protein concentration is usually in excess of 3.0 g/dl while the glucose concentration is low to normal. Routine direct examination (e.g. Gram stain) usually does not reveal the organism, but cytological preparations are useful in the diagnosis of blastomycosis, cryptococcus, and, to a lesser degree, coccidioidomycosis. Culture of synovial fluid or synovial tissue usually yields the organism.

Synovial tissue histopathology is variable but often non-specific, such as in infection resulting from *S. schenckii* in which the organisms are few and difficult to visualize. Often a granulomatous reaction is observed and the differential diagnosis which must be considered includes fungal infection, brucellosis, mycobacterial infection, syphilis, protothecosis, rheumatoid arthritis, pigmented villonodular synovitis, sarcoidosis, Crohn's disease, foreign body reaction, gout, pseudogout, and oxalosis (Schwarz 1984).

The use of additional diagnostic procedures, such as blood cultures, bone marrow examination and culture, antibody tests, or tests for the detection of fungal antigen in serum or other body fluids, depend upon the clinical setting and the suspected aetiology. Tests of delayed dermal hypersensitivity to fungal antigens are generally not useful for diagnostic purposes.

Some infections may cause tenosynovitis in the absence of osteomyelitis or arthritis. Tenosynovitis may occur as the result of haematogenous dissemination or of direct inoculation, and is most often associated with *S. schenkii* infection, as well as, to lesser degrees, with infections caused by *C. immitis* and *C. neoformans*.

During the primary pulmonary infection with *C. immitis*, an acute self-limited arthritis or periarthritis, commonly referred to as 'desert rheumatism', may be seen in association with erythema nodosum, erythema multiforme, and occasionally hilar adenopathy. Thus the clinical picture may resemble sarcoidosis. An immunological process, probably immune complex deposition, is thought to be aetiologic. Acute aseptic inflammatory arthritis may also be seen in histoplasmosis, as well as in acute blastomycosis.

Management

Amphotericin B remains the therapeutic agent of choice for most serious fungal infections with the exception of those resulting from *Pseudoallescheria boydii* infection in which azole therapy (usually miconazole) is preferred. Several newer oral antifungal agents which may provide similar efficacy with less toxicity are undergoing active investigation.

Amphotericin B, administered intravenously, penetrates into synovial fluid to some extent (Farrell *et al*. 1978). While several authors have advocated directly injecting or irrigating the joint space with amphotericin B, the necessity for this mode of therapy is unproven, and there is concern that a chemical synovitis and articular damage may result. The toxicities of amphotericin are well known and include fever, chills, nausea, vomiting, hypotension, renal dysfunction, hypokalaemia and hypomagnesaemia. The renal toxicity may be dose limiting. Preliminary studies of lipid-associated amphotericin B indicate that higher doses of this drug may be administered with less toxicity than with standard preparations. Whether this is beneficial, however, remains unproven since such preparations also have reduced antifungal activity.

Fluorocytosine (5-FC) enters susceptible fungal cells through a specific permease system and is then converted to 5-fluorouracil. It has a narrow spectrum of activity which includes most *Candida* spp. as well as *C. neoformans*. In most instances, fluorocytosine is not administered as a single agent because of the possibility of the development of drug resistance during therapy. The drug penetrates well into all body fluids, including synovial fluid, and is renally excreted. The major toxicity, bone marrow suppression and resultant cytopenias, is directly related to serum concentrations in excess of 100 µg/ml. Serum concentrations of fluorocytosine should be closely monitored during administration, and dose adjustments must be made in the presence of changing renal function.

The azoles, such as miconazole, ketoconazole, fluconazole and itraconazole, inhibit the C-14 demethylation of lanosterol, thus impairing fungal cell membrane assembly. They have variable pharmacokinetic, toxicity, and antifungal profiles (Fromptling 1988). The intravenous form of miconazole has limited utility because of toxicity and unfavourable pharmacokinetics.

Ketoconazole was the first of the azoles available for oral administration. It has efficacy in a variety of fungal infections including those due to most *Candida* spp. (*Candida* (*Torulopsis*) *glabrata* is resistant), *H. capsulatum, B. dermatiditis, C. neoformans*, and *C. immitis*. Absorption is impaired in the absence of gastric acid. Ketoconazole penetrates into synovial fluid from the bloodstream. Elimination is non-renal with a terminal half-life of approximately 7.5 h. The most common adverse effect is gastrointestinal. Modest elevations in hepatic transaminases are not uncommon but significant hepatic toxicity rarely occurs. Depressed cortisol and testosterone levels may occur as a result of interference with sterol synthesis, but symptomatic hypocortisolism is very rare. Drug interactions occur with cyclosporin, phenytoin, warfarin, and rifampin (Hawkins Van Tyle 1984), as well as rifabutin; astemizole and terfenadine are contraindicated in patients receiving ketoconazole.

Fluconazole is a water-soluble bis-triazole with a high degree of bioavailability and the ability to penetrate into body fluids, including synovial fluid, and achieves concentrations similar to those in serum (O'Meeghan *et al*. 1990). Protein binding is low (approximately 10 per cent) and the elimination half-life is approximately 22 h. Clearance is predominantly renal. Drug–drug interactions occur with phenytoin, rifampicin, and rifabutin, as well as with astemizole and terfenadine. Hepatic toxicity is rare; steroid hormone synthesis is not affected. The spectrum of activity is similar to that of ketoconazole (Galgiani 1990).

Another bis-triazole, itraconazole, has a broader spectrum of activity than either ketoconazole or fluconazole. It has now been approved by the United States Food and Drug Administration for use in both blastomycosis and histoplasmosis, and is also very active against *S. schenkii* and most *Aspergillus* species (Tucker *et al*. 1988). Itraconazole is lipophilic, highly protein bound, well absorbed, and has an elimination half-life of approximately 24 h. Elimination is non-renal. Body fluid penetration is less than that of fluconazole. Cortisol synthesis is not impaired by itraconazole. Pharmacokinetic interaction with cyclosporin occurs (Kramer *et al*. 1990); the same precautions regarding drug–drug interactions should be observed with itraconazole as ketoconazole. Gastrointestinal side-effects are common (Tucker *et al*. 1990). Itraconazole, which requires stomach acid for absorption, is not recommended for patients who are receiving antacids, H_2 blockers, or those who have achlorhydria, unless adequate serum levels can be demonstrated.

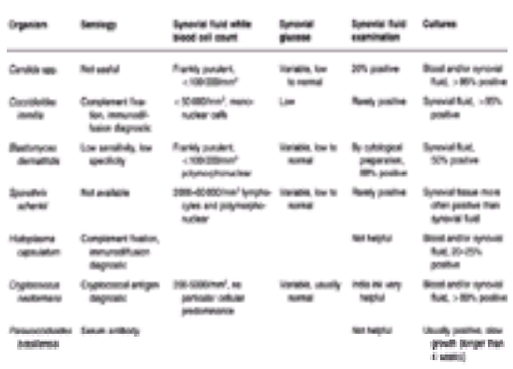
Individual mycoses

***Candida spp*.**

Candida albicans is a normal commensal of man and endogenous colonization is the source of most infections by *Candida* spp. Deep tissue infection generally occurs after amplification of colonization during an intervening immunodeficient state or during administration of broad-spectrum antibacterial therapy coupled with breaches in integumentary and mucosal barriers (Crislip and Edwards 1989). Candida infection of joints is typically the consequence of haematogenous dissemination (often from indwelling intravenous catheters in predisposed immunodeficient hosts or in intravenous drug users) (Marina *et al*. 1991). Joints previously afflicted by rheumatoid arthritis appear to be at increased risk of infection with candidal organisms (Campen *et al*. 1990). Less commonly, joint infection occurs secondary to direct inoculation of the organism into the joint during aspiration or injection of corticosteroids (Ginzler *et al*. 1979; Campen *et al*. 1990), trauma, or surgical intervention (including simple arthrotomy) (Arnold *et al*. 1981) (Table 1). Immunocompromise resulting from HIV infection does not appear to predispose to disseminated candidiasis or to candidal arthritis, except in those who use parenteral drugs (Edlestein and McCabe 1991; Munoz-Fernandez *et al*. 1991; Silveira *et al*. 1991). *Candida* spp. which have been implicated in septic arthritis include *C. albicans, Candida tropicalis, Candida parapsilosis, Candida guilliermondi, Candida krusei, Candida zeylanoides*, and *Torulopsis* (*Candida*) *glabrata*.

In contrast to many of the other fungal joint infections discussed here, the onset of disease caused by *Candida* spp. is acute in approximately two-thirds of cases (Bayer and Guze 1978). The remaining patients present with subacute disease. Instances of remarkably indolent presentations include those in which the arthritis was present for 9 months prior to diagnosis in a patient with acute myelocytic leukaemia (Gerster *et al*. 1980), and for 12 months in a patient receiving chronic haemodialysis (De Clerk *et al*. 1988). The large joints are most commonly affected.

Synovial fluid examination demonstrates a markedly elevated white blood-cell count (15 000 to 100 000 cells/mm ^{3}) with a predominance of polymorphonuclear leucocytes (Table 2) (Bayer and Guze 1978; Fainstein *et al*. 1982). The protein concentration is elevated while that of glucose is either low or normal. Histological examination of synovium reveals mononuclear cell infiltration but usually an absence of granulomata (Bayer and Guze 1978). The organism is visualized in only 20 per cent of cases on direct examination of synovial fluid by Gram stain or other methods. Synovial fluid or tissue consistently yield the organism in culture. Recovery of the organism from blood cultures may provide an important clue to the aetiology of the joint process.



Organism	Serology	Synovial fluid white blood cell count	Synovial glucose	Synovial fluid examination	Cultures
Candida spp.	Not useful	Frankly purulent, <100 000/mm^3	Variable, low to normal	20% positive	Blood and/or synovial fluid, >90% positive
Coccidioides immitis	Complement fixation, immunodiffusion diagnostic	<50 000/mm^3, mononuclear cells	Low	Rarely positive	Synovial fluid, <95% positive
Blastomyces dermatitidis	Low sensitivity, low specificity	Frankly purulent, <100 000/mm^3 polymorphonuclear	Variable, low to normal	By cytological preparation, 88% positive	Synovial fluid, 50% positive
Sporothrix schenkii	Not available	2000–50 000/mm^3 lymphocytes and polymorphonuclear	Variable, low to normal	Rarely positive	Synovial tissue more often positive than synovial fluid
Histoplasma capsulatum	Complement fixation, immunodiffusion diagnostic			Not helpful	Blood and/or synovial fluid, 20–25% positive
Cryptococcus neoformans	Cryptococcal antigen diagnostic	200–5000/mm^3, no particular cellular predominance	Variable, usually normal	India ink very helpful	Blood and/or synovial fluid, >80% positive
Paracoccidioides brasiliensis	Serum antibody			Not helpful	Usually positive, slow growth (longer than 4 weeks)

Table 2 Clinical and laboratory data helpful in the diagnosis of fungal joint infection

The cornerstone of management consists of systemic antifungal chemotherapy. Intravenously administered amphotericin B, with or without fluorocytosine, remains the standard treatment. However, based on limited data, the azoles (ketoconazole, fluconazole and itraconazole) appear to be as effective as amphotericin in the treatment of candidal infection caused by susceptible isolates, at least in non-neutropenic hosts. Several cases of candidal skeletal infection have been successfully managed with either ketoconazole (Gathe *et al*. 1987) or fluconazole alone (Sugar *et al*. 1990; Lafont *et al*. 1994). The potential role of lipopeptides, cilofungin, and lipid-associated amphotericin B is also unknown. Intra-articular amphotericin B has been utilized but the necessity or desirability of this method of treatment is questionable. Repeated joint aspiration is usually indicated. Surgical debridement may also be indicated (in addition to antifungal chemotherapy), particularly in cases of hip joint infection.

Candida spp. are causative in almost one-fifth of cases of nosocomial septic arthritis in neonates (Dan 1983; Ho *et al*. 1989), occurring exclusively in high-risk infants. Candida arthritis in neonates and in young infants is frequently associated with antibiotic therapy and parenteral administration of nutritional fluids (Yousefzadeh and Jackson 1980). Additional risk factors include prematurity, abdominal surgery, malnutrition, and immunosuppressive disease or therapy (Pope 1982).

In most neonatal cases, candida arthritis presents as just one facet of a systemic disease process and the organism may be recovered from a variety of extra-articular sites including blood, urine and spinal fluid. In the largest reported series (Dan 1983), joint aspirates yielded *C. albicans* in all but one instance. *C. tropicalis* was recovered from the remaining case. One or both knees were involved in 71 per cent of cases; polyarticular infection was seen in one-third of patients. The synovial fluid white blood-cell count was as high as 100 000/mm ^{3} with a predominance of polymorphonuclear leucocytes. The synovial membrane was hyperemic and purulent with erosion of cartilage. Radiographic evidence of adjacent osteomyelitis was seen in two-thirds of patients and in almost 90 per cent of joints, suggesting that in most cases, infection of the metaphysis was the original site of haematogenous dissemination with subsequent rupture into the articular cavity (Svirsky-Fein *et al*. 1979). Other radiographic findings included periarticular soft tissue swelling and joint effusion and, in the case of hip joint infection, subluxation of the femoral head. The mortality rate was 14 per cent. Major orthopaedic sequelas were seen in only one-tenth of survivors.

Fungal infection of prosthetic joints is exceedingly rare. In a series of reported cases, ten infections were identified in eight patients; all were caused by *Candida* spp. (Lambertus *et al*. 1988). One case was due each to *C. albicans* and *T. glabrata*, while three each were due to *C. tropicalis* and *C. parapsilosis*. Infection was probably the result of implantation of skin contaminants at the time of the original surgery. The infections were clinically low-grade, indolent, and presented 5 to 36 months after reconstructive arthroplasty. Pain and decreased range of motion were universally present and peri-articular swelling was common. Other signs of inflammation were absent. A sinus tract was seen in one patient. Radiographic examination revealed evidence of loosening and adjacent areas of osteolysis indicative of osteomyelitis. Technetium pyrophosphate and gallium nitrate scans are not useful in the setting of a loosened prosthesis (Fitzgerald and Kelly 1979). Synovial fluid white blood-cell counts were less inflammatory than that typically seen in native joint infections (4000 to 15 000/mm ^{3}); polymorphonuclear leucocytes were predominant.

Amphotericin B, with or without fluorocytosine, in combination with removal of the prosthesis and other foreign material, and debridement of affected tissue is the initial treatment of choice. Although no data are yet available, ketoconazole, fluconazole, and itraconazole may have a role in long-term 'maintenance' therapy of such cases. Reimplantation has been successfully reported in one patient 10 months after resection arthroplasty (Younkin *et al*. 1984).

Blastomycosis

Blastomycosis is an uncommonly encountered mycotic infection primarily endemic to parts of the midwestern, south-eastern, and Appalachian areas of the United States, but which is also seen in Africa and the Middle East (see Table 1). *B. dermatiditis* is a thermal dimorph whose mycelial phase is thought to reside in soil. Conversion to the yeast phase occurs after inhalation of spores. Primary pulmonary infection may be subclinical, acute, or subacute, but is usually self-limited; occasional cases may be chronic. Haematogenous dissemination is relatively frequent during the initial phase of the disease, leading to infection at almost any body site. While skin and bones are the most frequent sites of dissemination (25 to 60 per cent) in disseminated blastomycosis, only between 2.5 and 8 per cent of patients develop joint infection (Blastomycosis Cooperative Study of the Veterans Administration 1964; Witorsch and Utz 1968; George *et al*. 1985; McDonald *et al*. 1990). Those patients with particularly severe pulmonary disease, miliary involvement, or those who are immunocompromised are at the greatest risk for dissemination (Sarosi and Davies 1981; Recht *et al*. 1982). The risk of endogenous reactivation, which usually occurs during the first 2 to 3 years following the primary pulmonary infection, is small. Very rarely, joint infection is the result of direct inoculation secondary to trauma (Gnann *et al*. 1983).

While many patients with progressive or disseminated disease due to *B. dermatitidis* suffer from potentially predisposing conditions, such as diabetes, alcoholism, renal failure, and malignancy (Klein *et al*. 1986), this organism is not generally considered an opportunistic pathogen. Recht and colleagues described 78 patients with blastomycosis, 6 (13 per cent) of whom were immunocompromised, none, however, due to T-lymphocyte dysfunction (Recht *et al*. 1982). Those 6 patients had a similar clinical presentation and therapeutic response as the remaining patients who were not immunocompromised. Nevertheless, rapidly progressive and unusually

severe disease has been reported in patients with profoundly impaired immunity, such as patients who have had transplants and those with AIDS (Davies and Sarosi 1991).

Myalgias and arthralgias are common during the acute pulmonary phase of the disease, but erythema nodosum is not (Sarosi *et al*. 1974). A reactive arthritis, similar to that seen with coccidioidomycosis, has been reported (Berger and Kraman 1981).

The arthritis is monoarticular in 95 per cent of cases with the knee most commonly involved, followed by the ankle, elbow and wrist (see Table 2). Joint pain is often acute in onset and patients usually appear toxic. In contrast to coccidioidal arthritis, active pulmonary disease is present in more than 90 per cent of patients with joint involvement, and more than 70 per cent have evidence of additional dissemination to cutaneous or subcutaneous sites (Fountain 1973; Bayer *et al*. 1979b). In contrast to those with candidal or sporotrichal arthritis, less than one-third of patients have radiographic evidence of juxta-articular osteomyelitis (Bayer *et al*. 1979b).

Synovial fluid findings are similar to those seen in candida arthritis. The fluid is usually cloudy or frankly purulent with white blood-cell counts which may exceed 100 000/mm^3 and with a predominance of polymorphonuclear leucocytes. The concentration of protein in the synovial fluid exceeds 3.0 g/dl while the glucose is low to normal (Fountain 1973; Bayer *et al*. 1979b; Robert and Kauffman 1988). Cytological examination of synovial fluid may be more sensitive in detecting the organism than is culture. Bayer and colleagues described a series of nine patients who underwent joint fluid examination, eight (88 per cent) of whom had characteristic organisms detected by direct microscopy and seven (78 per cent) of whom had positive cultures of synovial fluid (Bayer *et al*. 1979b). In a study of five patients, cultures were positive in three (60 per cent) and KOH preparation were positive in two (40 per cent), but cytological examination of synovial fluid in four of the cases demonstrated the characteristic organisms in all (100 per cent) (George *et al*. 1985). The organism may also be recovered in culture or visualized on histopathology from synovial biopsy specimens. Histopathological examination of infected synovium reveals prominent polymorphonuclear leucocytes, often with microabscesses and occasional granulomata.

While most patients with acute self-limited pulmonary blastomycosis have demonstrable delayed dermal hypersensitivity to blastomycin, this reactivity wanes over time and is not of diagnostic value. Available serological tests have been disappointing with both false-negative and -positive results commonly seen.

Amphotericin B remains the drug of choice for many patients, particularly those who are critically ill, have evidence of progressive disease, or those who are immunosuppressed (Bradsher 1988). The total dose required is usually 1.0 to 2.0 g. In patients with joint infection and otherwise stable disease, itraconazole is very effective (approximate 90 per cent response rate) (Dismukes *et al*. 1992). The usual starting dose is 200 mg once daily, increasing to 400 mg daily as necessary. A loading dose of 200 mg three times daily can be given for the first 3 days. Therapy should be continued for approximately 6 months.

Ketoconazole also has some efficacy in patients without meningeal disease who are not severely immunocompromised (National Institute of Allergy and Infectious Disease Mycoses Study Group 1985; McManus and Jones 1986; Bradsher 1988). Ketoconazole should be initiated at a dose of 400 mg per day and continued for at least 6 months. The dose can be increased to 600 to 800 mg per day in those who are failing to respond to therapy or who develop a new focus of infection.

Fluconazole does not appear to be nearly as effective as itraconazole in the treatment of blastomycosis. There are, however, reported cases of its use in patients with pulmonary and meningeal disease (Pearson *et al*. 1992; Taillan *et al*. 1992).

Sporotrichosis

S. schenckii, a tissue dimorph, is commonly found on decaying vegetation and in soil in many areas of both hemispheres. Infections are both sporadic and epidemic. In contrast to the other soil fungi discussed here, cutaneous disease occurs secondary to inoculation of the organism as a result of trauma to the skin. The lymphocutaneous form, with the development of an ulcer at the site of cutaneous inoculation and proximal nodules in the area of lymphatic drainage, is the most common manifestation of infection (Belknap 1989). Persons at particular risk for this infection include rose cultivators and those who handle soil and spagnum moss (Kedes *et al*. 1964). Primary pulmonary infection may occur presumably as the result of inhalation of spores.

While arthralgias occur in approximately 2 per cent of those with acute cutaneous or lymphocutaneous disease, infection of the joint space with *S. schenckii* is rare, having occurred in only one of 3300 patients (0.03 per cent) in a large outbreak of sporotrichosis (Lurie 1963). Arthritis may occur in the presence of widespread dissemination to other sites, but is much more common as an isolated finding (Bayer *et al*. 1979a; Yao *et al*. 1986). Bayer and colleagues described 44 cases of sporotrichal joint infection, 20 per cent of which were associated with systemic and pulmonary disease (Bayer *et al*. 1979a). Most cases of sporotrichal arthritis are therefore believed to be caused by haematogenous dissemination of the organism, although some cases may be the result of articular extension of infection from an adjacent site of osteomyelitis or skin infection or, occasionally, from direct inoculation of the organism into the joint. More than 85 per cent of patients with systemic infection have predisposing underlying disease, including myeloproliferative disorders, malignancy, chronic corticosteroid use and alcoholism (Kedes *et al*. 1964; Bayer *et al*. 1979a).

Sporotrichal arthritis is an indolent and slowly progressive infectious process which predominantly affects the knee and other large weight-bearing joints, although the small joints of the hand and wrist are also commonly affected (Molstad and Strom 1978; Bayer *et al*. 1979a). Calhoun and colleagues described 11 cases of systemic sporotrichosis; 8 involved the skeletal system with a total of 12 joints being affected, including the wrist (63 per cent), knee (38 per cent), ankle (25 per cent), and elbow and phalanx (13 per cent each) (Calhoun *et al*. 1991). Monoarticular and polyarticular involvement occur with equal frequency. Most cases present as a slowly progressive synovitis or tenosynovitis with pain, warmth, swelling and restricted range of motion (Bayer *et al*. 1979a; Chang *et al*. 1984).

Radiographic abnormalities are seen in more than 90 per cent of cases, possibly reflecting the chronicity of infection prior to diagnosis. Osteoporosis of contiguous bone is the most common radiographic finding, followed by soft tissue swelling with effusion, 'punched out' osteolytic lesions, articular cartilage erosion and joint space narrowing (Bayer *et al*. 1979a) (see Fig. 2).

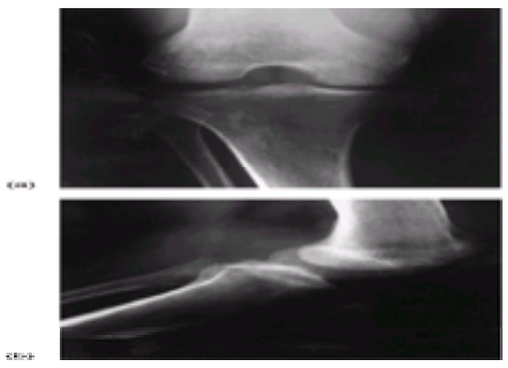


Fig. 2 (a) and (b). Radiographs of the left knee demonstrating only patchy osteopenia of the distal femur and proximal tibia from a patient with joint infection caused by *S. schenckii*. Bone scintigraphy of the same patient is shown in Fig. 1 (by courtesy of Jesse Hofflin MD).

Synovial fluid white blood-cell count is reported to range from 2800 to 60 000/mm^3. Both lymphocytes and polymorphonuclear leucocytes may be seen (see Table 2). The protein concentration is high while that of glucose is low to normal (Lesperance *et al*. 1988). The diagnosis may be delayed because of the isolated nature of the infection, the rarity of visualizing the organism on smears of synovial fluid, the often non-specific nature of synovial histopathology (which may resemble that of rheumatoid or tuberculous arthritis) (Stratton *et al*. 1981), and the paucity of organisms in tissue (Khan *et al*. 1983; Schwartz 1989) (Fig. 3). Asteroid bodies, often said to be pathognomonic of sporotrichosis, may, in fact, be seen in other infections. Isolation of the organism in culture is the cornerstone of diagnosis. Synovial tissue is more likely to yield the organism (usually within five days) than is synovial fluid. Skin tests are only useful for epidemiological surveys. A variety of serological tests have been utilized with varying results (Winn 1988).

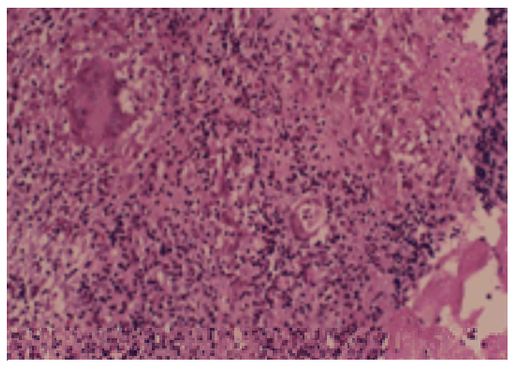


Fig. 3 Granulomatous reaction with typical giant cells in synovium obtained from the same patient described in Fig. 2. Organisms were not visualized with this stain or with Gomori–methanamine silver or periodic acid–Schiff stains, and the diagnosis was made by recovery of the organism in culture from the synovial tissue. Haematoxylin and eosin, original magnification × 200 (by courtesy of Jesse Hofflin MD).

While amphotericin B had been recommended for the treatment of skeletal sporotrichosis, newer data indicate that itraconazole is very effective in this infection. In a recent National Institute of Allergies and Infectious Diseases non-comparative clinical treatment trial of 30 patients with both lymphocutaneous and systemic sporotrichosis, one-half of whom had osseous or articular infection, itraconazole (100 to 600 mg daily for 3 to 18 months) was effective in 83 per cent (Sharkey-Mathis *et al*. 1993). However, 6 to 18 months after treatment was discontinued, 7 patients relapsed. Two of these patients have subsequently responded to a second course of therapy. Resolution of infection and normalization of joint mobility and function has been reported in 3 other patients who received itraconazole 200 mg daily (Winn *et al*. 1993).

Despite reasonably effective penetration by the drug into synovial fluid, ketoconazole has effected responses in only approximately two-thirds of patients with systemic sporotrichosis, including patients with joint infection (Dismukes *et al*. 1983; Graybill *et al*. 1983; Horsburgh *et al*. 1983; Pluss and Opal 1986; Calhoun *et al*. 1991). Fluconazole has been similarly disappointing in lymphocutaneous infection (Restrepo *et al*. 1986). Potassium iodide, which is effective in the lymphocutaneous form of the disease, has no role in the treatment of deep tissue infection, such as arthritis. Sporadic cases of skeletal disease have, however, reportedly responded to treatment with potassium iodide (Govender *et al*. 1989).

Intra-articular administration of amphotericin B has also been utilized, but this is unlikely to be necessary (Downs *et al*. 1989). Surgical debridement may also be necessary on occasion, but should be reserved for persistent culture positivity and in cases of tenosynovitis (Winn 1988).

Coccidioidomycosis

C. immitis is endemic to the soils of certain areas of the Lower Sonoran life zone of the Western hemisphere, with most cases resulting from exposure to airborne arthroconidia in Arizona and the the southern central valley of California (see Table 1). Upon reaching the alveoli of the infected host, the organism, a tissue dimorph, converts to the spherule–endospore phase. Approximately one-half of infected patients become symptomatic and, in the vast majority of these, the infection is self-limited with influenza-like symptoms. Transient arthralgias or aseptic inflammatory arthritis, which probably represents an immunologically mediated inflammatory process similar to erythema nodosum (which is also often seen), occur in 3 to 5 per cent of patients with primary pulmonary coccidioidomycosis (Fiese 1958). Treatment consists of giving non-steroidal anti-inflammatory agents.

Clinically important extrapulmonary dissemination occurs in fewer than 0.5 per cent of cases, although certain groups are at greater risk for dissemination. While many patients with disseminated disease have no impairment in immune function, approximately one-half are immunocompromised by corticosteroids, diabetes, renal failure, or other immunosuppresive therapy. HIV-infection increases both the risk of more frequent and more severe coccidioidomycosis (Galgiani and Ampel 1990).

Joint space infection occurs in up to 25 to 30 per cent of patients with disseminated disease, with occasional extension into adjacent bony areas (Fiese 1958; Deresinski 1980; Deresinski 1994). Monoarticular arthritis occurs in more that 90 per cent of cases, with large weight-bearing joints, particularly the knee and ankle, being most frequently affected (Deresinski 1980; Deresinski 1994). At the time of presentation with joint disease, occult sites of dissemination are present in up to 25 per cent of cases (Winter *et al*. 1975). Extrapulmonary sites of infection, including meningeal, bone, and joint infection, should therefore be avidly sought for in any patient with disseminated coccidioidomycosis.

While some patients may initially present with an acutely inflamed joint, most infections are indolent with progressive effusion and synovial thickening. The diagnosis of joint infection is often delayed, and chronic infection frequently results in significant articular and bony destruction with resultant loss of joint function (Fig. 4). Occasionally, chronic arthrocutaneous fistulas develop with drainage of synovial fluid (Fig. 5). Baker's cysts may occur as a consequence of knee involvement. Effusion and erosion of articular cortex and adjacent osteoporosis are commonly seen on radiographic examination (Carter 1934; Bayer and Guze 1979). Technetium pyrophosphate radioisotope scans usually localize to the affected joints.

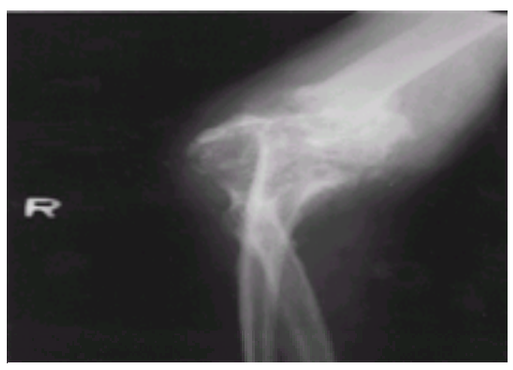


Fig. 4 Radiograph of the right elbow demonstrating destruction of the articular cortex and osteomyelitis of contiguous bone of an elderly women with chronic coccidioidal arthritis of many years duration despite multiple courses of antifungal therapy.

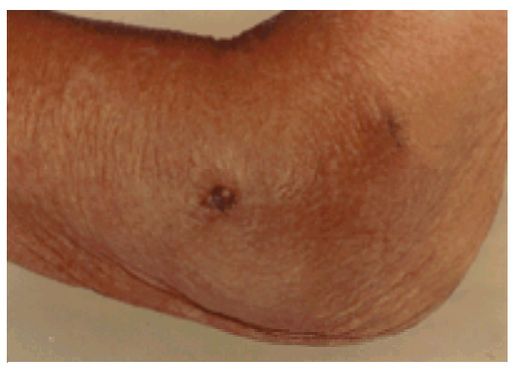


Fig. 5 Chronic coccidioidal arthritis of the same patient as in Fig. 4 demonstrating the right elbow joint fixed in flexion. The sinus tracts intermittently drain material from which *Coccidioides immitis* is recoverable in culture (by courtesy of John S. Hostetler MD).

Synovial fluid is inflammatory with total white blood-cell counts as high as 50 000/mm ^{3}(Table 2). Mononuclear cells usually, but not always, predominate. Protein is greater than 3.0 g/dl, glucose is low and mucin clot is poor (Aidem 1968; Deresinski and Stevens 1974). Culture of synovial fluid yields the organism in approximately 50 per cent of cases, usually within 3 to 6 days. Greater yield is seen with culture and histological examination of synovial tissue (Greenman *et al*. 1975). The affected proliferative synovium, which often invades cartilage and articular surfaces, exhibits granulomatous villonodular inflammatory changes with the characteristic endosporulating spherules visible on microscopic examination (Haug and Merrifield 1959). Most importantly, if coccidioidomycosis infection is suspected, the microbiology laboratory must be notified because of the significant biohazard represented by this organism in culture.

Serum complement-fixing antibody to coccidioidin is almost universally present, with the height of the titre reflecting the extent of dissemination, as in other manifestations of disseminated infection with this organism (Deresinski 1980). Delayed dermal hypersensitivity to coccidioidin may be absent.

Patient prognosis depends upon the extent of dissemination to other sites, particularly the central nervous system. Treatment consists of systemic administration of antifungal agents, and amphotericin B remains the treatment of choice in many cases. Patients with disseminated disease often receive a total of 1.0 to 2.5 g of amphotericin B. Continued therapy is indicated until remission has been achieved, as defined by objective clinical measures, and improvement in serological and radiographic data. Amphotericin B has also been administered intra-articularly, but the therapeutic necessity or advisability of this is uncertain. Arthrodesis is generally effective, but not desirable.

The need for synovectomy and debridement of infected bone and tissue remain controversial. Despite appropriate medical and surgical intervention, the joint infection often remains progressive and disabling (Lantz *et al*. 1988). In one study, 7 of 14 patients who received amphotericin alone failed therapy, whereas none of the 14 patients who were treated with a combination of medical and surgical approaches relapsed (Bried and Galgiani 1986). In another similar study, 7 of 9 patients who received amphotericin B and who underwent surgical debridement remained disease-free at least 4 years later (Bisla and Taber 1976). The two remaining patients developed recurrent disease, despite having received more than 3.0 g of amphotericin B each. Patients with complement fixation titres greater than or equal to 1:128 were most likely to fail in response to medical therapy alone (Bried and Galgiani 1986).

Ketoconazole (400 to 800 mg per day) has some efficacy, but the relapse rate is high (approximately 30 per cent) (Galgiani *et al*. 1988). Although both fluconazole and itraconazole have been used in cases of pulmonary disease and meningitis, relapses are frequent. The optimal dose and duration of therapy remain under study. An unusual case of prosthetic hip joint infection caused by *C. immitis* responsed to long-term therapy with fluconazole (Nomura nd Ruskin 1994).

Histoplasmosis

H. capsulatum is endemic to many areas within the temperate zones of the world, but is most heavily concentrated in the Ohio, Mississippi and Missouri River valleys of the United States (Table 1). The organism is a thermal dimorph with the mycelial phase existing in soil, generally in association with bird and bat guano. Large outbreaks occur in urban endemic areas. Speleologists throughout the world may be at risk.

Upon inhalation by the human or animal host, microconidia reach the alveoli where they convert to the yeast phase. While greater than 95 per cent of infections are subclinical, an influenza-like respiratory illness may result from infection. Haematogenous dissemination is rare and occurs most commonly in patients with impaired cellular immunity. Disseminated histoplasmosis is reported in approximately one-third of AIDS patients in Kansas City, Missouri (McKinsey *et al*. 1989). Persons with HIV infection who travel to or have previously lived in an endemic area are at risk for reactivation disease (Minamoto and Armstrong 1988; Salzman *et al*. 1988).

Immunologically mediated arthralgias and aseptic inflammatory arthritis, similar to that reported for coccidioidomycosis, are common in primary histoplasmosis (Class and Casio 1972; Rosenthal *et al*. 1983). Based on previously published reports, one review found that arthralgias occurred in 3 to 21 per cent, and erythema nodosum and erythema multiforme occurred in 1 to 42 per cent of patients with documented histoplasmosis (Schwarz 1984). During a single outbreak of acute histoplasmosis in 381 symptomatic patients, 16 (4.1 per cent) developed arthralgias and 6 (1.6 per cent) developed aseptic arthritis (Rosenthal *et al*. 1983). The knees, ankles, wrists, and small joints of the hands were the most common sites of involvement; approximately 50 per cent of the cases were polyarticular. The joint involvement may be additive or migratory, and is often symmetric. Synovial fluid is inflammatory. This clinical problem is self-limited and is treated with non-steroidal anti-inflammatory agents (Sellers *et al*. 1965).

In contrast to candida and coccidioidomycosis, infection of the synovium or joint space by *H. capsulatum* is exceedingly rare. It is usually monoarticular and has been reported in both apparently immunologically normal (Key and Large 1942; Omer *et al*. 1963; Van Der Schee *et al*. 1990) and compromised hosts (Jones 1985). Juxta-articular osteomyelitis may be present. The diagnosis of histoplasmosis can be made by culture of both blood and infected sites, including synovial fluid, and histological demonstration of the infecting organism. The organism is readily cultivated on a variety of media. The lysis–centrifugation technique hastens recovery from the blood of patients with active dissemination (Paya *et al*. 1987).

Detection of antigen in serum or urine has been utilized in the diagnosis of disseminated histoplasmosis (Wheat *et al*. 1986). Although both false-positive and -negative results occur, antibody tests are diagnostically useful. Detection of serum complement-fixing antibody to the yeast phase of the organism of 1:32 or greater should be regarded as presumptive evidence of histoplasmosis. Titres of 1:8 or greater to mycelial-phase antigens or the presence of 'M' or 'H' bands by immunodiffusion are also highly suggestive of histoplasmosis (Kaufmann 1971; Wheat *et al*. 1982). Histoplasmin skin testing is useful only for epidemiological purposes.

Amphotericin B remains the treatment of choice for severe, life-threatening forms of histoplasmosis. Inadequate information is available concerning the usefulness of the azoles in joint infection caused by *H. capsulatum*. However, both itraconazole (Sharkey-Mathis *et al*. 1993) and ketoconazole (National Institute of Allergy and Infectious Disease Mycoses Study Group 1985; Dismukes *et al*. 1992) have been effective in non-immunocompromised patients with other forms of this disease. In a recent non-comparative treatment trial, itraconazole (200 to 400 mg per day) was effective in 81 per cent of patients with histoplasmosis (Sharkey-Mathis *et al*. 1993). Patients with chronic pulmonary disease or less than 2 months of therapy were more likely to fail. In patients with AIDS, amphotericin B is often used to suppress the acute infection, but the majority of cases will recur without chronic suppressive therapy. For HIV-infected patients with milder disease, itraconazole alone is effective in approximately 80 per cent (Wheat *et al*. 1990)).

Cryptococcosis

C. neoformans is worldwide in distribution. Skin test surveys suggest that subclinical infection is quite common in normal hosts. Clinical disease occurs predominantly, but not exclusively, in individuals with defects in cellular immunity (Table 1). The incidence of cryptococcal disease has greatly increased because of the frequency with which patients with HIV are infected with this encapsulated yeast.

The primary pulmonary infection may be subclinical or acute, and haematogenous dissemination may result in diffuse organ involvement. Almost any organ can be involved, but the organism has a particular predilection for the brain and meninges (Perfect *et al*. 1983). Although osteomyelitis occurs in up to 10 per cent of patients with systemic disease, cryptococcal arthritis, frequently associated with areas of contiguous osteomyelitis, is rare with fewer than 25 cases reported in the literature (Bayer *et al*. 1980; Ricciardi *et al*. 1986; Stead *et al*. 1988). Most patients have severe deficits in cellular immunity, such as those with AIDS, sarcoidosis, diabetes, and renal allograft recipients (Bosch *et al*. 1994). Both gout and calcium pyrophosphate disease appear to increase the risk of cryptococcal infection in affected joints (Sinnott and Holt 1982; Ricciardi *et al*. 1986).

Although soft-tissue swelling, inflammation and frank cellulitis have been reported in cases of cryptococcal joint infection (Bunning and Barth 1984), most cases are indolent in presentation. The knee is involved in approximately 60 per cent of reported cases, followed by an equal number of cases in the sternoclavicular and acromial-clavicular joints, elbow, wrist and ankle (Table 2). Approximately one-third of the cases are polyarticular. Radiographs demonstrate an erosive arthritis and juxta-articular osteomyelitis, and computed tomographic scans often show evidence of parasynovial inflammation. Examination of the synovial fluid reveals a white blood-cell count of 200 to 20 000/mm^3, with a predominance of mononuclear cells. However, the peripheral white blood-cell count and erythrocyte sedimentation rate is often normal (Adams and McDonald 1984; Bunning and Barth 1984; Brand *et al*. 1985).

Amphotericin B should be administered as initial therapy in most cases of disseminated disease. Many experts advocate the concomitant administration of fluorocytosine for approximately 4 weeks. In order to avoid undue toxicity, serum concentrations of fluorocytosine can be monitored weekly and maintained within a

range of approximately 50 to 80 µg/ml. Once the systemic disease is under control and the joint disease is improving, consideration can be given to completing treatment with fluconazole (400 mg per day). Itraconazole may also be effective, but the data is limited. Patients with AIDS should remain on life-long suppressive maintenance therapy.

Paracoccidioidomycosis

Paracoccidioidomycosis is endemic only to areas of Central and South America where it is the most commonly encountered respiratory mycosis. *P. brasiliensis* is thermally dimorphic. As is true for the other dimorphic fungi, conidia released by the mycelial phase of the fungus are inhaled and convert to the yeast phase in the alveoli. Acute, self-limited pulmonary infection may occur, although most patients present with chronic pulmonary disease and evidence of chronic haematogenous dissemination, including painful granulomata of the skin, lymphadenopathy, and ulceration of mucous membranes (Sugar 1988). Skeletal disease occurs, but is rare (Table 1); no cases were identified in two reviews describing a total of 66 cases (Londero and Rambos 1972; Murray *et al*. 1974). A case of joint infection, with soft tissue swelling and cartilagenous destruction, has been described (Castaneda *et al*. 1985). Typical budding yeast forms were seen on examination of the synovial fluid and cultures were positive. The diagnosis is usually made on the basis of visualization of the organisms in synovial fluid or tissues, or by culture. Serological tests have been utilized with varying success. Skin tests are useful only for epidemiological surveys.

Although the disease is rarely encountered outside of endemic areas, the diagnosis should be suspected in any individual at epidemiological risk. Paracoccidioidomycosis primarily occurs in persons without evidence of immune dysfunction, but cases of severe disseminated disease have been recently described in immunosuppressed patients (Londero and Rambos 1972; Murray *et al*. 1974; Restrepo *et al*. 1976; Sugar *et al*. 1984; Sugar 1988).

Amphotericin B is effective in the treatment of disseminated paracoccidioidomycosis, although itraconazole and ketoconazole appear as effective in the treatment of milder cases (Stevens and Vo 1982; Sugar 1988; Kwon-Chung and Bennett 1992). Relapses after treatment are common, and chronic suppressive therapy with one of the azoles or a sulphonamide are therefore currently recommended.

Miscellaneous mycoses

A variety of additional fungi have been implicated in joint infections. Various species of *Aspergillus* have resulted in cases of joint infection, often associated with contiguous osteomyelitis (Tack *et al*. 1982; Denning and Stevens 1990; Cosgarea *et al*. 1993). While haematogenous dissemination of the organism was implicated in some cases, introduction of the organism into the joint space has occurred during surgical or arthroscopic procedures, or as a result of trauma.

Skeletal infections caused by *Alternaria* spp. and *Bipolaris hawaiiensis* (Sharkey *et al*. 1990), *Acremonium* spp. (Szombathy *et al*. 1988), *Cunninghamella bertholletiae* (Mostaza *et al*. 1989), *Exiophiala jeanselmei* (Roncoroni and Smayevsky 1988), *Exiophila spinifera* (Sharkey *et al*. 1990), *Fusarium solani* (Jakle *et al*. 1983), *Madurella mycetomi* (Yagi *et al*. 1983; McGinnis and Fader 1988), *Phialophora parasitica* (Kaell and Weitzman 1983), *Saccharomyces* spp. (Feld *et al*. 1982), *P. boydii* (Kemp *et al*. 1982; Ansara *et al*. 1987; Rippon 1988), and *Trichosporon beigelii* (Gardella *et al*. 1985) have also been reported.

Chapter References

Adams, R. and McDonald, M. (1984). Cryptococcal arthritis of the acromio-clavicular joint. *North Carolina Medical Journal*, **45**, 23–4.

Aidem, H.P. (1968). Intra-articular amphotericin B in the treatment of cocciodioidal synovitis of the knee. Case report. *Journal of Bone and Joint Surgery*, **50A**, 1663–8.

Ansara, R.A., Hindson, D.A., Stevens, D.L., and Kloss, J.G. (1987). *Pseudoallescheria boydii* arthritis and osteomyelitis in a patient with Cushing's disease. *Southern Medical Journal*, **80**, 90–2.

Arnold, H.J., Dini, A., Jonas, G., and Zorn, E.L. (1981). *Candida albicans* arthritis in a healthy adult. *Southern Medical Journal*, **74**, 84–5.

Bayer, A.S. and Guze, L.B. (1978). Fungal arthritis. I. Candida arthritis: diagnostic and prognostic implication and therapeutic considerations. *Seminars in Arthritis and Rheumatism*, **8**, 142–50.

Bayer, A.S. and Guze, L.B. (1979). Fungal arthritis. II. Coccidioidal synovitis: clinical, diagnostic, therapeutic, and prognostic considerations. *Seminars in Arthritis and Rheumatism*, **8**, 200–11.

Bayer, A.S., Scott, V.J., and Guze, L.B. (1979*a*). Fungal arthritis. III. Sporothrichal arthritis. *Seminars in Arthritis and Rheumatism*, **9**, 66–74.

Bayer, A.S., Scott, V.J., and Guze, L.B. (1979*b*). Fungal arthritis. IV. Blastomycotic arthritis. *Seminars in Arthritis and Rheumatism*, **9**, 145–51.

Bayer, A.S., Choi, C., Tilman, D.B., and Guze, L.B. (1980). Fungal arthritis. V. Cryptococcal and histoplasmal arthritis. *Seminars in Arthritis and Rheumatism*, **9**, 218–27.

Belknap, B.S. (1989). Sporotrichosis. *Dermatology Clinic*, **7**, 193–202.

Berger, R. and Kraman, S. (1981). Acute miliary blastomycosis after 'short-course' corticosteroid treatment. *Archives of Internal Medicine*, **141**, 1223–5.

Bisla, R.S. and Taber, T.H. (1976). Coccidioidomycosis of bone and joints. *Journal of Clinical Orthopedics and Related Research*, **121**, 196–204.

Blastomycosis Cooperative Study of the Veterans Administration (1964). Blastomycosis. 1: A review of 198 collected cases in Veteran Administration Hospitals. *American Review of Respiratory Disease*, **89**, 658–72.

Bosch, X., Roman, R., Font, J., Alemany, S., and Coca, A. (1994). Bilateral cryptococcosis of the hip. *Journal of Bone and Joint Surgery*, **76A**, 1234–8.

Bradsher, R.W. (1988). Blastomycosis. *Infectious Disease Clinics of North America*, **1**, 877–98.

Brand, C., Warren, R., Luxton, M., and Barraclough, D. (1985). Cryptococcal sacroiliitis. *Annals of the Rheumatic Diseases*, **44**, 126–7.

Bried, J.M. and Galgiani, J.N. (1986). *Coccidioides immitis* infections of bones and joints. *Clinical Orthopedics*, **211**, 235–43.

Brown, D.G., Edwards, N.L., Greer, J.M., Longley, S., Gillespy, T.I., and Panush, R.S. (1990). Magnetic resonance imaging in patients with inflammatory arthritis of the knee. *Clinical Rheumatology*, **9**, 73–83.

Bunning, R.D. and Barth, W.F. (1984). Cryptococcal arthritis and cellulitis. *Annals of the Rheumatic Diseases*, **43**, 508–10.

Calhoun, D.L. *et al*. (1991). Treatment of systemic sporotrichosis with ketoconazole. *Reviews of Infectious Diseases*, **13**, 47–51.

Campen, D.H., Kaufman, R.L., and Beardmore, T.D. (1990). Candida septic arthritis in rheumatoid arthritis. *Journal of Rheumatology*, **17**, 86–8.

Carter, R.A. (1934). Infectious granulomas of bones and joints with special reference to coccidioidal granuloma. *Radiology*, **23**, 1–16.

Castaneda, O.J., Alarcon, G.S., Garcia, M.T., and Lumbreras, H. (1985). *Paracoccidioides brasiliensis* arthritis. Report of a case and review of the literature. *Journal of Rheumatology*, **12**, 356–8.

Chang, A.C., Destouet, J.M., and Murphy, W.A. (1984). Musculoskeletal sporotrichosis. *Skeletal Radiology*, **12**, 23–8.

Class, R.N. and Casio, F.S. (1972). Histoplasmosis presenting as acute polyarthritis. *New England Journal of Medicine*, **287**, 1133–4.

Cosgarea, A.J., Tejani, N., and Jones, J.A. (1993). Carpal aspergillus osteomyelitis: case report and review of the literature. *Journal of Hand Surgery*, **18A**, 722–6.

Crislip, M.A. and Edwards, J.E. (1989). Candidiasis. *Infectious Disease Clinics of North America*, **3**, 103–33.

Cuellar, M.L., Silveira, L.H., and Espinoza, L.R. (1992). Fungal arthritis. *Annals of the Rheumatic Diseases*, **51**, 690–7.

Cuellar, M.L., Silveira, L.H., Citera, G., Cabrera, G.E., and Valle, R. (1993). Other fungal arthritides. *Rheumatic Disease Clinics of North America*, **19**, 439–55.

Dan, M. (1983). Neonatal septic arthritis. *Israel Journal of Medical Science*, **19**, 967–71.

Davies, S.F. and Sarosi, G.A. (1991). Clinical manifestations and management of blastomycosis in the compromised patient. In *Fungal infection in the compromised patient*, (2nd edn) (ed. D.W.

Warnock and M.D. Richardson), pp. 215–29. John Wiley, Chichester.

De Clerk, L., Dequaker, J., Westhovens, R., and Hauglustaine, D. (1988). *Candida parapsilosis* in a patient receiving chronic hemodialysis. *Journal of Rheumatology*, **15**, 372–4.

Denning, D.W. and Stevens, D.A. (1990). Antifungal and surgical treatment of invasive aspergillosis: Review of 2121 published cases. *Reviews of Infectious Diease*, **12**, 1147–201.

Deresinski, S.C. (1980). Coccidioidomycosis of the musculoskeletal system. In *Coccidioidomycosis* (ed. D.A. Stevens), pp. 195–212. Plenum Press, New York.

Deresinski, S.C. (1994). *Coccidioides immitis*. In *Infectious diseases* (ed. J.A. Bartlett, S.L. Gorback, and N.R. Blacklow), pp. 1912–23. W.B. Saunders, Philadelphia.

Deresinski, S.C. and Stevens, D.A. (1974). Coccidioidomycosis in compromised hosts. *Medicine*, **54**, 377–95.

Dismukes, W.E. *et al*. (1992). Itraconazole therapy for blastomycosis and histoplasmosis. National Institute for Allergy and Infectious Disease Mycoses Study Group. *American Journal of Medicine*, **93**, 489–97.

Dismukes, W.E. *et al*. (1983). Treatment of systemic mycoses with ketoconazole: emphasis on toxicity and clinical response in 52 patients. *Annals of Internal Medicine*, **98**, 13–20.

Downs, N.J., Hinthorn, D.R., Mhatre, V.R., and Liu, C. (1989). Intra-articular amphotericin B treatment of *Sporothrix schenkii* arthritis. *Archives of Internal Medicine*, **149**, 954–5.

Edlestein, H. and McCabe, R. (1991). *Candida albicans* septic arthritis and osteomyletitis of the sternoclavicular joint in a patient with human immunodeficiency virus infection. *Journal of Rheumatology*, **18**, 110–11.

Fader, R.C. and McGinnis, M.R. (1988). Infections causes by dermatiaceous fungi: chromoblastomycosis and phaeohyphomycosis. *Infectious Disease Clinics of North America*, **1**, 925–38.

Fainstein, V., Gilmore, C., Hopfer, R.L., Maksymiuk, A., and Bodey, G.P. (1982). Septic arthritis due to *Candida* species in patients with cancer: report of five cases and review of the literature. *Review of Infectious Disease*, **4**, 78–85.

Farrell, J.B., Person, D.A., Lidsky, M.D., Hopfer, R.L., and Musher, D.M. (1978). *Candidal tropicalis* arthritis—assessment of amphotericin B therapy. *Journal of Rheumatology*, **5**, 267–71.

Feld, R., Fornasier, V.L., Bombardier, C., and Hastings, D.E. (1982). Septic arthritis due to *Saccharomyces* species in a patient with chronic rheumatoid arthritis. *Journal of Rheumatology*, **9**, 637–40.

Fiese, M. (1958). *Coccidioidomycosis*. Charles.C. Thomas, Springfield, IL.

Fitzgerald, R.H., Jr. and Kelly, P.J. (1979). Total joint arthroplasty: biologic causes of failure. *Mayo Clinic Proceedings*, **54**, 590–6.

Fountain, F.F., Jr. (1973). Acute blastomycotic arthritis. *Archives of Internal Medicine*, **132**, 684–8.

Fromptling, R.A. (1988). Overview of medically important antifungal azole derivatives. *Clinical Microbiology Review*, **1**, 187–217.

Galgiani, J.N. (1990). Fluconazole, a new antifungal agent. *Annals of Internal Medicine*, **113**, 177–9.

Galgiani, J.N. and Ampel, N.M. (1990). Coccidioidomycosis in human immunodeficiency virus-infected patients. *Journal of Infectious Disease*, **162**, 1165–9.

Galgiani, J.N., Stevens, D.A., Graybill, J.R., Dismukes, W.E., and Cloud, G.A. (1988). Ketoconazole therapy of progressive coccidioidomycosis. Comparison of 400- and 800-mg doses and observations at higher doses. *American Journal of Medicine*, **84**, 603–10.

Gardella, S. *et al*. (1985). Fatal fungemia with arthritic involvement caused by *Trichosporon beigelii* in a bone marrow transplant recipient. *Journal of Infectious Disease*, **151**, 566.

Gathe, J.C., Harris, R.L., Garland, B., Bradshaw, M.W., and Williams, T.W. (1987). Candida osteomyelitis. Report of 5 cases and review of the literature. *American Journal of Medicine*, **82**, 927–37.

George, A.L., Jr., Hays, J.T., and Graham, B.S. (1985). Blastomycosis presenting as monoarticular arthritis: the role of synovial fluid cytology. *Arthritis and Rheumatism*, **28**, 516–21.

Gerster, J.C., Glauser, M.P., Delacretaz, F., and Nguyen, T. (1980). Erosive candida arthritis in a patient with disseminated candidiasis. *Journal of Rheumatology*, **7**, 911–14.

Ginzler, E., Meisel, A.D., Munters, M., and Kaplan, D. (1979). Candida arthritis secondary to repeated intra-articular corticosteroids. *New York State Journal of Medicine*, **79**, 392–4.

Gnann, J.W., Bressler, G.S., Bodet, C.A., and Avent, C.K. (1983). Human blastomycosis after a dog bite. *Annals of Internal Medicine*, **98**, 48–9.

Govender, S., Rasool, M.N., and Ngcelwane, M. (1989). Osseous sporotrichosis. *Journal of Infection*, **19**, 273–6.

Graybill, J.R., Craven, P.C., Donovan, W., and Matthew, E.B. (1983). Ketoconazole therapy of systemic fungal infections. Inadequacy of standard dosage regimens. *American Review of Respiratory Disease*, **12**, 171–4.

Greenman, R., Becker, J., Campbell, G., and Remington, J. (1975). Coccidioidal synovitis of the knee. *Archives of Internal Medicine*, **135**, 526–30.

Haug, W.A. and Merrifield, R.C. (1959). Coccidioidal villous synovitis. *American Journal of Clinical Pathology*, **31**, 165–71.

Hawkins Van Tyle, J. (1984). Ketoconazole: mechanism of action, spectrum of activity, pharmacokinetics, drug interactions, adverse reactions and therapeutic use. *Pharmacotherapy*, **4**, 343–73.

Ho, N.K., Low, Y.P., and See, H.F. (1989). Septic arthritis in the newborn—a 17 years' experience. *Singapore Medical Journal*, **30**, 356–8.

Horsburgh, C.R., Jr., Cannady, P.B., Jr., and Kirkpatrick, C.H. (1983). Treatment of fungal infections in the bones and joints with ketoconazole. *Journal of Infectious Disease*, **147**, 1064–9.

Jakle, C., Leek, J.C., Olson, D.A., and Robbins, D.L. (1983). Septic arthritis due to *Fusarium solani*. *Journal of Rheumatology*, **10**, 151–3.

Jones, P.G. (1985). Septic arthritis due to *Histoplasmosis capsulatum* in a leukaemic patient. *Annals of the Rheumatic Diseases*, **44**, 128–9.

Kaell, A.T. and Weitzman, I. (1983). Acute monoarticular arthritis due to *Phialophora parasitica*. *American Journal of Medicine*, **74**, 519–22.

Kaufmann, L. (1971). Serological tests for histoplasmosis: their use and interpretation. In *Histoplasmosis: Proceedings of the Second National Conference*, Centers for Disease Controls, Atlanta Georgia, (ed. L. Ajello, E.W. Chick, and M.L. Furcolow), pp. 321–6. Charles C. Thomas, Springfield, IL.

Kedes, L.H., Siemienski, J., and Braude, A.I. (1964). The syndrome of the alcoholic rose gardener. *Annals of Internal Medicine*, **61**, 1139–41.

Kemp, H.B.S., Bedford, A.F., and Fincham, W.J. (1982). *Petriellidium boydii* infection of the knee: a case report. *Skeletal Radiology*, **9**, 114–17.

Key, J.A. and Large, A.M. (1942). Histoplasmosis of the knee. *Journal of Bone and Joint Surgery*, **24**, 281–90.

Khan, M.I., Goss, G., Gotsman, A., and Asvat, M.S. (1983). Sporotrichosis arthritis: a case presentation and review of the literature. *South African Medical Journal*, **64**, 1099–1101.

Klein, B.S. *et al*. (1986). Isolation of *Blastomyces dermatitidis* in soil associated with a large outbreak of blastomycosis in Wisconsin. *New England Journal of Medicine*, **314**, 529–34.

Kramer, M. *et al*. (1990). Cyclosporine and itraconazole interaction in heart and lung transplant recipients. *Annals of Internal Medicine*, **113**, 327–8.

Kwon-Chung, K.J. and Bennett, J.E. (1992). Paracoccidioidomycosis. In *Medical mycology*, (ed. K.J. Kwon-Chung and J.E. Bennett), pp. 594–619. Lea and Febiger, Philadelphia.

Lafont, A. *et al*. (1994). *Candida albicans* spondylodiscitis and vertebral osteomyelitis in patients with intravenous heroin drug addiction. Report of 3 new cases. *Journal of Rheumatology*, **21**, 953–6.

Lambertus, M., Thordarson, D., and Goetz, M.B. (1988). Fungal prosthetic arthritis: presentation of two cases and review of the literature. *Review of Infectious Disease*, **10**, 1038–43.

Lantz, B., Selakovich, W.G., Collins, D.N., and Garvin, K.L. (1988). Coccidioidomycosis of the knee with 26-year follow-up examination. *Clinical Orthopedics*, **234**, 183–7.

Lesperance, M.L., Baumgartner, D., and Kauffman, C.A. (1988). Polyarticular arthritis due to *Sporothrix schenckii*. *Mycoses*, **31**, 599–603.

Londero, A.T. and Rambos, C.D. (1972). Paracoccidioidomycosis. A clinical and mycologic study of forty-one cases observed in Santa Maria, RS, Brazil. *American Journal of Medicine*, **52**, 771–5.

Lurie, H.I. (1963). Five unusual cases of sporotrichosis from South Africa showing lesions in muscles, bone and viscera. *British Journal of Surgery*, **50**, 585–91.

MacKenzie, T.R., Perry, C., Pearson, R., and Gilula, L.A. (1988). Magnetic resonance imaging in patients with inflammatory arthritis of the knee. *Orthopaedics Review*, **17**, 709–19.

Marina, N., Flynn, P., Rivera, G., and Hughes, W. (1991). *Candida tropicalis* and *Candida albicans* fungemia in children with leukemia. *Cancer*, **68**, 594–9.

McDonald, P.B., Black, G.B., and MacKenzie, R. (1990). Orthopaedic manifestations of blastomycosis. *Journal of Bone and Joint Surgery*, **72A**, 860–4.

McGinnis, M.R. and Fader, R.C. (1988). Mycetoma: A contemporary concept. *Infectious Disease Clinics of North America*, **1**, 938–54.

McKinsey, D.S., Gupta, M.R., Riddler, S.A., Driks, M., Smith, D.L., and Kurtin, P.J. (1989). Long-term amphotericin B therapy for disseminated histoplasmosis in patients with the acquired immunodeficiency syndrome (AIDS). *Annals of Internal Medicine*, **111**, 655–9.

McManus, E.J. and Jones, J.M. (1986). The use of ketoconazole in the treatment of blastomycosis. *American Review of Respiratory Disease*, **133**, 141–3.

Minamoto, G. and Armstrong, D. (1988). Fungal infection in AIDS. Histoplasmosis and coccidioidomycosis. *Infectious Disease Clinics of North America*, **2**, 447–56.

Molstad, B. and Strom, R. (1978). Multiarticular sporotrichosis. *Journal of the American Medical Association*, **240**, 556–7.

Mostaza, J.M., Barabao, F.J., Fernandez-Martin, J., Pena-Yandez, J., and Vasquez-Rodriguez, J.J. (1989). Cuneoarticular mucormycosis due to *Cunninghamella bertholletiae* in a patient with AIDS. *Reviews of Infectious Disease*, **11**, 316–18.

Munoz-Fernandez, S. *et al*. (1991). Rheumatic manifestations in 556 patients with human immunodeficiency virus infection. *Seminars in Arthritis and Rheumatism*, **21**, 30–9.

Murray, H.W., Littman, M.L., and Roberts, R.B. (1974). Disseminated paracoccidioidomycosis (South American blastomycosis) in the United States. *American Journal of Medicine*, **56**, 209–20.

National Institute of Allergy and Infectious Disease Mycoses Study Group (1985). Treatment of blastomycosis and histoplasmosis with ketoconazole. *Annals of Internal Medicine*, **103**, 861–72.

Nomura, J. and Ruskin, J. (1994). The prosthetic joint and disseminated coccidioidomycosis [abstract 32]. Centennial Conference on Coccidioidomycosis, Stanford, California,

O'Meeghan, T., Varcoe, R., Thomas, M., and Ellis-Preger, R. (1990). Fluconazole concentration in joint fluid during successful treatment of *Candida albicans* arthritis. *Journal of Antimicrobial Chemotherapy*, **26**, 601–2.

Omer, G.E., Jr., Lockwood, R.S., and Travis, L.O. (1963). Histoplasmosis involving the carpal joint. A case report. *Journal of Bone and Joint Surgery*, **45**, 1699–703.

Paya, C.V., Roberts, G.D., and Cockerill, F.R., III (1987). Laboratory methods for the diagnosis of disseminated histoplasmosis: clinical importance of the lysis-centrifugation blood culture technique. *Mayo Clinic Proceedings*, **62**, 480–5.

Pearson, G.J., Chin, T.W., and Fong, I.W. (1992). Case report: treatment of blastomycosis with fluconazole. *American Journal of Medical Science*, **303**, 313–15.

Perfect, J.R., Durack, D.T., and Gallis, H.A. (1983). Cryptococcemia. *Medicine*, **62**, 98–109.

Pluss, J.L. and Opal, S.M. (1986). Pulmonary sporotrichosis: review of treatment and outcome. *Medicine*, **65**, 143–53.

Pope, T.L., Jr. (1982). Pediatric *Candida albicans* arthritis: case report of hip involvement with a review of the literature. *Progress in Pediatric Surgery*, **15**, 271–83.

Recht, A.D., Davies, S.F., and Eckman, M.R. (1982). Blastomycosis in immunosuppressed patients. *American Review of Respiratory Disease*, **125**, 359–62.

Restrepo, A. *et al*. (1976). The gamut of paracoccidioidomycosis. *American Journal of Medicine*, **61**, 33–42.

Restrepo, A., Robledo, J., Gomez, I., Tabares, A.M., and Gutierrez, R. (1986). Itraconazole therapy in lymphangitic and cutaneous sporotrichosis. *Archives of Dermatology*, **122**, 413–17.

Ricciardi, D.D., Sepkowitz, D.V., Berkowitz, L.B., Bienenstock, H., and Maslow, M. (1986). Cryptococcal arthritis in a patient with acquired immune deficiency syndrome. Case report and review of the literature. *Journal of Rheumatology*, **13**, 455–8.

Rippon, J.W. (1988). In *Pseudoallescheriasis*, (ed. M. Wonsiewicz), pp. 651–80. W.B. Saunders, Philadelphia.

Robert, M.E. and Kauffman, C.A. (1988). Blastomycosis presenting as polyarticular septic arthritis. *Journal of Rheumatology*, **15**, 1138–42.

Roncoroni, A.J. and Smayevsky, J. (1988). Arthritis and endocarditis from *Exiophiala jeanselmei* infection. *Annals of Internal Medicine*, **108**, 773.

Rosenthal, J., Brandt, K.D., Wheat, J.L., and Slama, T.G. (1983). Rheumatologic manifestations of histoplasmosis in the recent Indianapolis epidemic. *Arthritis and Rheumatism*, **26**, 1065–70.

Salzman, S.H., Smith, R.L., and Aranda, C.P. (1988). Histoplasmosis in patients at risk for the acquired immunodeficiency syndrome in a nonendemic setting. *Chest*, **93**, 916–21.

Sarosi, G.A. and Davies, S.F. (1981). The clinical spectrum of blastomycosis. *Internal Medicine*, **2**, 64–70.

Sarosi, G.A, Hammerman, K.J., Tosh, F.E., and Kronenberg, R.S. (1974). Clinical features of acute pulmonary blastomycosis. *New England Journal of Medicine*, **290**, 540–3.

Schwartz, D.A. (1989). Sporothrix tenosynovitis—differential diagnosis of granulomatous inflammatory disease of the joints. *Journal of Rheumatology*, **16**, 550–3.

Schwarz, J. (1984). What's new in mycotic bone and joint diseases. *Pathology Research Practice*, **178**, 617–34.

Sellers, T.F., Price, W.N., and Newberry, W.M., Jr. (1965). An epidemic of erythema multiforme and erythema nodosum caused by histoplasmosis. *Annals of Internal Medicine*, **62**, 1244–62.

Sharkey, P.K. *et al*. (1990). Itraconazole treatment of phaeohyphomycosis. *Journal of the American Academy of Dermatology*, **23**, 577–86.

Sharkey-Mathis, P. *et al*. (1993). Treatment of sporotrichosis with itraconazole. National Institute of Allergy and Infectious Disease Mycosis Study Group. *American Journal of Medicine*, **95**, 279–85.

Silveira, L.H., Seleznick, M.J., Jara, L.J., Martinez-Osuna, P., and Espinoza, L. R. (1991). Musculoskeletal manifestations of human immunodeficiency virus infection. *Journal of Intensive Care Medicine*, **6**, 601–2.

Sinnott, J.T., IV. and Holt, D.A. (1982). Cryptococcal pyarthrosis complicating gouty arthritis. *Southern Medical Journal*, **82**, 1555–6.

Stead, K.J., Klugman, K.P., Painter, M.L., and Koornhof, H.J. (1988). Septic arthritis due to *Cryptococcus neoformans*. *Journal of Infection*, **17**, 139–45.

Stevens, D.A. and Vo, P.T. (1982). Synergistic interaction of trimethoprim and sulfamethoxazole on *Paracoccidioides brasiliensis*. *Antimicrobial Agents and Chemotherapy*, **21**, 852–4.

Stratton, C.W., Lichtenstein, K.A., Lowenstein, S.R., Phelps, D.B., and Reller, L.B. (1981). Granulomatous tenosynovitis and carpal tunnel syndrome caused by *Sporothrix schenckii*. *American Journal of Medicine*, **71**, 161–4.

Sugar, A.M. (1988). Paracoccioidomycosis. *Infectious Disease Clinics of North America*, **1**, 913–24.

Sugar, A.M., Restrepo, A.A., and Stevens, D.A. (1984). Paracoccidioidomycosis in the immunosuppressed host: report of a case and review of the literature. *American Review of Respiratory Disease*, **129**, 340–42.

Sugar, A.M., Saunders, C., and Diamond, R.D. (1990). Successful treatment of *Candida* osteomyelitis with fluconazole. A noncomparative study of two patients. *Microbiology and Infectious Disease*, **13**, 517–20.

Svirsky-Fein, S., Langer, L., Mibauer, B., Khermosh, O., and Rubinstein, E. (1979). Neonatal osteomyelitis caused by *Candida tropicalis*. Report of two cases and review of the literature. *Journal of Bone and Joint Surgery*, **61A**, 455–9.

Szombathy, S.P., Chez, M.G., and Laxer, R.M. (1988). Acute septic arthritis due to *Acremonium*. *Journal of Rheumatology*, **15**, 714–15.

Tack, K.J., Phame, F.S., Brown, B., and Thompson, R.C., Jr. (1982). *Aspergillus* osteomyelitis: report of four cases and review of the literature. *American Journal of Medicine*, **73**, 295–300.

Taillan, B. *et al*. (1992). Favourable outcome of blastomycosis of the brain stem with fluconazole and flucytosine treatment. *Annals of Medicine*, **24**, 71–2.

Tucker, R., Williams, P., Arathoon, E., and Stevens, D. (1988). Treatment of mycoses with itraconazole. *Annals of the New York Academy of Sciences*, **544**, 451–71.

Tucker, R.M., Denning, D.W., Arathoon, E.G., Rinaldi, M.G., and Stevens, D.A. (1990). Itraconazole therapy for nonmeningeal coccidioidomycosis: clinical and laboratory observations. *Journal of the American Academy of Dermatology*, **23**, 593–601.

Van Der Schee, A.C., Dinkla, B.A., and Festen, J.J.M. (1990). Gonarthritis as only manifestation of chronic disseminated histoplasmosis. *Clinical Rheumatology*, **9**, 92–4.

Wheat, L.J. *et al*. (1982). The diagnostic laboratory tests for histoplasmosis: analysis of experience in a large urban outbreak. *Annals of Internal Medicine*, **97**, 680–5.

Wheat, L.J., Kohler, R.B. and Tewari, R.P. (1986). Diagnosis of disseminated histoplasmosis by detection of *Histoplasmosis capsulatum* antigen in serum and urine specimens. *New England Journal of Medicine*, **314**, 83–8.

Wheat, L.J. *et al*. (1990). Disseminated histoplasmosis in the acquired immune deficiency syndrome: clinical findings, diagnosis and treatment, and review of the literature. *Medicine (Baltimore)*, **69**, 361–74.

Winn, R.E. (1988). Sporotrichosis. *Infectious Disease Clinics of North America*, **1**, 899–911.

Winn, R., Anderson, J., Piper, J., Aronson, N., and Pluss, J. (1993). Systemic sporotrichosis treated with itraconazole. *Clinical Infectious Disease*, **17**, 210–17.

Winter, W.J., Jr., Larson, R.K., Honeggar, M.M., Jacobsen, D.T., Pappagianis, D., and Huntington, R.W. (1975). Coccidioidal arthritis and its treatment. *Journal of Bone and Joint Surgery*, **57A**, 1152–7.

Witorsch, P. and Utz, J.P. (1968). North American blastomycosis: a study of 40 patients. *Medicine*, **47**, 169–200.

Yagi, K.I., Abbas, K., and Prabhu, S.R. (1983). Temperomandibular joint ankylosis due to maduromyecetoma caused by *Madurella mycetomi*. *Journal of Oral Pathology and Medicine*, **38**, 71–3.

Yao, J., Penn, R.G., and Ray, S. (1986). Articular sporotrichosis. *Clinical Orthopedics*, **204**, 207–14.

Younkin, S., Evarts, C.M., and Steigbigel, R.T. (1984). *Candida parapsilosis* infection of a total hip-joint replacement: successful reimplantation after treatment with amphotericin B and 5-fluorocytosine. A case report. *Journal of Bone and Joint Surgery*, **66A**, 142–3.

Yousefzadeh, D.K. and Jackson, J.H. (1980). Neonatal and infantile candidal arthritis with or without osteomyelitis: a clinical and radiographical review of 21 cases. *Skeletal Radiology*, **5**, 77–90.

5.3.11 Immunodeficiency

A. D. B. Webster

Introduction

Joint disease is a relatively common complication of immunodeficiency, and in the majority of cases it can be shown to be due to infection. Immunodeficiency is classified into 'primary' and 'secondary', most cases in the former category being due to inherited single or multiple gene defects, while the common forms of secondary immunodeficiency are associated with lymphoreticular neoplasia, particularly chronic lymphatic leukaemia and myeloma, or infection with the human immunodeficiency virus. It is useful to subclassify patients into those that have a predominantly antibody or cell-mediated immune deficiency, and those with severe combined immunodeficiency (see Table 1).

Table 1 Classification of immunodeficiency

Antibody deficiency

The principal diseases of the joint that may be associated with primary hypogammaglobulinaemia are shown in Table 2.

Disease	Comments
Mycoplasma arthritis	Mainly large joints but any joint can be affected
Monoarthritis of knee	Mainly in children, self-limiting
Tenosynovitis/arthralgia	Usually affects hands and feet; rapid response to immunoglobulin therapy
Echovirus disease	Flexion contractures of elbows and knees
Rheumatoid arthritis	Drugs used to treat rheumatoid arthritis may cause hypogammaglobulinaemia

Table 2 Joint disease in primary hypogammaglobulinaemia

Mycoplasma arthritis

Mycoplasmas are prokaryotic organisms which frequently infect mammals but which also occur in fish and reptiles (Taylor-Robinson 1990). There are many different types in humans, the majority colonizing mucosal surfaces and, in general, behaving as commensals. *Mycoplasma pneumoniae* is an exception, being a recognized pathogen that infects the respiratory tract causing pneumonia. Ureaplasmas, so called because they metabolize urea to ammonia, account for about a third of cases of 'non-specific' urethritis in otherwise healthy men, although the infection is usually self-limiting (Taylor-Robinson and McCormack 1980).

Mycoplasmas are normally found on mucosal surfaces and do not penetrate epithelial cells, although they may be taken up by phagocytes and remain viable within such cells (see below). When present in large numbers they do cause inflammation, possibly through direct activation of complement. Different strains have become adapted to different systems within the body: for instance ureaplasmas are mainly found in the urinary tract.

The first descriptions of mycoplasmas in the joints of patients with hypogammaglobulinaemia were published in 1978 (Stuckey *et al*. 1978; Webster *et al*. 1978), but even now there are some sceptics who suggest that the organisms may not cause the arthritis, but are merely 'passengers' in an immunocompromised host. This view stems from the early claim that patients with hypogammaglobulinaemia are prone to rheumatoid arthritis, as well as other autoimmune rheumatic disorders (Good *et al*. 1957). Most of the investigators involved in these early studies now concede that they were probably dealing with an infective arthritis. Surprisingly, mycoplasma arthritis has only been described so far in the primary hypogammaglobulinaemias (e.g. X-linked agammaglobulinaemia and 'common variable' immunodeficiency; see 'rheumatoid arthritis' below), and not in the secondary types associated with chronic lymphatic leukaemia and myeloma. Although cases may have been missed, the severity of the antibody deficiency may not be enough to predispose to systemic mycoplasma infection in the secondary immunodeficiencies. There are anecdotal reports of mycoplasma septicaemia and/or arthritis in patients on immunosuppressive drugs or following severe trauma, but, surprisingly, most of these patients were not investigated for immunoglobulin deficiency. Although there are claims that mycoplasmas may act as cofactors in the mechanism of the immune deficiency in acquired immune deficiency syndrome (**AIDS**), joint disease has not been described, although chronic mycoplasma infection of the kidneys may contribute to the characteristic nephropathy seen in AIDS (Bauer *et al*. 1991).

The evidence that mycoplasmas are a common cause of arthritis in hypogammaglobulinaemia can be summarized as follows. The organisms can consistently be

isolated in high titre from the joint during the active phase; the symptoms usually improve rapidly when antibiotics are given to which the organism is sensitive, and chronic arthritis with joint destruction occurs when antibiotic-resistant organisms are involved. Finally, some patients with severe, generalized infections develop discharging sinuses, often communicating with joints, from which mycoplasmas can be isolated (Taylor-Robinson *et al*. 1985).

Origin of infection

Ureaplasmas are commonly present in the vagina of pregnant women, and it is thought that most neonates become colonized in the respiratory tract shortly after birth, although the organisms then disappear through mechanisms that are not understood (Taylor-Robinson and McCormack 1980). In later life, colonization of the urinary tract may occur from sexual intercourse, and since about a third of healthy women are persistently colonized, most men are exposed to these organisms. Colonization is usually transient and asymptomatic, although a minority of men will develop urethritis. In contrast, there is a high incidence of symptomatic ureaplasma urethritis in both men and women with hypogammaglobulinaemia, which may progress to a chronic cystitis and occasionally pyelonephritis (Webster *et al*. 1981). Fibrosis of the bladder wall with contraction may be the end result of chronic infection. Many patients with hypogammaglobulinaemia who develop ureaplasma arthritis have previously suffered from an episode of urethritis, so it is likely that the origin of the infection is in the urinary tract. Trauma to a joint, which may be minor, often triggers the septic arthritis, suggesting that the organisms are present in small numbers in the circulation and become established at an inflammatory focus within the joint. Once one joint has been infected, there is a tendency for other joints to follow if the infection is not eliminated with appropriate antibiotics. Roifman *et al*. (1986) have suggested that ureaplasmas may cause chronic bronchitis in antibody deficient patients, and although this issue is controversial, the lungs may be a source of systemic infection.

Apart from *M. pneumoniae*, which is acquired by droplet inoculation from an infected person, two other mycoplasmas have been cultured from the inflamed joints of patients with hypogammaglobulinaemia. *M. hominis*, like ureaplasmas, is a commensal in the urinary tract, and is usually acquired through sexual contact. The 10 per cent of 'healthy' women with bacterial vaginosis have particularly heavy colonization of their vaginas with *M. hominis* (Hillier and Holmes 1990); antibody deficient women with this problem may have an increased risk of *M. hominis* arthritis. Women presenting with *M. hominis* arthritis should therefore be tested for bacterial vaginosis and if necessary be given metronidazole to eliminate the bacteria; this should reduce the risk of further systemic spread of the mycoplasma infection. We have seen dual infection in a joint with *M.hominis* and *Ureaplasma urealyticum* in one patient. *M. salivarium* has also been isolated from the joints of a few patients, presumably originating from the upper respiratory tract where it is found in the saliva of about 80 per cent of healthy individuals (So *et al*. 1982). This organism is unequivocally regarded as a commensal and is a good example of commensal overgrowth leading to disease in an immunocompromised host.

Clinical features

Large joints are usually affected, particularly the knees, although the ankles, hips, shoulders, and wrists (including carpal bones) are frequently involved in persistent infections. The fingers and toes are rarely affected, and when this does occur it usually involves a single interphalangeal joint. The initial symptoms are swelling and stiffness of the affected joint, usually with an obvious effusion when the knee is involved. Nodules may occur on the elbows that have the same histological features as classical rheumatoid nodules. Systemic symptoms are rare and there is usually no fever or blood leucocytosis. Joint pain increases over a few weeks or months, and if the infection persists the synovium will gradually deteriorate through chronic inflammation, leading eventually to fibrosis and fixation of the joint (Fig. 1). This sequence usually takes months, or sometimes years, to reach a conclusion if inappropriate or no treatment is offered.

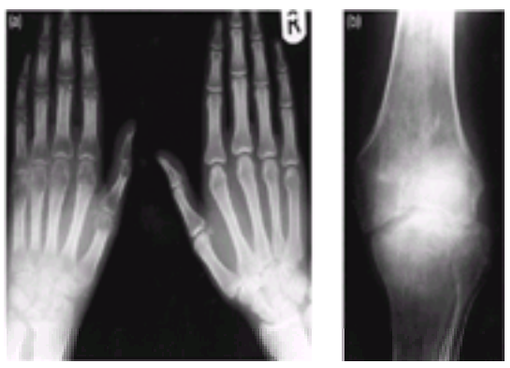


Fig. 1 (a) Chronic infection with *Ureaplasma urealyticum* of the wrist in a 30-year-old man with X-linked agammaglobulinaemia, showing destruction of the left distal ulna, the medial radius, and lunate, with osteoporosis of all joint levels related to disuse. (b) The left knee of the same patient showing destruction of the cartilage, obliteration of the joint space, and disorganization of the joint.

Some patients may enter a severe phase, presumably reflecting high levels of circulating organisms. Subcutaneous abscesses may then occur, sometimes at sites of minor trauma (e.g. injection sites), but more often adjacent to joints; the skin then breaks down leaving a chronically discharging sinus in communication with the joint space. Even at this relatively late stage, the patient may show no systemic effects of chronic infection, apart from being immobilized and in considerable pain.

Diagnosis

Organisms can be cultured from the synovial fluid, provided that appropriate techniques are used, although cultures may be negative if the patient has recently been treated with antibiotics, particularly erythromycin and tetracyclines. Unfortunately, there are only a few laboratories in the world that are in a position routinely to culture mycoplasmas, despite the fact that the techniques required are relatively straightforward. It should be be remembered that the diagnosis of *M. pneumoniae* infection in immunocompetent patients is usually made retrospectively by positive serology, which is obviously inappropriate in patients with hypogammaglobulinaemia. There are some routine laboratories that will culture for *M. pneumoniae*, although these organisms are slow growing and may take up to 2 weeks to show a positive result. Ureaplasmas, *M. hominis*, and *M. salivarium* need different media, and a positive result can be obtained in 48 h. Organisms can also be isolated from the pus of discharging sinuses, and from synovial tissue removed at biopsy.

Synovial fluid should always be sent for the routine culture of common pathogens, such as staphylococci and *Haemophilus influenzae*, which may rarely cause arthritis in patients with hyogammaglobulinaemia. Microscopic examination of the fluid is not very helpful unless it is clear with a predominant lymphocytosis, in which case mycoplasma arthritis is unlikely. The fluid is usually yellow and/or turbid, and contains many neutrophils. If the routine culture is negative, then the working diagnosis should be mycoplasma arthritis until proved otherwise. Occasionally, there may be confusion between mycoplasma arthritis and other types of arthropathy (see below) that have a raised incidence in patients with hypogammaglobulinaemia. In particular, rheumatoid arthritis in its early stages can be confused with mycoplasma infection, although multiple involvement of finger joints is very much against the latter.

Management

Patients should be given doxycycline as soon as synovial fluid has been aspirated. We use doxycycline intravenously (Pfizer) at a dose of 200 mg at once, followed by 100 mg/day until there is improvement in joint swelling and the level of serum C-reactive protein has become normal, which is a useful marker of disease activity. Treatment should then continue with oral doxycycline at 100 mg/day for at least 3 months. Fortunately, most mycoplasma strains are sensitive to doxycycline and there will be obvious improvement in the joint symptoms within a few days. However, it is useful to arrange sensitivity tests against a range of antibiotics as soon as possible, because occasionally the organism may be resistant. Sensitivity tests should include erythromycin, streptomycin, doxycycline, kanamycin, clindamycin, spectinomycin, netilmicin, ciprofloxacin, and azithromycin, the last being a new macrolide that accumulates in phagocytes, a property which may be of particular advantage in eradicating mycoplasmas. In our experience, patients who are infected with a doxycycline-resistant strain are very difficult to manage and usually progress to multi-joint involvement despite various combinations of other antibiotics. However, it is important to test regularly the antibiotic sensitivities of new isolates from joint aspirates in order to keep pace with any changes. Although this can be time consuming for the laboratory, there is a good chance that eventually the infection can be eradicated.

All antibiotics so far tested are static, and consequently long-term therapy is required (Escalante *et al*. 1985). Surgical interference of the joint should be kept to a minimum; arthroscopy and washing out the joint with saline is unhelpful, although drainage under vacuum of a very tense effusion for 24 h may rarely be necessary.

Trauma from surgical interference appears to increase the growth of mycoplasmas, as well as increasing the risk of sinus formation. The joints should be immobilized until there is no longer any swelling, and then if there is damage to the ligaments, the joint must be stabilized by splinting until there is full recovery. Provided the diagnosis is made early and the mycoplasmas are sensitive to tetracyclines, there is usually full recovery.

Patients who steadily deteriorate in spite of treatment with a wide range of antibiotics should be considered for therapy with hyperimmune serum. Unfortunately, regular treatment with intravenous human immunoglobulin has little effect on eradicating mycoplasma infection, because the specific antibody levels are so low in these preparations. However, intravenous human immunoglobulin may have some protective effect against infection. In our series of 18 patients with primary immunodeficiency and mycoplasma arthritis, most presented before treatment with intravenous human immunoglobulin, and in those who developed infection after treatment their trough levels of serum IgG were below 5 g/l (Franz *et al*. 1997). It is therefore reasonable to treat infected patients with high-dose, intravenous, human immunoglobulin (i.e. 400 mg/kg per week for 8 weeks), and as a prophylactic measure it is our policy to keep serum IgG levels above 7 g/l.

We have successfully treated two patients with goat serum taken from animals hyperimmunized with the patient's particular strain (Taylor-Robinson *et al*. 1985). Fortunately, serum sickness does not occur in patients with hypogammaglobulinaemia, who can tolerate repeated infusions of animal serum. The amount of serum given is somewhat arbitrary, and there are no standard immunization schedules for goats. We have raised titres in excess of 1:5000 by first subcutaneously injecting a mycoplasma concentrate in Freund's complete adjuvant, followed by intravenous injections of the concentrate alone at 4-week intervals. It may be better to immunize rabbits using published schedules that are known to produce very high titres of antisera (i.e. $1:10^{5}$), which might compensate for relatively small amounts of serum obtainable.

Monoclonal antibodies are not yet available, and anyway may have to be raised specifically against the infecting organism because of considerable strain variation in exposed antigenic epitopes. Nevertheless, this is worth considering early in a patient who is difficult to treat, as it will take many months to raise enough antibody.

Role of antibodies in protection against mycoplasmas

The clinical observations on patients with hypogammaglobulinaemia clearly demonstrate that antibodies are important in protection against mycoplasma arthritis. *In vitro* experiments have shown that the growth of mycoplasmas is readily inhibited by specific antibody, and that, even in the absence of antibody, mycoplasmas will activate the first component of complement. In turn this will activate the complement cascade and split C3, generating opsonic complement, which enables the organisms to be taken up by neutrophils (Webster *et al*. 1988). However, once within the neutrophil phagolysosome the organisms remain viable, and it is likely that neutrophils transport organisms from mucosae to the joints. The factors that encourage mycoplasma growth within the joints are unknown. It is likely, therefore, that the main function of specific antibodies against mycoplasmas is to control their growth on the mucosal surface. Massive overgrowth of organisms occurs in the absence of antibody. The result is inflammation, with influx of neutrophils and macrophages, followed by phagocytosis of large numbers of mycoplasmas which initiate chronic infection in a susceptible joint.

Bacterial septic arthritis

Patients with hypogammaglobulinaemia are prone to septicaemia and pneumonia due to pneumococci, non-typeable *H. influenzae*, and staphylococci, any of which can occasionally cause septic arthritis (Asherson and Webster 1980). However, in contrast to mycoplasma arthritis, bacterial arthritis is usually more acute and painful. Bacterial arthritis is extremely rare in patients already established on replacement immunoglobulin therapy, whereas mycoplasma infection still occurs.

Rheumatoid arthritis

Early reports suggested that there was a high incidence of rheumatoid arthritis in patients with hypogammaglobulinaemia, including those with X-linked agammaglobulinaemia (Good *et al*. 1957). However, it now seems likely that many of these patients suffered from mycoplasma arthritis. Nevertheless, rheumatoid arthritis does occur in patients with 'common variable' immunodeficiency, at a frequency of about 2 per cent in our series. It is often difficult to ascertain whether the arthritis precedes or follows the onset of hypogammaglobulinaemia, and there is the added complication that a number of drugs used to treat rheumatoid arthritis (see below) are known to 'cause' or 'trigger' hypogammaglobulinaemia (So *et al*. 1984). Nevertheless, there are patients whose hypogammaglobulinaemia clearly precedes the onset of classical rheumatoid arthritis, with the typical involvement of small joints of the hand, rheumatoid nodules, and gradual destruction of small and large joints. These patients are seronegative for rheumatoid factor and other relevant autoantibodies. This in itself is interesting, because it demonstrates that severe rheumatoid arthritis can occur and progress in the absence of autoantibodies, and supports the view that the disease is driven by cellular interactions.

Drug-associated hypogammaglobulinaemia

Gold, penicillamine, sulphasalazine, and phenytoin have been associated with hypogammaglobulinaemia, and partly explain the apparent raised incidence of rheumatoid arthritis and Still's disease in patients with selective IgA deficiency and other types of hypogammaglobulinaemia (Barclay *et al*. 1979). The immunoglobulin levels usually return to normal after withdrawal of the drug, although this may take many months or even years. The mechanism is unclear, and may be different for the four drugs. In the case of gold, there is evidence that the drug inhibits T-cell proliferation *in vitro* at concentrations likely to be found in tissues (Lipsky and Ziff 1977), but there is no evidence that it affects B-cell differentiation and immunoglobulin synthesis *in vitro*, even when cells are used from patients who have recovered from drug-associated hypogammaglobulinaemia (S. Sukram and A. D. B. Webster, unpublished). One possibility is that the drugs induce 'common variable' immunodeficiency in patients who have genetic predisposing factors for this disease. In this context, it is not yet known whether drug-associated hypogammaglobulinaemia is linked to the same susceptibility genes for common variable immunodeficiency which are located in the class II and III region on chromosome 6 (Schaffer *et al*. 1989; Howe *et al*. 1991).

Miscellaneous arthritides associated with hypogammaglobulinaemia

Chronic arthritis of the knee

Children with hypogammaglobulinaemia, particularly those with X-linked agammaglobulinaemia, are prone to a chronic insidious arthritis of the knee, which is usually unilateral. There is a chronic and relatively painless effusion, which may persist for many years, sometimes with hypertrophy of the epiphyseal cartilage (Fig. 2). The condition usually recovers spontaneously after a few months or years, leaving very little permanent damage apart from the unsightly appearance of enlarged condyles. Some damage to the cartilage may necessitate menisectomy in later years.

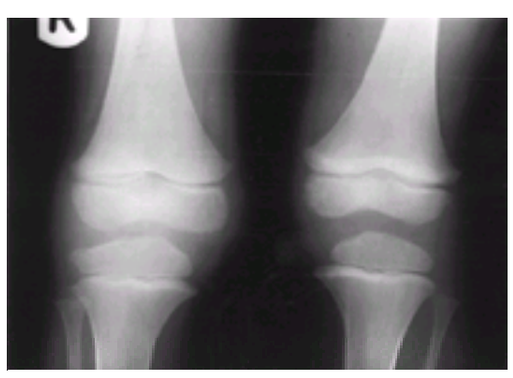


Fig. 2 Epiphyseal hypertrophy in the right knee of a 3-year-old boy with X-linked agammaglobulinaemia. There was a sterile effusion; the knee subsequently improved spontaneously.

The synovial fluid is colourless and contains mainly lymphocytes; it is sterile when cultured for bacteria, viruses, and mycoplasmas. Nevertheless, there is still a possibility that this is a low-grade mycoplasma infection, although antimycoplasma drugs such as doxycycline or erythromycin have no effect. Anti-inflammatory drugs

are often useful.

Our impression is that this complication is much less common since the introduction of intravenous immunoglobulin therapy in children with hypogammaglobulinaemia, and in one of our cases the arthritis improved after such therapy.

Chronic tenosynovitis/arthritis

This condition was relatively common in British patients with hypogammaglobulinaemia in the 1970s but is now very rare (Webster *et al*. 1976). It usually occurred in adult patients with 'common variable' immunodeficiency as one of the presenting features of the disease. The usual pattern was of a patient suffering from recurrent respiratory infections for many years, who then developed stiffness in the wrists, elbows, and ankles, with swelling of the tendon sheaths on the dorsal aspects of the wrist, hands, and feet. This sometimes led to tendon tethering with cystic swellings that required surgery. The arthritis was usually mild and fluctuating.

The condition was frequently confused with rheumatoid arthritis, but there were no bony erosions. Furthermore, the condition dramatically improved with intramuscular immunoglobulin therapy, suggesting an infectious cause, perhaps by a virus that was neutralized by antibodies in the immunoglobulin therapy. These cases are very rare nowadays in the United Kingdom, although some patients still suffer from mild, fluctuating arthropathy while on intramuscular or intravenous immunoglobulin therapy, which may represent a mild form of the same condition. One of our patients developed tenosynovitis and arthropathy while on intramuscular therapy, but rapidly improved when given intravenous therapy. Finally, it is possible that this condition is caused by low-level mycoplasma infection, as one of our patients initially responded to intravenous immunoglobulin and then relapsed with a confirmed mycoplasma infection which then responded to doxycycline.

Echovirus disease

Patients with severe primary antibody deficiency are prone to chronic echovirus infection of the central nervous system and muscles, and there are reports of arthritis in a few patients (McKinney *et al*. 1987). This is a 'slow' virus disease that mainly affects small vessels in the meninges, subcutaneous tissues, and muscles, leading in the muscles to fibrosis with a 'woody' sensation on palpation. The arms and legs are predominantly affected, and the muscle fibrosis can lead to flexion contractures at the knee and elbows, sometimes producing a characteristic stooped posture (Webster 1984). An erythematous rash may occur transiently, which, together with the subcutaneous oedema, was responsible for the early references to a 'dermatomyositis-like syndrome'. This was thought to provide further evidence that patients with hypogammaglobulinaemia had a predisposition to 'connective tissue disorders'.

The virus can usually be cultured from the cerebrospinal fluid, but occasionally cultures are repeatedly negative and molecular techniques are required to identify viral RNA (Webster *et al*. 1993). The treatment is unsatisfactory, although regular intravenous infusions with high titre, hyperimmune plasma can stabilize the disease, and in a few cases it may have eradicated the infection. Regular injection directly into the cerebrospinal fluid via an Omaya reservoir of pooled immunoglobulin, suitable for intravenous use and preferably with a reasonable titre to the relevant echovirus strain, has been associated with remission in a few patients (McKinney *et al*. 1987). However, most patients gradually deteriorate, sometimes with episodes of partial remission, and ultimately die from involvement of a critical centre in the brain.

Chronic echovirus disease is much less common in the United Kingdom than it was in the 1970s, probably reflecting the increased use of intravenous immunoglobulin during the 1980s. Because each batch of intravenous immunoglobulin is made from a pool of about 20 000 donors, it is likely that significant amounts of antibody to most enteroviral strains are present in these preparations. It is therefore important to maintain serum IgG levels above 7 g/l, particularly in young children with immunodeficiency who may be more susceptible to enteroviruses. This may be difficult to achieve in infants with poor venous access for regular intravenous immunoglobulin therapy. Regular subcutaneous infusions are becoming more popular in this age group but regular monitoring is needed to ensure an adequate serum level.

Defects in cellular immunity

There are a variety of primary defects in cellular immunity, ranging from highly selective T-cell defects to severe combined immunodeficiency. The survival of patients suffering from the latter depends on a successful bone marrow graft (Morgan *et al*. 1987). The best example of a secondary defect in cellular immunity is AIDS, which clinically closely resembles severe combined immunodeficiency. Although patients with selective T-cell defects are not prone to septic or other types of arthritis, those with severe combined immunodeficiency are prone to a wide range of infections, some leading to septic arthritis.

Complement deficiencies

Septic arthritis following pneumococcal, meningococcal, or gonococcal infection may occur in patients with genetically determined homozygous complement deficiencies (Rother 1986). Children with homozygous C2 deficiency, which occurs in about 1 in 10 000 of the population, are prone to pneumococcal and *H. influenzae* septicaemia, which may be associated with septic arthritis (see Chapter 5.3.2). However, some affected individuals are not prone to infection, presumably because they have other unknown host defence factors against these organisms. Once a patient has developed a septicaemia, it is usually appropriate to recommend regular, prophylactic, oral penicillin therapy, as well as advising the patient to take a broader-spectrum antibiotic (e.g. amoxicillin/clavulanic acid complex) in the event of illness. Patients with defects in other complement components, particularly the late components C7 and C8, are prone to neisserial septicaemia, and occasionally septic arthritis. Prophylactic penicillin may be appropriate for them.

Neutrophil defects

Chronic granulomatous disease (CGD) is one of the primary neutrophil defects described in which joints may be involved, although both primary and secondary neutropenia predisposes to septic arthritis and osteomyelitis, particularly due to staphylococci (White and Gallin 1986). In CGD affected patients are prone to chronic infection with catalase-positive organisms (e.g. salmonella, staphylococci, *Serratia marcescens*), usually causing chronic suppuration of lymph glands with draining sinuses. Staphylococcal and salmonellal osteomyelitis may sometimes occur, although involvement of the joints is very rare.

Chapter References

Asherson, G.L. and Webster, A.D.B. (1980). *Diagnosis and treatment of immunodeficiency diseases*. Blackwell, Oxford.

Barclay, D., Ansell, B., Howard, A., Hohermuth, H., and Webster, A.D.B. (1979). IgA deficiency in juvenile chronic polyarthritis. *Journal of Rheumatology*, **6**, 219–24.

Bauer, F.A., Wear, D.J., Angritt, P., and Lo, S.C. (1991). *Mycoplasma fermentans* (incognitus strain) infection in the kidneys of patients with acquired immunodeficiency syndrome and associated nephropathy: a light microscopic, immunohistochemical, and ultrastructural study. *Human Pathology*, **22**, 63–9.

Escalante, A., Aznar, J., De Miguel, C., and Perea, E.J. (1985). Activity of nine antimicrobial agents against *Mycoplasma hominis* and *Ureaplasma urealyticum*. *European Journal of Sexually Transmitted Diseases*, **2**, 85–7.

Franz, A., Webster, A.D.B., Furr, P.M., Taylor-Robinson, D. (1997). Mycoplasmal arthritis in patients with primary immunoglobulin deficiency: clinical features and outcome in 18 patients. *British Journal of Rheumatology* (in press).

Good, R.A., Rotstein, J., and Mazzitello, W.F. (1957). The simultaneous occurrence of rheumatoid arthritis and agammaglobulinaemia. *Journal of Laboratory and Clinical Medicine*, **49**, 343–57.

Hillier, S. and Holmes, K.K. (1990). Bacterial vaginosis. In *Sexually transmitted diseases*, (2nd edn) (ed. K.K. Holmes, P.-A. Mardh, P.F. Sparling, and P.J. Wiesner), pp. 547–59. McGraw Hill, New York.

Howe, H.S., So, A.K.L., Farrant, J., and Webster, A.D.B. (1991). Common variable immunodeficiency is associated with polymorphic markers in the human major histocompatibility complex. *Clinical and Experimental Immunology*, **83**, 387–90.

Lipsky, P.E. and Ziff, M. (1977). Inhibition of antigen- and mitogen-induced human lymphocyte proliferation by gold compounds. *Journal of Clinical Investigation*, **59**, 455–66.

McKinney, R.E., Jr, Katz, S.L., and Wilfert, C.M. (1987). Chronic enteroviral meningoencephalitis in agammaglobulinaemic patients. *Review of Infectious Diseases*, **9**, 334–56.

Morgan, G., Levinsky, R.J., Hugh-Jones, K., Fairbanks, L.D., Morris, G.S., and Simmonds, A. (1987). Heterogeneity of biochemical, clinical and immunological parameters in severe combined

immunodeficiency due to adenosine deaminase deficiency. *Clinical and Experimental Immunology*, **70**, 491–9.

Roifman, C.M., Rao, C.P., Lederman, H.M., Lavi, S., Quinn, P., and Gelfand, E.W. (1986). Increased susceptibility to *Mycoplasma* infection in patients with hypogammaglobulinaemia. *American Journal of Medicine*, **80**, 590–4.

Rother, K. (1986). Summary of reported deficiencies. In Hereditary and acquired complement deficiencies in animals and man, (ed. K. and U. Rother), *Progress in Allergy* series, pp. 202–11. Karger, Basle.

Schaffer, F.M., Palermos, J., Zhu, Z.B., Barger, B.O., Cooper, M.D., and Volanakis, J.E. (1989). Individuals with IgA deficiency and common variable immunodeficiency share polymorphisms of major histocompatibility complex class III genes. *Proceedings of the National Academy of Sciences* (*USA*), **86**, 8015–19.

So, A.K.L., Furr, P.M., Taylor-Robinson, D., and Webster, A.D.B. (1982). Arthritis caused by *Mycoplasma salivarium* in hypogammaglobulinaemia. *British Medical Journal*, **286**, 762–3.

So, A.K.L., Peskett, S.A., and Webster, A.D.B. (1984). Hypogammaglobulinaemia associated with gold therapy. *Annals of the Rheumatic Diseases*, **43**, 581–2.

Stuckey, M., Quinn, P.A., and Gelfand, E.W. (1978). Identification of *Ureaplasma urealyticum* (T-strain *Mycoplasma*) in patients with polyarthritis. *Lancet*, **i**, 198–201.

Taylor-Robinson, D. (1990). The mycoplasmatales: mycoplasma, ureaplasma, acholeplasma, spiroplasma and anaeroplasma. In *Topley and Wilson's principles of bacteriology*, (8th edn), Vol. 2, *Systematic bacteriology*, (ed. M.T. Parker and B.I. Duerden), pp. 663–81. Edward Arnold, London.

Taylor-Robinson, D. and McCormack, W.M. (1980). The genital mycoplasmas. *New England Journal of Medicine*, **302**, 1003 and 1063.

Taylor-Robinson, D., Furr, P.M., and Webster, A.D.B. (1985). *Ureaplasma urealyticum* causing persistent urethritis in a patient with hypogammaglobulinaemia. *Genitourinary Medicine*, **61**, 404–8.

Webster, A.D.B. (1984). Echovirus disease in hypogammaglobulinaemic patients. *Clinics in Rheumatology*, **10**, 189–203.

Webster, A.D.B., Loewi, G., Dourmashkin, R.D., Golding, D.N., Ward, D.J., and Asherson, G.L. (1976). Polyarthritis in adults with hypogammaglobulinaemia and its rapid response to immunoglobulin treatment. *British Medical Journal*, **1**, 1314–16.

Webster, A.D.B., Taylor-Robinson, D., Furr, P.M., and Asherson, G.L. (1978). Mycoplasmal (ureaplasma) septic arthritis in hypogammaglobulinaemia. *British Medical Journal*, **1**, 478–9.

Webster, A.D.B., Taylor-Robinson, D., Furr, P.M., and Asherson, G.L. (1981). Chronic cystitis and urethritis associated with ureaplasmal and mycoplasmal infection in primary hypogammaglobulinaemia. *British Journal of Urology*, **54**, 287–91.

Webster, A.D.B., Furr, P.M., Hughes-Jones, N.C., Gorick, B.D., and Taylor-Robinson, D. (1988). Critical dependence on antibody for defence against mycoplasmas. *Clinical and Experimental Immunology*, **71**, 383–7.

Webster, A.D.B., Rotbart, H.A., Warner, T., Rudge, P., and Hyman N. (1993). Diagnosis of enterovirus brain disease in hypogammaglobulinemic patients by polymerase chain reaction. *Clinical Infectious Diseases*, **17**, 657–61.

White, C.J. and Gallin, J.I. (1986). Phagocyte defects. *Clinical Immunology and Immunopathology*, **40**, 50–61.

WHO Scientific Group Report (1995). Primary immunodeficiency diseases. Report of a WHO Scientific Group. *Clinical and Experimental Immunology*, **99** (Suppl. 1).

5.3.12 Rheumatic fever

Allan Gibofsky and John B. Zabriskie

Introduction

Acute rheumatic fever is a delayed, non-suppurative sequel to a pharyngeal infection with the group A streptococcus. Following the initial streptococcal pharyngitis, there is a latent period of 2 to 3 weeks. The onset of disease is usually characterized by an acute febrile illness, which may show itself in one of three classical ways: (i) the patient may present with migratory arthritis predominantly involving the large joints; (ii) there may be concomitant clinical and laboratory signs of carditis and valvulitis; (iii) there may be involvement of the central nervous system, manifesting itself as Sydenham's chorea. The clinical episodes are self-limiting but damage to the valves may be chronic and progressive, resulting in cardiac decompensation and death.

Although there has been a dramatic decline in both the severity and fatality of the disease since the turn of the century, there are recent reports of its resurgence in the United States (Veasy *et al*. 1987) and in many military installations in the world, reminding us that it remains a public health problem even in developed countries. In addition, the disease continues essentially unabated in many of the developing countries: estimates suggest there will be 10 to 20 million new cases per year in those countries where two-thirds of the world population lives.

Epidemiology

The incidence of rheumatic fever actually began to decline long before the introduction of antibiotics into clinical practice, decreasing, for example, from 250 to 100 patients/100 000 population from 1862 to 1962 in Denmark (Gordis 1985). The introduction of antibiotics in 1950 caused a rapid acceleration of this decline, until in 1980 the incidence ranged from 0.23 to 1.88 patients per 100 000, primarily children and teenagers. A notable exception has been in the Hawaii and Maori populations (both of Polynesian ancestry), where the rate continues to be 13.4/100 000 children admitted to hospital per year (Pope 1989).

As reviewed by Markowitz (1987), only a few M-streptococcal serotypes (types 5, 14, 18, 24) have been implicated in outbreaks of rheumatic fever, suggesting there could be a particular 'rheumatogenic' potential of certain strains of group A streptococci. However, in Trinidad, types 41 and 11 have been the most common strains isolated from rheumatics. In our own series, gathered over 20 years (see Table 1), a large number of different M serotypes were isolated, including six strains that were not typable. Kaplan *et al*. (1989) found that several different M types were isolated from the patients seen during an outbreak of rheumatic fever, and that these strains were both mucoid and non-mucoid in character. Whether or not certain strains are more 'rheumatogenic' than others remains unresolved. What is true, however, is that a streptococcal strain capable of causing a well-documented pharyngitis is almost always potentially capable of causing rheumatic fever, although some notable exceptions have been recorded (reviewed by Whitnack and Bisno 1980).

M type	RHD	No RHD	Total
Non-typable	1	5	6
1	1	1	2
2	0	1	1
5	1	1	2
6	1	1	2
12	0	2	2
18	2	2	4
19	2	1	3
28	1	0	1
TOTAL	9	11	23

RHD, patients with rheumatic heart disease; no RHD, patients without rheumatic heart disease.

Table 1 Positive throat cultures—group A b-haemolytic streptococci (Rockefeller University Hospital, rheumatic fever patients; *n* = 87)

Pathogenesis

There is little evidence for the direct involvement of group A streptococci in the affected tissues of patients with acute rheumatic fever, but there is a large body of epidemiological and immunological evidence indirectly implicating the group A streptococcus in the initiation of the disease process. For example, (i) it is well known that outbreaks of rheumatic fever closely follow epidemics of either streptococcal sore throats or scarlet fever (Whitnack and Bisno 1980), (ii) adequate treatment of a documented streptococcal pharyngitis markedly reduces the incidence of subsequent rheumatic fever (Denny *et al*. 1950), (iii) approximate antimicrobial prophylaxis prevents recurrences of the disease in patients known to have had acute rheumatic fever (Markowitz and Gordis 1972), and (iv), if one tests the serum of the majority of patients with acute rheumatic fever for three antistreptococcal antibodies (streptolysin O, hyaluronidase, and streptokinase), the vast majority of samples (whether or not the patients recall an antecedent streptococcal sore throat) will have elevated antibody titres to these antigens (Stollerman *et al*. 1956).

A note of caution is necessary concerning the documentation (either clinical or microbiological) of an antecedent streptococcal infection. The rate of isolation of group A streptococci from the oropharynx is extremely low, even in populations who generally do not have access to microbial antibiotics. Further, there appears to be an age-related discrepancy in the clinical documentation of an antecedent sore throat. In older children and young adults, the recollection of a streptococcal sore throat approaches 70 per cent; in younger children, it approaches only 20 per cent (Veasy *et al*. 1987). Thus, it is important to have a high index of suspicion of acute rheumatic fever in children or young adults presenting with signs of arthritis and/or carditis, even in the absence of a clinically documented sore throat.

Another intriguing, and as yet unexplained, observation has been the invariable association of rheumatic fever only with streptococcal pharyngitis rather than other streptococcal lesions. While there have been many outbreaks of impetigo, rheumatic fever almost never follows infection with these impetigo strains. Furthermore, as Potter *et al*. (1978) have pointed out, in Trinidad, where both impetigo and rheumatic fever are concomitant infections, the strains colonizing the skin are different from those associated with rheumatic fever, and did not influence the incidence of acute rheumatic fever.

These observations remain inexplicable. It is clear that group A streptococci fall into two main classes based on differences in the C-repeat regions of the M protein (Bessen *et al*. 1989). One class is clearly associated with streptococcal pharyngeal infection, the other (with some exceptions) belongs to strains commonly associated with impetigo. Thus, the particular strain of streptococcus may be crucial in initiating the disease process. The pharyngeal site of infection, with its large repository of lymphoid tissue, may also be important in the initiation of the abnormal humoral response by the host to those antigens cross-reactive with target organs (see below). Finally, while impetigo strains do colonize the pharynx, they do not appear to elicit as strong an immunological response to the M-protein moiety as do the pharyngeal strains (Kaplan *et al*. 1970; Bisno and Nelson 1975). This may prove to be an important factor, especially in light of the known cross-reactions between various streptococcal structures and mammalian proteins.

Group A streptococcus

Figure 1 is a schematic cross-section of the group A streptococcus. The capsule is composed of equimolar concentrations of *N*-acetyl glucosamine and glucuronic acid, and is structurally identical to hyaluronic acid of mammalian tissues (Kendall *et al*. 1937).

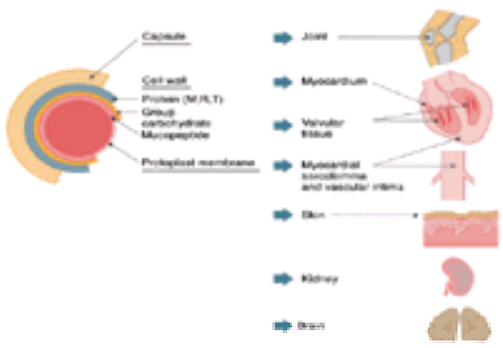


Fig. 1 Schematic representation of the various structures of the group A streptococcus. Note the wide variety of cross-reactions between its antigens and mammalian tissues.

Numerous past attempts to demonstrate antibodies to this capsule were unsuccessful (Seastone 1939; Quinn and Singh 1957). More recently, Fillit *et al*. (1986) successfully demonstrated high titres to hyaluronic acid, using techniques designed to detect non-precipitating antibodies in the serum of animals. Similar antibodies have been found in man (Faarber *et al*. 1984). Almost no published data implicate this capsule as important in human infections, although Stollerman (1975) commented on the presence of a large mucoid capsule as one of the more important characteristics of certain 'rheumatogenic' strains.

Investigations by Dr Rebecca Lancefield and others, spanning almost 70 years (reviewed by Fischetti 1989), have established that the M-protein molecule (at least 80 distinct serological types) is perhaps the most important virulence factor in human group A streptococcal infections. The protein is a helical, coiled-coil structure; it has striking structural homology with the cardiac cytoskeletal proteins tropomyosin and myosin, as well as with many other coil-coiled structures like keratin, DNA, lamin, and vimentin.

Once the amino acid sequence of a number of M proteins became known, it was possible to localize specifically those cross-reactive areas. The studies of Dale and Beachey (1985) showed that the part of the M protein involved in the opsonic reaction also cross-reacted with human sarcolemmal antigens. Sargent *et al*. (1987) more precisely localized this cross-reaction to the M-protein amino acid residues 164–197.

The evidence implicating these cross-reactions in the pathogenesis of acute rheumatic fever remains scant. Antibodies to myosin have been detected in the serum of patients with acute rheumatic fever, but they are also present in a large percentage of sera obtained from individuals who have had a streptococcal infection but did not subsequently develop acute rheumatic fever (Cunningham *et al*. 1988). The significance of this observation is unclear, since myosin is an internal protein of cardiac muscle cells and therefore not easily exposed to M-protein cross-reacting antibodies.

The group-specific carbohydrate of the streptococcus is a polysaccharide chain consisting of repeating units of rhamnose capped by *N*-acetyl glucosamine molecules. The *N*-acetyl glucosamine is immunodominant and gives rise to the serological group specificity of group A streptococci (McCarty 1970). The cross-reaction between group A carbohydrate and valvular glycoproteins was first described by Goldstein *et al*. (1968), and the reactivity was related to the *N*-acetyl glucosamine moiety present in both structures. Goldstein and Caravano (1967) noted that serum from patients with rheumatic fever reacted with the heart-valve glycoprotein. More recently, H. M. Fillit (personal communication) has observed strong reactivity of such sera with purified proteoglycan material. Thus, these cross-reactions could involve the sugar moiety present in both the proteoglycan portion of the glycoprotein and the carbohydrate.

It has always been assumed that group A anticarbohydrate antibodies did not play a part in the phagocytosis of group A streptococci. However, the studies of Salvadori *et al*. (1995) have demonstrated that human serum containing high titres of anti-group A carbohydrate antibody were opsonophagocytic for a number of different M-protein-specific strains, and that the opsonophagocytic properties were directed to the *N*-acetyl glucosamine moiety of the group A carbohydrate.

The mucopeptide portion of the cell wall is the 'backbone' of the organism and thus rather rigid in structure. It is composed of repeating units of muramic acid and *N*-acetyl glucosamine, cross-linked by peptide bridges (Chetty and Schwab 1984). It is particularly difficult to degrade and induces a wide variety of lesions when injected into various species, including arthritis in rats (Cromartie *et al*. 1977) and myocardial granulomas in mice that resemble (but are not identical to) lesions in rheumatic fever (Cromartie and Craddock 1966).

The connection between cell-wall mucopeptides and the pathogenesis of rheumatic fever remains obscure. Elevated titres of antimucopeptide antibody have been detected in the serum of patients with acute rheumatic fever, and also in the serum of patients with rheumatoid and juvenile rheumatoid arthritis (Heymer *et al*. 1976), but their pathogenetic relation to the clinical disease has been difficult to establish. There is no evidence that cell-wall antigens are present either in the Aschoff lesion or in the myocardial tissue obtained from patients with rheumatic fever.

Perhaps the most significant cross-reactions lie in the streptococcal membrane structure. We have shown (Zabriskie 1985) that immunization with membrane material elicited antibodies that bound to sections of heart in a pattern similar to that observed with serum from acute rheumatic fever (see Fig. 2).

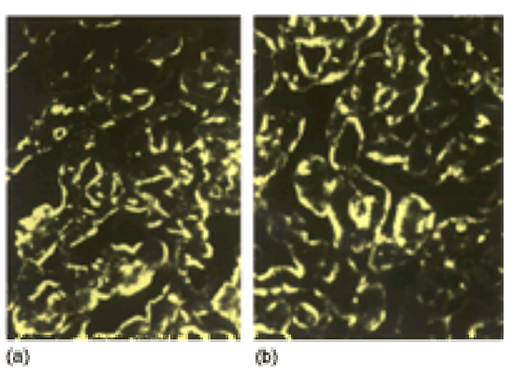


Fig. 2 Photomicrographs of immunofluorescent staining of heart sections with (a) rabbit serum immunized with group A streptococcal membranes and (b) serum obtained from a patient with acute rheumatic fever. Note the identical sarcolemmal staining patterns of both sera.

Kingston and Glynn (1971) were the first to show that animals immunized with streptococcal antigens develop serum antibodies that stain astrocytes. Husby *et al*. (1976) demonstrated that serum from patients with acute rheumatic fever with chorea contains antibodies that were specific for caudate cells. Absorption of the serum with streptococcal membrane antigens eliminated the reactivity with caudate cells.

Numerous other cross-reactions between streptococcal membranes and other organs have also been reported, for example renal basement membranes, basement membrane proteoglycans, and skin, particularly keratin. Here, space does not permit an exhaustive discussion of these cross-reactions, and the reader is referred to our recent review (Froude *et al*. 1989) for more detail. Whether or not these cross-reactions (especially those seen with basement membranes and skin) play a part in the disease awaits further study.

Genetics

The concept that rheumatic fever might be the result of a host genetic predisposition has intrigued investigators for over a century (Cheadle 1889). It has been variously suggested that the disease gene is transmitted in an autosomal-dominant fashion (Wilson *et al*. 1943), an autosomal-recessive fashion with limited penetrance (Taranta *et al*. 1959), or that it is possibly related to the genes conferring blood-group secretor status (Glynn and Holborow 1961).

Renewed interest in the genetics of rheumatic fever came with the recognition that gene products of the human major histocompatibility complex (**MHC**) were associated with certain clinical disease states. Using an alloserum from a multiparous donor, an increased frequency of a B-cell alloantigen was reported in several genetically distinct and ethnically diverse populations of individuals with rheumatic fever, and was not MHC-related (Patarroyo *et al*. 1979).

Most recently, studies were accomplished with a monoclonal antibody (D8/17) prepared by immunizing mice with B cells from a patient with rheumatic fever (Khanna *et al*. 1989). This B-cell antigen was expressed on increased numbers of B cells in 100 per cent of rheumatics of diverse ethnic origin, and only in 10 per cent of normal individuals. The antigen defined by this monoclonal antibody showed no association with, or linkage to, any of the known MHC haplotypes, nor did it appear to be related to B-cell activation antigens.

These findings are in contrast to reports of an increased frequency of HLA-DR4 and HLA-DR2 in white and black patients with rheumatic heart disease (Ayoub *et al*. 1986). Other studies have implicated HLA-DR1 and -DRW6 as susceptibility factors in Black South African patients with rheumatic heart disease (Maharaj *et al*. 1987). Most recently, Guilherme *et al*. (1991) have noted a close association of HLA-DR7 and -DW53 with rheumatic fever in Brazil. These apparently differing results concerning HLA antigens and susceptibility to rheumatic fever prompt speculation that the reported associations might involve genes close to, but not identical with, an unknown gene for that susceptibility. Alternatively, and more likely, susceptibility to acute rheumatic fever is polygenic, and the D8/17 antigen might be associated with only one of the genes (i.e., those of the MHC complex encoding for DR antigens) conferring susceptibility. While the explanation remains to be determined, it is none the less true that the presence of the D8/17 antigen appears to identify a population at special risk for contracting acute rheumatic fever.

Aetiological considerations

While a large body of evidence, both immunological and epidemiological, has implicated the group A streptococcus in the induction of the disease process, the exact pathological mechanisms involved still remain obscure. At least three main theories have been proposed.

The first is concerned with the question of whether persistence of the organism is important. Despite several controversial reports, no investigators have been able consistently to demonstrate live organisms in cardiac tissues or valves in rheumatic fever (Watson *et al*. 1961).

The second theory revolves around the question of whether the deposition of toxic products is required. Although an attractive hypothesis, little or no experimental evidence has been obtained in its support. For example, Halbert *et al*. (1961) suggested that streptolysin O (an extracellular product of group A streptococci) is cardiotoxic and might be carried to the site by circulating complexes containing streptolysin O and antibody. However, in spite of an intensive search for these products (J. B. Zabriskie, unpublished data), no such complexes have been identified *in situ* (Wagner 1960).

Renewed interest in these extracellular toxins has recently emerged with the observation by Schlievert *et al*. (1987) that certain streptococcal pyrogenic toxins (A and C) may act as superantigens. These antigens may stimulate large numbers of T cells through their unique interaction between MHC class II and T-cell receptors of specific V_b types. This interaction does not involve the usual concept of antigen presentation in the context of the MHC complex. Once activated, these cells induce the production of tumour necrosis factor, interferon-g, and a number of interleukin moieties, thereby contributing to the initiation of pathological damage. Furthermore, it has been suggested (Paliard *et al*. 1991) that in certain disease states such as rheumatoid arthritis, autoreactive cells of specific V_b lineage may 'home' to the target organ. Although an attractive hypothesis, no data on the role of these superantigens in rheumatic fever have as yet emerged.

Perhaps the best evidence to date favours the concept that, in the genetically susceptible individual, there is an abnormal host immune response (both humoral and cellular) to those streptococcal antigens cross-reactive with mammalian tissues. The evidence supporting this concept may be divided into three broad categories as follows.

1. Employing a wide variety of methods, numerous investigators have documented the presence of heart-reactive antibodies in serum from rheumatic fever. The incidence of these antibodies has varied from a low of 33 per cent to a high or 85 per cent in various series. While these antibodies are seen in other individuals (notably those with uncomplicated streptococcal infections and patients with post-streptococcal glomerulonephritis), the titres are always lower than in rheumatic fever and decrease with time during the convalescent period (Zabriskie 1985).
2. Serum in rheumatic fever also contains higher titres of antibodies to both myosin and tropomyosin than serum from patients with uncomplicated streptococcal infections. These myosin affinity-purified antibodies also cross-react with M-protein moieties, suggesting this molecule could be the antigenic stimulus for the production of myosin antibodies in these sera (Cunningham *et al*. 1988).
3. Finally, as indicated above, autoimmune antibodies are a prominent finding in another major clinical manifestation of acute rheumatic fever, chorea, and these antibodies are directed against the cells of the caudate nucleus. The titre of this antibody corresponds with clinical disease activity (Husby *et al*. 1976).

While not necessarily autoimmune in nature, the presence of elevated amounts of immune complexes has been well documented both in serum and joints in acute rheumatic fever (van de Rijn *et al*. 1978). These amounts, which may be as high as those seen in classical post-streptococcal glomerulonephritis, may be responsible for immune-complex vasculitis seen in acute rheumatic fever and may provide the initial impetus for vascular damage, followed by the secondary penetration of autoreactive antibodies. Support for the concept is found in the close clinical similarity of arthritis in rheumatic fever to experimentally induced serum sickness in animals or the arthritis secondary to drug hypersensitivity. Deposition of host immunoglobulin and complement is also seen in the cardiac tissues of patients with acute rheumatic fever, suggesting autoimmune deposition of immunoglobulins in or near the Aschoff lesions.

At a cellular level, there is now ample evidence for the presence of both lymphocytes and macrophages at the site of pathological damage in the heart in patients with acute rheumatic fever (Kemeny *et al*. 1989). The cells are predominantly CD4+ helper lymphocytes during acute stages of the disease (4:1). The ratio of CD4+/CD8+ lymphocytes (2:1) more closely approximates the normal ratio in valvular specimens in chronic rheumatics (see Table 2). A majority of these cells express Ia antigens. A potentially important finding has been the observation that macrophage-like fibroblasts present in the diseased valves express Ia antigens (Amoils *et al*. 1986) and might be the antigen-presenting cells for the CD4+ lymphocytes.

Patients	Type of valve	Type of valvulitis*	HLA-DR+	6003b-2	Leu 15+2	Leu 4+2	Leu 3a+2	Leu 8a+2	Leu 3a: Leu 4/5
Acute valvulitis									
1	Mitral	Acute	58.0	42.0	5.1	45.5	75.6	23.9	3.1
2	Mitral	Acute	49.8	43.1	8.9	43.1	58.7	34.3	1.9
	Aortic	Acute	52.7	51.0	9.9	38.1	65.9	26.5	2.5
3	Mitral	Acute	63.9	42.0	5.5	52.4	75.4	18.9	4.0
4	Aortic	Acute	68.1	56.0	7.4	33.7	71.6	22.0	3.3
Chronic valvulitis									
4	Mitral	Chronic active	69.4	67.4	7.4	44.3	52.7	36.8	1.4
5	Mitral	Chronic active	48.8	39.1	1.4	53.9	45.2	51.5	0.9
	Aortic	Chronic active	67.9	35.0	4.0	36.8	47.5	46.1	1.0
6	Mitral	Chronic active	41.8	23.4	8.0	65.8	57.3	33.3	1.7
	Aortic	Chronic active	59.6	48.7	6.2	39.1	59.2	32.6	1.8
7	Mitral	Chronic active	55.4	24.2	8.1	55.8	64.8	24.7	2.6
8	Mitral	Chronic active	80.4	34.1	13.4	44.6	44.8	50.9	0.9
9	Mitral	Chronic active	48.1	29.6	6.8	65.8	61.8	33.3	1.8

*Determined in the frozen-section samples studied.
1, mononuclear macrophages; 2, B cells; 3, pan-T cells; 4, helper cells; 5, suppressor cells.

Table 2 Composition of mononuclear cellular infiltrates in acute and chronic active rheumatic valvulitis

There was greater reactivity to streptococcal antigens in preparations of mononuclear cells from peripheral blood of patients with acute rheumatic fever than in these cells isolated from patients with nephritis (Read *et al*. 1986). This abnormal reactivity peaked at 6 months after the attack but could persist for as long as 2 years after the initial episode. Once again the reactivity was specific only for those strains associated with acute rheumatic fever, suggesting an abnormal humoral and cellular response to streptococcal antigens unique to rheumatic fever-associated streptococci.

Support for the potential pathological importance of these T cells is further strengthened by the observation that lymphocytes obtained from experimental animals sensitized to cell membranes but not cell walls are specifically cytotoxic for syngeneic embryonic cardiac myofibrils in vitro (Yang *et al*. 1977). In humans, normal mononuclear cells primed in vitro by M-protein molecules from a rheumatic fever-associated strain were also cytotoxic for myofibrils but specificity solely for cardiac cells was lacking (Dale and Beachey 1987). Similar studies have not yet been done with lymphocytes from patients with active acute rheumatic fever.

Clinical features of acute rheumatic fever

The clinical presentation of acute rheumatic fever is rather variable, and the lack of a single pathognomonic feature has resulted in the development of the revised Jones criteria, as illustrated in Table 3 (Jones Criteria Update 1992), which are used to establish a diagnosis. It should be noted that these criteria were established only as guidelines for the diagnosis and were never intended to be 'etched in stone'. Thus, depending on the age, geographical location, and ethnic population, emphasis on one or the other criterion for the diagnosis of acute rheumatic fever may be more or less important. Manifestations of rheumatic fever that are not clearly expressed pose a dilemma because of the importance of clearly identifying a first rheumatic attack in order to establish the need for prophylaxis (see below). Some of the isolated manifestations, particularly polyarthritis, may be difficult or impossible to distinguish from other diseases, especially at their onset. The diagnosis can be made, however, when 'pure' chorea is the sole manifestation, because of the rarity with which this syndrome is due to any other cause.

Major manifestations	Minor manifestations
Carditis	Fever
Polyarthritis	Arthralgia
Chorea	Previous rheumatic fever or rheumatic heart disease
Erythema marginatum	
Subcutaneous nodules	
Laboratory findings	
(1) Elevated acute-phase reactants:	
(a) C-reactive protein	
(b) erythrocyte sedimentation rate	
(2) Prolonged P–R interval rate	
Supporting evidence of preceding streptococcal infection	
(a) Increased ASO or other streptococcal antibodies	
(b) Positive throat culture for group A haemolytic streptococci	
(c) Recent scarlet fever	

Jones Criteria Update (1992)
ASO, antistreptolysin O.

Table 3 Revised Jones criteria for diagnosis of acute rheumatic fever

Arthritis

In the classic, untreated case the arthritis of rheumatic fever affects several joints in quick succession, each for a short time. The legs are usually affected first and later the arms. The terms 'migrating' or 'migratory' are often used to describe the polyarthritis of rheumatic fever, but these designations are not meant to signify that the inflammation necessarily disappears in one joint when it appears in another. Rather, the various localizations usually overlap in time, and the onset, as opposed to the full course of the arthritis, 'migrates' from joint to joint.

Joint involvement is more common, and also more severe, in teenagers and young adults than in children. It occurs early in the rheumatic illness, and is usually the earliest symptomatic manifestation of the disease, although asymptomatic carditis may precede it. Rheumatic polyarthritis may be excruciatingly painful, but is almost always transient. The pain is usually more prominent than the objective signs of inflammation.

When the disease was allowed to express itself fully, unchecked by anti-inflammatory treatment, over half of patients studied show a true polyarthritis, with inflammation in any of from 6 to 16 joints. Classically, each joint is maximally inflamed for only a few days, or a week at the most: the inflammation decreases, perhaps lingering for another week or so, and then disappears completely. Radiographs taken at this point may show a slight effusion but most probably will be unremarkable.

In routine practice, however, many patients with arthritis and/or arthralgias are treated empirically with salicylates or other non-steroidal anti-inflammatory drugs. Accordingly, arthritis subsides quickly in the joint(s) already affected and does not 'migrate' to new joints. Thus, therapy may deprive the diagnostician of a useful sign. In a large series of patients with rheumatic fever and associated arthritis, most of whom had been treated, involvement of only a single large joint was common (25 per cent). One or both knees were affected in 76 per cent, and one or both ankles in 50 per cent. Elbows, wrists, hips, or small joints of the feet were involved in 12 to 15 per cent of patients, and shoulder or small joints of the head were affected in 7 to 8 per cent. Joints rarely affected were the lumbosacral (2 per cent), cervical (1 per cent), sternoclavicular (0.5 per cent), and temporomandibular (0.5 per cent). Involvement of the small joints of the hands or feet alone occurred in only 1 per cent of these patients (Feinstein and Spagnulo 1962).

Analysis of the synovial fluid in well-documented cases of rheumatic fever with arthritis generally reveals a sterile inflammatory fluid. There may be a decrease of the complement components C1q, C3 and C4, indicating their consumption by immune complexes in the joint fluid (Svartman *et al*. 1975).

Post-streptococcal reactive arthritis

A number of investigators (Goldsmith and Long 1982; Arnold and Tyndall 1989; Fink 1991) have raised the question of whether post-streptococcal migratory arthritis, in the absence of carditis both in adults and children, is really acute rheumatic fever, for the following reasons.

1. The latent period between the antecedent streptococcal infection and the onset of acute rheumatic fever is shorter (1–2 weeks) than the 3 to 4 weeks usually seen in classical acute rheumatic fever.
2. The response of the arthritis to aspirin and other non-steroidal medications is poor in comparison to the dramatic response seen in classical acute rheumatic fever.
3. Evidence of carditis is not usually seen in these patients and the arthritis is rather severe.

4. Extra-articular manifestations such as tenosynovitis and renal abnormalities are often seen in these patients.

While these cases (admittedly rare) do exist, migratory arthritis without evidence of other major Jones criteria, if supported by two minor manifestations (see Table 3), must still be considered acute rheumatic fever, especially in children. Variations in the response to aspirin in these children often are not recorded with serum salicylate concentrations, and an unusual clinical course is not sufficient to exclude the diagnosis of acute rheumatic fever; appropriate prophylactic measures should be therefore taken (reviewed by Gibofsky and Zabriskie 1994). Support for this concept may be found in the work of Crea and Mortimer (1959), in which 50 per cent of the children with signs of migratory arthritis alone went on to develop significant valvular damage after a long follow-up.

Rheumatic fever also occurs in adults. Although migratory arthritis is a common presenting symptom, a recent outbreak in San Diego Naval Training Camp (Wallace *et al*. 1989) revealed a 30 per cent incidence of valvular damage in these patients.

The importance of clearly defining this reactive arthritis as a variant of rheumatic fever has obvious implications for secondary prophylactic treatment. As suggested by some investigators, post-streptococcal reactive arthritis is a benign condition without need for prophylaxis. Yet as these patients by and large do fulfil the Jones criteria (one major, two minor), they should be considered as having rheumatic fever and treated as such.

Carditis

Cardiac valvular and muscle damage can manifest in a variety of signs or symptoms, including organic heart murmurs, cardiomegaly, congestive heart failure or pericarditis. Mild to moderate chest discomfort, pleuritic chest pain or a pericardial friction rub are indications of pericarditis. On clinical examination, the patient can have new or changing organic murmurs, most commonly mitral regurgitant murmurs, and occasionally aortic regurgitant murmurs and/or systolic ejection murmurs, caused by acute valvular inflammation and deformity. Rarely, a Carey Coombs mid-diastolic murmur caused by rapid flow over the mitral valve is heard. If the valvular damage is severe enough, together with concurrent cardiac dysfunction, congestive heart failure can ensue, which is the most life-threatening clinical syndrome of acute rheumatic fever. Congestive heart failure needs to be treated intensively and quickly with a combination of anti-inflammatory drugs, diuretics and, occasionally, steroids to acutely decrease the cardiac inflammation. Electrocardiographic abnormalities include all degrees of heart block, including atrioventricular dissociation, but first-degree heart block is not associated with a poor prognosis. Second- or third-degree heart block can occasionally be symptomatic. If heart block is associated with congestive heart failure, a temporary pacemaker can be placed if indicated. The most common manifestation of carditis is cardiomegaly, as seen on radiographs.

In the population of patients recently reviewed from our institution, The Rockefeller University Hospital, who were diagnosed with acute rheumatic fever between 1950 and 1970, with an average of 20 years of follow-up, 90 per cent had evidence of carditis at diagnosis (Table 4). In a classic review of 1000 patients (Bland and Jones 1951), only 65 per cent were diagnosed with carditis. The addition of Doppler sonography to the clinical evaluation of patients during the recent Utah outbreak increased the proportion diagnosed with carditis from 72 to 91 per cent (Veasy *et al*. 1987), indicating that, with more sensitive measurements of cardiac dysfunction, almost all patients with acute rheumatic fever have signs of acute carditis.

	RHD (*n*=60) %	No RHD (*n*=47) %	Total (*n*=87) %	Bland and Jones %
Carditis	100	83.0	90.1	65.3
Arthritis	67.5	68.1	67.8	41.0
Epistaxis	0.0	10.6	5.7	27.4
Chorea	5.0	2.1	3.4	51.8
Pericarditis	2.5	4.3	3.4	13.0
Nodules	7.5	0.0	3.4	8.8
Erythema marginatum	0.0	4.3	2.3	7.1

Table 4 Physical signs and symptoms of acute rheumatic fever, Rockefeller University Hospital 1950–1970

Rheumatic heart disease

Rheumatic heart disease is the most severe outcome of acute rheumatic fever. Usually occurring 10 to 20 years after the original attack, it is the major cause of acquired valvular disease in the world. The mitral valve is mainly involved and the aortic valve less often. Mitral stenosis is a classic finding in rheumatic heart disease and can manifest as a combination of mitral insufficiency and stenosis, secondary to severe calcification of the mitral valve. When symptoms of left atrial enlargement are present, mitral valve replacement may become necessary.

In various studies, the incidence of rheumatic heart disease in patients with a history of acute rheumatic fever has varied. In the classic study of Bland and Jones (1951), after 20 years, one-third of patients had no murmur, another one-third had died, and the remaining one-third was alive with rheumatic heart disease. A majority of the patients who died had rheumatic heart disease. While the dogma is that patients with rheumatic heart disease invariably have had more than one attack of acute rheumatic fever, recent analysis of our patients at the Rockefeller University Hospital disproves this. The population studied was 87 patients who had had only one documented attack of acute rheumatic fever, without any evidence (clinical or laboratory) of a recurrence during a 20-year follow-up under close supervision. Over 80 per cent had carditis at admission and approx. 50 per cent now have organic murmurs (Table 4). Thus, valvular damage manifesting as organic murmurs later in life is still likely to occur in 50 per cent of the patients if they presented with evidence of carditis at initial diagnosis. All of the patients in our population who ended up with rheumatic heart disease had carditis at diagnosis.

Chorea

Sydenham's chorea, chorea minor, or 'St. Vitus dance' is a neurological disorder consisting of abrupt, purposeless, non-rhythmic involuntary movements, muscular weakness, and emotional disturbances. They disappear during sleep, but may occur at rest and may interfere with voluntary activity. Initially, it may be possible to suppress these movements, which may affect all voluntary muscles, with the hands and face usually the most obvious. Grimaces and inappropriate smiles are common. Handwriting usually becomes clumsy and provides a convenient way of following the patient's course. Speech is often slurred. The movements are commonly more marked on one side and are occasionally completely unilateral (hemichorea).

The muscular weakness is best revealed by asking the patient to squeeze the examiner's hands: the pressure of the patient's grip increases and decreases continuously and capriciously, a phenomenon known as relapsing grip, or milking sign.

The emotional changes manifest themselves in outbursts of inappropriate behaviour, including crying and restlessness. In rare cases, the psychological manifestations may be severe and may result in transient psychosis.

The neurological examination fails to reveal sensory losses or involvement of the pyramidal tract. Diffuse hypotonia may be present.

Chorea may follow streptococcal infections after a latent period, which is longer, on the average, than the latent period of other rheumatic manifestations. Some patients with chorea have no other symptoms, but other patients develop chorea weeks or months after arthritis. In both cases, examination of the heart may reveal murmurs.

Skin lesions

Subcutaneous nodules

The subcutaneous nodules of rheumatic fever are firm and painless. The overlying skin is not inflamed and can usually be moved over the nodules. The diameter of these round lesions varies from a few millimetres to 1 or even 2 cm. They are located over bony surface or prominences, or near tendons; their number varies from a single nodule to a few dozen and averages three or four; when numerous, they are usually symmetrical. These nodules are present for one or more weeks, rarely for more than a month. They are smaller and more short-lived than the nodules of rheumatoid arthritis. Although in both diseases the elbows are most frequently involved, the rheumatic nodules are more common on the olecranon, and the rheumatoid nodules are usually found 3 or 4 cm distal to it. Rheumatic subcutaneous nodules generally appear only after the first few weeks of illness, usually only in patients with carditis.

Erythema marginatum

Erythema marginatum is an evanescent, non-pruritic skin rash, pink or faintly red, usually affecting the trunk, sometimes the proximal parts or the limbs, but not the face. The lesion extends centrifugally while the skin in the centre returns gradually to normal; hence, the name 'erythema marginatum'. The outer edge of the lesion is sharp, whereas the inner edge is diffuse. Because the margin of the lesion is usually continuous, making a ring, it is also known as 'erythema annulare'.

The individual lesions may appear and disappear in a matter of hours, usually to return. A hot bath or shower may make them more evident or may even reveal them for the first time.

Erythema marginatum usually occurs in the early phase of the disease. It often persists or recurs, even when all other manifestations of disease have disappeared. Occasionally, the lesions appear for the first time or, more probably, are noticed for the first time, late in the course of the illness or even during convalescence. This disorder usually occurs only in patients with carditis.

The minor manifestations of rheumatic fever

Temperature is increased in almost all rheumatic attacks and ranges from 38.4 to 40°C. Usually fever decreases in about a week without antipyretic treatment, and may become low grade for another week or two. Fever rarely lasts for more than 3 to 4 weeks.

Abdominal pain

The abdominal pain of rheumatic fever resembles that of other conditions associated with acute microvascular mesenteric inflammation and is non-specific. It usually occurs at or near the onset of the rheumatic attack, so that other manifestations may not yet be present to clarify the diagnosis. In many cases, it may mimic acute appendicitis.

Epistaxis

In the past, epistaxis occurred most prominently and severely in patients with severe and protracted rheumatic carditis. Early clinical studies reported a frequency as high as 48 per cent, but it probably occurs even less frequently now (see Table 3). Although epistaxis has been correlated in the past with the severity of rheumatic inflammation, it is difficult to assess retrospectively the possible thrombasthenic effect of large doses of salicylates administered for prolonged periods in protracted attacks.

Rheumatic pneumonia

Rheumatic pneumonia may appear during the course of severe rheumatic carditis. This inflammatory process is difficult or impossible to distinguish from pulmonary oedema or the alveolitis associated with respiratory distress syndrome due to a variety of pathophysiological states.

Laboratory findings

The diagnosis of rheumatic fever cannot readily be established by laboratory tests. Nevertheless, tests may be helpful in two ways: first in demonstrating that an antecedent streptococcal infection has occurred and second in documenting the presence or persistence of an inflammatory process. Serial chest radiographs may be helpful in following the course of carditis and the electrocardiogram may reflect the inflammatory process on the conduction system.

Throat cultures are usually negative by the time rheumatic fever appears but an attempt should be made to isolate the organism. It is our practice to take three throat cultures during the first 24 h, before giving antibiotics. Streptococcal antibodies are more useful because (i) they reach a peak titre at about the time of onset of rheumatic fever, (ii) they indicate true infection rather than transient carriage, and (ii) by performing several tests for different antibodies, any significant recent streptococcal infection can be detected. To demonstrate a rising titre, it is useful to take a serum specimen when the patient is first seen and another one 2 weeks later for comparison.

The specific antibody tests most frequently used to diagnosis streptococcal infections are those directed against extracellular products. They include antistreptolysin O, anti-DNAse B, antihyaluronidase, anti-NADase (anti-DPNase), and antistreptokinase. Antistreptolysin O has been the most widely used test and is generally available in hospitals in the United States.

Titres of antistreptolysin O vary with age, season, and geographical region. They reach peak levels in the young, school-age population. Titres of 200 to 300 Todd units/ml are common, therefore, in healthy children of elementary-school age. After a streptococcal pharyngitis, the antibody response peaks at about 4 to 5 weeks, which is usually during the second or third week of rheumatic fever (depending on how early it is detected). Thereafter, antibody titres fall off rapidly in the next several months, and, after 6 months, they decline more slowly. Since only 80 per cent of documented rheumatics exhibit a rise in the titre of antistreptolysin O, it is recommended that other antistreptococcal antibody tests be performed in the absence of a positive titre. These include anti-DNAse B, hyaluronidase or streptozyme, which is a combination of various streptococcal antigens.

Streptococcal antibodies, when increased, support but do not prove the diagnosis of acute rheumatic fever, nor are they a measure of rheumatic activity. Even in the absence of intercurrent streptococcal infection, titres decline during the rheumatic attack despite the persistence or severity of rheumatic activity.

Acute-phase reactants

Acute-phase reactants are elevated during acute rheumatic fever, just as they are during other inflammatory conditions. Both the C-reactive protein and erythrocyte sedimentation rate are almost invariably abnormal during the active rheumatic process, if it is not suppressed by antirheumatic drugs. Pure chorea and persistent erythema marginatum are exceptions. Particularly when treatment has been discontinued or is being tapered off, the C-reactive protein or erythrocyte sedimentation rate are useful in monitoring 'rebounds' of rheumatic inflammation, which indicate that the rheumatic process is still active. If either remains normal a few weeks after discontinuing antirheumatic therapy, the attack may be considered ended unless chorea appears. Even then, usually, there will be no exacerbation of the systemic inflammation and chorea will be present as an isolated manifestation.

Anaemia

A mild normochromic normocytic anaemia of chronic infection or inflammation may be seen during acute rheumatic fever. Suppressing the inflammation usually improves the anaemia, thus iron therapy is usually not indicated.

Clinical course and treatment of acute rheumatic fever

The mainstay of treatment for acute rheumatic fever has always been anti-inflammatory agents, most commonly aspirin. Dramatic improvement in symptoms is usually seen after the start of therapy. Usually 80 to 100 mg/kg per day in children and 4 to 8 g/day in adults is required for an effect to be seen. Aspirin concentrations can be measured and 20 to 30 mg/dl is the therapeutic range. The duration of anti-inflammatory therapy can vary but the treatment should be maintained until all symptoms are absent and laboratory values are normal. If severe carditis is also present, as indicated by significant cardiomegaly, congestive heart failure or third-degree heart

block, steroid therapy can be instituted. The usual dosage is 2 mg/kg per day of oral prednisone during the first 1 to 2 weeks. Depending on clinical and laboratory improvement, the dosage is then tapered over the next 2 weeks and during the last week aspirin may be added in the dosage recommended above sufficient to achieve 20 to 30 mg/dl.

Whether or not signs of pharyngitis are present at the time of diagnosis, antibiotic therapy with penicillin should be started and maintained for at least 10 days, in doses recommended for the eradication of a streptococcal pharyngitis. In addition, all family contacts should be cultured and treated for streptococcal infection, if positive. If compliance is an issue, depot penicillins, such as benzathine penicillin G 600 000 units in children, 1.2 million units in adults, should be given. Recurrences of acute rheumatic fever are most common within 2 years of the original attack but can occur at any time. The risk of recurrence decreases with age. Recurrence rates have been decreasing, from 20 to 2 to 4 per cent in recent outbreaks. This might be due to better surveillance and treatment.

Prophylaxis

Antibiotic prophylaxis with penicillin should be started immediately after the resolution of the acute episode. The optimal regimen consists of oral penicillin VK 250 000 units twice a day, or depot intramuscular injection of 1.2 million units penicillin G every 4 weeks. Recent data suggest, however, that injections every 3 weeks are more effective than every 4 weeks in preventing recurrences of acute rheumatic fever (Lue *et al*. 1986). If the patient is allergic to penicillin, erythromycin 250 mg/day can be substituted.

The end-point of prophylaxis is unclear; most believe it should continue at least until the patient is a young adult, which is usually 10 years from an acute attack with no recurrence. In our opinion, individuals with documented evidence of rheumatic heart disease should be on continuous prophylaxis indefinitely since our experience has been that rheumatic fever can recur even in the fifth or sixth decades. Another potential problem for recurrences is the presence in the household of young children who could transmit new group A streptococcal infections to rheumatic-susceptible individuals.

Obviously the alternative to long-term prophylaxis in an individual with rheumatic fever will be the introduction of streptococcal vaccines designed not only to prevent recurrent infections in rheumatic-susceptible individuals but also to prevent streptococcal disease in general. While it is not within the scope of this chapter to discuss the prospects for these vaccines (see review by Fischetti 1989), at least a few words are appropriate. Immunization of mice with either C-repeat peptides of M protein or 'cloned' M protein in a vaccinia virus vector protected them against intranasal infection with homologous or heterologous strains of group A streptococci. Whether or not these antigens and/or vectors are protective in man is being investigated. One of the major problems will be to avoid using those parts of the molecule that are cross-reactive with mammalian tissues.

Conclusion

In spite of its disappearance in many areas, rheumatic fever continues to be a serious problem in those geographical areas where two-thirds of the world's population lives. Even in developed countries with full access to medical care, and better nutrition and housing, the recent resurgence of rheumatic fever emphasizes the need for continued vigilance by physicians and other health officials in both diagnosing and treating the disease. The importance of early diagnosis and therapy cannot be overemphasized. Although the joint manifestations are transient and self-limiting, the cardiac sequelae are chronic and life-threatening. Whether the resurgence represents a change in the virulence of the organism or failure to recognize the importance and need for adequate treatment of an antecedent streptococcal infection is an area of intense debate and will therefore require careful and controlled epidemiological surveillance.

Nevertheless, rheumatic fever remains one of the few autoimmune disorders known to occur as a result of infection with a specific organism. The confirmed observation of an increased frequency of a B-cell alloantigen in several populations of rheumatics suggests that it might be possible to identify rheumatic fever-susceptible individuals at birth. If so, then from a public health standpoint, (i) these individuals would be prime candidates for immunization with any streptococcal vaccine that might be developed in the future, (ii) careful monitoring of streptococcal disease in the susceptible population could lead to early and effective antibiotic strategies, resulting in disease prevention, and (iii), in individuals previously infected, who later present with subtle or non-specific manifestations of the disease, the presence or absence of the marker could be of value in arriving at a diagnosis.

The continued study of rheumatic fever as a prime example of microbial–host interactions also has important implications for the study of autoimmune diseases in general and rheumatic diseases in particular. Further insights into this intriguing host–parasite relation may shed additional light on those diseases where infection is assumed to have occurred but has not as yet been identified.

Chapter References

Amoils, B. *et al*. (1986). Aberrant expression of HLA-DR antigen on valvular fibroblasts from patients with acute rheumatic carditis. *Clinical and Experimental Immunology*, **66**, 84–94.

Arnold, M. H. and Tyndall, A. (1989). Post-streptococcal reactive arthritis. *Annals of the Rheumatic Diseases*, **48**, 681–8.

Ayoub, E. A., Barrett, D. J., Maclaren, N. K., and Krischer, J. P. (1986). Association of class II human histocompatibility leucocyte antigens with rheumatic fever. *Journal of Clinical Investigation*, **77**, 2019–26.

Bessen, D., Jones, K. F., and Fischetti, V. A. (1989). Evidence for the distinct classes of streptococcal M protein and their relationship to rheumatic fever. *Journal of Experimental Medicine*, **169**, 269–83.

Bisno, A. L. and Nelson, K. E. (1975). Type-specific opsonic antibodies in streptococcal pyoderma. *Infection and Immunity*, **10**, 1356–61.

Bland, E. F. and Jones, T. D. (1951). Rheumatic fever and rheumatic heart disease: a twenty year report on 1,000 patients followed since childhood. *Circulation*, **4**, 836–43.

Cheadle, W. B. (1889). Harvean lectures on the various manifestations of the rheumatic state as exemplified in childhood and early life. *Lancet*, **i**, 821–32.

Chetty, C. and Schwab, J. H. (1984). Chemistry of endotoxins. In *Handbook of endotoxin*, Vol. 1. (ed. E. T. Rietschel Elsenier), pp. 376–410. Elsevier Science Publishers BV, Amsterdam.

Crea, M. A. and Mortimer, E. A. (1959). The nature of scarlatinal arthritis. *Pediatrics*, **23**, 879–84.

Cromartie, W. J. and Craddock, J. B. (1966). Rheumatic-like cardiac lesions in mice. *Science*, **154**, 285–7.

Cromartie, W. J., Craddock, J. B., Schwab, J. H., Anderle, S. K., and Yang, C. H. (1977). Arthritis in rats after systemic injection of streptococcal cells or cell walls. *Journal of Experimental Medicine*, **146**, 1585–602.

Cunningham, M. W. *et al*. (1988). Human monoclonal antibodies reactive with antigens of the group A streptococcus and human heart. *Journal of Immunology*, **141**, 2760–6.

Dale, J. B. and Beachey, E. H. (1985). Multiple cross reactive epitopes of streptococcal M proteins. *Journal of Experimental Medicine*, **161**, 113–22.

Dale, J. B. and Beachey, E. H. (1987). Human cytotoxic T lymphocytes evoked by group A streptococcal M proteins. *Journal of Experimental Medicine*, **166**, 1825–35.

Denny, F. W., Jr, Wannamaker, L. W., Brink, W. R., Rammelcamp, C. H., and Custer, E. A. (1950). Prevention of rheumatic fever: treatment of the preceding streptococcal infection. *Journal of the American Medical Association*, **143**, 151–3.

Faarber, P. *et al*. (1984). Cross reactivity of anti DNA antibodies with proteoglycans. *Clinical and Experimental Immunology*, **55**, 402–12.

Feinstein, A. R. and Spagnulo, M. (1962). The clinical patterns of rheumatic fever: a reappraisal *Medicine*, **41**, 279–305.

Fillit, H. M., McCarty, M., and Blake, M. (1986). Induction of antibodies to hyaluronic acid by immunization of rabbits with encapsulated streptococci. Journal of Experimental Medicine, **164**, 762–76.

Fink, C. W. (1991). The role of streptococcus in post streptococcal reactive arthritis and childhood polyarteritis nodosa. *Journal of Rheumatology*, **18**, 14–20.

Fischetti, V. A. (1989). Streptococcal M protein: molecular design and biological behavior. *Clinical and Microbiology Reviews*, **2**, 285–314.

Froude, J., Gibofsky, A., Buskirk, D. R., Khanna, A., and Zabriskie, J. B. (1989). Cross reactivity between streptococcus and human tissue: a model of molecular mimicry and autoimmunity. *Current*

Topics in Microbiology and Immunology, **145**, 5–26.

Gibofsky, A. and Zabriskie, J. B. (1994). Rheumatic fever: new insights into an old disease. *Bulletin on the Rheumatic Diseases: Arthritis Foundation*, **42**, 5–7.

Glynn, L. E. and Holborow, E. J. (1961). Relationship between blood groups, secretion status and susceptibility to rheumatic fever. *Arthritis and Rheumatism*, **4**, 203.

Goldsmith, D. P. and Long, S. S. (1982) Poststreptococcal disease of childhood—a changing syndrome. *Arthritis and Rheumatism*, **25**, S18 (Abstr.).

Goldstein, I. and Caravano, R. (1967). Determination of anti group A streptococcal polysaccharide antibodies in human sera by an hemagglutination technique. *Proceedings of the Society of Experimental Biology and Medicine*, **124**, 1209–12.

Goldstein, I., Rebeyrotte, P., Parlebas, J., and Halpern, B., (1968). Isolation from heart valves of glycopeptides which share immunological properties with *Streptococcus haemolyticus* group A polysaccharides. *Nature*, **219**, 866–8.

Gordis, L. (1985). The virtual disappearance of rheumatic fever in the United States: lessons in the rise and fall of disease. *Circulation*, **72**, 1155–62.

Guilherme, L., Weidenbach, W., Kiss, M. H., Snitcowsky, R., and Kalil, J. (1991). Association of human leucocyte class II antigens with rheumatic fever or rheumatic heart disease in a Brazilian population. *Circulation*, **83**, 1995–8.

Halbert, S. P., Bircher, R., and Dahle, E. (1961). The analysis of streptococcal infections. V. Cardiotoxicity of streptolysin O for rabbits *in vivo*. *Journal of Experimental Medicine*, **113**, 759–84.

Heymer, B., Schleifer, K. H., Read, S. E., Zabriskie, J. B., and Krause, R. M. (1976). Detection of antibodies to bacterial cell wall peptidoglycan in human sera. *Journal of Immunology*, **117**, 23–6.

Husby, G., van de Rijn, I., Zabriskie, J. B., Abdin, Z. H., and Williams, R. C., Jr (1976). Antibodies reacting with cytoplasm of subthalmic and caudate nuclei neurons in chorea and acute rheumatic fever. *Journal of Experimental Medicine*, **144**, 1094–110.

Jones Criteria Update (1992). Guidelines for diagnosis of rheumatic fever. *Journal of the American Medical Association*, **268**, 2069–70.

Kaplan, E. L., Anthony, B. F., Chapman, S. S., Ayoub, E. M., and Wannamaker, L. W. (1970). The influence of the site of infection on the immune response to group A streptococci. *Journal of Clinical Investigation*, **49**, 1405–14.

Kaplan, E. L., Johnson, D. R., and Cleary, P. P. (1989). Group A streptococcal serotypes isolated from patients and sibling contacts during the resurgence of rheumatic fever in the United States in the mid 1980's. *Journal of Infectious Diseases*, **159**, 101–3.

Kemeny, E., Grieve, T., Marcus, R., Sareli, P., and Zabriskie, J. B. (1989). Identification of mononuclear cells and T cell subsets in rheumatic valvulitis. *Clinical Immunology and Immunopathology*, **52**, 225–37.

Kendall, F., Heidelberger, M., and Dawson, M. (1937). A serologically inactive polysaccharide elaborated by mucoid strains of group A hemolytic streptococcus. *Journal of Biological Chemistry*, **118**, 61–82.

Khanna, A. K., Buskirk, D. R., Williams, R. C., Jr, Gibofsky, A., Crow, M. K., and Menon, A. (1989). Presence of a non-HLA B cell antigen in rheumatic fever patients and their families as defined by a monoclonal antibody. *Journal of Clinical Investigation*, **83**, 1710–16.

Kingston, D. and Glynn, L. E. (1971). A cross-reaction between *Streptococcus pyogenes* and human fibroblasts, endothelial cells and astrocytes. *Immunology*, **21**, 1003–16.

Lue, H. C., Mil-Wham, W., Hsieh, K. H., Lin, G. J., Hsieh, R. P., and Chou, F. F. (1986). Rheumatic fever recurrences: controlled study of 3 week versus 4 week benzathine penicillin prevention programs. *Journal of Pediatrics*, **108**, 299–304.

McCarty, M. (1970). The streptococcal cell wall. *The Harvey Lectures Series*, **65**, 73–96.

Maharaj, B., Hammond, M. G., Appadoo, B., Leary, W. P., and Pudifin, D. J. (1987). HLA-A, B, DR and DQ antigens in black patients with severe chronic rheumatic heart disease. *Circulation*, **765**, 259–61.

Markowitz, M. (1987). Rheumatic fever: recent outbreaks of an old disease. *Connecticut Medicine*, **51**, 229–33.

Markowitz, M. and Gordis, L. (1972). *Rheumatic fever*, (2nd edn). Saunders, Philadelphia.

Paliard, X. *et al*. (1991). Evidence for the effects of superantigen in rheumatoid arthritis. *Science*, **253**, 325–9.

Patarroyo, M. E. *et al*. (1979). Association of a B cell alloantigen with susceptibility to rheumatic fever. *Nature*, **278**, 173–4.

Pope, R. M. (1989). Rheumatic fever in the 1980s. *Bulletin on the Rheumatic Diseases, Arthritis Foundation*, **38**, 1–8.

Potter, E. V., Svartman, M., Mohammed, I., Cox, R., Poon-King, T., and Earle, D. P. (1978). Tropical acute rheumatic fever and associated streptococcal infections compared with concurrent acute glomerulonephritis., *Journal of Pediatrics*, **92**, 325–33.

Quinn, R. W. and Singh, K. P. (1957). Antigenicity of hyaluronic acid. *Biochemical Journal*, **95**, 290–301.

Read, S. E. *et al*. (1986). Serial studies on the cellular immune response to streptococcal antigens in acute and convalescent rheumatic fever patients in Trinidad. *Journal of Clinical Immunology*, **6**, 433–41.

Salvadori, L. G., Blake, M. S., McCarty, M., Tai, J. Y., and Zabriskie, J. B. (1995). Group A streptococcus-liposome ELISA antibody titers to group A polysaccharide and opsonophagocytic capabilities of the antibodies. *Journal of Infectious Diseases*, **171**, 593–600.

Sargent, S. J., Beachey, E. H., Corbett, C. E., and Dale, J. B. (1987). Sequence of protective epitopes of streptococcal M proteins shared with cardiac sarcolemmal membranes. *Journal of Immunology*, **139**, 1285–90.

Schlievert, P. M., Johnson, L. P., Tomai, M. A., and Handley, J. P. (1987). Characterization and genetics of group A streptococcal pyrogenic exotoxins. In *Streptococcal genetics* (ed. J. Ferretti and R. Curtisis), pp. 136–42. ASM, Washington DC.

Seastone, C. V. (1939). The virulence of group C hemolytic streptococci of animal origin. *Journal of Experimental Medicine*, **70**, 361–78.

Stollerman, G. H. (ed.) (1975). In *Rheumatic fever and streptococcal infection*, p.70. Grune and Stratton, New York.

Stollerman, G. H., Lewis, A. J., Schultz, I., and Taranta, A. (1956). Relationship of the immune response to group A streptococci to the cause of acute, chronic and recurrent rheumatic fever. *American Journal of Medicine*, **20**, 163–9.

Svartman, M., Potter, E. V., and Poon-King, T. (1975). Immunoglobulin components in synovial fluids of patients with acute rheumatic fever. *Journal of Clinical Investigation*, **56**, 111–17.

Taranta, A. *et al*. (1959). Rheumatic fever in monozygotic and dizygotic twins. *Circulation*, **20**, 778–92.

van de Rijn, I. *et al*. (1978). Serial studies on circulating immune complexes in post-streptococcal sequelae. *Clinical and Experimental Immunology*, **34**, 318–25.

Veasy, L. G. *et al*. (1987). Resurgence of acute rheumatic fever in the intermountain area of the United States. *New England Journal of Medicine*, **316**, 421–7.

Wagner, B. M. (1960). Studies in rheumatic fever. III. Histochemical reactivity of the Aschoff body. *Annals of the New York Academy of Sciences*, **86**, 992–1008.

Wallace, M. R., Garst, P. D., Papadimos, T. J., and Oldfield, E. C. (1989). The return of acute rheumatic fever in young adults. *Journal of the American Medical Association*, **262**, 2557–61.

Watson, R. F., Hirst, G. K., and Lancefield, R. C. (1961). Bacteriological studies of cardiac tissues obtained at autopsy from eleven patients dying with rheumatic fever. *Arthritis and Rheumatism*, **4**, 74–85.

Whitnack, E. and Bisno, A. L. (1980). Rheumatic fever and other immunologically mediated cardiac diseases. In *Clinical immunology*, Vol. II. (ed. C. Parker), pp. 894–929. Saunders, Philadelphia.

Wilson, M. G., Schweitzr, M. D., and Lubschez, R. (1943). The familial epidemiology of rheumatic fever. *Journal of Pediatrics*, **22**, 468–82.

Yang, L. C., Soprey, P. R., Wittner, M. K., and Fox, E. N. (1977). Streptococcal induced cell mediated immune destruction of cardiac myofibers in vitro. *Journal of Experimental Medicine*, **146**,

344–60.

Zabriskie, J. B. (1985). Rheumatic fever: The interplay between host genetics and microbe. *Circulation*, **71**, 1077–86.

5.4.1 Immunopathogenesis of rheumatoid arthritis

Ravinder N. Maini and Marc Feldmann

Historical review

The delineation of rheumatoid arthritis as a disease entity in the contemporary medical literature began to emerge in the eighteenth century. Initially, clinical observations sought to distinguish the disorder from other prevalent joint diseases, such as gout and rheumatic fever, and emphasized distinctive features, for example, its chronicity, joint deformities, female sex distribution, and disability. Thus Sydenham (1676), Landry-Beauvais (1800), Brodie, and others were in all probability describing rheumatoid arthritis in their writings, but it was Alfred Baring Garrod (1859) who first used the term 'rheumatoid' arthritis (Garrod 1859; and for reviews see Short 1974 and Fraser 1982).

The definition of rheumatoid arthritis and its separation from other forms of chronic polyarthritis did not end with Garrod, and has continued to evolve since. A proportion of patients who might previously have been diagnosed as having rheumatoid arthritis would now be readily reclassified as having polyarthritis seen in the context of ankylosing spondylitis, Reiter's syndrome, psoriatic arthritis, or inflammatory bowel disease. The uniform lack of rheumatoid factor (seronegativity), spinal involvement, and HLA-B27 positivity has linked these disorders into the so-called seronegative spondylarthropathies. It is of special historical interest to note that until the late 1950s, despite striking differences in clinical features, ankylosing spondylitis was termed rheumatoid spondylitis by North American physicians in the belief that it was part of the disease spectrum of rheumatoid arthritis. This view was essentially based on the striking histopathological similarity to rheumatoid arthritis of the synovitis and erosive arthropathy of diarthrodial joints in ankylosing spondylitis. However, the inflammatory lesion at the point of tendon and ligamentous attachments to bone (enthesopathy) and the association with HLA B27 proved to be sufficiently distinctive to constitute a basis for differentiation from rheumatoid arthritis.

The possibility that rheumatoid arthritis may result from an infection has had its proponents since the early part of the century. In the early days of modern medicine, rheumatoid arthritis, like other diseases of unknown causes, was thought to result from foci of infection (Hunter 1901; Wilcox 1935). The belief that rheumatoid arthritis may result from infection with *Mycobacterium tuberculosis* is alleged to have motivated Forrestier in France (Forrestier 1935) to use gold salts, which have some antimicrobial activity, in its therapy. An alternative concept of the aetiology of rheumatoid arthritis arose from microscopical observations of 'fibrinoid' change in rheumatoid joints and nodules. Fibrinoid change in connective tissue in systemic lupus and systemic sclerosis had prompted Klemperer *et al*. (1942) to consider that these diseases might result from diffuse primary degeneration of collagen. This led to the inclusion of rheumatoid arthritis in the group of 'collagen diseases'; however, the development of this new theory was hampered by the observation that the hydroxyproline and collagen content of subcutaneous nodules in rheumatoid arthritis was normal (Ziff *et al*. 1953).

The discovery by Waaler half a century ago (Waaler 1940) of IgM rheumatoid factor in the blood of patients with rheumatoid arthritis was the first immunological marker of rheumatic disease to be recognized and served to distinguish it from other forms of arthritis. However, a proportion of patients with the features of rheumatoid arthritis are persistently seronegative for rheumatoid factor, and this has formed the basis for a subdivision of rheumatoid arthritis, from which the recognized entity of seronegative spondylarthropathies arose. The patients who still remain in the category 'seronegative rheumatoid arthritis' may eventually prove to be a distinct subgroup with a different aetiological basis.

The finding of other autoantibodies circulating in the blood of patients with chronic polyarthritis has provided a continuing impetus for recognition of distinct rheumatic disorders. Notable, and first among these, was the LE cell test described by Hargraves *et al*. (1948), dependent on the presence of antinuclear antibodies, which served to distinguish patients with the polyarthritis of systemic lupus from those with rheumatoid arthritis. The better definition of the antigens with which such autoantibodies react has led to greater confidence in their diagnostic specificity; thus, antibodies to double-stranded-DNA and Sm antigen will distinguish systemic lupus from rheumatoid arthritis with some certainty. Similarly, more contemporary laboratory tests have contributed to the serological basis of disease classification of a patient with polyarthritis superficially resembling rheumatoid arthritis. The presence of anti-La (anti-SS-B) antibodies raises the strong probability of a diagnosis of primary Sjögren's syndrome, the presence of anticentromere or anti-Scl-70 antibodies suggests a diagnosis of systemic sclerosis, of anti-nRNP and anti-Jo-1 a diagnosis of overlap or mixed connective tissue disease, and of antineutrophil cytoplasmic antibody a diagnosis of Wegener's granulomatosis.

The relatively recent historical description of rheumatoid arthritis has prompted the speculation that it is a disease of modern times. A great deal of interest has therefore focused on seeking evidence of the occurrence of rheumatoid arthritis in mediaeval and ancient times. Examination of ancient medical writings and mediaeval paintings has yielded evidence which has satisfied some researchers that rheumatoid arthritis was indeed prevalent in these periods (Short 1974; Dequeker 1987; Dieppe 1988). However, subjectivity of judgement is an obvious problem in assessing such evidence and there appears to be a dearth of convincing descriptions, given that nowadays rheumatoid arthritis is such a common and ubiquitous cause of disability and pain. Another approach in establishing the antiquity of rheumatoid arthritis has been an attempt to gauge its prevalence by using palaeontographic methods. Fossil remains of archaic Indian skeletons found in Alabama and Kentucky in the United States, dated as several thousand years old, have been described as exhibiting changes consistent with an erosive arthritis compatible with rheumatoid arthritis (Rothschild and Woods 1990). However, the basis of ascribing the changes to rheumatoid arthritis has been challenged and the lack of rheumatoid pathology in a study of 800 skeletons excavated in the West Country in England has been used as an argument highlighting the lack of this disorder in ancient times (Rogers and Dieppe 1990).

The implication of the claim that rheumatoid arthritis is a relatively modern disease is the possibility that it might have become widespread as a result of an environmental trigger factor which in itself was new. A report from South Africa purporting to demonstrate an increased prevalence of rheumatoid arthritis in Xhosa

tribesmen living in urban surroundings compared with their cousins in rural areas (Solomon *et al*. 1975) has been interpreted as a recapitulation in a contemporary setting of a global scenario unfolding over the past two centuries. The insight that the epidemiology of acquired immunodeficiency disease (**AIDS**) has provided into how new diseases of man become established has provided an arena for renewed interest in the possibility that the environmental factor responsible could be an infectious agent.

Aetiology

Rheumatoid arthritis is a disease of unknown cause, but current thinking favours the notion that interplay among genetic factors, sex hormones, and an infectious agent initiates an autoimmune disease mechanism that culminates in a disease with inflammatory and destructive features.

Genetic factors

Genetic factors were implicated by population studies that showed a slight increase in the frequency of rheumatoid arthritis in first-degree relatives of patients with rheumatoid arthritis, especially if seropositive for rheumatoid factor (Lawrence 1970). Concordance rates of disease in identical twins in hospital-based studies were estimated to be of the order of 30 per cent, compared with 5 per cent in non-identical twins (Lawrence 1970), although the figures are lower in community-based studies (Silman *et al*. 1989), again supporting the concept of a genetic contribution, but arguing against the proposition that rheumatoid arthritis results from a dominant single-gene disorder. The rates of prevalence in the general population, families, and twins have in fact led to the conclusion that rheumatoid arthritis is a polygenic disease, and that non-inherited factors are also of great importance.

Attempts at identifying the genes involved in predisposition to rheumatoid arthritis took a step forward when tissue typing for HLA class II antigen of Caucasian patients showed that 60 to 70 per cent of patients with rheumatoid arthritis were HLA-DR4 positive by cellular or serological techniques compared with 20 to 25 per cent of control populations (Statsny 1976; Statsny 1978; Panayi *et al*. 1979). The patients with more severe rheumatoid arthritis, especially those with systemic complications such as vasculitis and Felty's syndrome, were even more likely to have HLA DR4 than patients with less aggressive disease confined to joints (Ollier *et al*. 1984; Westedt *et al*. 1986).

The increased frequency of HLA DR4 has also been reported in American black subjects, Japanese, Asian North Indians, and Latin Americans. In Israeli Jews and an Indian immigrant community in the United Kingdom an increased frequency of HLA DR1 has been found. The increase of HLA DR4 or DRl cannot, however, be found in all races and ethnic groups, and a notable exception was a study of Greek patients in whom no HLA associations could be discovered, irrespective of disease severity or serological status (reviewed by Goldstein and Arnett 1987).

Typing by mixed leucocyte culture has defined several HLA-DR4 subtypes. It is of considerable interest that while the subtypes HLA Dw4 and Dw14 are associated with rheumatoid arthritis in several studies, Dw15 is only associated with rheumatoid arthritis in the Japanese, while Dw10 and Dw13 are not associated with rheumatoid arthritis in any ethnic group. A recent study has shown that in *DR4* homozygotes, *Dw4*/*Dw14* individuals were at greater risk of developing rheumatoid arthritis than *Dw4*/*Dw4* (Wordsworth *et al*. 1992). The significance of this is discussed below.

At the phenotypic level, the importance of HLA class II molecules lies in their participation in a trimolecular reaction involving the HLA antigen-binding cleft formed by the a and b chains of an antigen-presenting cell binding to a processed linear peptide antigen of at least nine amino acids, and the HLA–antigen complex in turn binding to the variable portion of the T-cell receptor. Several research techniques have been used in an attempt to define the similarity of HLA class II molecules common to all patients with rheumatoid arthritis, including patients who are not necessarily HLA-DR4 positive. These include, for example, genotyping of DNA and nucleotide sequencing, using the polymerase chain reaction and enzymatic digestion for restriction fragment length polymorphisms. At the level of expressed surface proteins, HLA epitopes have been sought by using monoclonal antibodies and alloreactive T-cell clones (see Goldstein and Arnett 1987). These studies have lent support to the concept that susceptibility to rheumatoid arthritis is related to a 'shared epitope' on the HLA molecules (Gregersen *et al*. 1987; Hammer *et al*. 1995).

Nucleotide sequencing of *HLA-DR* b_1 exons coding amino acid residues 70 to 74 has revealed that HLA DR4 subtypes Dw4, Dw14, and Dw15 share similarities with each other (with a conservative substitution of glutamine with lysine at position 71 in Dw4) and with HLA DRl (Table 1) (Winchester and Gregersen 1988). The sequence predicts susceptibility to rheumatoid arthritis and, for example, is associated with rheumatoid arthritis in 83 per cent of Caucasians in Britain (Wordsworth *et al*. 1989). In contrast, negative associations are observed in individuals who are DR4w10, in whom the charged basic amino acids glutamine and arginine in positions 70 and 71 are replaced by the acidic amino acids aspartic and glutamic acid. In Dw13 individuals, in whom a negative association is also observed, arginine is substituted for glutamic acid in position 74. Molecular modelling studies suggest that amino acid residues 70 to 74 are located in the a-helix forming the wall of the peptide-binding groove, and thus likely to be involved in antigen binding and subsequent interaction with T-cell receptors (Fig. 1). Acidic substitutions could profoundly alter protein structures and thereby alter affinity for peptide antigens. The predictions that protein structures on the HLA molecule are important in susceptibility to rheumatoid arthritis are supported by serotyping with alloantisera and monoclonal antibodies as well as reactivity with homozygous T cells and T-cell clones (Goronzy *et al*. 1986; Winchester and Gregersen 1988). However, whether the susceptibility to rheumatoid arthritis is due to permissive binding of specific peptides such as those on autoantigens or on environmental antigens, whether superantigens may initiate disease by binding specifically to the HLA molecules (Herman *et al*. 1991), or whether selection or tolerance of the T-cell repertoire are also involved, remains to be elucidated.

DR type	Sequence					Association
	70	71	72	73	74	
DR4–W4	Q	K	R	A	A	Positive
–W14	Q	R	R	A	A	Positive
–W15	Q	R	R	A	A	Positive
DR1	Q	R	R	A	A	Positive
DR4–W10	D	E	R	A	A	Negative
–W13	Q	R	R	A	E	Negative

Q, glutamine; K, lysine; R, arginine; A, alanine; D, aspartic acid; E, glutamic acid.

Table 1 HLA-DR associations with rheumatoid arthritis defined by DR b_1 sequence position 70 to 74

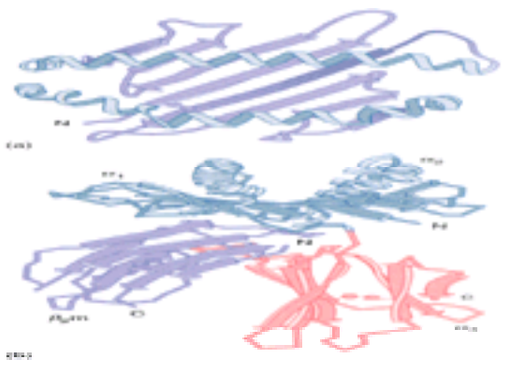


Fig. 1 Structure of HLA. (a) View of peptide-binding groove as seen by T-cell receptor. (b) From side.

The evidence that *Dw4*/*Dw14* heterozygotes are more likely to develop rheumatoid arthritis than are *Dw4* homozygotes (Wordsworth *et al*. 1992), and the evidence that individuals expressing DR4 (and not DR1) are more likely to have severe disease, challenges the concept that sequences 70 to 74 are the only HLA-D regions

that influence disease expression. This is despite the sequence identity of DR1 and DR4w4 in these positions, and a conservative substitution in Dw14 (Table 1). It has been hypothesized that the severity of disease and extra-articular complications are related to homozygosity and the density of disease-associated MHC molecules which critically influence the selection of the T-cell repertoire and tolerance to antigens (Weyand *et al*. 1992; Goronzy and Weyand 1993).

Although HLA genes are of obvious importance in rheumatoid arthritis, it has been calculated from studies of HLA in multicase families that they may only account for 37 per cent of the genetic factors involved (Deighton *et al*. 1989). However, a recent reanalysis has suggested that up to 60 per cent of susceptibility could be determined genetically (Macgregor and Silman 1994). Other susceptibility genes have been sought, and associations with T-cell receptors in gene polymorphisms and deletions of immunoglobulin genes have been observed (Olee *et al*. 1991). The picture is currently incomplete and their relative importance has not been documented.

Infectious agents

Population and twin studies strongly suggest that non-inherited, presumably environmental, factors such as infections may play a part in the aetiology of rheumatoid arthritis, although rearrangement of the a b genes on T-cell receptors may also contribute to the non-inherited component. As discussed above, although prevalence rates of rheumatoid arthritis from worldwide studies are similar, a population survey designed to seek effects of the environment showed that urbanized tribal South African black subjects suffered from rheumatoid arthritis more than their rural cousins. However, the classical clue of case clustering suggesting an infectious background has not been found for rheumatoid arthritis in any study, unlike Lyme disease. Attempts at demonstrating microbial organisms directly from joints have had a chequered history, since in all instances claims of positive findings—for example, of mycoplasma, diphtheroids, and viruses—have either been attributed to laboratory contamination or have been refuted on grounds of a lack of reproducibility.

Other studies have sought to implicate microbes in the aetiology of rheumatoid arthritis by seeking evidence of immune hyperreactivity to microbial antigens. Increased antibody titres to Epstein–Barr virus (**EBV**) antigens and an induced rheumatoid arthritis nuclear antigen have suggested that this ubiquitous virus, known to infect the majority of people by the late teens, may be of aetiological importance (Venables *et al*. 1981; Venables 1988). Its persistence in B lymphocytes of patients with rheumatoid arthritis in greater than normal amounts because of impaired cellular immunity could lead to the hyperreactivity of B cells, and autoantibody production typical of the disease. Sequence similarity of an EBV capsid antigen to the HLA-DR b susceptibility sequence QKRAA (Table 1) (Roudier *et al*. 1988) has suggested a possible explanation for the increased persistence of EBV in rheumatoid arthritis, while immunological cross-reactivity of EBV nuclear antigens to the autoantigens collagen, actin, and cytokeratin and to an antigen in the synovial membrane have suggested mechanisms for the induction of autoimmunity and for localization of immune cells to joints (Baboonian *et al*. 1991). However, these data fail to explain why only a small proportion of individuals infected with EBV might develop rheumatoid arthritis, and conversely, that there are well-documented patients with rheumatoid arthritis who have not been infected with EBV (Venables *et al*. 1981).

At least three bacteria have attracted attention in recent years as candidate agents in the aetiology of rheumatoid arthritis. The first, *M. tuberculosis*, gained current interest following the studies of Cohen *et al*. (1985) on an animal model of rheumatoid arthritis (see adjuvant arthritis in Animal models below). In these studies, mycobacterial protein showed immunological cross-reactivity and sequence similarity to a cartilage link protein, a finding which suggested a possible reason for the localization of the immune response to joints (van Eden *et al*. 1985). The link with human rheumatoid arthritis was suggested by the demonstration of reactivity of synovial T cells to mycobacterial antigens in rheumatoid arthritis (Holoshitz *et al*. 1986). The mycobacterial 65-kDa protein was subsequently shown to belong to the family of heat-shock proteins (**hsp65**) which are expressed in a variety of bacteria and also in the inflamed synovium of rheumatoid arthritis (van Eden *et al*. 1988; de Graeff-Meeder *et al*. 1990). Whether human hsp is a major target of T-cell autoreactivity in rheumatoid arthritis, however, is unproven, as there are significant differences in sequence and epitopes expressed by bacterial and human hsp (Gaston *et al*. 1989). Arguing against a role is the fact that identical responses to hsp65 are found in other inflammatory sites, for example, pleural effusion (Res *et al*. 1990). Attempts to suppress rheumatoid arthritis by vaccination with T cells derived from joints, in a protocol similar to that successfully used in adjuvant arthritis, could have provided support for the importance of mycobacterial immunity, but preliminary attempts have not been impressively successful (Van Laar *et al*. 1991).

In a further study seeking a mechanism dependent on molecular mimicry, antigens homologous to the amino acid sequence QKRAA (see Table 1) present in the hsp DNAj of *Escherichia coli* were reported to elicit T-cell responses only in patients with rheumatoid arthritis (Albani *et al*. 1995). It was suggested by these workers that activated T cells may cross-react with autologous DNAj heat-shock proteins that are expressed in the joints.

The third bacterium proposed as a candidate aetiological agent is *Proteus mirabilis*. Increased levels of IgG antibody to the organism have been detected in patients with rheumatoid arthritis but not ankylosing spondylitis or control subjects (Ebringer *et al*. 1989). It has been claimed that persistence of the organism in the urinary tracts, especially of women, may provide the nidus of infection that triggers a deleterious immune response culminating in rheumatoid arthritis.

The similarity of retrovirus-induced caprine arthritis to rheumatoid arthritis has attracted interest in the possibility that retroviruses may be of aetiological importance (Trabandt *et al*. 1992). Retroviral GAG proteins have been demonstrated immunohistochemically in the synovium of patients with rheumatoid arthritis (Ziegler *et al*. 1989), and a transgenic mouse carrying the human T-cell leukaemia virus type I that developed chronic arthritis with synovial inflammation and joint erosion similar to rheumatoid arthritis has been described (see Animal models).

Other aetiological factors

Apart from the possible role that infectious agents may play, the predominance of rheumatoid arthritis in females in the premenopausal period compared with males and the protective effect of the contraceptive pill, presumably because of its progesterone content, have suggested that sex hormones may accelerate or retard its onset (Lahita 1990). Other aetiological factors that have been considered include diet (Buchanan *et al*. 1991) and stress (Adler 1985), but their role in initiating disease is debatable, and may be more significant in altering disease expression and outcome.

Autoimmunity

Autoantibodies

The discovery of the autoantibody, IgM rheumatoid factor, in the blood of patients with rheumatoid arthritis was the principal reason for the inclusion in the group of autoimmune diseases. Although high-titre IgM rheumatoid factor is relatively specific for a diagnosis of rheumatoid arthritis in the context of chronic polyarthritis, its occurrence in many autoimmune rheumatic diseases without arthritis and in chronic infections has raised doubts about the role it might play in the pathogenesis of rheumatoid arthritis. However, rheumatoid factor-secreting plasma cells of IgG, IgA, and IgM class can be demonstrated in the rheumatoid synovium (reviewed by Maini *et al*. 1987), thus implicating them at the site of disease. Indeed, cells of the B-lymphocyte lineage constitute 10 to 15 per cent of the population of mononuclear cells in rheumatoid arthritis, produce autoantibodies, are a source of immune complexes that fix complement, and can act as efficient antigen-presenting cells. It seems likely that they contribute to perpetuation of the disease (Andrew *et al*. 1991).

IgG in patients with rheumatoid arthritis shows markedly reduced glycosylation, with a galactose 'pocket' in the Fc region (Parekh *et al*. 1985), in association with low levels of B-cell galactosyl transferase (Axford *et al*. 1987). It has been suggested that this glycosylation defect could result in conformational changes in the Fc region, rheumatoid factors more readily aggregating such molecules. Passive transfer of an acute synovitis in T-cell-primed mice has been shown to be enhanced using IgG containing autoantibodies to type II collagen when the antibodies are present as the agalactosyl glycoform. (Rademacher *et al*. 1994), demonstrating that agalactosyl IgG glycoforms are directly associated with pathogenicity in murine collagen-induced arthritis. However, the role of such glycosylation defects in the aetiology or pathogenesis of rheumatoid arthritis has yet to be established.

Other 'autoantibodies' that occur in rheumatoid arthritis include natural autoantibodies, antinuclear antibodies, anticollagen antibodies, antikeratin antibodies, and an IgG perinuclear factor. Of these, antibodies directed to two distinct epidermal antigens appear to show high diagnostic specificity for rheumatoid arthritis with a sensitivity of about 50 per cent. These are:

1. IgG antikeratin antibodies, which are present in 36 to 60 per cent of patients with rheumatoid arthritis, show a specificity of over 95 per cent in most studies. The antibody activity is directed against an antigen in the keratinized stratified epithelium of the rat oesophagus and is demonstrated by indirect immunofluorescence (Young *et al*. 1979). As a proportion of patients with rheumatoid arthritis without rheumatoid factor were positive in this test, it may be viewed as an additional serological marker of rheumatoid arthritis. Although it has been claimed that the antigen is an epidermal cytokeratin, the supporting data are poorly substantiated, and the identity of the antigen remains unknown.

2. An antibody directed against another epidermal antigen is known as antiperinuclear factor: this is demonstrated by an indirect immunofluorescence technique using buccal mucosal epithelial cells as substrate (Nienhuis and Mandema 1964). These IgG antibodies in rheumatoid arthritis are directed against spherical cytoplasmic granules and, when undiluted serum is used, show a diagnostic specificity of 98 per cent, with a sensitivity of 52 per cent (Westgeest *et al*. 1987). However, antiperinuclear factor has been found by some in a significant proportion of patients with Sjögren's syndrome, systemic lupus, systemic sclerosis, infectious mononucleosis, and metastatic lung cancer. It has been claimed that a positive test for antiperinuclear factor occurs in rheumatoid arthritis in the absence of IgM rheumatoid factor and therefore is of value in diagnosis. The biochemical and molecular properties of the antigen reactive with antibodies to antiperinuclear factor are also poorly characterized. Recent studies have shown colocalization of antiperinuclear factor and profilaggrin in human buccal cells, and evidence has accumulated suggesting that antibodies to antiperinuclear factor and stratified epithelial keratin recognize epitopes on profilaggrin (Berthelot *et al*. 1994; Sebbag *et al*. 1995).

Antinuclear antibodies detected by indirect immunofluorescence occur in up to 40 per cent of sera from patients with rheumatoid arthritis. Antibodies to histones, which also react as rheumatoid factors as a result of an epitope shared with IgG-Fc, have also been described in rheumatoid sera (Hannestad and Stollar 1978). Precipitating antibodies to soluble cellular antigen have been described in rheumatoid vasculitis (Venables *et al*. 1979). An antibody detected by Western blotting to a ribonucleoprotein termed RA33 has been found in 36 per cent of rheumatoid sera, including early in disease, but also occurs in sera from mixed connective tissue disease and systemic lupus. Partial sequencing of RA33 shows it to be identical to the A2 protein of the heterogeneous nuclear ribonucleoprotein (**hnRNP**) complex (Steiner *et al*. 1992; Hassfeld *et al*. 1995).

All the foregoing examples of antibodies appear to be associated specifically with rheumatoid arthritis, but react with antigens that are not restricted to the site of disease in joints. However, another set of antibodies in rheumatoid arthritis is directed against antigens present in cartilage only, such as collagen type II, IX, and XI, and chondrocyte-specific antigens. The published data on the frequency of these antibodies are variable in different series, possibly as a result of the differing derivation of collagen (both homologous and heterologous collagens are used and are not identical), lack of purity of antigens, interference from serum factors, and the wide variety of techniques for detection. In one study, antibodies to collagen II occurred in 29 per cent of patients with rheumatoid arthritis, while antibodies to type IX and XI were present in 40 per cent (Charriere *et al*. 1988). However, antibodies to collagen II and XI were also equally frequent in osteoporosis and Paget's disease, whereas anticollagen IX was relatively restricted to rheumatoid arthritis. Antibodies to collagen I and II are produced locally in rheumatoid joints (Tarkowski *et al*. 1989). Antibodies to chondrocyte membrane antigens occur in rheumatoid arthritis but have been poorly characterized so far (Mollenhauer *et al*. 1988), but a preliminary report suggests that a glycoprotein synthesized by chondrocytes is a specific T-cell autoantigen in rheumatoid arthritis (Rijnders *et al*. 1996).

Induction

In the context of autoimmune diseases in general, as in rheumatoid arthritis, environmental agents are seen as triggers rather than as being directly involved in the disease process. However, how environmental agents induce autoimmunity is not understood. Various hypotheses have been proposed of which the concept of 'antigenic mimicry' is the most popular.

'Antigenic mimicry' implies that an immune response to an extrinsic antigen (usually microbial), closely resembling an autoantigen, induces an immune response that cross-reacts with the autoantigen. If the response is to be long lasting, then the autoantigen must perpetuate it as the extrinsic antigen is eliminated. Despite the popularity of this concept, there are as yet no definite examples in human autoimmunity. Mimicry can occur in autoimmunity, it is the mechanism by which heterogeneous or chemically treated autoantigens can induce experimental autoimmune diseases, for example thyroiditis by using thyroglobulin, or collagen arthritis (see below).

Another concept, proposed by Bottazzo, Feldmann, and colleagues (reviewed by Feldmann 1987; Feldmann 1989), was that a local immune response, to any environmental agents, may release enough cytokines into the environment to upregulate local antigen-presenting capacity, so allowing autoantigens, otherwise 'hidden' from the immune system because of lack of HLA class II expression, to be presented to immunocompetent T cells that have escaped elimination or induction of tolerance. This was first proposed for endocrine autoimmune diseases, with the suggestion that the endocrine epithelium becomes the critical source of (atypical) antigen-presenting capacity and of autoantigen. Substantial evidence has since accumulated that this scheme may apply in both experimental models and human disease. Transgenic mice, producing interferon-g in their islets of Langerhans under the control of the insulin promoter, develop an immune, T-cell-dependent diabetes, with autoreactive T cells lysing islets and rejecting transplanted islets (Sarvetnick *et al*. 1990). In human Graves' thyroiditis, the antigen-presenting capacity of thyrocytes has been documented, as well as the presence of activated autoantigen-reactive T cells, and of local cytokines needed to maintain both antigen-presenting function and T-cell activation (reviewed Feldmann *et al*. 1991). In rheumatoid arthritis, abundant antigen-presenting function resides in macrophages, dendritic cells, B cells, endothelium, and possibly activated T cells, although which of these is most deeply involved in antigen presentation in rheumatoid arthritis is not known. The presence of CD5+ B lymphocytes and their descendants may contribute significantly to local antigen-presenting function by binding to autoantibody containing immune complexes in their immunoglobulin receptor (Andrew *et al*. 1991).

What are the important autoantigens in rheumatoid arthritis? In a local autoimmune disease the autoimmune response is localized by the restricted distribution of critical autoantigens. This can be shown in Graves' disease where antigens synthesized by thyroid epithelial cells—thyroglobulin, thyroid peroxidase, and thyroid-stimulating hormone receptor—are targets of both T- and B-cell recognition (Dayan *et al*. 1991). In rheumatoid arthritis, cartilage autoantigens such as collagen type II, type IX, and type XI recognized by T and B cells would fulfil this role. These antigens as well as other cartilage- or chondrocyte-specific antigens could be of importance in the initial localization to synovial joints. A report of benefit following daily intake of a preparation of purified chicken type II collagen, has excited interest in the possibility of induction of 'by-stander' T-cell tolerance by regulating T cells to joints from the gut lymphoid system (Trentham *et al*. 1993). T cells recognizing hsp65 or the antigen implicated in the autologous mixed lymphocyte reaction, which have been described in rheumatoid arthritis, could not have this role because the antigens are ubiquitous in cell types in most tissues, but may be of importance in maintaining the disease process, and in the extra-articular manifestations.

It is not clear whether rheumatoid arthritis should be considered as a single disease, with all cases having the same aetiology, or whether it should be viewed as a syndrome, with a range of aetiological factors initiating the same pathogenetic mechanism, and so producing a similar constellation of features.

Pathology

Introduction

The most pronounced and invariant pathology is in the synovial joints. There is a typical distribution, the small joints of the hands and feet, knees, and hips being most often implicated, symmetrically. In the different joints there are minor differences in pathology, but there is an overall pattern. There are also extra-articular manifestations, such as nodules and systemic disease.

Involvement of synovial joints

While attempts have been made to study the early events in rheumatoid arthritis, this is difficult, and so the pathology that is well known is from established cases. The involvement of synovial joints in rheumatoid arthritis is both of the synovial fluid and membrane (Zvaifler *et al*. 1994). Synovial fluid volumes are increased, and the cellularity increased; the predominant cell is the polymorph, which is only rarely seen in the lining layer of the synovial membrane. The other major cells in the synovial fluid and membrane are T cells and macrophages, with dendritic cells and cells of the B-lymphocyte lineage in small numbers. Typical numbers in acute cases are about 10^6/ml of polymorphs, and 1 to 3 × 10^5/ml of mononuclear cells. The exact relationship of the cells in the fluid to those in the membrane is not clear. Those in the fluid originate from the membrane, but how they reach the fluid is not clear. Whether they can re-enter the membrane or directly damage cartilage is also not known.

The involvement of synovial membrane is summarized in Fig. 2. There are several key features:

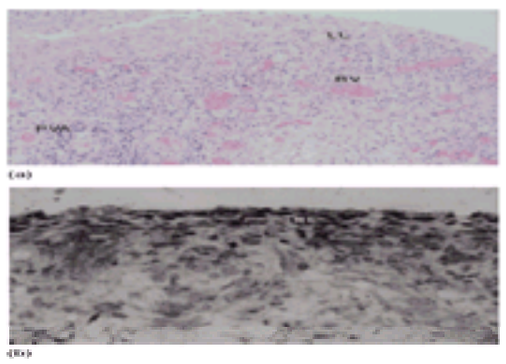


Fig. 2 Synovial membrane from a rheumatoid joint. (a) Haematoxylin–eosin staining of paraffin-embedded tissue (original magnification × 100) showing lining layer (LL) hypercellularity, prominent blood vessels (BV), and perivascular aggregates (PVA) of lymphocytes. The perivascular T lymphocytes are predominantly CD4+, CD45RO+, and CD29+ and a proportion bear activation markers HLA DR and IL-2 receptors. CD8+ T cells are distributed in interaggregate areas, as are plasma cells. (b) Tumour necrosis factor-a (TNF-a), IL-1a, b, and IL-6 are located in LL and in deeper layers: this cryostat section stained with $F(ab')_2$anti-TNF-a and developed with an immunoperoxidase method shows intracytoplasmic TNF-a.

1. The lining layer, normally two cells thick, is much thickened with increased numbers of both type A (macrophage-like) cells and type B (fibroblast-like) cells, both expressing activation markers.
2. The deeper layers are of increased cellularity, with perivascular accumulations and follicles. These are rich in T cells particularly CD4+ cells. CD8+ T cells are more frequently found in between perivascular accumulations, as are the abundant plasma cells and infrequent B cells. Macrophages are found in the follicles and in between. There are few polymorphs and dendritic cells in the membrane, the majority of which accumulate in fluid.
3. The rheumatoid synovium is particularly vascular. There are markedly increased numbers of vessels, and in some instances high-endothelial venules develop, as in lymph nodes.
4. Many of the cells, of all types, in the rheumatoid joint are activated. Thus HLA class II expression is found on nearly all the cell types, at an increased level, compared with that in normal or osteoarthritic joints. T cells are about 50 per cent class II positive, providing strong evidence of their activation status. B cells are positive, but typically plasma cells are not, as these lose the capacity to express class II. Macrophages express class II, as is often the case when activated, and class II-expressing fibroblasts and endothelial cells can also be seen.

Of interest is the HLA-DQ expression in rheumatoid arthritis, which is significantly greater than in other types of joint inflammation, for example Reiter's syndrome (Barkley *et al*. 1989a). The meaning of this difference is not clearly understood, as the relative roles of the commonest class II antigen, HLA DR, compared with the less common DQ and DP molecules are not known. Certain evidence allies HLA-DQ-restricted T cells to the suppressive immunoregulatory lineage (Sasazuki *et al*. 1986).

Other markers of activation abound. On the macrophage lineage, expression of CD11b (CR3) is increased, as is the related CD11c (p150/95). CD11a (lymphocyte function associated antigen-1) is increased on macrophages and many cell types. On T cells, expression of very late antigen (**VLA**) is increased, as is class II on a major proportion. In contrast interleukin (**IL**)-2 receptor is much less apparent. Endothelial cell expression of the adhesion molecules (**AMs**) **ICAM-1**(intercellular), VLA-1, and **ELAM-1** (endothelium–leucocyte) is increased. Tumour necrosis factor (**TNF**) receptors, also markers of activation, are upregulated in rheumatoid joints and are detectable on more than 80 per cent of T cells, on cells of the lining layer, and on cells at the cartilage–pannus junction (Deleuran *et al*. 1992; Brennan *et al*. 1995).

A common feature of activated cells is their increased production of cytokines and expression of cytokine receptors. This is the case in rheumatoid synovium, and the details will be discussed under pathogenesis.

Pannus

The junction between synovial tissue, cartilage, and bone is the site of early erosive damage in rheumatoid arthritis. This site becomes filled and overlaid by vascular tissue termed pannus. The lining layer of pannus is in continuity with the lining layer of hypercellular synovium and has been regarded as being derived from it. The cellular pannus forms a distinct junction with underlying cartilage (see Fig. 3), which shows many characteristics of degradation, including loss of matrix and water content, and chondrocyte depletion. The conventional view is that pannus has an invasive degradative effect on underlying cartilage, mediated by the secretions of enzymes such as metalloproteinases. This is associated with further loss of cartilage as a result of enzymatic destruction of matrix by chondrocytes themselves, coupled with a lack of synthesis of newly formed matrix. Pannus also appears to erode adjacent bone by a similar process involving degradation of bone matrix, but in addition involving active bone resorption by osteoclasts.

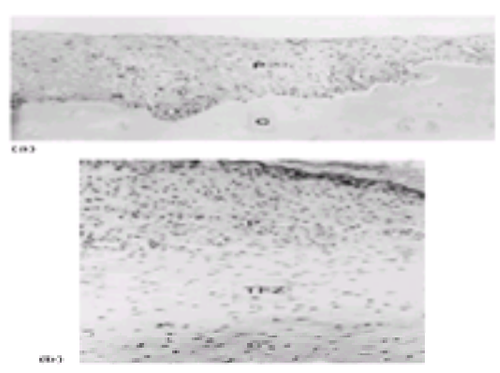


Fig. 3 Two types of cartilage–pannus junction seen in rheumatoid joints. (a) A distinct, well-defined margin can be seen between pannus (P) and cartilage (C). (b) A transitional fibroblastic zone (TFZ) separates a cellular, vascular pannus (P) from the underlying cartilage (C). Safranin O stain; original magnification × 46. (Reproduced with permission from Allard *et al*. 1987.)

A second type of pannus may also be observed, especially in the marginal cartilage area of weight-bearing joints. This consists of vascular pannus overlying cartilage with an indistinct, intervening, multilayered zone of fibroblast-like cells (Fig. 3). In contrast, the underlying cartilage of this pannus does not show degradative changes with loss of matrix. This type of pannus could represent a fibrotic healing phase, but has alternatively been termed a 'transitional' fibroblastic zone because the cytoplasm of these cells contains cartilage components such as keratan sulphate, chondroitin sulphate, and collagen type II (Allard *et al*. 1987; Allard *et al*. 1991). These cells may be derived from chondrocytes as a result of metaplastic change. However, the possibility has been raised that cells with the same phenotype resident in the subperiosteum, contiguous to synovium in normal joints, may give rise to the transitional fibroblastic zone in rheumatoid arthritis (Allard *et al*. 1990). The finding of proinflammatory cytokines capable of degrading cartilage (such as IL-l and TNF) in pannus cells contiguous with cartilage in the invasive type of erosion contrasts with absence of these cytokines in the transitional fibroblastic zone (Chu *et al*. 1991a; Chu *et al*. 1991b). Instead, the latter type of pannus shows the presence of only transforming growth factor-b (**TGF-b**), as this factor stimulates collagen and matrix production its presence is compatible with the proposal that the tissue is in an anabolic state of healing or differentiation.

Extra-articular manifestations

Local

These are more common in long-standing and severe cases. Rheumatoid nodules are the most common, and are found in areas susceptible to trauma, such as elbows. They consist of a palisade of macrophages surrounding fibrous tissue.

Systemic

There are disagreements about the extent and frequency of systemic manifestations, and whether rheumatoid arthritis is always manifest systemically. Elevated concentrations of acute-phase proteins such as C-reactive protein, serum amyloid A, or complement components are found in most cases, and these suggest that their production in liver is increased. IL-6 can activate the liver to produce many acute-phase proteins, and it is currently assumed that the increased production of IL-6 (and IL-1) in rheumatoid arthritis (reviewed Feldmann *et al*. 1996) is responsible.

In more severe cases there may be:

i. vasculitis;
ii. fibrosis of the lungs which may progress to significant fibrotic impairment of lung function;
iii. granuloma formation as characterized by nodule formation;
iv. serositis as characterized by pericarditis and pleurisy, commonly asymptomatic;
v. Felty's syndrome: enlargement of the spleen with lymphadenopathy, fever, leg ulcers, and susceptibility to bacterial infections.

Pathogenesis

Introduction

Describing the pathogenesis of a chronic disease, such as rheumatoid arthritis, for which there are no very accurate animal models, is difficult. It is possible to describe, on the basis of human studies, the events occurring when the disease is well established. Accurate description of early events is not possible, only informed speculation can be made. As the pathology—the morphological description of what has happened—has been discussed, consideration of the pathogenesis will be itemized in relation to how these changes may have evolved.

Cell recruitment

The vast majority of the increased number of cells in the rheumatoid joint are of lymphohaemopoietic origin, as shown by immunostaining techniques, and their presence in the rheumatoid joint implies that there are mechanisms for increasing cell input, and also for increasing retention. This increase in cellularity is accompanied by angiogenesis in the synovial membrane, thus increasing delivery of cells and molecules to areas of inflammation (Folkman 1995). Neovascularization involves angiogenic cytokines such as **VEGF** (vascular endothelial growth factor), an endothelial-specific mitogen which promotes the growth of new blood vessels (Colville-Nash and Scott 1992) and also renders the vasculature hyperpermeable *in vivo* (Ferrara *et al*. 1991). Much work has focused on the endothelium in rheumatoid arthritis, as blood-borne cells would first have to adhere and migrate through endothelium. Augmented expression of adhesion molecules capable of binding lymphocytes, polymorphs, and monocytes has been noted: ICAM-1, E-selectin, and **VCAM-1** (vascular cell adhesion molecule) are all increased at various pathological sites. Isolated rheumatoid synovial endothelial cells constitutively express ICAM-1 and E-selectin and this expression is upregulated by IL-1 and TNF (Abbot *et al*. 1992). Immunohistochemical techniques have shown that VCAM-1, ICAM-1, and E-selectin are highly expressed by rheumatoid synovial vascular endothelial cells and cells in the lining layer (Koch *et al*. 1992; Morales-Ducret *et al*. 1992; Wilkinson *et al*. 1993).

A local differentiation of T cells in rheumatoid arthritis synovial membrane has been suggested by the predominance of T cells with the phenotype of memory cells (CD45RO+, CD29). Increased expression of VLA-4 on CD45RO+ T cells could indicate selective migration of memory T cells into the inflamed synovial membrane. In fact CD45RO+ T cells have been shown to have a better adherence to endothelial cells than CD45RO- T cells (Pitzalis *et al*. 1987). In addition, synovial T cells in rheumatoid arthritis have a significantly greater capacity to migrate transendothelially compared with those from normal or rheumatoid arthritis peripheral blood (Cush *et al*. 1992).

Equally important in cell recruitment is the action of chemotactic factors, which promote the migration of cells into a site. Some of these mediators have been identified within rheumatoid joints including representatives of the two families of chemokines, the C-X-C(a) chemokines such as IL-8 (Brennan *et al*. 1990a), and the C-C(b) chemokines such as **RANTES** (regulated upon activation, T-cell expressed, and secreted) (Rathanaswami *et al*. 1993), **MCP-1** (monocyte chemoattractant protein 1) (Koch *et al*. 1992; Akahoshi *et al*. 1993), **MIP-1a**(macrophage inflammatory protein-a) (Koch *et al*. 1994), and MIP-1b (Villiger *et al*. 1992). As the majority of the cells in the rheumatoid arthritis synovium are macrophages and T lymphocytes, b-chemokines are likely to be important but neutrophil chemoattractants such as IL-8, **GROa** (melanoma growth-stimulating activity), and **ENA-78** (epithelial neutrophil activating peptide) are likely to play a role in neutrophil accumulation within the joint fluid. Split complement components C3a and C5a present in rheumatoid arthritis joints (Jose *et al*. 1990; Abbink *et al*. 1992) are also chemotactic for neutrophils.

T-cell activation in rheumatoid arthritis

The T lymphocyte is one of the most common cells in active rheumatoid arthritis, with an abundance ranging from 20 to 50 per cent of the cells extracted from synovial membrane. CD4+ cells are more abundant than CD8+ in the membrane, but not necessarily in the synovial fluid. The CD4+ cells tend to concentrate in perivascular nodules, whereas the CD8+ are more diffusely scattered. CD4+ cells have been subdivided into subsets, depending on their CD45 expression. In normal blood, about one-half are CD45RA+, indicating a 'virgin state'. Essentially all the cells in the rheumatoid joint lack CD45RA and express CD45RO/CD29, indicating a 'primed' or 'memory' state (Pitzalis *et al*. 1987). This is not surprising, as there is evidence for an ongoing immune response, as judged by the expression of T-cell activation markers, such as HLA class II (on 50 per cent), and IL-2 receptors on fewer cells (2 to 12 per cent) (Brennan *et al*. 1988a; Londei *et al*. 1989).

T lymphocytes may also be classified according to their T-cell receptor for antigen. In normal blood the great majority (more than 95 per cent) express a heterodimer of a and b chains, whereas a minority use g and d chains. Of interest was the observation (Brennan *et al*. 1988b) that there was selective enrichment of gd T cells in active rheumatoid joints. and that some of the gd T cells recognize mycobacterial antigens (Holoshitz *et al*. 1989). However, elevated gd cell numbers have not been confirmed in all studies. In patients who have augmented gd T cells in their blood there is a trend towards increased amounts of CD5+ B cells (Brennan *et al*. 1989c).

An important question is whether T cells have a critical role in rheumatoid arthritis. Firestein and Zvaifler (1990) , based on low or absent levels of T-cell cytokines in the rheumatoid synovial environment, have proposed that T cells may not be important in the chronic established phase of disease. Indeed, in controlled clinical trials of rheumatoid arthritis with anti-T-cell (for example anti-CD4, anti-CD5) monoclonal antibodies, no beneficial effects were observed (Olsen *et al*. 1994; Van der Lubbe *et al*. 1995). However, our opinion is that in a prolonged, chronic, asynchronous disease with profound immunoregulation, the quantity of cytokines detected need not reflect their importance. Various lines of evidence support this possibility. First is the abundance of T cells in rheumatoid joints. Virtually none are present in normal joints. Second is their partially activated status and proximity to antigen-presenting cells (see above). Third is the fact that the proportions of different types of T cells present in rheumatoid joints are not the same as in blood (as discussed above), indicating that it is not a reflection of passive trafficking in an inflammatory response. Fourth is the observation that antigen-specific T cells are present, are activated, and persist in rheumatoid joints. For example we found that collagen type II-specific T cells were present, and expressing IL-2 receptors, in three operative specimens in a patient with rheumatoid arthritis, over a period of more than 4 years (Fig. 4) (Londei *et al*. 1989). Finally, **Th1** cells (T-helper 1) appear to predominate in the joint and interferon-g and IL-2 are expressed, albeit at low levels, but with unexpectedly high IL-10 production (Buchan *et al*. 1988b; Simon *et al*. 1994; Cohen *et al*. 1995).

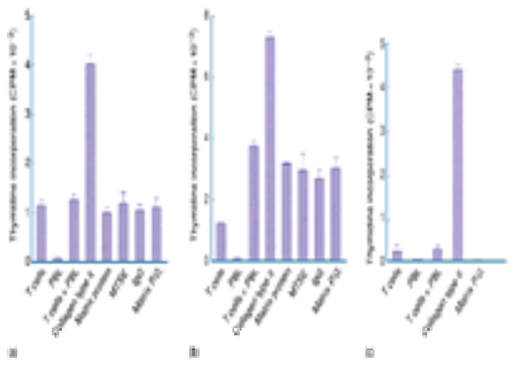


Fig. 4 Proliferative response of collagen type II-specific clones. (a) Clone 4 from the first synovial membrane. (b) Clone 55 with autologous mixed lymphocyte reactivity, from the first synovial membrane. (c) Clone B8 from the second synovial preparation. Results are the arithmetic means ± SEM of a representative experiment from each clone. In other experiments, collagen type I was also used, but there was no response from any of the clones. Cloned T cells (10^4), 2×10^4 irradiated, autologous, peripheral blood, mononuclear cells were cultured with or without the antigen indicated. Antigens were used at 100 µg/ml, except MTSE which was at 1 µg/ml. These were the optimal concentrations. PBL, peripheral blood lymphocytes; PG, proteoglycan; MTSE, mycobacterial tuberculosis soluble extract. (Reproduced with permission from Londei *et al*. 1989.)

A restricted pattern of V_b chains of the T-cell receptor was observed in rheumatoid joints by one group of workers, suggesting the possibility that activation of lymphocytes was mediated by superantigens (Paliard *et al*. 1991) However, other laboratories have not confirmed this observation. Thus, whereas Paliard *et al*. (1991) reported low V_b14 in blood of patients with rheumatoid arthritis compared with the level found in joints, Howell *et al*. (1991) reported that multiple *V*bgene families were transcribed in patients, although sequence similarities were found, in keeping with the hypothesis that superantigen may play a role in rheumatoid arthritis. Clearly much more work is needed in this area, but noting that all known autoimmune diseases and models are T-cell dependent, it is very likely that rheumatoid arthritis is T-cell dependent even in its later stages.

Whilst the function of T cells in rheumatoid arthritis is unresolved, much can be learned from investigations that define the role of T cells in induction and perpetuation of experimental models of arthritis. Collagen-induced arthritis is a CD4+ T-cell dependent disease as demonstrated by T-cell depletion (Ranges *et al*. 1985). However, unlike adjuvant arthritis, it is not readily transferred with T cells or T-cell clones. In the case of collagen-induced arthritis, transfer from histoincompatible DBA/1 mice to mice with subacute combined immunodeficiency has demonstrated that T and B cells act in synergy in the full expression of disease (Williams *et al*. 1992b; Taylor *et al*. 1995).

B-cell lineage

Plasma cells are abundant in rheumatoid arthritis. Some, but not all are involved in the production of rheumatoid factors. Rheumatoid factor immune complexes have been shown to induce the production of cytokines such as IL-1 and TNF (Fig. 5) (Chantry *et al*. 1989).

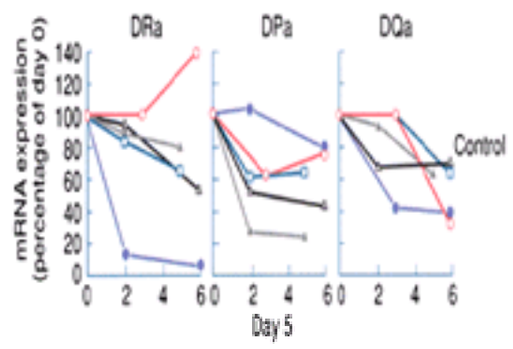


Fig. 5 Fresh isolated synovial membrane and synovial fluid cells obtained from five patients were placed in culture at 1×10^6/ml in the presence or absence of mediators (IFN-g and IL-2). Cells were harvested at the times indicated for the determination of cytoplasmic RNA. The results are expressed as percentages of the basal level. (●), Patient 1; (△), Patient 2; (▲), Patient 3; (□), Patient 4; (○), Patient 5. (Reproduced with permission from Kissonerghis *et al*. 1989.)

The specificity of the antibodies produced in rheumatoid joints has been investigated by the cell fusion technique using the human B-cell fusion partner, SPAZ4. Large numbers of hybridomas producing IgM and IgG were detected (Maini 1989). While a few of these were 'polyreactive', and some produced rheumatoid factors, the majority did not bind to a battery of autoantigens tested. However, it has been claimed that the majority of rheumatoid joints contain B cells producing antibody to collagen type II and IgG Fc (Tarkowski *et al*. 1989).

In examining the B-cell repertoire activated in rheumatoid joints, attempts are being made to ascertain whether there is any evidence of a restricted use of certain genes selected from among the multiple heavy-chain *V* genes available in the genome. In one such study, analysed by northern blotting, an overrepresentation of V_H4 was noted (Brown *et al*. 1992). Such overrepresentation may result from dominance of a B-cell subset in diseased tissues or from selection pressures created by specific antigens or superantigens. Alternatively, regulatory elements in flanking regions active in rheumatoid arthritis may favour recombination of particular individual gene elements and so skew the B-cell repertoire activated (Brown *et al*. 1995). As primed B cells recognizing antigen present antigen to T cells more efficiently than do macrophages (Lanzavecchia *et al*. 1985) and are probably important in the development and maintenance of the immune network in neonatal and adult life (reviewed by Plater-Zyberk *et al*. 1992), a greater understanding of the role of B cells should illuminate the pathogenesis of rheumatoid arthritis.

Antigen-presenting cells

There are abundant cells with antigen-presenting capacity in human rheumatoid joints. Which of these are of major importance in different stages of disease is a controversial question. Macrophages and monocytes represent some 30 to 50 per cent of the cell pool, and there is evidence for their activation, for example increased expression of HLA DQ, and diminished CD14. Dendritic cells are present in increased numbers. Regrettably, due to lack of specific markers for human dendritic cells, their numbers are not easy to quantify. However, cell separation studies by several groups have all demonstrated increased numbers of dendritic cells in rheumatoid synovial fluid, comprising up to 5 to 7 per cent of the mononuclear cells, whereas synovial tissue contained few dendritic cells (March 1987; Tsai *et al*. 1989). Dendritic cells from rheumatoid synovial fluid were potent antigen-presenting cells, but not more so than normal dendritic cells.

CD5+ B cells have Fc receptors and many produce rheumatoid factors. These may permit CD5 B cells to take up immune complexes and present the relevant antigens. The possible importance of CD5+ B cells in antigen presentation in rheumatoid arthritis has been discussed (Maini 1989). Chondrocytes can be activated to express HLA class II and may be critical in the early events of rheumatoid arthritis. ICAM-l expression in chondrocytes, which facilitates antigen-presenting cell function, has been reported (Davies *et al*. 1992).

Cell interaction

The importance of cell interactions in the rheumatoid joint can be inferred from the immunohistological studies, which show close apposition of T cells and antigen-presenting cells in nodules and in other sites throughout the synovial membrane. However, there are very few T cells in the pannus, suggesting that different

interactions may prevail in this specialized site.

Dissociated cells from rheumatoid joints, placed in tissue culture, in the absence of any extrinsic stimulus, rapidly reform into aggregates. This suggests that interactions are of critical importance in the disease process. Experimentally, one can demonstrate that these cell interactions are of importance *in vitro*. We have noted that rheumatoid synovial cells, placed in culture and in the absence of extrinsic stimulation, retain many features of active rheumatoid arthritis. Thus expression of HLA class II persists *in vitro*, at both the protein and mRNA levels, provided that the whole mixture of joint cells is cultured (Fig. 5). If only the adherent cells (chiefly fibroblasts) are cultured, class II expression apparently does not persist in culture (Teyton *et al*. 1987). Below, the persistence of cytokine production in cell cultures from rheumatoid joint not extrinsically stimulated is discussed.

The role of T cells in the persistence of class II expression has been studied by depleting T cells using a combination of lysis with antibody and complement, and antibody-coated magnetic beads. Even with an incomplete depletion of cells, a marked reduction in class II expression was noted after 6 days in culture (C.M. Hawrylowicz *et al*., unpublished observations). This emphasizes the importance of cell interactions, but does not clarify which T cells are of critical importance, nor which are the critical antigen-presenting cells.

Cytokine expression

As rheumatoid arthritis is mostly manifest in synovial joints, which are the sites of inflammation and destruction, cytokine production in the joints has been investigated by several groups. However, cytokines can also be detected in blood cells by immunostaining, for example IL-1a (Barkley *et al*. 1989b). Whether other cytokines can also be detected in the blood cells remains to be established. Elevated serum concentrations of cytokines have been reported, for example, IL-l b (Eastgate *et al*. 1988), but their reproducibility and significance remain to be established in view of the presence of serum cytokine inhibitors.

Rheumatoid joints contain a wide variety of activated cell types, and so it would be expected that many cytokines would be produced locally in the joint. When we began studies on the expression of cytokines in rheumatoid joints in 1985, slot blotting and cDNA hybridization were used to obtain maximum data on cytokine expression from a small number of cells. With these techniques 2×10^6 cells were used, and could provide data on about 6, or sometimes up to 10 cytokines. By densitometry, relative quantification was possible. Further advantages of this technique are its specificity for individual cytokines, ease of performance (same technique for all cytokines), and its resistance to artefacts caused by rheumatoid factors (a problem in binding assays), and to 'toxic' components of synovial fluid in bioassays (Buchan *et al*. 1988a; Buchan *et al*. 1988b). A disadvantage is that the amounts of cytokine mRNA being measured do not always correlate with the amounts of cytokine protein produced. This is especially a problem with cytokines known to be regulated post-transcriptionally, for example TNF-a. If this type of work was to be begun again, obviously the polymerase chain reaction (**PCR**) would be used, for its much greater sensitivity (Brenner *et al*. 1989). However, the use of PCR is not without problems; for example quantification is very difficult and contamination frequent. Using PCR to explore cytokine expression in rheumatoid arthritis has yielded the same results as slot blotting, that is, predominance of IL-1a and abundance of TNF-a (Brennan *et al*. 1989a). *In situ* hybridization has yielded analogous results (Firestein *et al*. 1990), and also provides information about localization, but quantification is difficult.

It is not surprising that virtually all the cytokines sought have been detected because of the wide variety of activated cells. Table 2 summarizes the cytokine expression in the rheumatoid joint. There are some interesting generalizations that can be made. For example, cytokines that are predominantly macrophage products are abundant, at both the mRNA and protein level, for example IL-1, IL-6, TNF, and IL-8. In contrast, cytokines produced by T cells are detectable at the mRNA level, but barely detectable at the protein level, for example interferon-g, lymphotoxin, and IL-2. The reasons for this discrepancy are not yet known, but TGF-b which inhibits cytokine production post-transcriptionally may be responsible. Local consumption of cytokines by cells with high-affinity receptors may also contribute, as, for example, there are free and cell-bound IL-2 receptors in rheumatoid joint cells (Symons *et al*. 1988).

Cytokine	mRNA	Protein
IL-1α	+	+
IL-1β	+	+
TNF-α	+	+
LT	+	–
IL-2	+	–
IL-3	–	–
IL-4	?	–
IFN-γ	+	–
GM-CSF	+	+
G-CSF	+	?
M-CSF	–	+
TGF-β	+	+
EGF	+	+
PDGF-A	+	+
PDGF-B	+	+

EGF, epidermal growth factor; G-CSF, granulocyte colony-stimulating factor; GM-CSF, granulocyte–macrophage CSF; IL, interleukin; LT, lymphotoxin; M-CSF, macrophage CSF; PDGF, platelet-derived growth factor; TGF, transforming growth factor; TNF, tumour necrosis factor.
After Brennan *et al*. (1992).

Table 2 Summary of cytokines produced by rheumatoid synovial cells

Cytokines are essential for many processes in rheumatoid arthritis, such as cell growth and expression of HLA class II, reviewed in Feldmann *et al*. (1996). However, it is not clear which cytokines are of major importance in different processes. A critical step in the generation of an immune or inflammatory reaction is activation of macrophages and induction of HLA class II expression. Interferon-g is potentially the most effective cytokine at inducing such expression in the absence of other factors (Portillo *et al*. 1989). However, negligible amounts of interferon-g (or other T-cell lymphokines) are produced by rheumatoid synovial cells (Firestein and Zvaifler 1987; Brennan *et al*. 1989a) suggesting that other factors alone or in combination with this interferon-g are involved. One possible candidate is the haemopoietic growth factor **GM-CSF** (granulocyte–macrophage colony-stimulating factor) which induces HLA-DR expression on human monocytes (Chantry *et al*. 1990) and which could be an important macrophage activator and induce HLA class II expression in the rheumatoid joint (Alvaro-Garcia *et al*. 1989). However, the most significant inhibition of that expression which we observed in the rheumatoid synovial cultures was with anti-TNF antibody (unpublished observation) and was greater than that with antibodies to interferon-g or GM-CSF. This is unlikely to be a direct effect, as TNF by itself does not induce the expression of HLA class II (for example Pujol-Borell *et al*. 1987). This suggests that many different cytokines may work together to induce this expression (Sadeghi *et al*. 1992a; Sadeghi *et al*. 1992b) or that other, as yet undefined, molecules may be involved. Alternatively (or in addition) cell–cell interactions through cell adhesion molecules may be necessary to maintain this. Of interest is the observation that TNF-a is a potent inducer of many adhesion molecules including ICAM-1 and VCAM-1 (Pober *et al*. 1986; Rice and Bevilacqua 1989).

The activation and differentiation of B cells is also mediated by cytokines, of which IL-4 and IL-6 are the most important. IL-4 is a potent B-cell growth factor but is detected in negligible amounts in rheumatoid synovial cells (unpublished observation) or in synovial fluid. In contrast, high levels of IL-6 have been detected both in rheumatoid synovial fluid and in cells from rheumatoid synovial membrane (Hirano *et al*. 1988; Field *et al*. 1991). The presence of high levels of IL-6 in rheumatoid joints may explain the large numbers of plasma cells and few B cells in the synovium and the production of autoantibodies including rheumatoid factors. Lymphotoxin and TNF-a can also act as a B-cell growth factor (Kehrl *et al*. 1987). The presence of immune complexes containing rheumatoid factor may further contribute to the pathogenesis of rheumatoid arthritis by inducing the production of IL-1 as shown by Chantry *et al*. (1989) (Fig. 6). The 'cytokine synthesis inhibitor', IL-10, which inhibits T-cell production of interferon-g, is also a potent B-cell stimulator (Moore *et al*. 1990).

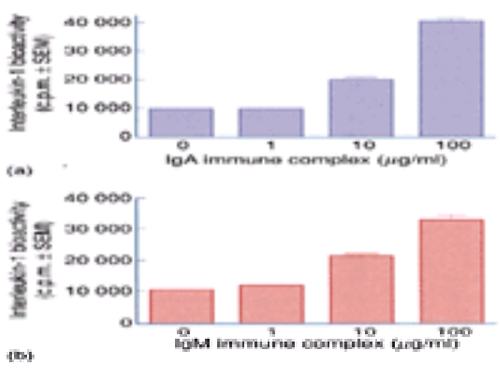


Fig. 6 Both IgA-containing immune complexes (a) and IgM-containing immune complexes (b) induce interleukin 1. Supernatants from monocytes cultured with various concentrations of immune complex for 24 h were assayed for interleukin1 bioactivity using the thymocyte comitogenic assay. Data is shown as [^{3}H]-thymidine incorporation (mean ± SEM of triplicate cultures). Proliferation due to phytohaemagglutinin antigen alone was 3486 ± 450 c.p.m. For both immune complexes significant (p<0.01) interleukin could be detected at concentrations as low as 10 µg/ml. (Reproduced with permission from Chantry *et al*. 1989.)

T-cell growth is controlled by cytokines. For many years, after the discovery of 'T-cell growth factor' and the purification, cloning, and expression of IL-2, it was thought that all T-cell growth was mediated by IL-2. Subsequent work has shown that the position is much more complex. IL-4, initially, described as a B-cell stimulating factor, is also a potent growth factor for many T cells, and IL-7, described as a pre-B growth factor, is also highly active (for example Londei *et al*. 1990). T-cell activation is found in rheumatoid arthritis, with many cells expressing HLA class II (20 to 50 per cent), and a few (2 to 12 per cent) expressing IL-2 receptors. However, the mechanism of T-cell growth is unclear, as while IL-2 mRNA is found (Buchan *et al*. 1988b), the protein is not readily detectable. This could be due to absorption by cell-bound receptors, to IL-2 inhibitors such as the soluble IL-2 receptor (Symons *et al*. 1991), or to post-transcriptional regulation. IL-4 is also not readily detectable, for possibly the same reasons. Currently it is unknown whether there is IL-7 in rheumatoid joints, but it is clearly an important candidate for T-cell growth regulation in rheumatoid arthritis. A synergy between all these cytokines could permit T-cell growth in the presence of low protein levels of each of these mediators. This possibility requires investigation using cells from rheumatoid joints in culture.

Fibrosis is an important component and complication of rheumatoid arthritis. It participates in deformation of joints, but pulmonary fibrosis can be a damaging systemic complication. Which cytokine drives the fibrosis in the rheumatoid joint (or other tissues) is not currently known. There are abundant candidates present in the rheumatoid joint, for example IL-1a and b, TNF-a, which may act indirectly via induction by platelet-derived growth factor (Raines *et al*. 1989), and TGF-a and -b. The presence in rheumatoid joints of members of the fibroblast growth family is not known, but is likely.

Cytokine regulation

Initial studies of the expression of IL-1 in rheumatoid joints revealed that all samples contained IL-1 mRNA. After the experimental activation of normal cells *in vitro*, that of IL-1 mRNA (and the expression of other cytokines) is brief (24 to 48 h), so the fact that all samples from rheumatoid arthritis were positive suggested that cytokine production in the rheumatoid joint may be relatively stable and persistent. In a chronic disease only persistent features can be relevant to the maintenance of the disease process, so the consistence and persistence of cytokine production suggested that it was of importance in the pathogenesis (Buchan *et al*. 1988a).

Cytokine persistence was directly tested *in vitro* by culturing dissociated cells from rheumatoid joints in the absence of extrinsic stimulation. The initial results showed that both IL-1a and IL-1b mRNA persisted for up to the 5-day culture period (Fig. 7) (Buchan *et al*. 1988a). This indicates that the signals necessary to regulate cytokine production are present in the culture, and so can be analysed.

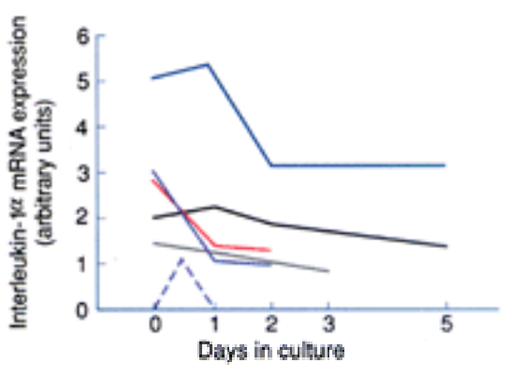


Fig. 7 Persistence of interleukin-1a mRNA production in rheumatoid joint cells in culture. Slot blot analysis of interleukin-1a production by rheumatoid arthritis synovial fluid mononuclear cells cultured in the absence of extrinsic antigen. SF and SM mononuclear cells were cultured for 0 to 5 days, the RNA extracted, blotted on to nitrocellulose, and probed with IL-1a. Integral values were calculated. Different symbols represent different patients. .....is interleukin 1a mRNA from mitogen-activated peripheral blood mononuclear cells. (Reproduced with modification from Buchan *et al*. 1988a.)

Neutralizing antibodies were chosen as the tool to investigate the signals involved in regulating the production of IL-1. The strongest non-microbial signals for the regulation of this production were the cytokines TNF-a and TNF-b (lymphotoxin), so neutralizing antibodies to these two cytokines were used. The results were clear cut; anti-TNF-a but not anti-TNF-b or control rabbit Ig inhibited IL-1 production after the first day of culture (Fig. 8(a)) (Brennan *et al*. 1989b). Assays at the mRNA level show more rapid kinetics, but the lack of an early effect on the amount of IL-1 protein indicates that already ongoing synthesis of IL-1 was not affected, but that subsequent activation was blocked. As a control, the same antibodies were used on cultures of cells from osteoarthritic joints. Despite the presence of immunoreactive TNF-a there was no effect of anti-TNF-a on the low levels of IL-1 in osteoarthritis (Fig. 8(b)). This is now known to be caused by the TNF in osteoarthritis not being biologically active.

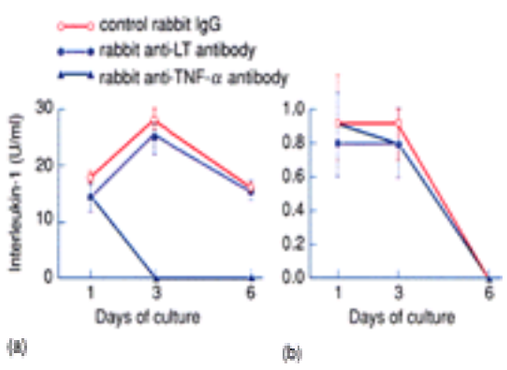


Fig. 8 (a) Effect of anti-TNF-a on rheumatoid arthritis joint cell culture. (b) Lack of effect of anti-TNF-a on osteoarthritis joint cell cultures. (Reproduced with modification from Brennan *et al*. 1989a.)

The finding that TNF-a was the single dominant signal regulating the production of IL-1 was surprising. It had been anticipated that multiple signals may be of importance, including immune complexes and perhaps other non-cytokine signals. However, samples from the first seven patients behaved in this way, regardless of their therapy. This led us to investigate what other effects of anti-TNF-a on the disease process may be. It has been postulated that GM-CSF is an important cytokine in rheumatoid arthritis, as it is an inducer of class II on monocytes, and induces cytokine production and macrophage activation (Alvaro-Garcia *et al*. 1989). It was therefore of interest to determine which cytokine regulates the production of GM-CSF in cell cultures from rheumatoid joints. Anti-TNF-a markedly inhibited the production of GM-CSF, but more slowly than inhibition of IL-1, being virtually complete only by day 5 (Haworth *et al*. 1991). We have also found that anti-TNF partially inhibits class II expression and also the aggregation normally found in these cultures. A summary of the effects of TNF in rheumatoid joints is shown in Fig. 9.

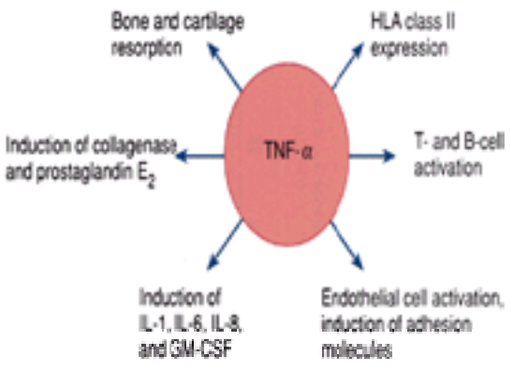


Fig. 9 Effects mediated by TNF-a in rheumatoid arthritis.

The results obtained in rheumatoid arthritis are analogous to those now described in the response of mice to systemic Gram-negative bacteria. Production of TNF-a, IL-1, and IL-6 was monitored; peaks of TNF-a preceded those of IL-1 then IL-6. Anti-TNF-a abrogated the production of IL-1 and IL-6 in this animal model (Fong *et al*. 1989). Thus it seems likely that in rheumatoid arthritis, the dominant position of TNF-a recapitulates the physiological situation.

Following the demonstration that anti-TNF can inhibit the production *in vitro* of IL-1 and other proinflammatory cytokines (IL-6, GM-CSF, IL-8) (Feldmann *et al*. 1996) and the successful amelioration of collagen-induced arthritis in DBA/1 mice by use of anti-TNF (Williams *et al*. 1992a), we formulated the hypothesis that TNF-a is at the apex of a cytokine cascade. This gave the rationale to blockade TNF in 20 patients with active rheumatoid arthritis in an open phase I/II trial lasting 8 weeks (Elliott *et al*. 1993).

The monoclonal antibody used was a chimeric (mouse Fv, human IgG1) neutralizing antibody produced by Centocor, Inc. The benefits of anti-TNF treatment were evident in all patients within a few days and lasted 8 to 26 weeks (median 12 weeks). Improvements in clinical parameters included reduction in pain and morning stiffness, falls in swollen and tender joint counts, increased erythrocyte sedimentation rate and reduced C-reactive protein and serum amyloid A. Following this initial success a randomized, double-blind, placebo-controlled, multicentre trial of anti-TNF in 73 patients was undertaken, the results of which confirmed the open study (Elliott *et al*. 1994) and supported the hypothesis that TNF is of major importance in the pathogenesis of rheumatoid arthritis (Fig. 10 and Fig. 11).

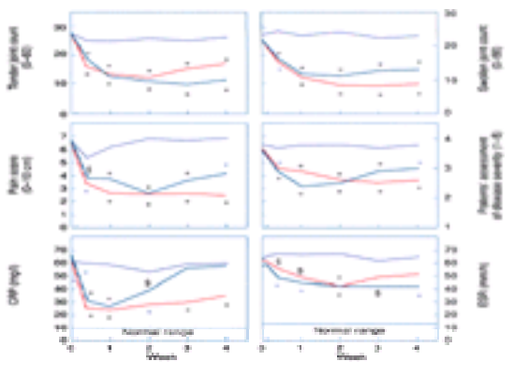


Fig. 10 Changes in clinical assessments in 73 patients treated with placebo (●), 1 mg/kg (▲), or 10 mg/kg (■) anti-TNF monoclonal antibody in a randomized, double-blind trial (*p* values represent significance versus placebo: + $p < 0.05$; § $p < 0.01$; * $p < 0.001$). (Reproduced by kind permission from an article by Elliott *et al*. 1994, *Lancet*, **344**, 1105–10.)

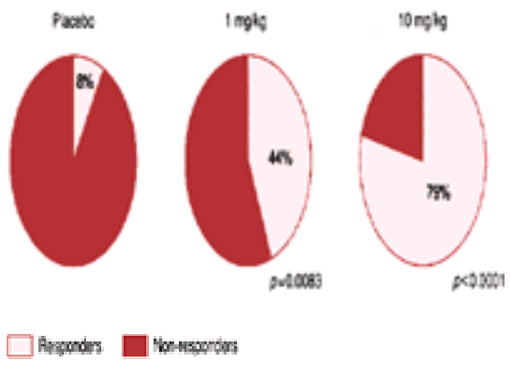


Fig. 11 Overall clinical responses to placebo, 1 mg/kg, or 10 mg/kg anti-TNF monoclonal antibody in 73 patients, 4 weeks after treatment, in a randomized, double-blind trial (*p* values represent significance versus placebo). (Reproduced by kind permission from articles by Elliott *et al*. 1994, *Lancet*, **344**, 1105–10 and Maini *et al*. 1995, *Immunology Reviews*, **144**, 95–223.)

The mechanism of the anti-inflammatory action of anti-TNF-a antibody is under examination in current studies. A rapid decrease of serum IL-6 following anti-TNF demonstrates the effect of anti-TNF antibody on downregulation of other cytokines (Maini *et al*. 1995). A second, and possibly more important, effect of TNF blockade is in reducing the cellularity of the synovium (Maini *et al*. 1995). This is accompanied by an increase in peripheral blood lymphocyte count and a decrease in expression of adhesion molecules in synovial biopsies taken before and after therapy (Tak *et al*. 1996). There is an associated decrease in circulating levels of E-selectin and ICAM-1 (Paleolog *et al*. 1996). Viewed together, these data suggest that anti-TNF therapy reduces cell traffic into joints by reducing leucocyte–endothelium interactions, thereby reducing the mass of inflammatory tissue and its clinicopathological consequences.

Cytokine antagonists

In 1989, TNF-binding protein, capable of inhibiting the action of TNF, was discovered in the blood and urine of febrile patients (Englemann *et al*. 1989; Seckinger *et al*. 1989). This was subsequently found to be derived from the extracellular domain of the two TNF receptors, probably by proteolytic cleavage. On account of the proposed role of TNF-a in the pathogenesis of rheumatoid arthritis, the role of TNF inhibitor was explored, with the expectation that levels of soluble TNF receptors may be low.

Analysis of serum samples from a variety of arthritic patients has revealed that in rheumatoid arthritis the TNF inhibitor system is enhanced, and there are elevated levels of both soluble TNF receptors (p55 and p75) in serum and in the joint fluids, with an intermediate rise in levels in seronegative arthritis and osteoarthritis (Cope *et al*. 1992). However, these upregulated levels of soluble TNF receptors do not neutralize fully the TNF-a produced by cells from rheumatoid joints in culture, whereas they generally appear to be sufficient to neutralize TNF-a produced by cells from osteoarthritic joints in culture (Brennan *et al*. 1995). Thus, in arthritis there appears to be an attempt at homeostasis, which, however, is inadequate (Cope and Maini 1995). Other soluble cytokine receptors that in the fluid phase would act as cytokine antagonists have been described, for example, soluble IL-1 (Symons *et al*. 1991), –2 (p55), –4, –6, and –7 receptors and soluble interferon-g receptor, which may act as regulators of the cytokine network.

There is so far only one cytokine inhibitor described which acts as a receptor antagonist, i.e. the IL-1 receptor antagonist, which is produced in rheumatoid joints

(Arend 1991). This is a member of the IL-1 family, with 30 per cent homology to IL-1b. The physiological role of this molecule is unclear, as quantities greatly in excess of those of IL-1 are necessary to exert inhibitory effects, far larger than are physiologically present *in vivo*. Its role may be simply to localize the effects of IL-1 in the environment in which the cytokine is produced. Whatever the exact physiological roles of IL-1 receptor antagonist and soluble TNF receptor, these natural agents with the capacity to interfere with cytokine action are potential therapeutic agents in rheumatoid arthritis (Elliott and Maini 1996).

Immune suppression in rheumatoid arthritis

Despite the evidence for an ongoing autoimmune response at both the T- and B-cell level, there is also considerable evidence that the systemic immune response is suppressed in patients with rheumatoid arthritis. This has been demonstrated as a reduced response to tuberculin (purified protein derivative) testing (Kingsley *et al*. 1987), in IL-2 production, or in T-cell proliferative or cytokine production. Serial studies have shown that the degree of immune suppression is more severe as the patients are clinically more ill. The mechanisms of this suppression are not understood, but could lead eventually to new forms of therapy aimed at reinforcing the endogenous mechanisms of immune suppression.

A number of molecules found in the joints in rheumatoid arthritis may be important contributors to this process. These include prostaglandins and TGF-b. Other potential candidates not yet known to be present include oncostatin M, IL-10, and IL-4. Prostaglandins are produced in the inflammation of rheumatoid arthritis. However, most non-steroidal anti-inflammatory drugs interfere with the production of prostaglandins, and this does not overcome the endogenous suppression, which accordingly is mostly caused by other agents. TGF-b is found in large amounts in supernatants of synovial cell cultures (10 to 20 ng/ml) of which 1 to 2 ng/ml is bioactive (Brennan *et al*. 1990b). There is thus sufficient TGF-b to influence immune functions of T and B cells, and cytokine production. However, appropriate neutralizing experiments on synovial cell cultures remain to be done. Oncostatin M is growth inhibitory for a variety of cell types, is produced by activated T cells and macrophages, and so may be expected to be present. IL-10, a product of T cells and B cells, can interfere with antigen-presenting capacity and with the production of interferon-g. IL-10 has been demonstrated in rheumatoid joints and *in vitro* studies reveal that it is apparently exerting a suppressive effect on endogenous production of TNF-a and IL-1 (Katsikis *et al*. 1993). The beneficial effect of administration of recombinant IL-10 in established collagen-induced arthritis supports its possible therapeutic potential (Walmsley *et al*. 1996)

Animal models (see Chapter 3.4)

Research in unravelling the factors that initiate rheumatoid arthritis, understanding the perpetuation of disease, and devising new strategies for therapy or prevention has to some extent depended on concepts developed and validated in animal models. Recent clinical studies have raised questions that have increased, rather than diminished, the complementary value of the use of animal models. However, there is as yet no ideal animal model of rheumatoid arthritis that exhibits all key features, namely:

1. predictable and spontaneous development of an erosive, chronic, symmetrical arthritis punctuated by flares;
2. female preponderance;
3. association with the MHC homologue of HLA DR4 or DR1;
4. high frequency of circulating IgM rheumatoid factor;
5. synovitis with a cellular response, profile of local production of cytokines, proteases, and inflammatory mediators identical to that observed in rheumatoid arthritis;
6. cartilage and bone degeneration with pannus formation;
7. response to disease-modifying antirheumatoid drugs akin to that observed in rheumatoid arthritis.

Despite reservations, the ensuing section gives examples of existing and new models of rheumatoid arthritis, which have contributed to our understanding of this disease.

Collagen-induced arthritis

The distribution of collagen II is essentially restricted to cartilage and the vitreous humour of the eye. Intradermal injection of native collagen II in Freund's adjuvant (but not of collagen types I and III or denatured type II) induces a polyarthritis in rats (Trentham *et al*. 1977), mice (Courtenay *et al*. 1980), and monkeys (Cathcart *et al*. 1986). Heterologous or autologous type II collagens are effective but the former leads to a destructive yet self-limiting disorder, whereas the latter is characterized by a chronic remitting and exacerbating course of disease (Holmdahl *et al*. 1986). T- and B-cell responses to multiple epitopes on collagen II occur, and disease of a milder variety than in the immunized animals has been transferred into syngeneic animals by serum and/or T cells. Rheumatoid factor is detectable and villous synovitis with increased cellularity of the lining layer, infiltration of deeper layers with mononuclear cells (predominantly CD4+ T cells), and pannus formation echo the changes observed in rheumatoid arthritis. The best-documented susceptible mouse, the DBA/l, bears the *H-2q*haplotype; H-2^{r} mice are also susceptible, but H-2^{d} mice are resistant to collagen arthritis. Although a polygenic disease, it is of considerable interest that the HLA class II molecules mapping to I-A (the mouse homologue of human HLA DQ) appear to be the element controlling susceptibility and immune responses to type II collagen in DBA/1 (Holmdahl *et al*. 1989). The importance of non-MHC genes, for example, genes regulating complement synthesis and the expression of IgG subclass isotypes, has been deduced from other studies.

The collagen arthritis has significant similarities to rheumatoid arthritis and its importance lies in the ability of an immune response to a constituent of cartilage to induce disease. As B- and T-cell-specific responses to collagen II occur in a proportion of patients with rheumatoid arthritis, especially, and sometimes exclusively, when lymphocytes from the synovial membrane are studied, the model provides evidence that collagen immunity might perpetuate rheumatoid disease. The model has provided useful data on the arthritogenic epitopes on type II collagen and therapeutic manipulations have provided evidence that antibodies directed against CD4+ T cells (Ranges *et al*. 1985), B cells (Helfgott *et al*. 1984), and TNF-a (Williams *et al*. 1992a) are effective in ameliorating established disease.

Adjuvant arthritis

A single intradermal injection of Freund's complete adjuvant (containing *Mycobacterium tuberculosis*) in the footpad or tail of rats induces a severe arthropathy involving the wrists, ankles, paws, and caudal part of the spine and tail (Pearson 1956). The arthropathy in its developed stage consists of synovitis with villous formation, pannus eroding cartilage and bone, marked periostitis with new bone formation, and inflammation and fibrosis of periarticular tissues. After peaking, the inflammatory arthritis declines and is followed by fibrous and bony ankylosis of joints. Extra-articular features can be prominent and include balanitis, conjunctivitis, and cutaneous lesions resembling psoriasis. Although the disease in diarthrodial joints has some similarity to rheumatoid arthritis, the other features are reminiscent of the clinical spectrum of spondylarthropathies, especially Reiter's syndrome. This is further suggested by a consistent lack of IgM rheumatoid factor. Susceptibility is strain dependent; for example, Lewis rats are most susceptible, and Fisher rats are less so. Susceptibility is believed to involve multiple genes with no convincing role for the MHC genes.

The major interest in the model springs from the demonstration that the disease is mediated by T cells that recognize mycobacterial peptides (Cohen *et al*. 1985). Furthermore, there is evidence for molecular mimicry between a mycobacterial antigen and cartilage antigens (van Eden *et al*. 1985), and this is believed to be the key factor in localization of the disease to joints. The relevant antigen has been defined and is a mycobacterial nonapeptide present in a 65-kDa mycobacterial protein that belongs to the family of hsp65 (van Eden *et al*. 1988). The nonapeptide stimulates T-cell clones of the CD4 phenotype, termed A2b and A2c, derived from a parent line, A2, obtained from a rat with adjuvant arthritis. Following *in vivo* inoculation into irradiated syngeneic Lewis rats, A2b causes a severe arthritis, whereas A2c protects from disease induction and causes a rapid remission (Cohen *et al*. 1985). When hsp65 or the nonapeptide are given before Freund's adjuvant, the rats are protected from the disease.

The possibility that T cells equivalent to the suppressive clones isolated from adjuvant arthritis are present in the inflammatory exudate of rheumatoid arthritis has prompted optimism that T-cell vaccination may prove to be a promising therapy. Activated T cells, treated with hydrostatic pressure, or T-cell receptors cross-linked with glutaraldehyde are used as surrogate suppressor–inducers in vaccination protocols. However, preliminary trials have not shown any benefit in rheumatoid arthritis (Van Laar *et al*. 1991).

Streptococcal cell-wall arthritis

A single injection intraperitoneally of an aqueous suspension of group A streptococcal cell-wall fragments into rats induces a polyarthritis (Cromartie *et al*. 1977). The arthritis involves wrists, ankles, and other joints, spares the axial skeleton, and is biphasic with an early phase reaching its maximum at 3 days, followed by the onset

of a chronic arthritis 2 to 4 weeks later. Lewis (LEW/N) female rats are the most susceptible to this form of arthritis and exhibit many pathological features of rheumatoid arthritis—a villous synovial thickening with surface fibrin, thickening of the synovial lining layer, polymorph exudation into joint fluid, mononuclear cell infiltrates with a predominance of CD4+ T cells, angiogenesis, and fibroblast proliferation with pannus formation and associated erosion of underlying cartilage and bone. Low titres of IgM rheumatoid factor are detectable. The active proinflammatory constituent of streptococcal cell-wall is its peptidoglycan component, which has extensive pathophysiological effects involving many cell types; the smallest active subunit of peptidoglycan is muramyl dipeptide, which is itself an activator of macrophages and endothelial cells. Persistence of streptococcal cell-wall owing to its protective carbohydrate side chains, is believed to contribute to the initiation and perpetuation of disease.

The importance of T cells in the pathogenesis of the disease has been demonstrated by transfer of arthritis to nude, T-cell deficient, inbred Lewis rats. Like rheumatoid arthritis, the T-cell abnormalities include depressed responses to mitogen and defective production of IL-2. As hsp65 protects against arthritis induced by streptococcal cell wall (van den Broek *et al*. 1989), it has been suggested that this protein may be the host protein target for the T-cell response. However, the molecular basis of this has not been resolved. Although HLA class II molecules are rapidly induced in endothelial cells in inflamed tissues, the role of MHC antigen-associated susceptibility is ambiguous because a related strain of rats (the Fisher strain) with the same histocompatibility locus is relatively resistant to arthritis.

Of considerable interest are the observations on the hypothalamoadrenal axis in the arthritis induced by streptococcal cell wall. In susceptible female Lewis rats there is an abnormally low gene expression at the mRNA level of the corticotrophin-releasing hormone and encephalin, with a deficient response of adrenal corticotrophic hormone and adrenal corticosteroid (Sternberg *et al*. 1989a; Sternberg *et al*. 1989b). In contrast, Fisher rats, resistant to arthritis, show relatively rapid and efficient responses from the hypothalamoadrenal axis. That these neuroendocrine responses are important in the pathogenesis of disease is suggested by the observation that giving corticosteroid in small doses simultaneously administered with streptococcal cell wall improves the course of the induced arthritis and, conversely, blockade of the glucocorticoid receptor with RU 486 accelerates disease in resistant Fisher rats (Sternberg *et al*. 1989a).

Other models of arthritis

The transient inflammatory arthritis of serum sickness in rabbits induced by antigen and mediated by antigen–antibody complexes (Dixon *et al*. 1958) had antedated the description of chronic arthritis in the rabbit knee induced by intra-articular injection of protein antigens such as fibrin, heterologous gammaglobulin, and ovalbumin into previously sensitized animals that had received antigen in Freund's adjuvant (Dumonde and Glynn 1962). One aspect of interest in the latter model was the demonstration that the arthritis was dependent on antigen–antibody complexes sequestered in cartilaginous tissues (Jasin 1975) which could act as a depot of persistent immunogen and gave rise to a cellular immune response. The induction of arthritis by a similar protocol in neonatally thymectomized or bursectomized chickens demonstrated that arthritis was inducible by both thymus-dependent T and bursa-dependent B cells; however, the fully developed lesion required an intact thymus and bursa (Oates *et al*. 1972).

Infective agents as a cause of arthritis have attracted much interest; and mycoplasmas (Decker and Barden 1975) and erysipelotherix (Drew 1972) are both well-described causes of chronic arthritis in swine. In the former model, a chronic arthritis persisted long after viable organisms could be cultured from joints, blood, or lymph nodes and non-viable antigen persisted for longer periods, and it was suggested thst this might have been responsible for the destructive arthritis. In a model of swine arthritis resembling rheumatoid arthritis studied in Sweden, introduction of fish meal in the diet was causative of arthritis and evidence was obtained that this was associated with population of the gut by *Clostridium perfringens* (Mansson *et al*. 1971). Immune responses to the clostridium were demonstrable, but the organism could not be isolated from the joints and as such represented a form of reactive arthritis.

Caprine arthritis, mainly involving large joints, has generated interest in the arthritogenic potential of the causative lentivirus, which is a lentiform retrovirus, as is the human immunodeficiency virus (Crawford *et al*. 1980). In this disease, possibly acquired by ingestion of milk by goat kids, the virus is harboured by mononuclear phagocytes. Encephalitis also occurs and this feature makes it distinct from rheumatoid arthritis. Chronic destructive joint lesions are described, and these resemble rheumatoid arthritis, as does the mononuclear cell infiltrate in the synovium. However, the lentivirus cannot be isolated from joints. In contrast to rheumatoid arthritis, mononuclear cells rather than polymorphs are dominant in joint fluids and, also unlike rheumatoid arthritis, high levels of interferon activity are demonstrable. Studies of the tropism for joints and the pathogenesis of the inflammatory reaction provide insight into the pathways that might prove important in devising investigations of the possibility that retroviruses may cause rheumatoid arthritis.

The MRL/lpr mouse is generally regarded as a model of systemic lupus and develops a multisystem disease characterized by glomerulonephritis, vasculitis, and antibodies to double-stranded DNA and Sm antigen, associated with marked lymphoproliferation involving a T cell with an ab heterodime receptor, but lacking CD4 and CD8 antigens (Andrews *et al*. 1978). This mouse strain, however, also develops an arthritis of the hind limbs with invasion of cartilage by pannus, high levels of rheumatoid factor, and anticollagen type II antibodies; therefore in some respects it shows features of rheumatoid arthritis. Production of rheumatoid factor appears to be under the control of the lymphoproliferation gene. The early destruction of articular tissue is at the marginal junction of the synovium with cartilage and bone, in association with proliferation of fibroblastic cells, and antedates an inflammatory response (O'Sullivan *et al*. 1985). This spontaneous model of arthritis appears to be of importance in delineating the relationship between autoimmunity and non-immune cellular responses in the synovium and pannus invasion of cartilage and bone, as well as understanding the genetic regulation of rheumatoid factor.

Transgenic mice offer the possibility of assessing *in vivo* the effect of introduced genes on the development of disease. Two models have recently been published which may shed insight into the mechanism of arthritis. The simplest to evaluate is the introduction of a modified human TNF-a gene, under its own promoter, into fertilized ova. The modification of the gene was replacement of the TNF-a 3' untranslated region, which has been shown to confer mRNA instability, with the 3' untranslated region of b-globin, which has a very stable mRNA. With this deregulated TNF-a production, it was found that the mice developed a progressive arthritis by 4 weeks of age. The arthritis was preventable by the injection of antihuman TNF-a monoclonal antibody from birth onwards. The arthritis is characterized by subchondral erosions and frequent fibrosis, but further analysis is necessary to establish how closely this disease resembles rheumatoid arthritis (Keffer *et al*. 1991). This model, however, confirms the hypothesis that TNF-a is intimately involved in the arthritic process.

Transgenic mice carrying the human T-cell leukaemia virus-1 genome also develop a chronic erosive arthritis (Iwakura *et al*. 1991). In this model, approximately one-third of the mice highly expressing the transgene developed arthritis with synovial inflammation and cartilage erosion, closely resembling a pannus. Low levels of rheumatoid factor were occasionally detected. It was of interest that the mRNA of the *Tax* gene, a transacting transcriptional activator, was highly expressed in the joints. Thus it is likely that the arthritis is the result of increased cytokine expression in the joints. Supporting this is the preliminary observation, cited in Iwakura *et al*. (1991), that IL-1a mRNA is expressed in the joints of these mice, as it is in rheumatoid arthritis.

Conclusions

Our understanding of the pathogenesis of rheumatoid arthritis in molecular terms has progressed rapidly in the past few years. This has provided a number of molecular targets, and therapeutic trials based on these targets have been initiated. The first of these was CD4, and monoclonal anti-CD4 has been used by a number of groups, with beneficial though transient results in a proportion of patients.

Attempts are being made to devise peptide-based therapies which will selectively block the critical HLA peptide-presenting genomes. Neutralization of cytokines such as TNF-a and IL-1 with antibodies or soluble receptors has been successfully applied in clinical trials and has set the stage for an era of promising new therapeutic interventions.

These (and more) therapeutic trials will have an important benefit in helping to evaluate ideas concerning the pathogenesis of rheumatoid arthritis, and even if (as is likely) they are not totally successful, they will contribute to refining our concepts of the pathogenesis of the disease and to more effective therapies.

Chapter References

Abbink, J.J., Kamp, A.M., Niujens, J.S., Erenberg, A.J., and Hack, C.E. (1992). Relative contribution of contact and complement activation to inflammatory reactions in arthritic joints. *Annals of the Rheumatic Diseases*, **51**, 1123–8.

Abbot, S.E., Kaul, A., Stevens, C.R., and Blake, D.R. (1992). Isolation and culture of synovial microvascular endothelial cells: characterization and assessment of adhesion molecule expression. *Arthritis and Rheumatism*, **35**, 401–6.

Adler, R. (1985). Psychoneuroimmunologic contributions to the study of rheumatic diseases. In *Immunology of rheumatic diseases* (ed. S. Gupta and N. Talal), pp. 669–96. Plenum Medical Book

Company, New York.

Akahoshi, T. *et al*. (1993). Expression of monocyte chemotactic and activating factor in rheumatoid arthritis. *Arthritis and Rheumatism*, **36**, 762–71.

Albani, S. *et al*. (1995). Positive selection in autoimmunity: abnormal immune responses to a bacterial DNAJ antigenic determinant in patients with early rheumatoid arthritis. *Nature New Medicine*, **1**, 448–52.

Allard, S.A., Muirden, K.D., Camplejohn, K.L., and Maini, R.N. (1987). Chondrocyte-derived cells and matrix at the rheumatoid cartilage–pannus junction identified with monoclonal antibodies. *Rheumatology International*, **7**, 153–9.

Allard, S.A., Bayliss, M.T., and Maini, R.N. (1990). The synovial–cartilage junction of the normal human knee: implications for joint destruction and repair. *Arthritis and Rheumatism*, **33**, 1170–9.

Allard, S.A., Bayliss, M.T., and Maini, R.N. (1991). Correlation of histopathological features of pannus with patterns of damage in different joints in rheumatoid arthritis. *Annals of the Rheumatic Disease*, **50**, 278–83.

Alvaro-Garcia, J.M., Zvaifler, N.J., and Firestein, G.S. (1989). Cytokines in chronic inflammatory arthritis. IV. Granulocyte/macrophage colony stimulating factor-mediated induction of class II MHC antigen on human monocytes: a possible role in rheumatoid arthritis. *Journal of Experimental Medicine*, **170**, 865–75.

Andrew, E.M., Plater Zyberk, C., Brown, C.M.S., Williams, D.G., and Maini, R.N. (1991). The potential role of B lymphocytes in the pathogenesis of rheumatoid arthritis. *British Journal of Rheumatology*, **30** (Suppl. l), 47–52.

Andrews, B.S. *et al*. (1978). Spontaneous murine lupus-like syndrome. *Journal of Experimental Medicine*, **l48**, 1198–215.

Arend, W.F. (1991). Interleukin-l receptor antagonist—a new member of the IL-l family. *Journal of Clinical Investigation*, **88**, 1445–51.

Axford, J.S., Mackenzie L., Lydyard, P.M., Hay, F.C., Isenberg, D.A., and Roitt, I.M. (1987). Reduced B-cell galactosyltransferase activity in rheumatoid arthritis. *Lancet*, **ii**, 1486–8.

Baboonian, C., Venables, P.J.W., Williams, D.G., Williams, R.O., and Maini, R.N. (1991). Cross-reaction of antibodies to a glycine–alanine repeat sequence of Epstein–Barr virus nuclear antigen-l with collagen, cytokeratin, and actin. *Annals of the Rheumatic Diseases*, **50**, 772–5.

Barkley, D., Allard, S.A., Feldmann, M., and Maini, R.N. (1989 *a*). Increased expression of HLA-DQ antigens by interstitial cells and endothelium in the synovial membrane of rheumatoid arthritis patients compared with reactive arthritis patients. *Arthritis and Rheumatism*, **32**, 955–63.

Barkley, D., Feldmann, M., and Maini, R.N. (1989 *b*). The detection by immunofluorescence of distinct cell populations producing interleukin-1a and interleukin-1b in activated human peripheral blood. *Journal of Immunological Methods*, **120**, 277–83.

Berthelot, J.-M., Vincent, C., Serre, G., and Youinou, P. (1994). APF (antiperinuclear factor). In *Manual of biological markers of disease* (ed. R.N. Maini and W.J. Van Venrooij), B1.2, pp. 1–9. Kluwer, Dordrecht.

Brennan, F.M. *et al*. (1988*a*). Heterogeneity of T cell receptor idiotypes in rheumatoid arthritis. *Clinical and Experimental Immunology*, **73**, 417–23.

Brennan, F.M. *et al*.(1988*b*). T cells expressing gd chain receptors in rheumatoid arthritis. *Journal of Autoimmunity*, **1**, 319–26.

Brennan, F.M., Chantry, D., Jackson, A.M., Maini, R.N., and Feldmann, M. (1989 *a*). Cytokine production in culture by cells isolated from the synovial membrane. *Journal of Autoimmunity*, **2**, (Suppl.), 177–86.

Brennan, F.M., Chantry, D., Jackson, A., Maini, R.N., and Feldmann, M. (1989 *b*). Inhibitory effect of TNFa antibodies on synovial cell interleukin-1 production in rheumatoid arthritis. *Lancet*, **ii**, 244–7.

Brennan, F.M., Plater Zyberk, C., Maini, R.N., and Feldmann, M. (1989 *c*). Co-ordinate expansion of 'fetal type' lymphocytes (TCR gd+T and CD5+b) in rheumatoid arthritis and primary Sjögren's syndrome. *Clinical and Experimental Immunology*, **77**, 175–8.

Brennan, F.M. *et al*.(1990*a*). Detection of interleukin 8 biological activity in synovial fluids from patients with rheumatoid arthritis and production of IL-8 mRNA by isolated synovial cells. *European Journal of Immunology*, **20**, 2141.

Brennan, F.M., Chantry, D., Turner, M., Foxwell, B., Maini, R.N., and Feldmann, M. (1990 *b*). Detection of transforming growth factor b in rheumatoid arthritis synovial tissue: lack of effect on spontaneous cytokine production in joint cell cultures. *Clinical and Experimental Immunology*, **81**, 278–85.

Brennan, F.M., Gibbons, D.L., Cope, A.P., Katsikis, P., Maini, R.N., and Feldmann, M. (1995). TNF inhibitors are produced spontaneously by rheumatoid and osteoarthritis synovial joint cell cultures: evidence of feedback control of TNF action. *Scandinavian Journal of Immunology*, **42**, 158–65.

Brenner, C.A., Tam, A.W., Nelson, D.A., Suzuki, N., Fry, K.E., and Larrick, J.W. (1989). Message amplification phenotyping (MAPPing): a technique to simultaneously measure multiple mRNAs from small number of cells. *BioTechniques*, **7**, 1096–2003.

Brown, C.M.S, Longhurst, C., Haynes, G., Plater-Zyberk, C., Malcolm, A., and Maini, R.N. (1992). Immunoglobulin heavy chain variable region gene utilization by B cell hybridomas derived from rheumatoid synovial tissue. *Clinical and Experimental Immunology*, **89**, 230–8.

Brown, C.M.S., Fitzgerald, K.J., Moyes, S.P., Mageed, R.A., Williams, D.G., and Maini, R.N. (1995). Sequence analysis of immunoglobulin heavy-chain variable region genes from the synovium of a rheumatoid arthritis patient shows little evidence of mutation but diverse CDR3. *Immunology*, **84**, 367–74.

Buchan, G., Barrett, K., Turner, M., Chantry, D., Maini, R.N., and Feldmann, M. (1988 *a*). Interleukin-1 and tumour necrosis factor mRNA expression in rheumatoid arthritis: prolonged production of IL-1a. *Clinical and Experimental Immunology*, **73**, 449–55.

Buchan, G., Barrett, K., Fujita, T., Taniguchi, T., Maini, R.N., and Feldmann, M. (1988 *b*). Detection of activated T cell products in the rheumatoid joint using cDNA probes to interleukin 2, IL-2 receptor and interferon g. *Clinical and Experimental Immunology*, **71**, 295–301.

Buchanan, H.M., Preston, S.J., Brooks, P.M., and Buchanan, W.W. (1991). Is diet important in rheumatoid arthritis? *British Journal of Rheumatology*, **30**, 125–34.

Cathcart, E.S., Hayes, K.C., Gonnerman, W.A., Lazzari, A.A., and Franzblau, C. (1986). Experimental arthritis in a non-human primate. l. Induction by bovine type II collagen. *Laboratory Investigation*, **54**, 26–31.

Chantry, D., Winearls, C.G., Maini, R.N., and Feldmann, M. (1989). Mechanism of immune complex mediated damage: induction of interleukin 1 by immune complexes and synergy with interferon g and tumour necrosis factor a. *European Journal of Immunology*, **19**, 189–92.

Chantry, D., Turner, M., Brennan, F., Kingsbury, A., and Feldmann, M. (1990). Granulocyte–macrophage colony stimulating factor induces both HLA-DR expression and cytokine production by human monocytes. *Cytokine*, **2**, 60–7.

Charriere, G., Hartmann, D.J., Vignon, E., Ronziere, M.C., Herbage, D., and Ville, G. (1988). Antibodies to type I, II, IX, XI collagen in the serum of patients with rheumatic diseases. *Arthritis and Rheumatism*, **3l**, 325–32.

Chu, C.Q., Field, M., Feldmann, M., and Maini, R.N. (1991 *a*). Localization of tumor necrosis factor a in synovial tissues and at the cartilage–pannus junction in patients with rheumatoid arthritis. *Arthritis and Rheumatism*, **34**, 1125–32.

Chu, C.Q. *et al*. (1991*b*). Transforming growth factor b_l in rheumatoid synovial membrane and cartilage/pannus junction. *Clinical and Experimental Immunology*, **86**, 380–6.

Cohen, I.R., Holoshitz, J., van Eden, W., and Frenkel, A. (1985). T lymphocyte clones illuminate pathogenesis and effect therapy of experimental arthritis. *Arthritis and Rheumatism*, **28**, 84l–5.

Cohen, S.B.A. *et al*. (1995). High level of interleukin-10 production by the activated T cell population within the rheumatoid synovial membrane. *Arthritis and Rheumatism*, **38**, 946–52.

Colville-Nash, P.R. and Scott, D.L. (1992). Angiogenesis and rheumatoid arthritis: pathogenic and therapeutic implications. *Annals of the Rheumatic Diseases*, **51**, 919–25.

Cope, A.P. and Maini, R.N. (1995). Soluble tumor necrosis factor receptors in arthritis. *Journal of Rheumatology*, **22**, 382–4.

Cope, A. *et al*. (1992). Increased levels of soluble tumor necrosis factor receptors in the sera and synovial fluids of patients with rheumatic diseases. *Arthritis and Rheumatism*, **35**, 1160–9.

Courtenay, J.S., Dallman, M.J., Dayan, A.D., Martin, A., and Mosedale, B. (1980). Immunisation against heterologous type II collagen induces arthritis in mice. *Nature*, **283**, 666–8.

Crawford, T.B., Adams, D.S., Cheevers, W.P., and Cork, L.V. (1980). Chronic arthritis in goats caused by a retrovirus. *Science*, **207**, 997–9.

Cromartie, W.J., Craddock, J.C., Schwab, J.H., Anderle, S.K., and Yang, C.H. (1977). Arthritis in rats after systemic injection of streptococcal cells or cell walls. *Journal of Experimental Medicine*, **146**, 1585–602.

Cush, J.J., Pietschmann, P., Oppenheimer-Marks, N., and Lipsky, P.E. (1992). The intrinsic migratory capacity of memory T cells contributes to their accumulation in rheumatoid synovium. *Arthritis and Rheumatism*, **35**, 1434–44.

Davies, M.E., Sharma, H., and Pigott, R. (1992). ICAM-1 expression on chondrocytes in rheumatoid arthritis: induction by synovial cytokines. *Mediators of Inflammation*, **1**, 71–4.

Dayan, C.M. *et al*. (1991). Autoantigen recognition by thyroid-infiltrating T cells in Graves' disease. *Proceedings of the National Academy of Sciences* (*USA*), **88**, 7415–19.

de Graeff-Meeder, E.R. *et al*.(1990). Antibodies to the mycobacterium-65 kDa heat-shock protein are reactive with synovial tissue of adjuvant arthritic rats and patients with rheumatoid arthritis and osteoarthritis. *American Journal of Pathology*, **137**, 1013–17.

Decker, J.L. and Barden, J.A. (1975). *Mycoplasma hyortinis* of swine: a model for rheumatoid arthritis? *Rheumatology*, **6**, 338–45.

Deighton, C.M., Walker, D.J., Griffiths, I.D., and Roberts, D.F. (1989). The contribution of HLA to rheumatoid arthritis. *Clinical Genetics*, **36**, 178–82.

Deleuran, B.W. *et al*. (1992). Localization of tumor necrosis factor receptors in the synovial tissue and cartilage–pannus junction in patients with rheumatoid arthritis: implications for local actions of tumor necrosis factor a. *Arthritis and Rheumatism*, **35**, 1170–8.

Dequeker, J. (1987). Rheumatic diseases in visual arts. (General review.) In *Art history and antiquity of rheumatic diseases* (ed. T. Appelboom), p. 84. Elsevier, Brussels.

Dieppe, P.A. (1988). Did Galen describe rheumatoid arthritis? *Annals of the Rheumatic Diseases*, **47**, 84–5.

Dixon, F.J., Vasquez, J.J., Weigle, W.D., and Cochrane, C.G. (1958). Pathogenesis of serum sickness. *Archives of Pathology*, **65**, 18–22.

Drew, R.A. (1972). Erysipelothrix arthritis in pigs as a model of rheumatoid arthritis. *Proceedings of the Royal Society of Medicine*, **65**, 42–6.

Dumonde, D.C. and Glynn, L.E. (1962). The production of arthritis in rabbits by an immunological reaction to fibrin. *British Journal of Experimental Pathology*, **43**, 373–83.

Eastgate, J.A., Symons, J.A., Wood, N.C., Grinlinton, F.M., Di Giovine, F.S., and Duff, G.W. (1988). Correlation of plasma interleukin-1 levels with disease activity in rheumatoid arthritis. *Lancet*, **ii**, 706–9.

Ebringer, A., Khalafpour, S., and Wilson C. (1989). Rheumatoid arthritis and *Proteus*: a possible aetiological association. *Rheumatism International*, **9**, 223–8.

Elliott, M.J. and Maini, R.N. (1996). What are the prospects for therapy based on cytokines and anticytokines in rheumatoid arthritis? In *Cytokines in autoimmunity* (ed. F.M. Brennan and M. Feldmann), pp. 239–56. R.G. Landes, Austin, TX.

Elliott, M.J. *et al*. (1993). Treatment of rheumatoid arthritis with chimeric monoclonal antibodies to tumor necrosis factor a. *Arthritis and Rheumatism*, **36**, 1681–90.

Elliott, M.J. *et al*. (1994). Randomised double-blind comparison of chimeric monoclonal antibody to tumour necrosis factor a (cA2) versus placebo in rheumatoid arthritis. *Lancet*, **344**, 1105–10.

Engelmann, H., Novick, D., and Wallach, D. (1990). Two tumour necrosis factor-binding proteins purified from human urine. *Journal of Biological Chemistry*, **265**(3), 153l–6.

Feldmann, M. (1987). Regulation of HLA class II expression and its role in autoimmune disease. In *Autoimmunity and autoimmune disease*, Ciba Foundation Symposium 129, pp. 88–108. Wiley, Chichester.

Feldmann, M. (1989). Molecular mechanisms involved in human autoimmune diseases: relevance of chronic antigen presentation, class II expression and cytokine production. *Immunology*, Suppl. **2**, 66–71.

Feldmann, M. *et al*. (1991). Cytokine assays: role in evaluation of the pathogenesis of autoimmunity. *Immunological Reviews*, **119**, 105–23.

Feldmann, M., Brennan, F.M., and Maini, R.N. (1996). Role of cytokines in rheumatoid arthritis. *Annual Review of Immunology*, **14**, 397–440.

Ferrara, N., Houck, K.A., Jakeman, L.B., Winer, J., and Leung, D.W. (1991). The vascular endothelial growth factor family of polypeptides. *Journal of Cell Biochemistry*, **47**, 211–8.

Field, M., Chu, C., Feldmann, M., and Maini, R.N. (1991). Interleukin-6 in the synovial membrane in rheumatoid arthritis. *Rheumatism International*, **11**, 45–50.

Firestein, G.S. and Zvaifler, N.J. (1987). Peripheral blood and synovial fluid monocyte activation in inflammatory arthritis. II. Low levels of synovial fluid and synovial tissue interferon suggest that g-interferon is not the primary macrophage activating factor. *Arthritis and Rheumatism*, **30**, 864–71.

Firestein, G.S. and Zvaifler, N.J. (1990). How important are T cells in chronic rheumatoid synovitis? *Arthritis and Rheumatism*, **33**, 768–73.

Firestein, G.S., Alvaro-Garcia, J.M., and Maki, R. (1990). Quantitative analysis of cytokine gene expression in rheumatoid arthritis. *Journal of Immunology*, **144**, 3347–53.

Folkman, J. (1995). Angiogenesis in cancer, vascular, rheumatoid and other disease. *Nature Medicine*, **1**, 27–30.

Fong, Y. *et al*. (1989). Antibodies to cachectin/tumour necrosis factor reduce interleukin 1b and interleukin 6 appearance during lethal bacteremia. *Journal of Experimental Medicine*, **170**, 1627–33.

Forrestier, J. (1935). Rheumatoid arthritis and its treatment by gold salts. *Journal of Laboratory and Clinical Medicine*, **20**, 827–40.

Fraser, K.J. (1982). Anglo-French contributions to the recognition of rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **41**, 335–43.

Garrod, A.B. (1859). *Nature and treatment of gout and rheumatic gout*. Walton and Maberly, London.

Gaston, J.S., Life, P.F., Bailey, L.C., and Bacon, P.A. (1989). *In vitro* response to a 65 kDa mycobacterial protein by synovial T cells from inflammatory arthritis patients. *Journal of Immunology*, **143**, 2494–500.

Goldstein, R. and Arnett, F.C. (1987). The genetics of rheumatic disease in man. *Rheumatic Disease Clinics of North America*, **13** (3), 487–510.

Goronzy, J. and Weyand, P.M., (1993). Interplay of T lymphocytes and HLA-DR molecules in rheumatoid arthritis. *Current Opinion in Rheumatology*, **5**, 169–77.

Goronzy, J., Weyand, P.M., and Fathman, C.G. (1986). Shared T cell recognition sites on human histocompatibility leukocyte antigen class II molecules of patients with seropositive rheumatoid arthritis. *Journal of Clinical Investigation*, **77**, 1042–9.

Gregersen, P.K., Silver, J., and Winchester, R.J. (1987). The shared epitope hypothesis. An approach to understanding the molecular genetics of susceptibility to rheumatoid arthritis. *Arthritis and Rheumatism*, **30**, 1205–13.

Hammer, J. *et al*.(1995). Peptide binding specificity of HLA-DR4 molecules: correlation with rheumatoid arthritis association. *Journal of Experimental Medicine*, **181**, 1847–55.

Hannestad, K. and Stollar, B.D. (1978). Certain rheumatoid factors react with nucleosomes. *Nature*, **275**, 671–3.

Hargraves, M.M., Richmond, H., and Morton, R. (1948). Presentation of two bone marrow elements, the 'Tart' cell and 'L-E' cell. *Proceedings of Staff Meetings of the Mayo Clinic*, **23**, 25–8.

Hassfeld, W. *et al*. (1995). Autoimmune response to the spliceosome: an immunologic link between rheumatoid arthritis, mixed connective tissue disease and systemic lupus erythematosus. *Arthritis and Rheumatism*, **38**, 777–85.

Haworth, C., Brennan, F.M., Chantry, D., Turner, M., Maini, R.N., and Feldmann, M. (1991). Expression of granulocyte-macrophage colony-stimulating factor in rheumatoid arthritis: regulation by tumour necrosis factor a. *European Journal of Immunology*, **21**, 2575–9.

Helfgott, S.M., Bazin, H., Dessein, A., and Trentham, D.E. (1984). Suppressive effects of anti-μ serum on the development of collagen arthritis in rats. *Clinical Immunology and Immunopathology*, **31**, 403–11.

Herman, A., Kappler, J.W., Marrack, P., and Pullen, A.M. (1991). Superantigens: mechanism of T-cell stimulation and role in immune responses. *Annual Review of Immunology*, **9**, 745–72.

Hirano, T. *et al*. (1988). Excessive production of interleukin 6/B cell stimulatory factor-2 in rheumatoid arthritis. *European Journal of Immunology*, **18**, 1797–801.

Holmdahl, R., Jansson, L., Larsson, E., Rubin, K., and Klareskog, L. (1986). Homologous type II collagen induces chronic and progressive arthritis in mice. *Arthritis and Rheumatism*, **29**, 106–13.

Holmdahl, R., Karlsson, M., Andersson, M.E., Rask, L., and Andersson, L. (1989). Localisation of a critical restriction site on the 1-A beta chain which determines susceptibility to collagen-induced arthritis. *Proceedings of the National Academy of Sciences* (*USA*), **86**, 9475–9.

Holoshitz, J. *et al*.(1986). T lymphocytes of rheumatoid arthritis patients show increased reactivity to a fraction of mycobacteria cross-reactive with cartilage. *Lancet*, **11**, 305–9.

Holoshitz, J., Koning, F., Coligan, J.E., De Bruyn, J., and Strober, S. (1989). Isolation of CD4–CD8-mycobacteria-reactive T lymphocyte clones from rheumatoid arthritis synovial fluid. *Nature*, **339**, 226–9.

Howell, M.D. *et al*. (1991). Limited T-cell receptor b chain heterogeneity among interleukin-2 receptor-positive T cells suggests a role for superantigen in rheumatoid arthritis. *Proceedings of the National Academy of Sciences*(*USA*),**88**, 10921–5.

Hunter, W. (1901). *Oral sepsis as a cause of septic conditions*. Cassell, London.

Iwakura, Y. *et al*.(1991). Induction of inflammatory arthropathy resembling rheumatoid arthritis in mice transgenic for HTLV. *Science*, **253**, 1026–8.

Jasin, H.E. (1975). Mechanism of trapping of immune complexes into joint collagenous tissues. *Clinical and Experimental Immunology*, **22**, 473–85.

Jose, P.J., Moss, I.K., Maini, R.N., and Williams, T.J. (1990). Measurement of the chemotactic complement fragment C5a in rheumatoid synovial fluids by radioimmunoassay: role of C5a in the acute inflammatory phase. *Annals of the Rheumatic Diseases*, **49**, 747–52.

Katsikis, P., Chu, C.Q., Brennan, F.M., Maini, R.N., and Feldmann, M. (1993). Immunoregulatory role of interleukin 10 (IL-10) in rheumatoid arthritis. *Journal of Experimental Medicine*, **179**, 1517–27.

Keffer, J. *et al*. (1991). Transgenic mice expressing human tumour necrosis factor: a predictive gene model of arthritis. *EMBO Journal*, **10**, 4025–31.

Kehrl, J.H., Alvarez-Mon, M., Delsing, G.A., and Fauci, A.S. (1987). Lymphotoxin is an important T cell derived growth factor for human B cells. *Science*, **238**, 1144–7.

Kingsley, G.M., Pitzalis, C., and Panayi, G.S. (1987). Abnormal lymphocyte reactivity to self—major histocompatibility antigens in rheumatoid arthritis. *Journal of Rheumatology*, **14**, 667–73.

Kissonerghis, A.M., Maini, R.N., and Feldmann, M. (1989). High rate of HLA class II mRNA synthesis in rheumatoid arthritis joints and its persistence in culture: down-regulation by recombinent interleulin 2. *Scandinavian Journal of Immunology*, **29**, 73–82.

Klemperer, P., Pollack, A.D., and Baehr, G. (1942). Diffuse collagen disease: acute disseminated lupus erythematosus and diffuse systemic sclerosis. *Journal of the American Medical Association*, **119**, 331–2.

Koch, A.E. *et al*. (1992). Enhanced production of monocyte chemoattractant protein-1 in rheumatoid arthritis. *Journal of Clinical Investigation*, **90**, 772–9.

Koch, A.E. *et al*.(1994). Macrophage inflammatory protein-1a. *Journal of Clinical Investigation*, **93**, 921–8.

Lahita, R.G. (1990). Sex hormones and the immune system. Part 1: Human data. *Baillière's Clinical Rheumatology*, **4**, 1–12.

Lanzavecchia, A. *et al*. (1985). Antigen-specific interaction between T and B cells. *Nature*, **314**, 537–9.

Lawrence, J.S. (1970). Rheumatoid arthritis: nature or nurture? *Annals of the Rheumatic Diseases*, **29**, 357–69.

Londei, M. *et al*.(1989). Persistence of collagen type II specific T cell clones in the synovial membrane of a patient with rheumatoid arthritis. *Proceedings of the National Academy of Sciences* (*USA*), **86**, 636–40.

Londei, M., Verhoef, A., Hawrylowicz, C., Groves, J., De Berardinis, P., and Feldmann, M. (1990). Interleukin 7 is a growth factor for mature human T cells. *European Journal of Immunology*, **20**, 425–8.

Macgregor, A.J. and Silman, A. (1994). An analysis of the relative contribution of genetic and environmental factors to rheumatoid arthritis susceptibility. *Arthritis and Rheumatism*, **37** (Suppl.), S169.

Maini, R.N. (1989). Exploring immune pathways in rheumatoid arthritis. *British Journal of Rheumatology*, **28**, 466–79.

Maini, R.N., Plater-Zyberk, C., and Andrew, E.M. (1987). Autoimmunity in rheumatoid arthritis: an approach via a study of B lymphocytes. *Rheumatic Diseases Clinics of North America*, **13**, 319–38.

Maini, R.N. *et al*. (1995). Monoclonal anti-TNFa antibody as a probe of pathogenesis and therapy of rheumatoid disease. *Immunology Reviews*, **144**, 195–223.

Mansson, I., Norberg, R., Olhagen, B., and Bjorklund, N.E. (1971). Arthritis in pigs induced by dietary factors: microbiological, clinical, and histologic studies. *Clinical and Experimental Immunology*, **9**, 677–93.

March, L.M. (1987). Dendritic cells in the pathogenesis of rheumatoid arthritis. *Rheumatism International*, **7**, 93–100.

Mollenhauer, J., von der Mark, K., Burmester, G., Gluckert, K., Lütjen-Drecoll, E., and Brune, K. (1988). Serum antibodies against chondrocyte cell surface proteins in osteoarthritis and rheumatoid arthritis. *Journal of Rheumatology*, **15**, 1811–17.

Moore, K.W., Vieiva, P., Fiorentino, D.F., Trounstine, M.L., Khan, T.A., and Mosmann, T.R. (1990). Homology of cytokine synthesis inhibitory factor (IL-10) to the Epstein–Barr virus gene BCRF 1. *Science*, **248**, 1230–4.

Morales-Ducret, J., Wayner, E., Elices, M.J., Alvaro-Garcia, J.M., Zvaifler, N.J., and Firestein, G.S. (1992). Alpha 4/beta 1 integrin (VLA-4) ligands in arthritis: I. Vascular cell adhesion molecule 1 expression in synovium and on fibroblast-like synoviocytes. *Journal of Immunology*, **149**, 1424–31.

Nienhuis, R.L.F. and Mandema, E. (1964). A new serum factor in patients with rheumatoid arthritis: the perinuclear factor. *Annals of the Rheumatic Diseases*, **23**, 302–5.

Oates, C.M., Maini, R.N., Payne, L.N., and Dumonde, D.C. (1972). Possible role of lymphokines in the development of ectopic lymphoid foci in the chicken. *Advances in Experimental Medicine and Biology*, **29**, 611–18.

Olee, T. *et al*.(1991). Molecular basis of an autoantibody-associated restriction fragment length polymorphism that confers susceptibility to autoimmune diseases. *Journal of Clinical Investigation*, **88**, 193–203.

Ollier, W. *et al*. (1984). HLA antigen associations with extra-articular rheumatoid arthritis. *Tissue Antigens*, **24**, 279–91.

Olsen, N.J. *et al*.(1994). Multicenter trial of an anti-CD5 immuno-conjugate in rheumatoid arthritis (RA). *Arthritis and Rheumatism*, **37**, S295.

O'Sullivan, F.X., Fassbender, H.G., Gay, S., and Koopman, W.J. (1985). Etiopathogenesis of the rheumatoid arthritis-like disease in MRL/lpr mice. 1. The histomorphologic basis of joint destruction. *Arthritis and Rheumatism*, **28**, 529–36.

Paleolog, E.M., Hunt, M., Elliott, M.J., Woody, J.N., Feldmann, M., and Maini, R.N. (1996). Monoclonal anti-tumour necrosis factor a antibody deactivates vascular endothelium in rheumatoid arthritis. *Arthritis and Rheumatism*, **39**, 1082–91.

Paliard, X. *et al*. (1991). Evidence for the effects of a superantigen in rheumatoid arthritis. *Science*, **253**, 325–9.

Panayi, G.S., Woolley, P.H., and Batchelor, J.H. (1979). HLA-DRW4 and rheumatoid arthritis. *Lancet*, **i**, 730–4.

Parekh, R.B. *et al*. (1985). Association of rheumatoid arthritis and primary osteoarthritis with changes in the glycosylation pattern of total serum IgG. *Nature*, **316**, 452–7.

Pearson, C.M. (1956). Development of arthritis, periarthritis, and periostitis in rats given adjuvants. *Proceedings of the Society of Experimental Biology and Medicine*, **91**, 95–101.

Pitzalis, C., Kingsley, G., Murphy, J., and Panayi, G. (1987). Abnormal distribution of the helper-inducer and suppressor-inducer T lymphocyte subsets in the rheumatoid joint. *Clinical Immunology and*

Immunopathology, **45**, 252–8.

Plater-Zyberk, C., Maini, R.N., Brennan, F.M., and Feldmann, M. (1992). CD5+ B and double-negative T cells in rheumatoid arthritis. In *Rheumatoid arthritis* (ed. J. Smolen, J. Kalden, and R.N. Maini), pp. 122–36. Springer, Berlin.

Pober, J.S. *et al*. (1986). Overlapping patterns of activation of human endothelial cells by interleukin 1, tumour necrosis factor, and immune interferon. *Journal of Immunology*, **137**, 1893–6.

Portillo, G., Turner, M., Chantry, D., and Feldmann, M. (1989). Effect of cytokines on HLA-DR and IL-1 production by a monocytic tumour, THP-1. *Immunology*, **66**, 170–5.

Pujol-Borrell, R. *et al*. (1987). HLA class II induction in human islet cells by interferon-g plus tumour necrosis factor of lymphotoxin. *Nature*, **326**, 304–6.

Rademacher, T.W., Williams, P., and Dwek, R.A.. (1994). Agalactosyl glycoforms of IgG autoantibodies are pathogenic. *Proceedings of the National Academy of Sciences* (*USA*), **91**, 6123–7.

Raines, E.W., Dower, S.K., and Ross, R. (1989). Interleukin 1 mitogenic activity for fibroblasts and smooth muscle cells is due to PDGK AA. *Science*, **243**, 393–7.

Ranges, G.E., Stiram, S., and Cooper, S.M. (1985). Prevention of type II collagen-induced arthritis by *in vivo* treatment with anti-L_3T_4. *Journal of Experimental Medicine*, **162**, 1105–10.

Rathanaswami, P., Hachicha, M., Sadick, M., Schall, T.J., and McColl, S.R. (1993). Expression of the cytokine RANTES in human rheumatoid synovial fibroblasts. Differential regulation of RANTES and interleukin-8 genes by inflammatory cytokines. *Journal of Biological Chemistry*, **268**, 5834–9.

Res, P.C., Telgt, D., van Laar, J.M., Pool, M.O., Breedveld, F.C., and De Vries, R.R. (1990). High antigen reactivity in mononuclear cells from sites of chronic inflammation. *Lancet*, **336**, 1406–8.

Rice, G.E. and Bevilacqua, M.P. (1989). An inducible endothelial cell surface glycoprotein mediates melanoma adhesion. *Science*, **326**, 1303–6.

Rijnders, A., Boots, A., Verheijden, G., de Keijser, F., and Veijs, E. (1996). Identification of a key autoantigen in rheumatoid arthritis. Sixteenth European Workshop for Rheumatology Research, Stockholm.

Rogers, J. and Dieppe, P.A. (1990). Skeletal paleopathology and the rheumatic diseases. Where are we now? *Annals of the Rheumatic Diseases*, **49**, 885–6.

Rothschild, B.M. and Woods, R.J. (1990). Symmetrical erosive peripheral polyarthritis in arctic Indians: the origin of rheumatoid in the New World. *Seminars in Arthritis and Rheumatism*, **19**, 278–84.

Roudier, J., Rhodes, G., Petersen, J., Vaughan, J., and Carson, D.A. (1988). The Epstein–Barr virus glycoprotein gp110, a molecular link between HLA-DR4, HLA-DR1, and rheumatoid arthritis. *Scandinavian Journal of Immunology*, **27**, 367–71.

Sadeghi, R., Hawrylowicz, C.M., Chernajovsky, Y., and Feldmann, M. (1992 *a*). Synergism of glucocorticoids with granulocyte macrophage colony stimulating factor (GM-CSF) but not interferon g (interferon–g) or interleukin-4 (IL-4) on induction of HLA class II expression on human monocytes. *Cytokine*, **4**, 287–97.

Sadeghi, R., Feldmann, M., and Hawrylowicz, C.M. (1992 *b*). Upregulation of HLA class II, but not intercellular adhesion molecule 1 (ICAM-1) by granulocyte-macrophage colony stimulating factor (GM-CSF) or interleukin-3 (IL-3) in synergy with dexamethasone. *European Cytokine Network*, **3**, 373–80.

Sarvetnick, N. *et al*.(1990). Loss of pancreatic islet tolerance induced by B cell expression of interferon-g. *Nature*, **346**, 844–7.

Sasazuki, T. *et al*. (1986). HLA-linked immune suppression maps within HLA-DQ subregion. In *Regulation of immune gene expression* (ed. M. Feldmann and A. McMichael), pp. 197–206. Humana Press, New Jersey.

Sebbag, M. *et al*. (1995). The anti-perinuclear factor and the so-called anti-keratin antibodies are the same rheumatoid arthritis-specific autoantibodies. *Journal of Clinical Investigation*, **95**, 2672–9.

Seckinger, P., Isaaz, S., and Dayer, J.M. (1989). Purification and biologic characterisation of a specific tumour necrosis factor inhibitor. *Journal of Biological Chemistry*, **264**, 11966–73.

Short, C.L. (1974). The antiquity of rheumatoid arthritis. *Arthritis and Rheumatism*, **17**, 193–205.

Silman, A.J., Ollier, W., Hayton, R.M., Holligan, S., and Smith, K. (1989). Twin concordance rates for rheumatoid arthritis: preliminary results from a nationwide study. *British Journal of Rheumatology*, **28** (Suppl. 2), 95.

Simon, A.K., Seipelt, E., and Sieper, J. (1994). Divergent T-cell cytokine patterns in inflammatory arthritis. *Proceedings of the National Academy of Sciences* (*USA*), **91**, 8562–6.

Solomon, L., Robin, G., and Valkenburg, H.A. (1975). Rheumatoid arthritis in an urban South African negro population. *Annals of the Rheumatic Diseases*, **34**, 128–35.

Statsny, P. (1976). Mixed lymphocyte cultures in rheumatoid arthritis. *Journal of Clinical Investigation*, **57**, 1148–57.

Statsny, P. (1978). Association of the B cell alloantigen DRW4 with rheumatoid arthritis. *New England Journal of Medicine*, **97**, 664–761.

Steiner, G. *et al*.(1992). Purification and partial sequencing of the nuclear autoantigen RA33 shows that it is indistinguishable from the A2 protein of the heterogenous nuclear ribonuclearprotein complex. *Journal of Clinical Investigation*, **90**, 1061–6.

Sternberg, E.M. *et al*.(1989*a*). Inflammatory mediator-induced hypothalamic–pituitary–adrenal axis activation is defective in streptococcal cell wall arthritis-susceptible Lewis rats. *Proceedings of the National Academy of Sciences* (*USA*), **86**, 2374–8.

Sternberg, E.M. *et al*.(1989*b*). A central nervous system defect in biosynthesis of corticotropin-releasing hormone is associated with susceptibility to streptococcal cell-wall induced arthritis in Lewis rats. *Proceedings of the National Academy of Sciences*(*USA*), **86**, 4771–5.

Symons, J.A., Wood, N.C., Di Giovine, F.S., and Duff, G.W. (1988). Soluble IL-2 receptor in rheumatoid arthritis. *Journal of Immunology*, **141**, 2612–18.

Symons, J.A., Eastgate, J.A., and Duff, G.W. (1991). Purification and characterization of a novel soluble receptor for interleukin 1. *Journal of Experimental Medicine*, **174**, 1251–4.

Tak, P.P. *et al*. (1996). Reduction in cellularity and expression of adhesion molecules in rheumatoid synovial tissue after anti-TNFa monoclonal antibody treatment. *Arthritis and Rheumatism*, **39**, 1077–81.

Tarkowski, A., Klareskog, L., Carlsten, H., Herberts, P., and Koopman, W.J. (1989). Secretion of antibodies to types I and II collagen by synovial tissue cells in patients with rheumatoid arthritis. *Arthritis and Rheumatism*, **32**, 1087–96.

Taylor, P.C., Maini, R.N., and Plater-Zyberk, C. (1995). The role of the B cells in the adoptive transfer of collagen-induced arthritis from DBA/1 (H-2^q) to SCID (H-2^d) mice. *European Journal of Immunology*, **25**, 763–9.

Teyton, L. *et al*. (1987). HLA DR, DQ and DP antigen expression in rheumatoid synovial cells: a biochemical and quantitative study. *Journal of Immunology*, **138**, 1730–8.

Trabandt, A., Gay, R.E., and Gay, S. (1992). Oncogene activation in rheumatoid synovium. *Apmis*, **100**, 861–75.

Trentham, D.E., Townes, A.S., and Kang, A.H. (1977). Autoimmunity to type II collagen: an experimental model of arthritis. *Journal of Experimental Medicine*, **146**, 857–68.

Trentham, D.E. *et al*. (1993). Effects of administration of oral type II collagen on rheumatoid arthritis. *Science*, **261**, 1727–30.

Tsai, V., Bergroth, V., and Zvaifler, N.J. (1989). Dendritic cells in health and disease. In *T cell activation in health and disease* (ed. M. Feldmann, R.N. Maini, and J.N. Woody), pp. 33–44. Academic Press, New York.

Van den Broek, M.F., Hogervast, E.J., van Bruggen, M.C., van Eden, W., van der Zee, R., and van den Berg, W.B. (1989). Protection against streptococcal cell-wall-induced arthritis by treatment with the 65 kDa mycobacterial heat shock protein. *Journal of Experimental Medicine*, **170**, 449–66.

Van der Lubbe, P.A., Djikmans, B.A.C., Markusse, H.M., Nassander, U., and Breedveld, F.C. (1995). A randomized, double-blind, placebo-controlled study of CD4 monoclonal antibody therapy in early rheumatoid arthritis. *Arthritis and Rheumatism*, **38**, 1097–106.

van Eden, W., Holoshitz, J., Nevo, Z., Frenkel, A., Klajman, A., and Cohen, I.R. (1985). Arthritis induced by a T lymphocyte clone that responds to mycobacterium tuberculosis and to cartilage proteoglycans. *Proceedings of the National Academy of Sciences*(*USA*), **82**, 5117–20.

van Eden, W. *et al*. (1988). Cloning of the mycobacterial epitope recognised by thymocytes in adjuvant arthritis. *Nature*, **33l**, 171–3.

Van Laar, J.M., Miltenburg, A.M., Verdonk, M.J., Daha, M.R., de Vries, R.R., and Breedveld, F.C. (1991). T cell vaccination in rheumatoid arthritis. *British Journal of Rheumatology*, **30** (Suppl. 2), 28–9.

Venables, P.J.W. (1988). Epstein–Barr virus infection and autoimmunity in rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **47**, 265–9.

Venables, P.J.W., Erhardt, C.C., Mumford, P., and Maini, R.N. (1979). The occurrence of antibodies to extractable nuclear antigens (ENA) in extra-articular disease in rheumatoid arthritis (RA). *Annals of the Rheumatic Diseases*, **39**, 146–53.

Venables, P.J.W., Roffe, L.M., Erhardt, C.C., Maini, R.N., Edwards, J.M.B., and Porter, A.D. (1981). Titers of antibodies to RANA in rheumatoid arthritis and normal sera. *Arthritis and Rheumatism*, **24**, l459–69.

Villiger, P.M., Terkeltaub, R., and Lotz, M. (1992). Production of monocyte chemoattractant protein 1 by inflamed synovial tissue and cultured synoviocytes. *Journal of Immunology*, **149**, 722–7.

Waaler, E. (1940). On the occurrence of a factor in human serum activating the specific agglutination of sheep blood corpuscles. *Acta Pathologica et Microbiologica Scandinavica*, **17**, 172–6.

Walmsley, M. *et al*. (1996). IL-10 inhibits progression of established collagen-induced arthritis. *Arthritis and Rheumatism*, **39**, 495–503.

Westedt, M.L., Breedveld, F.C., Schreuder, G.M.T., d'Amato, J., Cats, A., and de Vries, R.R.P. (1986). Immunogenetic heterogeneity of rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **45**, 534–8.

Westgeest, A.A.A., Boerbooms, A.M.T, Jongmans, M., Vandenbroucke, J.P., Vierwinden, G., and van de Putte, L.B.A. (1987). Anti-perinuclear factor: indicator of more severe disease in seronegative rheumatoid arthritis. *Journal of Rheumatology*, **14**, 893–97.

Weyand, C.M., Xie, C., and Goronzy, J.J. (1992). Homozygosity for the HLA-DRB1 allele selects for extra-articular manifestations in rheumatoid arthritis. *Journal of Clinical Investigation*, **89**, 2033–9.

Wilcox, W.H. (1935). *Reports on chronic rheumatic diseases*, **1**, 72. Lewis, London.

Wilkinson, L.S., Edwards, J.C.W., Poston, R.N., and Haskard, D.O. (1993). Expression of vascular cell adhesion molecule-1 in normal and inflamed synovium. *Laboratory Investigation*, **68**, 82–6.

Williams, R.O., Feldmann, M., and Maini, R.N. (1992 *a*). Anti-tumor necrosis factor ameliorates joint disease in murine collagen-induced arthritis. *Proceedings of the National Academy of Sciences* (*USA*), **89**, 9784–8.

Williams, R.O., Plater-Zyberk, C., Williams, D.G.. and Maini, R.N. (1992 *b*). Successful transfer of collagen-induced arthritis to severe combined immunodeficiency (SCID) mice. *Clinical and Experimental Immunology*, **88**, 455–60.

Winchester, R.J. and Gregersen, P.K. (1988). The molecular basis of susceptibility to rheumatoid arthritis: the conformational equivalence hypothesis. *Springer Seminars in Immunopathology*, **10**, 119–39.

Wordsworth, B.P., Lanchbury, J.S.S., Sakkas, L.I., Welsh, K.I., Panayi, G.S., and Bell, J.I. (1989). HLA-DR4 subtype frequencies in rheumatoid arthritis indicate that DRbl is the major susceptibility locus within the HLA class II region. *Proceedings of the National Academy of Sciences*(*USA*), **86**, 10049–53.

Wordsworth, B.P. *et al*. (1992). HLA heterozygosity contributes to susceptibility to rheumatoid arthritis. *American Journal of Human Genetics*, **51**, 585–91.

Young, B.J.J., Mallya, R.K., Leslie, R.D.G., Clark, C.J.M., and Hamlin, T.J. (1979). Anti-keratin antibodies in rheumatoid arthritis. *British Medical Journal*, **ii**, 97–9.

Ziegler, B., Gay, R.E., Huang, G., Fassbender, H.G., and Gay, S. (1989). Immunohistochemical localisation of HTLV-l pk and p24-related antigens in synovial joints of patients with rheumatoid arthritis. *American Journal of Pathology*, **135**, 1–5.

Ziff, M., Kantor, T., Bien, E., and Smith, A. (1953). Studies on the composition of fibrinoid material of subcutaneous nodule of rheumatoid arthritis. *Journal of Clinical Investigation*, **32**, 1253–9.

Zvaifler, N.J., Boyle, D., and Firestein, G.S. (1994). Early synovitis, synoviocytes and mononuclear cells. *Seminars in Arthritis and Rheumatism*, **23** (6, Suppl. 2) 11–16.

5.4.2 Rheumatoid arthritis–the clinical picture

Frank A. Wollheim

Introduction

Rheumatoid arthritis is a systemic disease with manifestations in many organs. However, in the majority of cases, involvement of the locomotor system dominates the clinical picture and forms the basis for diagnosis. In a classical textbook from 1957 it was stated that rheumatoid arthritis is 'a chronic systemic inflammatory disorder of unknown etiology characterized by the manner in which it involves the joints' (Short *et al*. 1957). There is no exact definition of rheumatoid arthritis, but several attempts to delineate criteria all are dominated by signs and symptoms from the locomotor system.

Criteria

Criteria are needed both in epidemiological work and in classification in the context of clinical trials. The original American Rheumatism Association (**ARA**) scheme of 1958 lists 11 criteria and no less than 20 exclusions (Ropes *et al*. 1958). With their help one could distinguish classical, definite, probable, and possible rheumatoid arthritis, characterized by decreasing number of criteria (Table 1). The main shortcoming of the ARA criteria is their complexity and low specificity (O'Sullivan and Cathcart 1972). Later, criteria 9, 10, and 11 were dropped to form the scheme of the Rome criteria (Kellgren 1962) (Table 2). The much simpler New York criteria were proposed in 1966 (Bennett and Burch 1967) (Table 3). These require a history of polyarthritis with clinical signs and either radiological erosions or presence of rheumatoid factor, and do not contain any exclusion criteria. The New York criteria performed better in the Sudbury study (O'Sullivan and Cathcart 1972) but they lack sensitivity for early and atypical cases. More recently the American College of Rheumatology (**ACR**) has developed criteria to replace the 1958 ARA criteria (Arnett *et al*. 1988). Rheumatoid arthritis is diagnosed if at least four of seven criteria are present. Only one category of rheumatoid arthritis is distinguished and exclusions are superfluous, owing to the detailed characterization of the criteria (Table 4). These criteria were developed by observing a number of 'typical' patients considered to be suffering from rheumatoid arthritis by experienced rheumatologists. The mean disease duration was 7.7 years. The performance of these criteria in epidemiological work has only recently been assessed but they may be best used in clinical trials aimed at patients with well-established disease.

Table 1 ARA criteria for rheumatoid arthritis (RA) of 1958 [a] (Ropes *et al*. 1958)

1. A past history of polyarthritis	
2. Symmetrical deformity of peripheral joints consisting of ankylosis or irreducible subluxation, especially of the lateral metatarsophalangeal or metacarpophalangeal joints; there must be some involvement of one hand or foot; involvement limited to large joints, such as the elbows or knees, does not satisfy this criterion	
3. Radiological changes of RA of grade 2 or more	
4. Positive serological test for rheumatoid factor	

[a] The diagnosis of definite RA requires 3 or 4 criteria fulfilled; one probable RA requires 2 criteria.

Table 2 The Rome criteria for the diagnosis of inactive rheumatoid arthritis (RA) [a]

1. History of an episode of three painful limb joints; each group of joints (e.g. proximal interphalangeal joints) is counted as one joint, scoring each side separately	
2. Swelling, limitation of motion, subluxation and/or ankylosis of three limb joints. Necessary inclusions: (1) at least one hand, wrist or foot; (2) symmetry of one joint pair. Exclusions: (1) distal interphalangeal joints; (2) fifth proximal interphalangeal joints; (3) first metatarsophalangeal joints; (4) hips	
3. Radiological changes (erosions)	
4. Serum positive for rheumatoid factor	

[a] RA is present if criterion 1 and 2, plus either 3 or 4, are met.

Table 3 The New York criteria for the diagnosis of rheumatoid arthritis (RA) [a]

Criterion	Comment
1. Morning stiffness	Duration > 1 h lasting >1 weeks
2. Arthritis of at least three areas[a]	Soft tissue swelling or exudation lasting >1 weeks
3. Arthritis of hand joints	Wrist, metacarpophalangeal joints or proximal interphalangeal joints lasting >1 weeks
4. Symmetrical arthritis	At least one area, lasting >1 weeks
5. Rheumatoid nodules	As observed by a physician
6. Serum rheumatoid factor	As assessed by a method positive in less than 5% of control subjects
7. Radiographic changes	As seen on anteroposterior films of wrists and hands

[a] Possible areas: proximal interphalangeal joints, metacarpophalangeal joints, wrist, elbow, knee, ankle, metatarsophalangeal joints. At least four criteria must be satisfied for inclusion.

Table 4 The revised criteria of 1987 (ARA/ACR)

Epidemiology

For lack of an obligate pathognomonic test, the diagnosis of rheumatoid arthritis rests on a composite of clinical and laboratory observations. In epidemiological work one has to rely on one of the criteria described above. Their limited precision was demonstrated in the Sudbury study, where only 12 of 40 cases of definite rheumatoid arthritis defined by the ARA criteria still fulfilled the criteria 4 years later. The New York criteria in the same study identified 15 patients with rheumatoid arthritis, and of these 11 were still positive at follow-up (O'Sullivan and Cathcart 1972). This uncertainty, if not evidence for transient disease, should be kept in mind when discussing the epidemiology of rheumatoid arthritis.

Worldwide prevalence

Rheumatoid arthritis has been identified in all populations that have been examined, and prevalence figures ranged between 0.2 and 5.3 per cent (Spector 1990) (Table 5). Nearly all studies indicate point prevalences of between 0.5 and 1 per cent. It is not known whether the high prevalence in Pima Indians is due to environment, genetics, or both and recent figures indicate a decreasing prevalence (Jacobsson *et al*. 1994). The age distribution is a confounding factor in developing countries, perhaps contributing to low figures in Africa.

Population	Reference	Size	Prevalence (%)
England	Lawrence (1961)	3000	0.9
Netherlands	Valkenburgh (1979)	6500	0.9
Sweden	Heligren (1970)	39 000	0.9
Bulgaria	Tzonchev (1966)	2695	0.8
Finland	Aho (1980)	8000	1.06
USA Tecumseh	Mikkelsen (1967)	6000	0.5
USA Sudbury	Cathcart (1970)	5552	0.9
USA Eskimos	Boyer (1968)	4600	1.0
USA Pima	del Puente (1989)	1449	5.3
USA Inuit	Oen (1986)	2055	0.6
Japan	Kato (1971)	11 393	0.6
South Africa	Myers (1978)	1601	0.7
	Bagshaw (1975)	1752	0.2
Liberia/Nigeria	Muller	1027	0.1
Israel	Adler (1957)	6769	0.3
Iraq	Al-Rawi (1976)	6998	1.0

Data from Spector (1990); where references may be found.

Table 5 Prevalence of rheumatoid arthritis in various adult populations

Incidence

Considering the difficulties involved in establishing early diagnosis of rheumatoid arthritis it is not surprising that only relatively few studies have addressed incidence. The Sudbury population study arrived at the high figure of 29 cases/10 000 per year based on only three observed cases (Table 6), whereas four other studies, using various methods, found annual figures between 9 and 2.9 cases/10 000. The 1987 ACR criteria were used in a prospective population-based registration of all new cases of arthritis in a population of 450 000 people in Norfolk, United Kingdom (Symmons *et al*. 1994), and showed an incidence of 3.4 in women and 1.4/10 000 in men. The incidence increased sharply with age in men from age 45. It increased in women until age 45, then plateaued and fell after age 75. No good longitudinal incidence data are available, although the study of Linos *et al*. (1980) indicated a decline from 9.2 to 4 in women but not men after the year 1964. This coincided with the advent of widespread use of contraceptive pills.

Population	Reference	Incidence	Method
Japan	Mirok [1967][1]	4.5/10 000	Population study
Holland	Dutcher [1968][2]	9/10 000	Hospital referrals
Sudbury USA	O'Sullivan [1968][3]	29/10 000	Population study
Finland	Isomäki [1978][4]	4.2/10 000	Health statistics
Olmsted USA	Linos [1980][5]	2.9/10 000	Hospital and physician records

[1] Arthritis and Rheumatism 10, 2–12. [2] Excerpta Medica (International Congress Series) 148, 85–105 and [3] 128–13. [4] Journal of Chronic Diseases 31, 691–4. [5] American Journal of Epidemiology 111, 87–98

Table 6 Annual incidence of rheumatoid arthritis in adults

Mortality

Evidence that mortality is increased in rheumatoid arthritis has accumulated since the 1950s, as shown in at least 16 studies (Spector and Scott 1988). Life expectancy is reduced by approximately 7 years in men and 3 years in women according to a Dutch study (Vandenbroucke *et al*. 1984). The increased mortality is mainly due to infections, renal disease, respiratory disease, and rheumatoid arthritis itself (Table 7). Two more recent reports from Japan and the United States confirm these data, and stress age, male sex, poor functional status, and low education as predictors of death (Suzuki *et al*. 1994; Wolfe *et al*. 1994).

Cause of death	RA (%)	Controls (%)
Cardiovascular	42.0	41.0
Cancer	14.0	20.0
Infection	9.5	1.0
Renal	8.0	1.0
Respiratory	7.0	4.0
RA	5.3	
Gastrointestinal	4.2	2.4

Data from Pincus and Callahan (1986)

Table 7 Mortality in rheumatoid arthritis (RA)

Clinical picture

Onset

The typical onset of insidious pain, stiffness, and symmetrical swelling of small joints in a middle-aged woman is but one of several presenting patterns. Up to one-third of patients have an acute or subacute onset. In some 10 per cent, palindromic rheumatism occurs for months or years before the onset of chronic disease. In several patients the first symptoms are not obviously located in the joints, but consist of fatigue, malaise, loss of weight, and subfebrility. Myalgia, morning stiffness, and depression in the absence of objective abnormalities may mimic functional, non-organic disease. This prodromal phase may last for weeks or months.

Triggering events such as infections, vaccinations, physical trauma, or psychological stress, although often suspected by lay persons, are not known to be involved. The onset of rheumatoid arthritis shows no seasonal variation (Eberhardt *et al*. 1990a).

The joints most commonly involved first were the finger (40 per cent), a shoulder (20 per cent), a foot joint (20 per cent), and a wrist (15 per cent). The knee joint was first involved in only 3 per cent of patients in our recent series (Eberhardt *et al*. 1990a). Asymmetrical onset is common.

The early events in rheumatoid arthritis are still largely unknown. A Finnish survey of the prevalence of both rheumatoid factor and antikeratin antibodies in the general population ('Minifinland') showed progressively increasing occurrences in the years preceding disease onset (Aho *et al*. 1989; Kurki *et al*. 1992), strongly indicating the existence of a presymptomatic phase of rheumatoid arthritis.

Articular involvement

Synovitis of the small joints of the hands and feet with symmetry and sparing of the distal interphalangeal joints may be the single most characteristic feature of rheumatoid arthritis, although the dominant side is often more severely affected (Owsianik *et al*. 1980). Lack of symmetry is seen in hemiparetic patients, in whom the paretic side is usually much less affected (Thompson and Bywaters 1961). The initial involvement may be confined to a few joints only. Spread occurs within months to years but the total number of affected joints tends to reach a plateau after the first years of illness. The number of joints involved in the early stages of rheumatoid arthritis is related to disease severity and functional outcome (Feigenbaum *et al*. 1979; Sherrer *et al*. 1986).

Although virtually any joint may be affected in rheumatoid arthritis, the dominating locations are the wrists, the metacarpophalangeal and proximal interphalangeal finger joints, the ankles and metatarsophalangeal joints, the knees, the shoulders, the hips, and the elbows (Fig. 1). The reason for involvement or sparing of joints in rheumatoid arthritis is poorly understood. Loading and overuse is related to severity and damage, but probably not to localization as such.

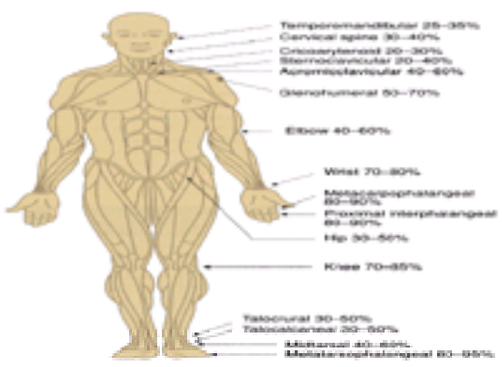


Fig. 1 Joint involvement in long-standing rheumatoid arthritis.

Hand

The hand joint affected by rheumatoid arthritis has many characteristic features which when present in combination may be almost pathognomonic. Boggy and tender swelling of proximal interphalangeal and metacarpophalangeal joints, wrists, and caput ulnae are common early signs. Tenosynovitis of extensors, and more often flexor, tendon sheets, with or without nodules are other features. Interosseal muscle atrophy is often seen early. Pain is prominent and may contribute to muscle atrophy. Limitation of motion, in particular dorsal flexion of wrist and finger flexion is very common. Compression of the median nerve in the carpal tunnel may be a presenting symptom of rheumatoid arthritis, and is frequent throughout the early development. Atrophy of the thenar muscle may be seen after long-standing compression of the median nerve.

As the disease progresses, signs of irreversible tissue damage appear: ulnar deviation of the fingers beginning with the index and fifth finger, volar subluxation of metacarpophalangeal joints, loosening of the distal radioulnar joint with dorsal protrusion of the ulnar head, as well as swan-neck and buttonhole deformities of the fingers occur in various combinations. Destruction of cartilage and bone as well as weakening and rupture of tendons and joint linings are contributing factors in the pathogenesis of hand deformities. Inability to make a fist, pinch thin objects, and loss of grip strength are important functional consequences of severe hand involvement in rheumatoid arthritis.

Elbow

Involvement of the elbow is present in half or more of the patients seen by specialists and it may become a major problem in 10 to 20 per cent. Extension defects usually develop, sometimes without the patients noticing anything abnormal. Active synovitis is readily palpated at the site of the groove normally present between the olecranon and the lateral epicondyle. Bulging, tenderness, and flexion contracture are typical findings, and all three may disappear after intra-articular injections of glucocorticoid. Epicondylitis and olecranon bursitis may be present in addition. Passive pronation and supination are painful in elbow synovitis. Destructive changes can be diagnosed by feeling crepitations when the radial head moves during this procedure.

Shoulder

The shoulder joints are involved in the majority of patients with rheumatoid arthritis, not least with increasing age. Synovitis of the glenohumoral joint causes tenderness on cranial palpation from below the axilla. It also results in pain with active and passive motion. Acromioclavicular synovitis is not uncommon giving local tenderness and a painful arch on abduction above 100°. Occasionally synovitis is present in the sternoclavicular joint. Rotator cuff tendinitis with varying degrees of supraspinatus injury is most common and causes considerable degrees of shoulder dysfunction, painful arch from 50° to 100° of abduction, and night pain. It may be difficult to distinguish inflammatory from degenerative abnormalities. Subacromial bursitis is also common, and may present with a palpable swelling.

Foot

Foot involvement is as common as hand involvement in rheumatoid arthritis, and 10 per cent of patients have the earliest erosions in the metatarsophalangeal joints (Eberhardt *et al*. 1990a). Pain in the forefoot elicited by walking is a well recognized presenting manifestation of rheumatoid arthritis (Luukkainen *et al*. 1983), and very often causes difficulties in ambulation. Synovitis is found in the metatarsophalangeal joints, giving rise to tenderness on direct palpation. Involvement is also seen in three more proximal foot joints. The joint between the distal tibia, fibula, and talus gives rise to tenderness on direct palpation distal to the malleoli. Talocalcaneal synovitis is diagnosed by causing pain on passive pronation or supination, and synovitis of the mid-tarsal joint gives rise to pain on rotating the foot while keeping the heel fixed. These joints may be involved individually. Foot deformities develop with time, e.g. lateral deviation of toes, hammer toes, cock-up toes, and valgus deformity of the ankle. Bursitis, corn formations, and tendinitis are also common.

Knee

The knee joint is the largest human joint. Although not often inflamed at the onset of rheumatoid arthritis, gonarthritis at some stage occurs in 80 per cent or more of all cases. Synovitis is usually easy to identify by tenderness and swelling below the patella. Other signs are exudation, which may be demonstrated by the patella click or the bulge sign. Gonarthritis often leads to quadriceps atrophy and flexion contracture, both of which should initiate prompt and vigilant therapy. The knee joint communicates with bursas in the fossa poplitea, which may become distended and merge into a large Baker's cyst. This may grow and dissect its way down into the calf muscle or rarely into the thigh and sometimes rupture, causing diffuse swelling and pain and may be mistaken for deep-vein thrombosis. Intra-articular pressure may increase to as much as 1100 mmHg in exudative knee joints (Geborek *et al*. 1989a). This will contribute to decreased muscular function, impede circulation through the synovial membrane, and lead to lactic acidosis and may be a pathogenic factor in the rupture of Baker's cysts. The natural course of long-standing knee involvement is often valgus instability, flexion contracture, and inability to walk. Increased intra-articular pressure may also contribute to bone erosion (Monsees *et al*. 1985).

Hip

Hip joint involvement is less common and was formerly considered as a late manifestation. However, once started the arthritis often leads to severe disability with pain on weight bearing and limitation of motion, in particular abduction and rotation. Functional limb shortening may sometimes be present secondary to adduction contracture. In a prospective study of recent-onset rheumatoid arthritis, hip involvement was found in 20 per cent. In 7 of 13 patients no symptoms were present despite abnormal swelling found on ultrasonography. Fifteen of 113 patients underwent total hip replacement after a median disease duration of 4 years (Eberhardt *et al*. 1995). Thus joint failure is more common in the hip than in the knee in early rheumatoid arthritis.

Involvement of the hip must be distinguished from periarticular bursitis or trochanteritis which is characterized by local tenderness, painful motion, and responsiveness to local infiltration with glucocorticoids (Raman and Haslock 1982).

Cervical spine

The cervical spine is involved in approximately 40 to 70 per cent of hospital patients with rheumatoid arthritis (Bland 1974). The dominating symptoms are occipital pain, sometimes worsened by motion, muscle spasm, and crepitation. Synovitis of the faceted joints is commonly present at C1 to C4. The symptoms in most patients improve with immobilization. However, in a small (but important) number of patients severe neurological complications develop secondary to subluxations. This may result in fatal outcome as shown in one study where 11 of 104 patients, dying in a geriatric hospital, had cervical cord compression at autopsy (Mikulowski *et al*. 1975). This complication will be dealt with below.

Temporomandibular joint

This is affected in one-quarter of patients, usually symmetrically and causing no major disability. In fact patients with unequivocal, palpable, synovial swelling or tenderness often do not report pain on chewing unless specifically asked. In severe destructive cases of rheumatoid arthritis, however, attrition of joint cartilage and bone causes malignment of the teeth with malocclusion (Chalmers and Blair 1973).

Synovitis of the cricoarytenoid joint

This is a well-recognized manifestation in rheumatoid arthritis, occurring in two-thirds of patients in a contemporary hospital-based study (Geterud 1991). The symptoms are sensation of a foreign body, hoarseness, and weak voice, as well as stridor, particularly at night. This may become severe and may eventually cause suffocation. The condition can be detected by palpating local tenderness and finding the vocal cords tender, red, swollen, and immovable on laryngoscopic examination. The laryngeal obstruction can be assessed by computed tomography. This condition is amenable to surgical correction (Geterud 1991).

Rheumatoid nodules

Rheumatoid nodules occur in 30 per cent or less of patients, and when present they are principally in extensor areas of the forearm and pressure areas throughout the skin. Although not specific for rheumatoid arthritis they constitute useful diagnostic and prognostic information, correlating strongly with seropositivity, somewhat with disease activity and progressive disease. Rarely they are the presenting symptom of rheumatoid arthritis, occasionally they are first detected in internal organs,

e.g. lungs, heart, and gallbladder.

The histological appearance is that of a central necrotic area surrounded by palisades of fibroblasts, histiocytes, and macrophages. Expression of HLA DR is present among these cells (Hedfors *et al*. 1983), and collagenase and proteinase production (Palmer *et al*. 1987) may explain the central necrosis.

Rheumatoid nodules develop in areas over spinal processes and over the neck in patients confined to bed, over the patellae in housewife's knee, and in the inside of the fingers in patients exposed to hand work, demonstrating pressure as an important pathogenetic factor.

It is of particular clinical significance to watch for skin ulcerations over rheumatoid nodules, as this indicates necrotizing vasculitis (Fig. 2).

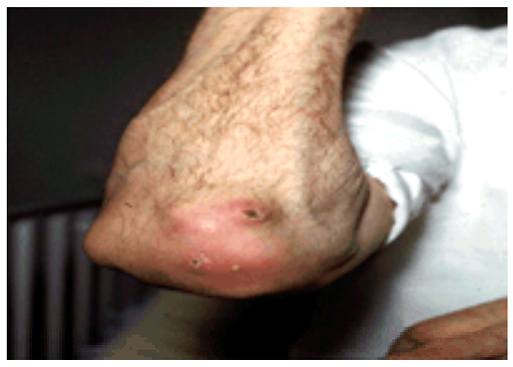


Fig. 2 Subcutaneous nodule over elbow with ulceration in a patient who developed necrotizing vasculitis and gangrene of toes.

Disease course

Rheumatoid arthritis is extremely heterogeneous with regard to severity and progression rate. Although early permanent remission may occur in some cases, this is rare once permanent joint damage has started. This does not rule out clinical improvement after initially severely impaired function (Eberhardt *et al*. 1990b). Long-term follow-up studies (Rasker and Cosh 1989), however, clearly indicate that functional deterioration occurs in most patients surviving 15 years or longer. It is suggested but not well documented that geographical differences may exist in this regard.

Some texts distinguish between cyclic types of disease with remissions and exacerbations and slow but relentless disease progression. In clinical practice, however, features of both types frequently are present in a majority of patients. A more useful distinction is that between widespread and more limited chronic joint involvement which can be assessed by careful joint examination (Feigenbaum *et al*. 1979).

Disease activity as manifested by number of swollen joints was low in half of patients with early rheumatoid arthritis followed for 6 years, whereas it was high in 17 per cent and fluctuating in the remainder (van Zeben *et al*. 1994). Disease activity can be monitored by simple measurements, such as joint swelling and tenderness. Erythrocyte sedimentation rate and C-reactive protein measured over time correlate with functional impairment (Hassell *et al*. 1993; Guillemin *et al*. 1994).

Non-articular manifestations

Rheumatoid arthritis is primarily a joint disease, but extra-articular manifestations can be detected in almost any organ system and may occasionally precede the onset of arthritis. Table 8 lists some of these. They are in general more prominent in seropositive and nodular disease.

Organ system	Manifestation	Cumulative frequency (%)
Lymph nodes	Enlargement, dysfunction	~ 50
Spleen	Splenomegaly	25
	Felty's syndrome	~ 1
Lungs	Pleuritis	~ 30
	Nodules	5
	Fibrosis	Rare
Heart	Pericarditis	~ 10
	Myopericarditis	~ 5
	Nodules	5
Muscles	Myositis	Common
	Atrophy	Common
Bone	Osteoporosis	Common
Skin	Nodules	~ 20
	Vasculitis	~ 1
Ophthalmic	Scleritis	1
	Nodules	~ 2
	Sjögren's syndrome	10
Neurological	Cord compression	~ 1
	Nerve entrapment	Common
	Polyneuropathy	Rare
	Mononeuritis multiplex	~ 1

Table 8 Some organ manifestations of rheumatoid arthritis

Lymph nodes

Lymph nodes are enlarged or of abnormal shape in the majority of patients but only rarely palpable. They show up on mammography and give rise to diagnostic concern (Andersson *et al*. 1980). However, in a few cases rheumatoid arthritis may start with widespread palpable lymphadenopathy and a histological picture mimicking Hodgkin's disease. It is of interest that total lymph-node irradiation has a transient suppressive effect on the synovitis in rheumatoid arthritis. Lymphopenia was seen in 15 per cent of cases in one study; it was related to severity but was not influenced by therapy (Symmons *et al*. 1989).

Pulmonary involvement

Pleuritis is most common but frequently asymptomatic. It may be associated with pericarditis and like other pulmonary and cardiac manifestations it is more common in older men. Rheumatoid nodules in the lungs are often asymptomatic and are only seen in seropositive cases. They may be single or multiple and can cause diagnostic problems. Smoking as well as exposure to, for example, silica are pathogenetic factors.

Diffuse interstitial fibrosis and fibrosing alveolitis are rare but serious manifestations, which individuals with HLA DR3 and carriers of the *PiZ* gene are more prone to develop (Geddes *et al*. 1977; Hyland *et al*. 1983). Smoking is also a contributing aetiological factor (Hyland *et al*. 1983).

Cardiac involvement

Pericarditis has been a common finding at autopsy, but causes symptoms in only a few patients. These may range from pain with friction rubs to severe exudative pericarditis with cardiac tamponade requiring surgical intervention (Hara *et al*. 1990). Most cases are, however, benign and self-limiting. Valvular insufficiency, conduction disturbances, and coronary occlusion may occasionally be due to endocarditis, nodulosis, or vasculitis.

Muscle involvement

Some degree of muscle atrophy is almost invariably present in rheumatoid arthritis, and there is electromyographic evidence of myositis (Moritz 1963). The type II or fast-twitch muscles are most affected (Halla *et al*. 1984). Focal necrosis is also seen. Creatine kinase levels are usually not elevated. Rheumatoid nodules may be

seen.

Bone

Generalized axial and appendicular bone loss is seen early in the course of rheumatoid arthritis and has been assessed by dual electron X-ray absorption (DEXA) measurements within 6 months from onset. Bone mineral loss is faster in postmenopausal women and in patients with active disease (Gough *et al*. 1994a). This is a strong argument for early disease suppressive therapy in severe rheumatoid arthritis. This bone loss should be distinguished from the juxta- and periarticular bone loss that probably is a consequence of local cytokine production. Increased bone resorption can also be measured indirectly by assaying the increased urinary excretion of pyridinoline crosslinks (Gough *et al*. 1994b).

Skin

Cold, clammy hands often accompany flaring rheumatoid arthritis. Palmar erythema is common and may reflect low androgen levels (Spector 1989). Various forms of arteritis manifestations are characteristic signs of systemic disease. Small splinter haemorrhages localize to the nailfolds and may disappear spontaneously. When larger arteries are affected, skin ulcerations develop, often on the lower extremities or where skin is exposed to pressure, e.g. the buttocks (Fig. 3). These ulcerations are very painful, may come in crops, and tend to grow and become chronic. Superinfection is a contributing factor. As mentioned, ulcerations may form over subcutaneous nodules in severe cases.

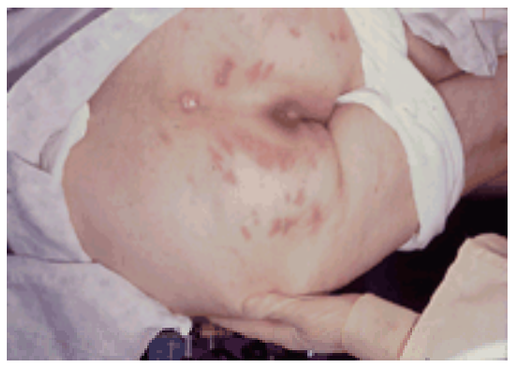


Fig. 3 Skin ulcers in the buttocks in a patient with rheumatoid vasculitis.

Leucocytoclastic vasculitis also occurs and is seen as a palpable purpura. This manifestation, as a rule, heals (Fig. 4).

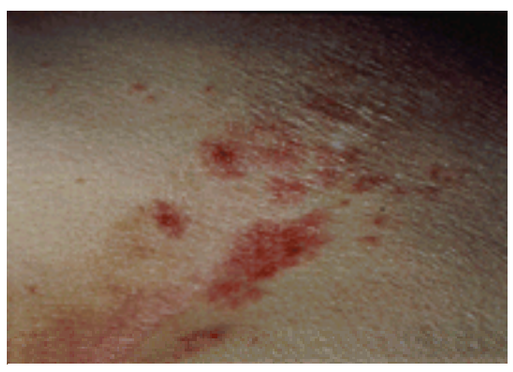


Fig. 4 Palpable purpura in a case of rheumatoid arthritis, where healing occurred without immunosuppressive therapy.

Ocular involvement

Rheumatoid vasculitis gives rise to a severe form of scleritis (Tessler 1985) and rheumatoid arthritis is one of the common causes of scleritis. Scleritis is very painful, gives blurred vision and may last months to years. Episcleritis is a benign condition, resolving within weeks and not highly associated with rheumatoid arthritis. Rheumatoid nodules are not unusual in scleritis, and may cause scleral thinning, secondary glaucoma, and even perforation (scleromalacia). Uveitis and conjunctivitis are not manifestations of rheumatoid arthritis.

Secondary Sjögren's syndrome

Sjögren's syndrome is described in Chapter 5.10. Secondary Sjögren's syndrome in rheumatoid arthritis is not uncommon. It is distinct from primary Sjögren's syndrome (Table 9), and may start many years after onset of the joint disease. Artificial tears, careful oral hygiene to avoid an increase in caries, and bromhexin (*N*-cyklohexyl-*N*-methyl-E-l-2-amino-3.5-dibromobenzyl amine, or simpler Bisolvon, Boehringer Ingelheim) in doses of 24 mg thrice daily may help in management.

Primary	Secondary
Keratoconjunctivitis sicca	Milder
Xerostomia	Less pronounced
HLA association	HLA DR4; not HLA DR3
Extraglandular manifestations	Not common
Response to therapy	Bromhexin

Table 9 Secondary Sjögren's syndrome in rheumatoid arthritis compared with primary Sjögren's syndrome

Influence of age, sex, and pregnancy

Age

Several studies have compared the onset of rheumatoid arthritis at older age with that at younger age (Deal *et al*. 1985; Sherrer *et al*. 1986). More recently, van der Heijde (1991) analysed certain features in a prospective 2-year study in patients with recent onset of disease. The only features distinguishing the older patients were

a higher biochemical activity, more joint involvement, and persisting marked disease activity after 2 years. Some previous studies have claimed that rheumatoid arthritis with onset in old age is associated with a milder form of the disease, but conversely, Sherrer *et al*. (1986) as well as Sjöblom *et al*. (1984) found age to correlate with a poor prognosis. Van Schaardenburg (1993) compared cases with disease onset over and under 60 years. The only difference was a poorer functional outcome in older seropositive cases, who also had a higher mortality. This is largely due to comorbidity and not to the rheumatoid arthritis as such. However, no population-based data exist and the divergent data may reflect differences in patient selection.

Sex

Female preponderance is present in most series of elderly patients, although to a lesser extent than among younger patients (Silman 1989; Goemaere *et al*. 1990; Jonsson and Larsson 1990). In women, disease onset occurs on average 5 years earlier than in men, and clusters around the menopause. Concomitant health problems, such as osteoporosis and arteriosclerosis, are likely to influence disease course in a negative way. The rate of progression as assessed radiographically in established rheumatoid arthritis did not differ appreciably in relation to gender (Sjöblom *et al*. 1984). Men have a greater risk of developing serious vasculitis complications (Geirson *et al*. 1987).

Pregnancy

The pregnancy-induced amelioration in rheumatoid arthritis, first described by Hench (1938), has been the starting point of both clinical and basic research (Table 10). The relation between sex hormones, the immune system, and autoimmune disease in man and animals is still a much-discussed enigma (Parke 1990). Attempts to ascribe the pregnancy-related remission to a variety of circulating factors, such as a-fetoprotein, adrenal glucocorticoids, or pregnancy-associated immunosuppressive plasma proteins have failed (Klipple and Cecere 1989). The intriguing finding that some 73 per cent of pregnancies induce remission but the remainder do not (Persellin 1977), is elucidated by the recent report (Nelson *et al*. 1993) that improvement correlates with HLA-DQ disparity between mother and fetus. This observation, if confirmed, indicates that an anti-HLA immune response in the mother is implicated. Earlier work in Montpellier (Sany 1994) had used gammaglobulin eluted from human placenta for therapy in rheumatoid arthritis, based on the hypothesis that this contained anti-DR antibodies, which would suppress HLA-DR-expressing cells in the rheumatoid inflammatory tissue. Although such antibodies have indeed been found, this interesting form of therapy has not been pursued.

1938	Hench	Pregnancy-induced improvement of RA
1968	Engel	Marital status as a risk factor for RA
1978	Wingrave and Kay	Reduced incidence of RA in consumers of oral contraceptives
1980	Linos *et al*.	Decreasing incidence of RA in women, 1955–1974
1993	Nelson *et al*.	Maternal–fetal HLA-DQ disparity associated with pregnancy-induced remission

Table 10 Some historical developments relating to rheumatoid arthritis (RA) and pregnancy

Sex hormones

Another line of investigation has focused on the role of sex hormones, greatly stimulated by the report on the apparent protective action of oral contraceptives (Wingrave and Kay 1978) and a declining incidence of the disease in women (Linos *et al*. 1980). A number of conflicting results have been published. Whereas most European reports have found relative risks of around 0.5 for developing rheumatoid arthritis in women who currently or have ever used oral contraceptives, the studies from North America have not shown any protective effect. A recent meta-analysis concluded that there was a small but not statistically significant protective effect (Romieu *et al*. 1989).

More detailed analyses indicate that oral contraceptives mainly reduce the risk for developing severe rheumatoid arthritis (van Zeben *et al*. 1990) explaining why, in general, hospital-based studies showed protection, whereas population-based studies did not (Hazes and van Zeben 1991). Postmenopausal use of hormones does not confer protection, according to two recent studies. The reason for this is not known.

In a study from the United States, married women had an increased risk of developing rheumatoid arthritis (Engel 1968), whereas nulliparous women in the United Kingdom had a higher incidence of rheumatoid arthritis in several studies (Spector *et al*. 1990). In women who were both nulliparous and non-users of oral contraceptives, the risk of developing rheumatoid arthritis was even higher. Whereas the correlations are convincing, the mechanisms are unclear.

Attempts to use high-dose oral contraceptives therapeutically in 10 women with rheumatoid arthritis were not successful (Hazes *et al*. 1989). Two controlled trials in postmenopausal women have shown improvements in bone mineral density as well as improved well being and Ritchie index of joint tenderness, although no changes in erthrocyte sedimentation rate were observed (Hall *et al*. 1994; MacDonald *et al*. 1994). Furthermore, the treatment was well tolerated.

Analysis of androgenic sex hormones in the blood of patients with rheumatoid arthritis has not revealed any abnormalities in women whereas men consistently have low levels, although still within the normal range (Spector 1989). It is possible that this is related to HLA, as the gene for 21-hydroxylase, one of the enzymes involved in the synthesis of sex hormones, is located on the short arm of the sixth chromosome, close to the *HLA-B* locus.

Oestrogen may bind to chondrocytes through oestrogen receptors, or indirectly influence chondrocyte function by inducing cytokine release. In experimental systems, proteoglycan synthesis was inhibited by pharmacological concentrations. T-cell driven models of arthritis are suppressed by oestrogen. It is, however, not known whether these effects are of any relevance in rheumatoid arthritis (Chander and Spector 1991).

Complications

Rheumatoid arthritis is not only a systemic disease with a host of involved organs (see Table 8) but also it is associated with a number of potentially serious complications that contribute to illness and death.

Infections

Patients with rheumatoid arthritis are more afflicted by general infections (Nived *et al*. 1985) and there is anecdotal evidence that respiratory and other infections may trigger flares in rheumatoid arthritis (Wollheim *et al*. 1984). Importantly, rheumatoid arthritis joints are more susceptible to septic arthritis. This may be due to altered T-cell responsiveness and compromised granulocyte function. Local glucocorticoid injections add to the risk for the development of this complication (Östensson and Geborek 1991).

Septic arthritis occurs more often in patients with long-standing disease who are taking oral glucocorticoids. More than one joint may be involved. Usual signs of sepsis, such as fever and leucocytosis, may be absent, and the diagnostic delay may be weeks.

Septic arthritis may easily be overlooked in patients with marked disease activity. The most frequent infecting organism is *Staphylococcus aureus*. The diagnosis may be confirmed by aspiration and positive culture of synovial fluid. However, negative cultures of aspirated fluids are not unusual and cultures of the synovium are a more sensitive and reliable diagnostic method (Kamme and Lindberg 1981). The increased use of immunosuppressive agents probably increases the risk for septic arthritis as well as for other infectious complications. Patients treated with low-dose methotrexate have more respiratory tract and skin infections with relative risks of

1.5 to 2 (van der Veen *et al*. 1994). Furthermore, the introduction of a foreign body in prosthetic joint replacement increases susceptibility to haematogenous septic arthritis.

Neurological involvement

The most common neurological complication in rheumatoid arthritis is entrapment neuropathy secondary to proliferative synovitis. It is most prevalent in the carpal tunnel and in early disease. Rheumatoid arthritis is probably the most common single cause of median-nerve compression (Chamberlain and Bruckner 1970). Carpal tunnel syndrome may be the first symptom that brings the patient to see a doctor. The signs of sensory median-nerve compression are easy to identify as nocturnal paraesthesia of the middle fingers, and may be provoked by forced dorsiflexion of the wrist or percussion over the carpal tunnel. Electromyography shows delayed nerve conduction but is often not more informative than careful clinical examination with testing of motor function. Although median-nerve entrapment in the carpal tunnel is most prevalent, this nerve may also become compressed at the elbow level under the pronator teres muscle. It is important to arrive at a correct diagnosis in these cases, as decompression at the right level will result in prompt restoration of function.

Other entrapments are of the ulnar nerve and rarely the radial nerve at the elbow, as well as the anterior and posterior tibial nerves at the fibular head and in the tarsal tunnel, respectively. The latter gives rise to burning feet and intrinsic foot weakness (Goodgold *et al*. 1965).

An autonomic neuropathy has been defined in rheumatoid arthritis by means of orthostatic electrocardiography (Edmonds *et al*. 1979). It may reflect a microvasculopathy and be related to disturbed skin circulation.

The most alarming form of peripheral neuropathy is called mononeuritis multiplex. It is usually bilateral and most common in the legs. The sudden onset of motor neuropathy signals the presence of aggressive rheumatoid vasculitis and poor prognosis (Geirson *et al*. 1987). The relative incidence of this rare condition is higher in men.

Cervical spine dislocation

Cervical myelopathy associated with rheumatoid arthritis is a well-characterized entity (Marks and Sharp 1981). The symptoms include paraesthesia, numbness, weakness, spastic paralysis, paraplegia, tetraplegia, sensory loss, loss of bladder control, faecal incontinence, and syncope, often in connection with cough or vomiting. Sudden death is also a distinct outcome, particularly if the cervical instability is not recognized (Mikulowski *et al*. 1975).

Cervical subluxation at the atlantoaxial level is present in approximately one-third of patients with rheumatoid arthritis admitted to hospital, and is more prevalent in long-standing disease with pronounced destruction of peripheral joints. The majority of cases are asymptomatic, or suffer from pain without manifest neurological signs. The pain may be mild but is often unbearably intensive. It is usually localized to the neck. Although the atlantoaxial dislocation is augmented by active forward flexion of the head, this manoeuvre rarely precipitates pain.

Some patients suffer from disturbing cracking sensations without pain, when moving the head. Shortening of the neck occurs in extreme cases and an asymmetric posture may be seen.

These patients are particularly vulnerable while under anaesthesia or when involved in accidents and such events may be the starting point of progressive myelopathy.

The diagnosis of unstable subluxation is established by conventional sagittal radiography of the neck in neutral position and anterior flexion, measuring the distance between the anterior odontoid process and atlas (Fig. 5). Subluxation at lower levels between C3 and 4 is another less common complication. However, when present it is much more likely to cause pain and neurological symptoms than atlantoaxial subluxation.

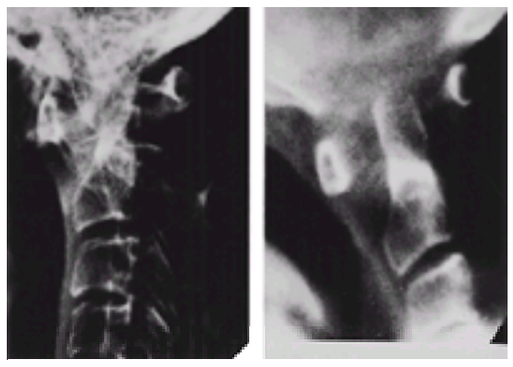


Fig. 5 (a) A lateral film of the cervical spine in flexion in a 68-year-old woman with a long history of rheumatoid arthritis shows a marked increase in the distance between the anterior arch of the atlas and odontoid process, measuring 10.2 mm; normally, it should not exceed 3 mm. (b) Trispiral tomogram demonstrates atlantoaxial subluxation in detail.

Fractures

Several mechanisms contribute to bone loss in severe rheumatoid arthritis (Woolf 1991). Cytokines, such as interleukin 1 and tumour necrosis factor-a, are generated in the inflammatory process and are known to induce osteoclast activity. Inactivity and nutritional deficiency in combination with frequent administration of glucocorticoids constitute additional risk factors. Postmenopausal women with rheumatoid arthritis are thus at particular risk of developing fragility fractures in connection with minor trauma. Stress fractures, occurring without noticed trauma, are also common (Hooyman *et al*. 1984) and often overlooked, as pain is naturally assumed to be related to rheumatoid inflammation as such (Lakhanpal *et al*. 1986). Whereas bone loss is well documented already in early disease (Gough *et al*. 1994a), the role of glucocorticoids and their proper management are still not settled (Woolf 1991). Although a lower bone-mineral content is found in patients on low doses of glucocorticoids, it is not clear whether this is due to disease activity or drug use (Butler *et al*. 1991). Methotrexate inhibits osteoblast activity and may be yet another risk factor in some patients. Pre-existing osteoporosis is probably a risk factor in patients developing rheumatoid arthritis at a late age. There is no evidence to support the likelihood of impaired healing of fractures, not even in patients with marked disease activity.

Tendon and ligament damage

The destructive process in rheumatoid arthritis involves tendons and ligaments and may cause 'spontaneous' rupture. The most common sites of clinically significant involvement are the hands, wrists, shoulders, neck, knees, and feet.

The flexor tendons of the fingers are affected in more than half of patients by tenosynovitis and rheumatoid nodules, causing stenosis, pain, and occasionally rupture. The extensor tendons are likewise commonly detached and ruptures are frequent, starting with the fifth and spreading to the fourth, third, and second in that order. Reconstructive surgery should be performed at an early stage of this development.

Weakening of ligaments causes joint instability and subluxation of metacarpophalangeal joints, allowing ulnar drift and volar subluxation. Loosening of the distal radioulnar ligament likewise gives instability with piano-key sign and volar luxation of the radius and carpus, and prominence of the ulnar head. Rupture of the rotator cuff and supraspinate tendon is common, resulting in 'shrugging' of shoulders when abduction is attempted. At the knee, damage to the lateral and cruciate ligaments causes instability and pain, and in the ankles, valgus deformity is common. Atlantoaxial subluxation only occurs in the presence of weakening or rupture of transverse ligaments.

Amyloidosis

Rheumatoid arthritis remains the most common cause of secondary amyloidosis (amyloid A; **AA**) in Europe, occurring in some 2 to 5 per cent of hospital cases (Lender and Wolf 1972). Geographical differences in prevalence may exist. Amyloidosis was found in 17 out of 81 Japanese autopsies performed between 1960 and 1990 (Suzuki *et al*. 1994). No genetic predisposition has been identified in rheumatoid arthritis. Although related to active and widespread disease with prolonged high levels of C-reactive protein and AA in serum, the onset of clinical signs vary between years and decades. The most common type of organ failure in amyloidosis is uraemia, although the skin, liver, and gastrointestinal tract are microscopically involved in a majority of cases. Five-year survival was reduced to 50 per cent in earlier studies (Wegelius *et al*. 1980), but in more recent series, the use of intensive antirheumatic therapy gives a more favourable outlook of 70 to 93 per cent survivors (Ahlmén *et al*. 1987; Berglund *et al*. 1993). It is therefore important to be aware of the occurrence of amyloidosis and perform diagnostic procedures in suspected cases, usually presenting with proteinuria or decreased renal function (see Chapter 5.13.1).

Felty's syndrome

In 1924 Felty observed the occurrence of splenomegaly and neutropenia in five patients with rheumatoid arthritis (Felty 1924). This triad, although showed by Still in juvenile arthritis several decades earlier, has since been called Felty's syndrome. Felty's syndrome is seen in 1 per cent or less of hospital patients with rheumatoid arthritis, and has the same female/male ratio as rheumatoid arthritis itself (Goldberg and Pinals 1980; Campion *et al*. 1990). It has, with few exceptions, only been observed in Caucasians. The rheumatoid arthritis is typically rather destructive, but signs of ongoing synovitis are not prominent despite radiological progression of erosions. Systemic manifestations are common, with subcutaneous nodules, weight loss, and secondary Sjögren's syndrome in over half of the cases. Hepatomegaly and lymphadenopathy are common. Fever, skin pigmentation, and leg ulcers are other characteristics. The diagnosis is established when splenomegaly and a white cell count of less than $2 \times 10^9/mm^3$ are found on three consecutive occasions. Anaemia and thrombocytopenia are usually not pronounced, and in contrast to systemic lupus erythematosus, there is no lymphopenia. The mean disease duration at diagnosis was 20 years on average in one study, but asymptomatic Felty's syndrome may be present and thus undiagnosed for several years.

Skin ulceration and severe systemic infections constitute the main clinical problems in these patients. Repeated upper and lower respiratory infections also contribute to mortality, which is substantial. Eight of 32 patients died in one centre during 5 years' follow-up, and five of these succumbed with fulminant pneumonia. Although neutropenia is related to susceptibility, there is no quantitative correlation to the degree of leucopenia.

Serological abnormalities are present in most cases (Table 11). Of particular interest are antinuclear antibodies that are granulocyte specific (Faber and Elling 1966) and fix complement (Wiik and Munthe 1974). Plasma from patients with Felty's syndrome induced transient granulocytopenia, and failed to stimulate colony formation from bone marrow cells in culture, indicating the pathogenetic importance of circulating factors. A pathogenetic role for granulocyte– macrophage colony-stimulating factor is indicated by a recent observation, where administration of this factor corrected the neutropenia in a patient, but also elicited release of interleukin 6 and flare up of arthritis (Hazenberg *et al*. 1989).

Antinuclear antibody >80% with HEp-2 cells
Rheumatoid factor >90%, high titre IgG
Cryoglobulins—immune complexes with complement
Complement-fixing granulocyte-specific antinuclear antibodies
Relative hypocomplementaemia

Table 11 Serological markers of Felty's syndrome

The multiple occurrence of rheumatoid arthritis in relatives is more common than expected. HLA DR4 is present in close to 100 per cent and homozygosity of HLA DR4 was present in 14 of 28 tested cases in a recent report (Campion *et al*. 1990). The complement C4B null allele was also associated with Felty's syndrome in 17 of 30 patients. Thus genetic factors clearly distinguish rheumatoid arthritis in Felty's syndrome from other types (Thomson *et al*. 1988).

The management of Felty's syndrome is still controversial. The effect of splenectomy is not convincing and often transient at best. It confers no protection against sepsis. Anecdotal evidence claims responses to parenteral gold, penicillamine, and methotrexate (Campion *et al*. 1990). Two cases have apparently responded to cyclosporin A in combination with methylprednisolone (Camp *et al*. 1991). In view of potential risks with all these regimens a conservative approach to the management of Felty's syndrome is warranted in uncomplicated cases.

Laboratory abnormalities

Ideally, laboratory methods, to be clinically useful, should have one or more of the features listed in Table 12. There is presently a proliferation of new tests, many with promising novel features. However, it is not clear which, if any, of these will stand the test of time (di Giovine *et al*. 1990; Wollheim and Eberhardt 1992).

1. Diagnostic tool
2. Measure of general disease activity
3. Measure of organ-specific activity
4. Prognostic marker
5. Monitoring of therapeutic effectiveness

Table 12 Desirable features of laboratory tests in rheumatoid arthritis

The rheumatoid factor remains the most important aid in establishing the diagnosis of rheumatoid arthritis. The classical Waaler–Rose test, using sheep cells sensitized with rabbit gammaglobulin, combines relative specificity with reasonable sensitivity. Titres have been replaced by units and the use of World Health Organization standard serum should facilitate comparison between different laboratories.

Another development has been the introduction of enzyme-linked immunosorbent assays (**ELISA**) for the IgG, IgA, and IgM subclasses of rheumatoid factor. These can be made highly sensitive. Thus in one series of patients with early rheumatoid arthritis, 67 per cent were positive in the Waaler–Rose test compared with 85 per cent for IgM rheumatoid factor by ELISA (Eberhardt *et al*. 1990b). All three classes of rheumatoid factors are present in the majority of seropositive cases. A minority of cases may have only two and occasionally only one immunoglobulin class in the serum. Much attention has been focused on the putative clinical correlations with early selective occurrence of IgA rheumatoid factors, and some investigators find them predictive of erosive disease whereas others do not. It has not been ruled out

that technical differences, for instance whether human or rabbit IgG has been used for coating the ELISA plates, may explain the discrepancies. At present, however, ELISAs still should be considered research tools.

A large number of antinuclear antibodies are present in increased amounts in rheumatoid arthritis, but this is of more theoretical than practical interest. Granulocyte-specific antinuclear antibodies are present in two-thirds of patients. Another antibody reacting with a soluble nuclear antigen, called anti-RA 33, was present in one-third and seemed rather specific for this disease (Hassfeld *et al*. 1989). Other frequently found antibodies that are specific for histones, single-stranded DNA, or Epstein–Barr virus-related antigens are not specific for rheumatoid arthritis.

In the 1960s, reduced amounts of haemolytic complement in synovial fluid were found to distinguish seropositive rheumatoid arthritis from other forms of chronic arthritis (Zvaifler and Pekin 1963; Hedberg *et al*. 1964). The mechanism behind this characteristic finding is activation and accelerated catabolism of complement products.

Surprisingly, assays for complement in synovial fluid have not become widely used. Complement synthesis is high in rheumatoid inflammatory tissue, and plasma concentrations are often raised or normal. However, the relative concentrations of C3 and C4 are low when compared with other acute-phase reactants, and absolute values may be subnormal in the presence of active vasculitis. Detection of cleavage products, e.g. C2a and C3a, is direct evidence of complement activation, and may be useful for investigative purposes.

The acute-phase reaction in rheumatoid arthritis can be measured by a number of methods (Table 13). The erythrocyte sedimentation rate is most widely used and cheapest. Analysis of individual plasma proteins adds little essential information, although C-reactive protein is a more sensitive indicator of activity. It often remains elevated despite clinical remission, and has been found in some studies to correlate with destructive joint changes (Dawes *et al*. 1986; van Leeuwen *et al*. 1993). The erythrocyte sedimentation rate is influenced by anaemia and immunoglobulins, but this does not limit its usefulness in longitudinal monitoring of rheumatoid arthritis.

Analysis	Degree of change	Oestrogen influence	Glucocorticoids	Other comments
ESR	++++	+	−	Best and cheapest screen test
CRP	+++++		−	Highest amplitude of change
Orosomucoid	++++	−	+	
α_1 Antitrypsin	++	++	−	Liver ++, PiZ gene low levels
Haptoglobin	++++	+	−	Liver −, haemolysis + +
Fibrinogen	++++		−	Requires plasma
Caeruloplasmin	+	++	−	
Viscometry	++++			Expensive, not better than ESR
Complement (C3)	+, ↑, or −		Variable	
Complement (C4)	+, ↑, or −		Variable	C4 null gives low levels

ESR, erythrocyte sedimentation rate; CRP, C-reactive protein.

Table 13 Acute-phase reactants and rheumatoid arthritis

In recent years an increasing number of more or less tissue-specific markers have been identified and assays constructed for clinical use (Table 14). This field holds promise for better future monitoring of the disease process and response to therapy in rheumatoid arthritis (Wollheim and Eberhardt 1992).

Cell	Marker	Compartment
Granulocytes	Lactoferrin	Synovial fluid
	Cytidine deaminase	Serum, synovial fluid
	L1 calprotectin	Serum, synovial fluid
	Elastase	Synovial fluid, serum
Synovial membrane (fibroblasts)	Neopterin	Synovial fluid
	IL-1, TNF, IL-6	Serum, synovial fluid
Lymphocytes	IL-2 receptor	Synovial fluid, serum
	β_2-Microglobulin	Synovial fluid, serum
	$P_{III}NP$	Serum
	Hyaluronan	Serum
	IL-6	Serum, synovial fluid
Endothelial cells	von Willebrand factor	Serum
Cartilage	Proteoglycan fragments	Synovial fluid, serum
	COMP	Synovial fluid, serum
	CMP (148 kDa protein)	Serum
Bone	Deoxypyridinoline	Urine
	Osteocalcin	Serum
	Bone sialoprotein	Serum

CMP, cartilage matrix protein; COMP, cartilage oligomeric protein; IL, interleukin; $P_{III}NP$, type III collagen *N*-terminal propeptide; TNF, tumour necrosis factor.

Table 14 Tissue-specific markets of activity in rheumatoid arthritis

Cytidine deaminase, although not entirely granulocyte specific, has been claimed to be a better marker for inflammation in rheumatoid arthritis than C-reactive protein (Thompson 1987). Calprotectin (L1) may have a similar profile (Berntzen *et al*. 1989). Cytokines and cytokine receptors, although present in trace amounts only, are the subject of numerous studies due to their central position in the pathogenesis of rheumatoid arthritis (di Giovine *et al*. 1996). Increased interleukin 6 levels in synovial fluid but not plasma (Sack *et al*. 1993) and increased amounts of tumour necrosis factor-a and its soluble receptors, and interleukin-2 receptors have been found (Barrera *et al*. 1993).

The concentration of hyaluronate in serum has been correlated with disease activity and morning stiffness in rheumatoid arthritis (Engström-Laurent and Hällgren 1985). Hyaluronate is also a predictor of radiological damage in rheumatoid arthritis (Paimela *et al*. 1991). The type-III collagen *N*-terminal propeptide, $P_{III}NP$, has been used to monitor therapeutic response to disease suppressive therapy and claimed to differ from C-reactive protein (Hørslev-Petersen *et al*. 1988). Cartilage-specific markers have been used and found to vary independently of acute-phase reactants and to correlate with destruction and possibly to stage of disease in rheumatoid arthritis (Wollheim and Saxne 1992; Poole 1994).

Anaemia is common in rheumatoid arthritis and in part related to disease activity as so-called anaemia of chronic disease. This is not responsive to iron therapy, and may be due to hyporesponsiveness to erythropoietin (Means *et al*. 1989). In addition, iron deficiency is present in 30 to 50 per cent of patients with anaemia, as shown by low mean corpuscular volume (Vreugdenhil *et al*. 1990). This may be confirmed by negative iron stains of bone marrow aspirates. Analysis of synovial fluid has traditionally not been much used in evaluation of rheumatoid arthritis. However, it has now been shown that synovial fluid acidosis is related to destruction (Geborek *et al*. 1989b) and that higher concentrations of acid-phosphatase total protein predicted destructive disease (Luukkainen *et al*. 1989). Cytological testing of synovial fluid has been proposed for diagnostic work (Freemont 1991); however, it is not clear whether such observations will become clinically useful in prognostic work.

Lymphopenia was seen in 15 per cent of patients with rheumatoid arthritis. Both CD4 and CD8 cells were equally affected, and changes did not correlate with disease activity or drug therapy (Symmons *et al*. 1989).

Imaging (see also Section 4.9)

As radiography approaches its 100th birthday it remains a cornerstone in the diagnosis and assessment of rheumatoid arthritis. In the last decade, however, scintigraphy, ultrasonography, computed tomography, and magnetic resonance imaging (**MRI**) have emerged as important complementary techniques. Used in concert, they allow rather comprehensive objective documentation of the anatomic disease progression. However, judicious use, based on well-defined questions, will save patients unnecessary exposure to radiation and society unnecessary costs.

Important issues in early rheumatoid arthritis are concerned with presence or absence of erosive changes and with their quantification in the course of disease. Two

methods have emerged to meet the latter need. The Larsen–Dale index (Larsen *et al*. 1977) uses a set of standard radiographs for each joint and distinguishes five stages of involvement, starting with soft tissue swelling and juxta-articular osteoporosis and ending with complete cartilage and advanced bone destruction. This method has the advantage of relative simplicity and has been most used in Europe (Wollheim *et al*. 1988). The Sharp method involves assessing joint space narrowing and counting erosions (Sharp *et al*. 1971). Both are reproducible, but less sensitive to change than the human eye.

Functional impairment correlates with the Larsen index, but is not closely predictable from radiographic changes (Fig. 6). A most impressive example of discrepancy between severe radiographic erosions and preserved function is called 'typus robustus', seen in men, often those employed in heavy manual work.

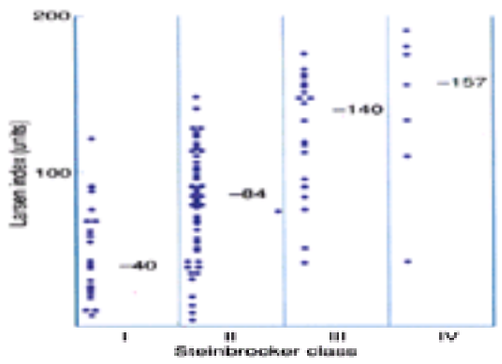


Fig. 6 Larsen index distribution in relation to functional classes according to Steinbrocker. Horizontal bars denote the median.

Scintigraphy with technetium pertechnetate is a simple and useful method to visualize synovial activity, and persistent activity predicts occurrence of radiographic erosions (Möttönen 1988). This promising technique needs to be tested as guide for therapy. Scintigraphy with bone-seeking technetium is useful in early detection of avascular necrosis. Other scintigraphic methods involve labelling of granulocytes with a radionuclide, using their localization after reinfusion to diagnose articular or extra-articular inflammation.

Ultrasonography is being used increasingly, as the technique improves. It has replaced arthrography in the diagnosis of Baker's cysts, and is useful for visualizing rotator-cuff lesions in the shoulder and hip joint effusion (Mitchell 1989).

MRI offers superior contrast resolution compared with other imaging techniques (Brahme 1991). It has been possible to visualize joint cartilage and MRI may be the most sensitive tool to detect early joint lesions. In the evaluation of instability of the neck, MRI allows visualization of granulation tissue around the odontoid process and its change after surgical fusion (Fig. 7). MRI also is a sensitive tool to detect effusion and tendon and ligament abnormality. The initial disadvantage of long exposure time has improved with newer, technically more sophisticated equipment.

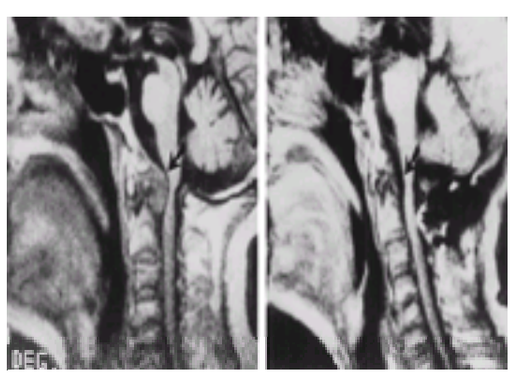


Fig. 7 Preoperative MRI (left) and 4 months postoperative MRI (right) after successful surgical fusion. Arrows indicate periodontoid pannus.

Treatment

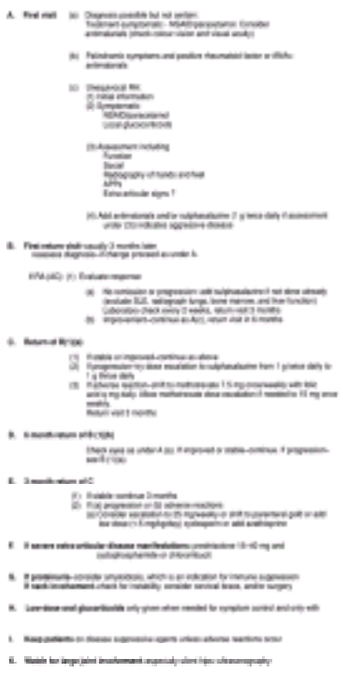
Assessment

Assessment of disease activity, joint damage, and change over time of these variables are essential tools to monitor intervention in rheumatoid arthritis. The traditional methodology of joint examination has been the subject of several evaluations and developments (Preevo *et al*. 1993). In essence, simple counts are better than weighted ones, swelling is a better indicator than tenderness, and a 28-joint count performs as well as more comprehensive 55- to 66-joint counts (Preevo *et al*. 1995). In December 1993 the European League Against Rheumatism published a handbook of standard methods for assessing disease activity. Simple self-assessment questionnaires are also available (Kazis *et al*. 1990). The radiological progression is still considered the gold standard to measure results of treatment. Based on average progression rates from different centres, it has been estimated that regardless of scoring system, due to great variability in progression rate, one would need patient groups of 150 to detect a 50 per cent slowing of radiological progression at a significance level of 0.05 per cent (Sharp *et al*. 1993). Obviously not many studies of disease suppressive agents meet such standards.

Individual goals

In addition to the physical and radiological assessment it is essential to identify and analyse individual disease-related problems and shortcomings, and list them in order of importance to the patient. This is often best accomplished in co-operation with the physiotherapist, occupational therapist, social worker, and metrologist or nurse, also called allied health professionals. Table 15 shows some of the main components of appropriate treatment of rheumatoid arthritis. Some or all of these therapeutic options may be indicated, according to variations in disease severity. It is, however, mandatory that the patient has easy access to a responsible physician with an overview of all the different aspects of management, and it is desirable that the health service provides the necessary organization. Box 1 has a schedule for management of drugs in rheumatoid arthritis.

Box 1 Management of rheumatoid arthritis (RA)



Patient education, adaption, and counselling

Physiotherapy—exercises, joint protection

Occupational therapy—adapt environment, personal aids, orthoses

Medication—pain control, disease suppression

Non-medical pain reduction—transcutaneous nerve stimulation, acupuncture, surgery

Rehabilitation—joint replacement, arthrodesis

Table 15 Main Components of Treatement for Rheumatoid Arthritis

Pharmacotherapy

The number of drugs used in the treatment of rheumatoid arthritis has increased steadily over the last decades, and there have been remarkable geographic differences in therapeutic traditions, often not scientifically based on proper trials. A new and more aggressive approach to early treatment of rheumatoid arthritis has yet to prove that it will change the long-term outcome. Originating in the feeling that most currently used drugs (Table 16) are only symptom modifying, a new terminology was proposed (Edmonds *et al*. 1993). It distinguishes between slow acting/disease modifying and 'disease controlling' drugs and requires sustained (2 year!) improvement of function and slowing or prevention of structural joint damage for the latter. The distinction obviously can only be made in retrospect and seems to be of little practical use at present.

Pain relief
Simple analgesics
Non-steroidal anti-inflammatory drugs

Disease suppression
Chloroquine—hydroxychloroquine
Gold—parenteral
D-Penicillamine
Sulphasalazine
Auranofin
Methotrexate
Azathioprine
Glucocorticoids
Cyclosporin A (now Neoral)

Disease complications
Anaemia: iron, erythropoietin
Osteoporosis: oestrogens, calcitonin, bisphosphonates
Vasculitis: glucocorticoids, chlorambucil, cyclophosphamide
Amyloidosis: chlorambucil, cyclophosphamide

Table 16 Pharmacology of rheumatoid arthritis

Non-steroidal anti-inflammatory drugs (NSAIDs)(see also Chapter 3.5.1)

NSAIDs are the dominating group of drugs and are prescribed periodically or continuously for the vast majority of patients. They offer reliable but limited symptomatic relief from pain and stiffness. Their mechanism of action is complex. NSAIDs inhibit constitutional Cox-1 and induced Cox-2 in varying proportions (Mitchell *et al*. 1994), which may have clinical implications. In addition to inhibition of cycloxygenase they also suppress neutrophil function and in vivo motility, and according to some investigators cause increase of suppressor lymphocyte function and reduced synthesis of rheumatoid factor (Brooks and Day 1991). They also inhibit platelet aggregation. In experimental systems high doses of NSAIDs have been found to inhibit chondrocyte synthesis of proteoglycans. This has caused speculation of potential harmful effects after long-term use in human disease. However, no convincing documentation of cartilage damage in rheumatoid arthritis has been produced.

It is hard to know with certainty if and how NSAID treatment affects long-term outcome in rheumatoid arthritis. There is, however, no doubt among most observers that they improve quality of life in an important way in most patients with rheumatoid arthritis for long periods.

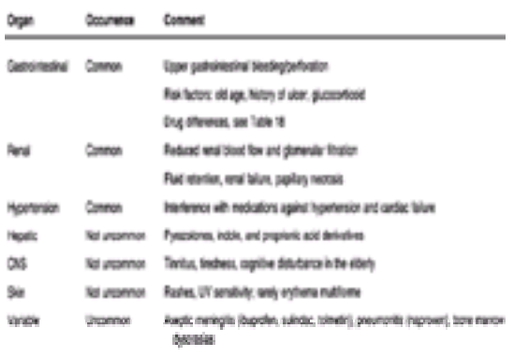
NSAIDs can be divided into two groups, those with a short and those with a long half-life. The former are often available in slow-release formulation, allowing them to be given once or twice daily, for which improved compliance is claimed. It is not possible to relate serum levels to effect, but a dose–response relationship has been established for some NSAIDs, e.g. naproxen and ibuprofen (Brooks and Day 1991).

NSAIDs contain an asymmetric carbon and inactive (*R*) and active (*S*) enantiomers. *In vivo* conversion occurs from (*R*) to (*S*) but not from (*S*) to (*R*). A variable rate of conversion perhaps contributes to variation in response and adverse effects. A large number of comparative short-term trials have on the whole failed to reveal relevant differences between NSAIDs, although individual patient preference has been suggested (Huskisson *et al*. 1976; Scott *et al*. 1982).

Aspirin was until recently the preferred first line NSAID, based on unsurpassed effectiveness and low price. However, a number of trials have shown it to be less well tolerated than other NSAIDs, and thus it is no longer considered a first choice. This does not apply to non-acetylated salicylates. If optimal dose of one NSAID has no satisfactory effect after 3 to 4 weeks, it is advisable to try another. Combined use of two or more NSAIDs is common practice but its rationale is not well founded; they should be prescribed one at a time. Indomethacin administration at night is popular, due to its combined analgesic and sedative effect.

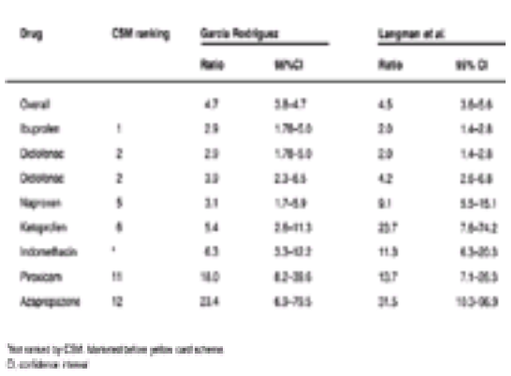
Adverse effects are very common, and may be life threatening (Fries *et al*. 1989; Langman 1989). Awareness and careful patient education are essential (Table 17). Gastric ulcer is more common that duodenal and may be asymptomatic. The overall risk factor is 3 to 4, which means that NSAIDs may be associated with one-quarter of all patients admitted to hospital with these conditions.. Patients at risk and requiring NSAIDs should receive prophylactic treatment with the prostaglandin E_1-analogue, misoprostol. Risk factors are old age, previous peptic ulcer, and concomitant use of glucocorticoids. There seems to be a hierarchy with

regard to upper gastointestinal toxicity among some NSAIDs (Bateman 1994; Table 18). Small bowel toxicity has also been suggested from biopsy studies, which show membranes and strictures as well as Crohn-like lesions (Bjarnason *et al*. 1988; Banerjee 1989). The practical importance of this is not known. Many patients also become constipated and develop nausea from NSAIDs.



Organ	Occurrence	Comment
Gastrointestinal	Common	Upper gastrointestinal bleeding/perforation
		Risk factors: old age, history of ulcer, glucocorticoid
		Drug differences, see Table 18
Renal	Common	Reduced renal blood flow and glomerular filtration
		Fluid retention, renal failure, papillary necrosis
Hypertension	Common	Interference with medications against hypertension and cardiac failure
Hepatic	Not uncommon	Pyrazolones, indole, and propionic acid derivatives
CNS	Not uncommon	Tinnitus, tiredness, cognitive disturbance in the elderly
Skin	Not uncommon	Rashes, UV sensitivity; rarely erythema multiforme
Variable	Uncommon	Aseptic meningitis (ibuprofen, sulindac, tolmetin); pneumonitis (naproxen); bone marrow dyscrasias

Table 17 Adverse reactions of non-steroidal anti-inflammatory drugs



Drug	CSM ranking	Garcia Rodriguez		Langman et al.	
		Ratio	95%CI	Ratio	95% CI
Overall		4.7	3.8–4.7	4.5	3.6–5.6
Ibuprofen	1	2.5	1.78–5.0	2.0	1.4–2.8
Diclofenac	2	2.3	1.78–5.0	2.0	1.4–2.8
Diclofenac	2	3.0	2.3–6.5	4.2	2.8–6.8
Naproxen	5	3.1	1.7–6.9	9.1	5.5–15.1
Ketoprofen	6	5.4	2.6–11.3	23.7	7.6–74.2
Indomethacin	*	6.3	3.3–12.2	11.3	6.3–20.3
Piroxicam	11	16.0	8.2–33.6	13.7	7.1–26.3
Azapropazone	12	23.4	6.3–75.5	31.5	10.3–96.9

*Rank reversed by CSM. Abbreviated before yellow card scheme
CI = confidence interval

Table 18 Odds ratio and 95 per cent CI for bleeding and perforation (Garcia Rodriguez) or acute gastrointestinal bleeding (Langman *et al.*), and CSM rank order of serious reports of gut toxicity expressed per million prescriptions in the first 5 years of marketing (after Bateman 1994)

Impaired renal blood flow is likely to occur in patients with pre-existing kidney diseases, and is thought to be related to the reduced presence of procirculatory prostaglandins. All NSAIDs have this effect, although two studies show sulindac to cause less impairment in short-term trials in patients with moderate renal failure (Eriksson *et al*. 1990; Whalton *et al*. 1990). It is important to avoid dehydration, which may precipitate severe renal failure in patients taking NSAIDs.

The effectiveness of thiazides and loop diuretics is reduced by NSAIDs with the possible exception of sulindac, and the same applies to b-blockers and angiotensin-converting enzyme inhibitors. A number of other pharmacokinetic as well as pharmacodynamic drug interactions have been identified (Brooks and Day 1991). Thus it is advisable to treat NSAIDs with respect, and only use them on strict indications of which rheumatoid arthritis, however, is one. Paracetamol (acetaminophen) in adequate dosage of 1 g three times a day is often a useful analgesic when NSAIDs are contraindicated.

Antimalarials

These are highly potent inhibitors of intracellular protein processing, and they accumulate in the cellular acid-vessel system for transport of macromolecules. They also inhibit monocyte production of interleukin1a and interleukin 6 in vitro (Sperber *et al*. 1993). They have a well-documented action of moderate disease suppression in rheumatoid arthritis, and cause, at an appropriate dosage, very few adverse reactions (Tett *et al*. 1990; Clark *et al*. 1993). Safety and little need for laboratory surveillance make them an inexpensive first choice among disease suppressing agents. Eye toxicity does not occur in patients without liver disease provided the dose of 4 mg/kg per day for chloroquine or 6 mg for hydroxychloroquine is not exceeded. It is also recommended to stop the drug for 4 to 6 weeks each year to check whether the disease may be in remission. Increased creatinine and 10 per cent lowered creatinine clearance may occur (Landewé *et al*. 1995). Antimalarials may diminish toxicity of methotrexate to the liver (Fries *et al*. 1991). Given in combination therapy with aureothiomalate, hydroxychloroqine improved efficacy without increasing adverse reactions (Scott *et al*. 1989). Antimalarials in combination with penicillamine, on the other hand, caused more adverse reactions and less therapeutic effect. Thus the antimalarials have a definite place in the management of rheumatoid arthritis of moderate severity, and may become of use in certain combination regimens.

Soluble gold salts

These have been given parenterally since the 1930s and are still considered among the most potent disease-suppressing agents in rheumatoid arthritis. The mechanism of action is unknown, but experimental data have shown inhibition of antigen presentation by macrophages, perhaps mediated by enzyme inhibition (Champion *et al*. 1990). This results in diminished lymphocyte function. It is remarkable that all clinically used preparations contain AuI and a sulphydryl ligand in an approximate ratio of 1:1. Aureothiomalate has been most studied, and proven superior to placebo and auranofin, an oral gold preparation, and equal to sulphasalazine in efficacy. Inefficacy, therapeutic escape, and not least, adverse reactions, however, seriously limit its use as a single drug (Sambrook *et al*. 1982). An impressive list of adverse reactions to parenteral gold compounds is shown in Table 19. Transient increase in joint pain, also called nitritoid reaction, rarely causes termination. A most serious, usually fatal complication, aplastic anaemia, may be treated successfully with bone marrow transplantation. Agranulocytosis and thrombocytopenia usually are reversible, but eliminate future use of the drug. Adverse reactions are interestingly not related to dose or circulating levels of gold. Aureothiomalate is administered once weekly in doses between 10 and 50 mg. When clinical remission is achieved, injection frequency is reduced to every 2 to 6 weeks. However, in practice, only half of the patients remain on therapy after 2 years and one-fifth after 4 years.

Table 19 Adverse reactions to parenteral gold

D-Penicillamine (dimethylcysteine)

In addition to its chelating effects on divalent metals, such as copper, D-penicillamine has a strongly reactive thiol group and forms mixed disulphides through oxidization. Thus free D-penicillamine has a very short half-life in vivo. It is excreted mainly in the form of penicillamine–cysteine mixed disulphide, but also forms disulphides with albumin and other proteins possessing reactive sulphydryl compounds. Oxidization and chelation occur readily in the gastrointestinal tract, which is why D-penicillamine should be administered without food or other drugs. Even then only half of the dose is absorbed.

The mechanism of action is not known, but D-penicillamine inhibits T-lymphocyte activation probably through a direct T-cell interaction (Lipsky and Ziff 1980). It also reduces the amount of circulating immune complexes and IgA–a_1-antitrypsin complexes as well as rheumatoid factor (Norberg *et al*. 1980). These effects are synchronous with the therapeutic response. Paradoxically, D-penicillamine causes a number of so-called autoimmune complications. An immune-complex glomerulonephritis, histologically identical with that of systemic lupus erythematosus, is accompanied by proteinuria which may last for a year. Occasionally full-blown systemic lupus develops with hypocomplementaemia and antinative DNA (Joyce 1990). Other autoimmune conditions occasionally induced by D-penicillamine are myasthenia, dermato-/polymyositis and Goodpasture's syndrome. After stopping D-penicillamine, these syndromes slowly remit although polymyositis may be fatal if not detected early.

Other side-effects are bone marrow suppression and various forms of skin lesions, including bullous pemphigoid eruptions. HLA DR3 is a risk factor for development of several of these complications. D-Penicillamine is contraindicated in pregnancy due to its teratogenic effect.

The use of D-penicillamine has diminished but it is still a useful drug in some patients. It is best used in doses from 125 to 750 mg/day, with slowly increasing dosage. Combination with sulphasalazine may be advantageous (Taggart *et al*. 1987).

Sulphasalazine

This was designed as a combined analgesic and antimicrobial agent against a putative intestinal pathogen causing rheumatoid arthritis (Svartz 1942). Although this mechanism of action is still hypothetical, its role as an effective disease-suppressing agent in rheumatoid arthritis is well documented (Porter and Capell 1990). The drug is largely but not completely split into 5-aminosalicylic acid and sulphapyridine by bacteria in the colon, and much of its adverse reactions are sulphapyridine related. However, the effect probably at least in part resides in the intact molecule. The onset of effect is 4 to 12 weeks. There is suggestive evidence for retardation of radiological progression if given in early rheumatoid arthritis (van der Heide *et al*. 1989; Hannonen *et al*. 1993).

The most common adverse reactions are nausea, abdominal pain, dizziness, and irritability, and this is only marginally improved by the use of enteric-coated tablets. An acute pneumonitis develops rarely, usually early. Thrombocytopenia occurs in 1:700 patients in the first 3 months of therapy (Keisu and Ekman 1992) but is rarer later on. Serious bone marrow toxicity is, however, uncommon. Spermiogenesis is affected, causing reversible subfertility (Birnie *et al*. 1981). Although some dose-related adverse reactions are more likely to occur in slow acetylators, fast acetylators are as likely to develop serious idiosyncratic reactions, and acetylator state does not influence response. Sulphasalazine is administered orally in doses starting with 0.5 g and increased to a maintenance dose of 1 g twice or three times daily.

Methotrexate

Methotrexate given once weekly, usually orally in doses of 5 to 15 mg, is widely used as a standard disease-suppressive agent, and placebo-controlled trials as well as follow-up studies of up to 90 month are favourable (Grosflam and Weinblatt 1991; Kremer and Phelps 1992), and superior to azathioprine (Jeurissen *et al*. 1991). Survival on the drug is also better than for parenteral gold, azathioprine, D-penicillamine, sulphasalazine, auranofin, and hydroxychloroquine (Morand *et al*. 1992).

The mechanism of action is not known. In addition to an antidihydrofolate reductase effect, methotrexate may have an immunosuppressive effect. Its toxicity may be reduced by administration of folate without loss of therapeutic effect (Morgan *et al*. 1994), indicating at least partly a folate-independent mode of action. The fast onset of clinical response suggests an anti-inflammatory mode of action.

The most common adverse reaction is gastrointestinal intolerance, which usually is not severe. Methotrexate may cause severe septic complications, both fungal and bacterial, and it may be advisable to stop the drug temporarily before arthroplastic surgery. Eosinophilia and pneumonitis arise in some 5 per cent of cases. Renal methotrexate clearance and creatinine clearance decreased by 10 per cent after 6 months (Kremer *et al*. 1995).

The question regarding long-term liver toxicity is still unsettled. Whereas biopsy studies reveal dose-dependent increased liver fibrosis, frank cirrhosis has not occurred with low-dose regimens. Routine liver biopsy is not recommended during initial years of therapy. While NSAIDs increase the toxicity from high-dose methotrexate, this is not reported with low-dose methotrexate, although patients with impaired renal function may be at higher risk.

Combination therapy of methotrexate with gold or azathioprine and hydroxychloroquine, or with hydroxychloroquine and sulphasalazine may be useful (Paulus 1990). A meta-analysis indicates the possibilities of a somewhat lower rate of radiologial progress than for other drugs except parenteral gold (Alarcón *et al*. 1992).

Auranofin

This is a lipophilic gold compound for oral use which is effective in rheumatoid arthritis, although weaker than parenteral gold (Champion *et al*. 1990). A 2-year placebo-controlled prospective study in early rheumatoid arthritis, indicated a retarding effect on radiological progression (Borg *et al*. 1988). The effective dose is 3 mg twice daily. The most common side-effects are diarrhoea and other mild gastrointestinal disturbances. Serious adverse reactions are unusual. The onset of effect is slow, which may be a disadvantage in short-term trials and explain its poor rating in a meta-analysis (Felson *et al*. 1990).

Azathioprine

The purine analogue azathioprine is an effective disease-suppressive drug in rheumatoid arthritis given in doses between 1.5 and 2.5 mg/kg per day, although onset of action is slow (Luqmani *et al*. 1990). Bone marrow toxicity and lymphoma induction are rare complications whereas nausea is common.

Cyclosporin (Neoral)

Cyclosporin is a fungal decapeptide with distinct actions on active T-lymphocytes and possibly antigen-presenting cells, lowering release of interleukin 2. It is an attractive candidate for immune modulation/suppression in rheumatoid arthritis. In doses of 2.5 to 5 mg it was better than placebo in several trials of active rheumatoid arthritis (van Rijthoven *et al*. 1986; Dougados *et al*. 1989; Tugwell *et al*. 1990) and not different from D-penicillamine, azathioprine, and hydroxychloroquine. Evidence for retardation of radiological progression is not convincing (Førre *et al*. 1994), but long-term trials in early rheumatoid arthritis to look into this crucial question are ongoing. A problem with this drug is that it induces hypertension and renal toxicity even in low doses. The long-term consequences of this are not known. Its effect on C-reactive protein is moderate or nil, and the erythrocyte sedimentation rate does not change. Disease activity increases rapidly after the drug is stopped. Cyclosporin A must, if used at all, be carefully monitored for adverse reactions, in particular hypertension and nephrotoxicity (Cash and Klippel 1994; Landewé *et al*. 1994).

Glucocorticoids

The introduction of glucocorticoid treatment of rheumatoid arthritis marks the beginning of modern rheumatology, in demonstrating reversibility of what was considered permanent disability (Hench *et al*. 1949). Glucocorticoids have a large number of effects on immune function and regulation, and cell traffic and adhesion. Inhibition of transcription of the interleukin 1b gene is one important effect (Lee *et al*. 1988). It is the most powerful and predictable remedy inducing immediate symptomatic relief in rheumatoid arthritis available to date, but its long-term effects are among the most controversial issues in the treatment of this disease (George and Kirwan 1990; McDougall *et al*. 1994).

High-dose systemic administration may prevent formation of erosions but adverse effects preclude its uninhibited use (Weiss 1989). The list of side-effects is long (George and Kirwan 1990) and diabetes, hypertension, excessive weight gain, cataract, arteriosclerosis, and not least osteoporosis are among them (Lukert and

Raisz 1990). New approaches to management of osteoporosis involving bisphosphonate, or oestrogen replacement in postmenopausal women are promising, and long-term trials involving low-dose glucocorticoids may become feasible. Patients with rheumatoid arthritis easily become addicted to oral glucocorticoids and it is therefore a serious decision to start such therapy.

Patients treated with NSAIDs have a low concentration of glucocorticoid receptors (Pelletier *et al*. 1994) as well as autoantibodies to lipocortin-1 with impaired responsiveness to glucocorticoids. Furthermore, the hypothalamic–hypophyseal–adrenal axis is hyporesponsive to, for example, surgical stress (Chikanza *et al*. 1992). Findings like these will have to be followed-up by formal clinical trials, before increased use of glucocorticoids can be advocated (Saag *et al*. 1994).

Pulsed intravenous megadoses of 1000 mg methyldprednisolone is no more effective than 100 mg, and this expensive form of administration may be overrated. Intra-articular use of crystalline preparations is, however, both effective and rational, since it delivers the disease-suppressive drug to the involved site. Triamcinolone hexacetonide has the longest lasting effect. If weight-bearing joints are injected a 24-h bed rest should be enforced in order to achieve better effect duration. Iatrogenic septic arthritis is no major problem when using a proper technique, but intra-articular glucocorticoids may predispose to later septic arthritis in the injected joint (Östensson and Geborek 1991).

Alkylating agents

Chlorambucil at 0.1 to 0.2 mg/kg per day and cyclophosphamide at 100 to 150 mg/day, are both active in the suppression of rheumatoid arthritis. No comparative studies have been published, but both drugs induce malignancies after prolonged use (Luqmani *et al*. 1994), and that limits their use in other than those with severe systemic disease (Table 16). Cyclophosphamide may be given in intravenous pulses of 15 mg/kg, which, however, may be no better than daily oral administration (Luqmani *et al*. 1994), except for lower risk of bladder toxicity. Both drugs may cause permanent gonadal failure in men and women. Bone marrow suppression is also a prominent feature, although cyclophosphamide may be slightly less toxic for megakaryocytes.

Combination therapy

This has been mentioned in connection with several drugs and is a matter of current interest (Paulus 1990; Brooks and Schwarzer 1991). Theoretically, benefits could be expected both from combining drugs with similar and different mechanisms of action, provided the adverse reactions differed and were dose related. Few controlled trials have been completed, however (Table 20). A promising combination was that of hydroxychloroquine, sulphasalazine, and methotrexate (O'Dell *et al*. 1994).

Table 20 Summary of studies using combination therapy in rheumatoid arthritis

Disease-suppressive strategies

Frustrated by lack of evidence for long-term disease modification by employment of established disease-suppressive drugs used according to the conventional 'pyramid', several authors have suggested alternative strategies. These include the early use of potent drug combinations, the inverted pyramid or step-down bridge treatment (Wilske and Healey 1990), the cyclic use of drug combinations (Schwarzer *et al*. 1990), and the successive use of various suppressive drugs according to monitored disability, the 'sawtooth strategy' (Fries 1990). The common feature is the aim for early, rapid control of disease activity and systemic glucocorticoid administration. Published data supporting these strategies are not available.

Experimental therapy

Limited success with established disease-suppressive agents in conjunction with increased knowledge of the rheumatoid disease process is stimulating active development of new therapies, targeting inflammatory mechanisms (Luqmani *et al*. 1994). Several lines of progress can be distinguished and an important breakthrough may be close (Table 21). It seems unlikely that one single substance will be able to stop the disease process; rather a multifaceted attack is more likely to be successful in the long term. Different targets depending on the stage and nature of the process may achieve the ultimate 'cure'. Escape mechanisms and redundancy are certainly very prevalent in the pathogenesis of rheumatoid arthritis, so the task is similar to that of Mr Sisyphus.

Target	Principle
The agent	Missing
Activated lymphocytes	Monoclonal antibodies Aphaeresis/photophoresis Oral tolerance induction/vaccination IL-2 fusion toxin, anti-CD5 ricin conjugate
Macrophage cytokine inhibition	Anti-IL-6 antibodies Anti-IL-1/sIL-1 receptor Anti-TNF/sTNF receptor
Cell adhesion molecules	Anti-ICAM-1
Proinflammation prostanoids	Fish- and plant-seed oil Selective Cox-2-inhibitors
Disease activity	Tetracycline Tenidap Bucillamine Podophyllotoxins Subneum Leflunomide Mycophenolic acid

IL, interleukin; TNF, tumour necrosis factor

Table 21 Experimental therapies in development for rheumatoid arthritis

Monoclonal anti-T-cell antibodies

A number of mouse antibodies which have been either chimeric or more recently humanized have been used in open as well as in a few placebo-controlled trials (Kalden 1994; Luqmani *et al*. 1994). The results have been equivocal or negative. Antibodies to CD7, although immunologically effective, have been uniformly ineffective did not influence disease symptoms, indicating that established disease is not dependent on CD7+ cells. Some of the anti-CD4 reagents have induced prolonged depletion from circulation of CD4+ helper cells but only transient clinical effects. Campath1H, a humanized antibody to CDw52, a cell surface structure on activated helper T cells, has induced depletion of CD4+ T cells lasting a year or longer, and a few patients died from infections that may have been drug related. In one study, clinical response correlated to diminished concentrations of tumour necrosis factor-a and interleukin 1. At present non-depleting anti-CD4 reagents are in trials with claims of more encouraging results.

Cytokine inhibition

Tissue destruction in rheumatoid arthritis is mediated by interleukin 6, tumour necrosis factor, and interleukin-1, and all three have been targets for immune therapy (Campion 1994). Interleukin-1 antagonism can be achieved by monoclonal antibodies, or by administering recombinant human interleukin-1 receptor. Trials are ongoing, the problem may be administering large enough doses, as the receptor is rapidly used up.

Remarkable short-term clinical remissions have been achieved with large doses of chimeric tumour necrosis factor-a antibody. This was originally developed for use in sepsis, where it, however, did not work. An open study on 20 patients was confirmed in a placebo-controlled trial. Response durations of 14 to 16 weeks were seen after administration of doses of 10 to 60 mg/kg but repeated treatment resulted in shorter periods of response, probably due to immune elimination of the antibody (Elliott *et al*. 1994a; Elliott *et al*. 1994b). Similar results are emerging using constructs of human Fc from IgG1 and two 55-kDa receptors for tumour necrosis factor. The striking observation has been made that the patients general well being improves within a day, indicating that this therapy really is affecting essential mechanisms. The effects are not lasting more than a few weeks, and immunization is a worry.

Diets and oral tolerance

Diets are of considerable interest and after many years of anecdotal accounts of carers, several controlled studies clearly show symptomatic effect of diets rich in unsaturated fish oil or plant-seed oil (Darlington 1994). One plausible mechanism behind the symptomatic effect seems to be reduced synthesis of inflammatory arachidonic acid products at the expense of eicosapentaenoic acid products. It has been hard to recruit patients for controlled trials and those entered have usually had relatively mild disease (Kjeldsen-Kragh *et al*. 1991). Elimination diets have also been attempted, usually with unconvincing effects (Darlington 1994).

Animal work has suggested oral tolerance induction in experimental autoimmune disease, such as experimental autoimmune encephalitis and adjuvant and collagen arthritis, by administering the antigen orally (Vischer and van Eden 1994). This principle, although claimed effective in initial open trials has not been confirmed in placebo-controlled trials. The suggested mechanism is that the antigen stimulates bystander suppressive pathways (Miller *et al*. 1992). Similar mechanisms may explain the effectiveness of Subreum (OM 8980), an orally administered Escherichia coli preparation (Vischer and van Eden 1994).

Other agents

Several substances are in phase III of clinical development or have actually been licensed in a few countries. Bucillamine is a disulphydryl amino acid similar to D-penicillamine and is much used in Japan. Tenidap is a cyclo-oxygenase inhibitor which also reduces C-reactive protein and is as effective as a NSAID–antimalarial combination (Littman *et al*. 1995). Minocyclin has shown moderate clinical efficacy in a double-blind study, where the interesting observation was a marked reduction in C-reactive protein combined with significant but less impressive clinical effects, indicating that the acute-phase response may not be directly linked to the rheumatoid process (Kloppenburg *et al*. 1994). Podophyllatoxin is a plant extract that in open studies induces prompt reduction in synovitis, acute-phase proteins, and rheumatoid factor concentration (Berglund *et al*. 1980). Unfortunately, it also causes severe abdominal pain or diarrhoea in most patients. A purified derivative called reumacon has weaker but significant clinical effects (Larsen *et al*. 1989). Intravenous immunoglobulin administration with or without apharesis, photophoresis, lymphapheresis, and infusion of retroplacental immunoglobulins have been tried as immunomodulating therapy. These interesting but extremely expensive procedures have yet to prove their efficacy (Sany 1994).

Surgical therapy

Surgery as a treatment in rheumatoid arthritis was made popular in the early 1960s when orthopaedic surgeons working at the Rheumatism Hospital in Heinola, Finland, showed that synovectomies were well tolerated even in active stages of the disease. Much of rheumatological endeavour since that period has been aimed at reducing the need for surgery. It need not be said that we are still far from this goal.

Synovectomy has become infrequent in most places, although pain relief may last for years. However, no retardation of radiological progression is achieved as was initially hoped (Gschwend *et al*. 1974; Doets *et al*. 1989). Tenosynovectomies have remained popular and the fastest and safest relief for nerve entrapment is still surgical decompression. The main indications for surgery are incapacitating pain and restoration or preservation of function. The state of surgical management was the subject of a recent publication (Kelly and Capell 1990).

The most frequent procedures are reconstructive arthroplasties of the hip. Approximately 10 per cent of a patient cohort with early rheumatoid arthritis were operated on within 5 years from onset in our unit (Eberhardt *et al*. 1995). Also common but much less frequently performed are total knee replacements. Shoulder arthroplasties have only recently become feasible and are still far from ideal (Bennett and Gerber 1994). Elbow reconstruction is also possible although loosening has been a major problem (Wadsworth 1993). Arthroplasty procedures of the wrist, fingers, and ankle joints have been devised but are at an experimental stage. Corrective arthrotomies of the metatarsal toe region are common and successful (Helal and Greiss 1984)

Surgical stabilization of the cervical spine to prevent or relieve compression of the medulla is an important task for highly specialized teams. With improved diagnostic and surgical techniques and results, the indications for this form of therapy can be widened (Zygmunt *et al*. 1988; Milbrink and Wingren 1989).

In order to achieve optimal results the rheumatologists and orthopaedic/hand surgeons must work together, preferably in joint clinics with the help of physiotherapists, occupational therapists, and other allied health professionals. Decisions should result from careful analysis of the problem(s) as seen by the patient and from weighing the possible gains and risks of the procedures. An operation always involves risks and is traumatic and tiresome for the patient. Psychological aspects also confound the issue; sometimes an operation is desired to demonstrate to the family how sick the patient is.

The majority of patients admitted to hospital are subjected to one or more operations with time. Two-thirds of these involve the lower limbs, with the often achieved goal of preserving ambulation. Patients with rheumatoid arthritis are at greater risk for acquiring infections, but less likely to have thromboembolic complications. Disease-suppressing therapy usually can be continued, but methotrexate may be an exception, owing to risk of infection. NSAIDs are given as required in most places, apparently with no important risk of bleeding. They have the added advantage of preventing postoperative calcifying myositis.

Surgical treatment is cost effective, since it reduces societies' costs for home service and may preserve productivity and independence (Goldie 1993).

Team rehabilitation

It should be emphasized that rheumatoid arthritis should never be managed without help of a professional team of allied health experts. Patient education is a neglected principle of adapting the individual to cope with a chronic incurable disease and may actually influence pain control and self-esteem favourably (Lindroth *et al*. 1995). Physical therapy may improve self-efficacy, physical capacity, and pain perception (Brighton *et al*. 1993; Stenström 1994). Proper use of technical aids for household and ambulation, adapted environment at home, at work, and in the car are other examples of important therapeutic measures that will reduce helplessness and improve quality of life for the patient. Thus, as mentioned, in most cases a team approach involving doctor(s), physiotherapist, occupational therapist, social worker, and nurse is best suited for the optimal management of rheumatoid arthritis.

Prognosis

There is a growing interest in developing early predictors of long-term outcome in rheumatoid arthritis, which would allow better patient selection for early intervention. Genetic markers and rheumatoid factors have a well-documented relation to disease susceptibility and development of erosions within a year (Gough *et al*. 1994b). The 'shared epitope' in double dose in particular, including DRB1*04 epitopes, also seems to predict a more destructive disease, but the correlation is far too weak to be of use for early therapeutic decisions (Wollheim *et al*. 1995). An in vitro test of Ig-synthesis induced by Epstein–Barr virus also can distinguish erosive from non-erosive cases (Jokinen *et al*. 1994). Early functional impairment is a crude but reliable indicator of poor prognosis (Rasker and Cosh 1989). Rheumatoid arthritis is still a severe disease in a large proportion of patients, although as mentioned (Heikkilä and Isomäki 1994) the proportion of milder cases may be increasing.

Chapter References

Ahlmén, M., Ahlmén, J., Svalander, C., and Bucht, H. (1987). Cytotoxic drug treatment of reactive amyloidosis in rheumatoid arthritis with special reference to renal insufficiency. *Clinical Rheumatology*, **6**, 27–38.

Aho, K., Heliövaara, M., Sievers, K., Maatela, J., and Isomäki, H. (1989). Clinical arthritis associated with positive radiological and serological findings in Finnish adults. *Rheumatology International*, **9**, 7–11.

Alarcón, G.S. *et al.* (1992). Radiographic evidence of disease progression in methotrexate treated and nonmethotrexate disease modifying antirheumatic drug treated rheumatoid arthritis patients: a meta-analysis. *Journal of Rheumatology*, **19**, 1868–72.

Andersson, I., Marsal, L., Nilsson, B., Sjöblom, K.G., and Wollheim, F.A. (1980). Abnormal axillary lymph nodes in rheumatoid arthritis. *Acta Radiologica Diagnostica*, **21**, 645–9.

Arnett, F.C. *et al.* (1988). The American Rheumatism Associaton 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis and Rheumatism*, **31**, 315–24.

Banarjee, A.K. (1989). Enteropathy induced by non-steroidal anti-inflammatory drugs. *British Medical Journal*, **298**, 1539–40.

Barrera, P. *et al.* (1993). Circulating soluble tumor necrosis factor receptors, interleukin-2 receptors, tumor necrosis factor a, and interlukin-6 levels in rheumatoid arthritis. Longitudinal evaluation during methotrexate and azathioprine therapy. *Arthritis and Rheumatism*, **36**, 1070–9.

Bateman, D.N. (1994). NSAIDs: time to re-evaluate gut toxicity. *Lancet*, **343**, 1051–2.

Bennett, P.H. and Burch, T.A. (1967). New York symposium on population studies in the rheumatic diseases; new diagnostic criteria. *Bulletin of the Rheumatic Diseases*, **17**, 453–8.

Bennett W.F. and Gerber C. (1994). Operative treatment of the rheumatoid shoulder. *Current Opinion in Rheumatology*, **6**, 177–82.

Berglund, K., Laurell, A.B., Nived, O., Sjöholm, A.G., and Sturfelt, G. (1980). Complement activation, circulating C1q binding substances and inflammatory activity in rheumatoid arthritis: relations and changes on suppression of inflammation. *Journal of Clinical and Laboratory Immunology*, **4**, 7–14.

Berglund, K., Thysell, H., and Keller, C. (1993). Results, principles and pitfalls in the management of renal AA-amyloidosis: a 10–21 year follow-up of 16 patients with rheumatic disease treated with alkylating cytostatics. *Journal of Rheumatology*, **20**, 2051–7.

Berntzen, H.B., Munthe, E., and Fagerhol, M.K. (1989). A longitudinal study of the leukocyte protein L1 as an indicator of disease activity in patients with rheumatoid arthritis. *Journal of Rheumatology*, **16**, 1416–20.

Birnie, G., McLeod, T., and Watkinson, G. (1981). Incidence of sulphasalazine induced male infertility. *Gut*, **22**, 452–5.

Bjarnason, I., Price, A.B., Gumpel, J.M., and Levi, A.J. (1988). Clinicopathological features of nonsteroidal antiinflammatory drug-induced small intestinal strictures. *Gastroenterology*, **94**, 1070–4.

Bland, J.H. (1974). Rheumatoid arthritis of the cervical spine. *Journal of Rheumatology*, **1**, 319–42.

Borg, G. *et al.* (1988). Auranofin improves outcome in early rheumatoid arthritis. Results from a 2 year, double blind, placebo controlled study. *Journal of Rheumatology*, **15**, 1747–54.

Brahme, S.K. (1991). The role of magnetic resonance imaging in rheumatic disorders. *European Journal of Rheumatological Inflammation*, **11**, 100–3.

Brighton S.W., Lubbe, J.E., and van der Merwe C.A. (1993). The effect of a long-term exercise programme on the rheumatoid hand. *British Journal of Rheumatology*, **32**, 392–5.

Brooks, P.M. and Day, R.O. (1991). Nonsteroidal antiinflammatory drugs—differences and similarities. *New England Journal of Medicine*, **324**, 1716–25.

Brooks, P.M. and Schwarzer, A.C. (1991). Combination therapy in rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **50**, 507–9.

Butler, R.C., Davie, M.W.J., Worsfold, M., and Sharp, C.A. (1991). Bone mineral content in patients with rheumatoid arthritis: relationship to low-dose steroid therapy. *British Journal of Rheumatology*, **30**, 86–90.

Camp, J., Sangro, G., Garcia, N., Subira, M.L., and Prieto, J. (1991). Felty's syndrome: response to cyclosporin A with disappaerance of neutrophil autoantibodies. (Letter.) *Arthritis and Rheumatism*, **34**, 353–5.

Campion, E.G. (1994). Leader: The prospect for cytokine based therapeutic strategies in rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **53**, 485–7.

Campion, G. *et al.*(1990). The Felty syndrome: a case-matched study of clinical manifestations and outcome, serologic features, and immunogenetic associations. *Medicine* (*Baltimore*), **69**, 69–80.

Cash, J.M., and Klippel, J.H. (1994). Second-line drug therapy for rheumatoid arthritis. *New England Journal of Medicine*, **330**, 1368–75.

Chalmers, I.M. and Blair, G.S. (1973). Rheumatoid arthritis of the temporomandibular joint. *Quarterly Journal of Medicine*, **42**, 369–86.

Chamberlain, M.A. and Bruckner, F.E. (1970). Rheumatoid neuropathy: clinical and electrophysiological features. *Annals of the Rheumatic Diseases*, **29**, 609–16.

Champion, G.D., Graham, G.C., and Ziegler, J.B. (1990). The gold complexes. *Baillière's Clinical Rheumatology*, **4**, 491–535.

Chander, C.L. and Spector, T.D. (1991). Oestrogens, joint disease and cartilage. *Annals of the Rheumatic Diseases*, **50**, 139–40.

Chikanza, I.C., Petrou, P., Kingsley, G, Chrousos, G., and Panayi G.S. (1992). Defective hypothalamic response to immune and inflammatory stimuli in patients with rheumatoid arthritis. *Arthritis and Rheumatism*, **35**, 1281–8.

Clark P. *et al.* (1993). Hydroxychloroquine compared with placebo in rheumatoid arthritis. A randomized controlled trial. *Annals of Internal Medicine*, **119**, 1067–71.

Darlington, L.G. (1994). Dietary therapy for rheumatoid arthritis. *Clinical and Experimental Rheumatology*, **12**, 235–9.

Dawes, P.T., Fowler, P.D., Clarke, S., Fisher, J., Lawton, A., and Shadforth, M.F. (1986). Rheumatoid arthritis: treatment which controls the C-reactive protein and erythrocyte sedimentation rate, reduces radiologic progression. *British Journal of Rheumatology*, **25**, 44–9.

Deal, C.L. *et al.* (1985). The clinical features of elderly-onset rheumatoid arthritis. *Arthritis and Rheumatism*, **28**, 987–94.

di Giovine, F.S., Ralston, S.H., and Duff, G.W. (1990). Laboratory and radiologic investigations in the diagnosis and evaluation of rheumatoid arthritis. *Current Opinion in Rheumatology*, **2**, 450–7.

Doets, H.C., Bierman, B.T., and von Soesbergen, R.M. (1989). Synovectomy of the rheumatoid knee does not prevent deterioration. 7-year follow-up of 83 cases. *Acta Orthopaedica Scandinavica*, **60**, 523–5.

Dougados, M., Duchesne, L., Awada, H., and Amor, B. (1989). Assessment of efficacy and acceptability of low dose cyclosporine in patients with rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **48**, 550–6.

Eberhardt, K.B., Rydgren, L.C., Pettersson, H., and Wollheim, F.A. (1990 *a*). Early rheumtoid arthritis—onset, course and prognosis over 2 years. *Rheumatology International*, **10**, 135–42.

Eberhardt, K.B. *et al.* (1990*b*). Disease activity and joint damage progression in early rheumatoid arthritis: relation to IgG, IgA, and IgM rheumatoid factor. *Annals of the Rheumatic Diseases*, **49**, 906–9.

Eberhardt, K.B., Fex, E., Johnsson, K., and Geborek, P. (1995). Hip involvement in early rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **54**, 45–8.

Edmonds, J.P., Scott, D.L., Furst, D.E., Brooks, P., and Paulus, H.E. (1993). Antirheumatic drugs: a proposed new classification. (Editorial.) *Arthritis and Rheumatism*, **36**, 336–9.

Edmonds, M.E., Jones, T.C., and Saunders, W.A. (1979). Autonomic neuropathy in rheumatoid arthritis. *British Medical Journal*, **3**, 173–5.

Elliot, M.J. *et al.* (1994*a*). Repeated therapy with monocloncal antibody to tumour necrosis factor a (cA2) in patients with rheumatoid arthritis. *Lancet*, **344**, 1125–7.

Elliot, M.J. *et al.* (1994*b*). Randomised double-blind comparison of chimeric monoclonal antibody to tumour necrosis factor a (cA2) versus placebo in rheumatoid arthritis. *Lancet*, **344**, 1105–10.

Engel, A. (1968). Rheumatoid arthritis in US adults 1960–62. In *Population studies of the rheumatic diseases* (ed. P.H. Bennet and P.H.N. Wood), pp. 83–8. Excerpta Medica, Amsterdam.

Engström-Laurent, A. and Hällgren, R. (1985). Circulating hyaluronate in rheumatoid arthritis: relationship to inflammatory activity and the effect of corticosteroid therapy. *Annals of the Rheumatic Diseases*, **44**, 83–8.

Eriksson, L.O., Sturfelt, G., Thysell, H., and Wollheim, F.A. (1990). Effects of sulindac and naproxen on prostaglandin excretion in patients with impaired renal function and rheumatoid arthritis. *American Journal of Medicine*, **89**, 313–21.

Faber, E. and Elling, P. (1966). Leucocyte-specific antinuclear factors in patients with Felty's syndrome, rheumatoid arthritis, systemic lupus erythematosus and other diseases. *Acta Medica Scandinavica*, **179**, 257–67.

Feigenbaum, S.L., Masi, A.T., and Kaplan, S.B. (1979). Prognosis in rheumatoid athritis. A longitudinal study of newly diagnosed younger adult patients. *American Journal of Medicine*, **66**, 377–84.

Felson, D.T., Anderson, J.J., and Meenan, R.F. (1990). The comparative efficacy and toxicity of second-line drugs in rheumatoid arthritis. Results of two metaanalyses. *Arthritis and Rheumatism*, **33**, 1449–61.

Felty, A.R. (1924). Chronic arthritis in the adult associated with splenomegaly and leukopenia; a report of five cases of an unusual clinical syndrome. *Bulletin of the Johns Hopkins Hospital*, **35**, 16–20.

Førre, ø. and the Norwegian Arthritis Study Group (1994). Radiologic evidence of disease modification in rheumatoid arthritis patients treated with cyclosporine. *Arthritis and Rheumatism*, **37**, 1506–12.

Freemont, A.J. (1991). Role of cytological analysis of synovial fluid in diagnosis and research. *Annals of the Rheumatic Diseases*, **50**, 120–3.

Fries, J.F. (1990). Reevaluating the therapeutic approach to rheumatoid arthritis: the 'sawtooth' strategy. *Journal of Rheumatology*, Suppl. **22**, 12–15.

Fries, J.F., Miller, S.R., Spitz, P.W., Williams, C.A., Hubert, H.B., and Block, D.A. (1989). Towards an epidemiology of gastropathy associated with nonsteroidal antiinflammatory drug use. *Gastroenterology*, **96** (Suppl. 2), 647–55.

Fries, J.F., Williams, C.A., and Bloch, D.A. (1991). The relative toxicity of nonsteroidal antiinflammatory drugs. *Arthritis and Rheumatism*, **34**, 1353–60.

Geborek, P., Moritz, U., and Wollheim, F.A. (1989 *a*). Joint capsular stiffness in knee arthritis. Relationship to intraarticular volume, hydrostatic pressures, and extensor muscle function. *Journal of Rheumatology*, **16**, 1351–8.

Geborek, P., Saxne, T., Pettersson, H., and Wollheim, F.A. (1989 *b*). Synovial fluid acidosis correlates with radiological joint destruction in rheumatoid arthritis knee joints. *Journal of Rheumatology*, **16**, 468–72.

Geddes, D.M. *et al.* (1977). a_1-Antitrypsin phenotypes in fibrosing alveolitis and rheumatoid arthritis. *Lancet*, **ii**, 1049–50.

Geirson, A.J., Sturfelt, G., and Truedsson, L. (1987). Clinical and serological features of severe vasculitis in rheumatoid arthritis—prognostic implications. *Annals of the Rheumatic Diseases*, **46**, 727–33.

George, E. and Kirwan, J.R. (1990). Corticosteroid therapy in rheumatoid arthritis. *Baillière's Clinical Rheumatology*, **4**, 621–47.

Geterud, Å. (1991). Rheumatoid arthritis in the larynx. A clinical and methodological study. Ph.D. thesis. University of Gothenburg, Sweden.

Goemaere, S. *et al.* (1990). Onset of symptoms of rheumatoid arthritis in relation to age, sex and menopausal transition. *Journal of Rheumatology*, **17**, 1620–2.

Goldberg, J. and Pinals, R.S. (1980). Felty's syndrome. *Seminars in Arthritis and Rheumatism*, **10**, 52–65.

Goldie, I. (1993). Is there any benefit in the surgery of rheumatoid arthritis? *Current Orthopaedics*, **7**, 120–6.

Goodgold, J., Kopell, H.P., and Speilholz, N.I. (1965). The tarsal tunnel syndrome. *New England Journal of Medicine*, **273**, 742–5.

Gough, A.K, Lilley, J., Eyre S., Holder, R.L, and Emery, P. (1994 *a*). Generalized bone loss in patients with early rheumatoid arthritis. *Lancet*, **344**, 23–7.

Gough, A. *et al.* (1994b). Genetic typing of patients with inflammatory arthritis at presentation can be used to predict outcome. *Arthritis and Rheumatism*, **37**, 1166–70.

Grosflam, J. and Weinblatt, M.E. (1991). Methotrexate: mechanism of action, pharmacokinetics, clinical indications, and toxicity. *Current Opinion in Rheumatology*, **3**, 363–8.

Gschwend, N., Winder, J., and Böni, A. (1974). Indikation und Ergebnisse der Synovektomi. *Therpaeutische Umschau*, **31**, 475.

Guillemin, F. *et al.* (1994). Functional disability in early rheumatoid arthritis: description and risk factors. *Journal of Rheumatology*, **21**, 1051–5.

Hall, G.M., Daniels, M., Huskisson E.C., and Spector, T.D. (1994). A randomised controlled trial of the effect of hormone replacement therapy on disease activity in postmenopausal rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **53**, 112–16.

Halla, J.T. *et al.* (1984). Rheumatoid myositis. *Arthritis and Rheumatism*, **27**, 737–43.

Hannonen, P., Möttönen, T., Hakola, M., and Oka, M. (1993). Sulfasalazine in early rheumatoid arthritis. *Arthritis and Rheumatism*, **36**, 1501–9.

Hara, K.S., Ballard, D.J., Ilstrup, D.M., Connolly, D.C., and Vollertsen, R.S. (1990). Rheumatoid pericarditis: clinical features and survival. *Medicine*, **69**, 81–91.

Hassell, A.B. *et al.* (1993). The relationship between serial measures of disease activity and outcome in rheumatoid arthritis. *Quarterly Journal of Medicine*, **86**, 601–7.

Hassfeld, W. *et al.* (1989). Demonstration of a new antinuclear antibody (anti-rheumatoid arthritis 33) that is highly specific for rheumatoid arthritis. *Arthritis and Rheumatism*, **32**, 1515–20.

Hazenberg, B.P., van Leewen, M.A., van Rijswijk, M.H., Stern, A.C., and Vellenga, E. (1989). Correction of granulocytopenia in Felty's syndrome by granulocyte–macrophage colony-stimulating factor. Simultaneous induction of interleukin-6 release and flare-up of the arthritis. *Blood*, **74**, 2769–70.

Hazes, J.M. and van Zeben, D. (1991). Oral contraception and its possible protection against rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **50**, 72–4.

Hazes, J.M.W., Dijkmans, B.A.C., Vandenbrouke, J.P., and Cuts, A. (1989). Oral contraceptive treatment for rheumatoid arthritis: an open study in 10 female patients. *British Journal of Rheumatology*, **28** (Suppl. I), 28–30.

Hedberg, H., Nordén, Å., Lundquist, A., and Afzelius, B. (1964). Depression of hemolytic complement activity of synovial fluid in adult rheumatoid arthritis. *Acta Medica Scandinavica*, **175**, 347–51.

Hedfors, E. Klareskog, L., Lindblad, S., Forsum, U., and Lindahl, G. (1983). Phenotypic characterization of cells within subcutaneous rheumatoid nodules. *Arthritis and Rheumatism*, **26**, 1333–9.

Heikkilä, S. and Isomäki, H. (1994). Long-term outcome of rheumatoid arthritis has improved. *Scandinavian Journal of Rheumatology*, **23**, 13–15.

Helal, B. and Greiss, M. (1984). Telescoping osteotomy for pressure metatarsalgia. *Journal of Bone and Joint Surgery*, **22**, 213–7.

Hench, P.S. (1938). The ameliorating effect of pregnancy on chronic atrophic (infectious rheumatoid) arthritis, fibrosis, and intermittent hydrarthrosis. *Proceedings of the Mayo Clinic*, **13**, 161–7.

Hench, P.S., Kendall, E.C., Slocumb, C.H., and Polley, H.F. (1949). The effect of a hormone of the adrenal cortex (17-hydroxy-11-dehydrocorticosterone compound E) and of pituitary adrenocorticotropic hormone on rheumatoid arthritis. *Proceedings of the Staff Meetings of the Mayo Clinic*, **24**, 181–97.

Hooyman, J.R., Melton, L.J., Nelson, A.M., O'Fallon, W.M., and Riggs, B.L. (1984). Fractures after rheumatoid arthritis. *Arthritis and Rheumatism*, **27**, 1353–61.

Hørslev-Petersen, K., Bendtzen, K.D., Engström-Laurent, A., Junker, P., Halberg, P., and Lorenzen, I. (1988). Serum amino terminal type III procollagen peptide and serum hyaluronan in rheumatoid arthritis: relation to clinical and serological parameters of inflammation during 8 and 24 months' treatment with levamisole, penicillamine, or azathioprine. *Annals of the Rheumatic Diseases*, **47**, 116–26.

Huskisson, E.C., Woolf, D.L., Balme, H.W., Scott, J., and Franklyn, S. (1976). Four new anti-inflammatory drugs: responses and variations. *British Medical Journal*, **1**, 1048–9.

Hyland, R.H. *et al.* (1983). A systemic controlled study of pulmonary abnormalities in rheumatoid arthritis. *Journal of Rheumatology*, **10**, 395–405.

Jacobsson, L.T. *et al.* (1994). Decreasing incidence and prevalence of rheumatoid arthritis in Pima Indians over a twenty-five-year period. *Arthritis and Rheumatism*, **37**, 1158–65.

Jeurissen, M.E.C. *et al.* (1991). Methotrexate versus azathioprine in the treatment of rheumatoid arthritis: a forty-eight-week, randomized, double-blind trial. *Arthritis and Rheumatism*, **34**, 961–72.

Jokinen, E.I., Möttönen, T.T., Hannonen, P.J., Mäkelä, M., and Arvilommi, H.S. (1994). Prediction of severe rheumatoid arthritis using Epstein–Barr virus. *British Journal of Rheumatology*, **33**, 917–22.

Jonsson, B. and Larsson, S.E. (1990). Rheumatoid arthritis evaluated by locomotion score. A population study. *Scandinavian Journal of Rheumatology*, **19**, 223–31.

Joyce, D.A. (1990). D-Penicillamine. *Ballière's Clinical Rheumatology*, **4**, 553–74.

Kalden, J.M. (1994). Biological agents in the therapy of inflammatory rheumatic diseases, including therapeutic antibodies, cytokines and cytokine antagonists. *Current Opinion in Rheumatology*, **6**, 281–6.

Kamme, C. and Lindberg, L. (1981). Aerobic and anaerobic bacteria in deep infections after total hip arthroplasty. Differential diagnosis between infectious and non-infectious loosening. *Clinical Orthopaedics and Related Research*, **154**, 201–7.

Kazis, L.E., Anderson, J.J., and Meenan, R.F. (1990). Health status as a predictor of mortality in rheumatoid arthritis: a five year study. *Journal of Rheumatology*, **17**, 609–13.

Keisu, M. and Ekman, E. (1992). Sulfasalazine associated agranulocytosis in Sweden 1972–1989. Clinical features, and estimation of its incidence. *European Journal of Clinical Pharmacology*, **43**, 215–18.

Kellgren, J.H. (1962). Diagnostic criteria for population studies. *Bulletin of the Rheumatic Diseases*, **12**, 291–2.

Kelly, I.G. and Capell, H.A. (1990). Surgical management of rheumatic diseases. *Annals of the Rheumatic Diseases*, **49** (Suppl 2), 823–82.

Kjeldsen-Kragh, J. *et al.* (1991). Controlled trial of fasting and one-year vegetarian diet in rheumatoid arthritis. *Lancet*, **338**, 899–902.

Klipple, G.L. and Cecere, F.A. (1989). Rheumatoid arthritis and pregnancy. *Rheumatic Disease Clinics of North America*, **15**, 213–39.

Kloppenburg, M., Breedveld, F.C., Terwiel, J.P., Mallee, C., and Dijkmans, B.A.C. (1994). Minocycline in active rheumtoid arthritis. A double-blind, placebo-controlled trial. *Arthritis and Rheumatism*, **37**, 629–36.

Kremer, J.M. and Phelps, C.T. (1992). Long-term prospective study of the use of methotrexate in the treatment of rheumatoid arthritis: update after a mean of 90 months. *Arthritis and Rheumatism*, **35**, 138–45.

Kremer, J.M., Petrillo, G.F., and Hamilton, R.A. (1995). Pharmacokinetics and renal function in patients with rheumatoid arthritis receiving a standard dose of oral weekly methotrexate: association with significant decreases in creatinine clearance and renal clearance of the drug after 6 months of therapy. *Journal of Rheumatology*, **22**, 38–40.

Kurki, P., Aho, K., Palosuo, T., and Heliövaara, M. (1992). Immunopathology of rheumatoid arthritis. Antikeratin antibodies precede the clinical disease. *Arthritis and Rheumatism*, **35**, 914–17.

Lakhanpal, S., McLeod, R.A., and Luthra, H.S. (1986). Insufficiency type stress fractures in rheumatoid arthritis: report of an interesting case and review of the literature. *Clinical and Experimental Rheumatology*, **4**, 151–4.

Landewé, R.B.M., Goei The, H.S., van Rijthoven, W.A.M., Rietveld, J.R., Breedveld, F.C., and Dijkmans, B.A.C. (1994). Cyclosporine in common clinical practice: an estimation of the benefit/risk ratio in patients with rheumatoid arthritis. *Journal of Rheumatology*, **21**, 1631–6.

Landewé, R.B.M. *et al.* (1995). Antimalarial drug induced decrease in creatinine clearance. *Journal of Rheumatology*, **22**, 34–7.

Langman, M.J.S. (1989). Epidemiologic evidence on the association between peptic ulceration and anti-inflammatory drug use. *Gastroenterology*, **96** (Suppl. 2), 640–6.

Larsen, A., Dale, K., and Eek, M. (1977). Radiographic evaluation of rheumatoid arthritis and related conditions by standard reference films. *Acta Radiologica Diagnostica*, **18**, 481–91.

Larsen, A., Petersson, I., and Svensson, B. (1989). Podophyllum derivatives (CPH 82) compared with placebo in the treatment of rheumatoid arthritis. *British Journal of Rheumatology*, **28**, 124–7.

Lee, S.W. *et al.* (1988). Glucocorticoids selectively inhibit the transcription of the interleukin 1b gene and decrease the stability of interleukin 1b mRNA. *Proceedings of the National Academy of Sciences* (*USA*), **85**, 1204–8.

Lender, M. and Wolf, E. (1972). Incidence of amyloidosis in rheumatoid arthritis. *Scandinavian Journal of Rheumatology*, **1**, 109–12.

Lindroth, Y., Bauman, A., Brooks, P.M., and Priestley, D. (1995). A five-year follow-up of a controlled trial of an arthritis education programme. *British Journal of Rheumatology*, **34**, 647–52.

Linos, A.D. *et al.* (1980). The epidemiology of rheumatoid arthritis in Rochester Minnesota. A study of its incidence, prevalence and mortality. *American Journal of Epidemiology*, **111**, 87–98.

Lipsky, P.E. and Ziff, M. (1980). Inhibition of human helper T-cell function in vitro by D-penicillamine and $CuSO_4$. *Journal of Clinical Investigation*, **65**, 1069–76.

Littman, B.H., Drury, C.E., Zimmerer, R.O., Stack, C.B., amd Law, C.G. (1995). Rheumatoid arthritis treated with tenidap and piroxicam. Clinical association with cytokine modulation by tenidap. *Arthritis and Rheumatism*, **38**, 29–37.

Lukert, B.P. and Raisz, L.G. (1990). Glucocorticoid-induced osteoporosis: pathogenesis and management. *Annals of Internal Medicine*, **112**, 352–64.

Luqmani, R.A., Palmer, R.G., and Bacon, P.A. (1990). Azathioprine, cyclophosphamide and chlorambucil. *Baillière's Clinical Rheumatology*, **4**, 595–619.

Luqmani, R., Gordon, C., and Bacon, P. (1994). Clinical pharmacology and modification of autoimmunity and inflammation in rheumatoid disease. *Drugs*, **47**, 259–85.

Luukkainen, R., Isomäki, H., and Kajander, A. (1983). Prognostic value of the type of onset of rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **42**, 274–5.

Luukkainen, R., Kaarela, K., Huhtala, H., Auerma, K., and Merilahti-Palo, R. (1989). Prognostic significance of synovial fluid analysis in rheumatoid arthritis. *Annals of Medicine*, **21**, 269–71.

MacDonald, A.G., Murphy, E.A., Capell, H.A., Bankowska, U.Z., and Ralson, S.H. (1994). Effects of hormone replacement therapy in rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **53**, 54–7.

Marks, J.S. and Sharp, J. (1981). Rheumatoid cervical myelopathy. *Quarterly Journal of Medicine*, **199**, 307–19.

McDougall, R,. Sibley, J., Haga, M., and Russell, A. (1994). Outcome in patients with rheumatoid arthritis receiving prednisone compared to matched controls. *Journal of Rheumatology*, **21**, 1207–13.

Means, R.T., Jr. *et al.* (1989). Treatment of the anemia of rheumatoid arthritis with recombinant human erythropoietin: clinical and in vitro studies. *Arthritis and Rheumatism*, **32**, 638–42.

Mikulowski, P., Wollheim, F.A., Rotmil, P., and Olsen, I. (1975). Sudden death in rheumatoid arthritis with atlanto-axial dislocation. *Acta Medica Scandinavica*, **198**, 445–51.

Milbrink, J. and Wingren, A. (1989). Surgical treatment of atlantoaxial subluxation in rheumatoid arthritis. *Journal of Orthopedic Rheumatology*, **2**, 191–9.

Miller, A., Lider, O., and Weiner, H.L. (1992). Antigen-driven bystander suppression after oral administration of antigens. *Journal of Experimental Medicine*, **174**, 791–8.

Mitchell, J.A. *et al.* (1994). Selectivity of nonsteroidal anti-inflammatory drugs as inhibitors of constitutive and inducible cyclooxygenase. *Proceedings of the National Academy of Science* (*USA*), **90**, 11693–7.

Mitchell, N. (1989). Imaging techniques in joint disease. *Current Opinion in Rheumatology*, **1**, 29–32.

Monsees, B., Destouet, J.M., Murphy, W.A., and Resnick, D. (1985). Pressure erosions of bone in rheumatoid arthritis: a subject review. *Radiology*, **155**, 53–9.

Morand, E.F., McCloud P.I., and Littlejohn, G.O. (1992). Life table analysis of 879 treatment episodes with slow acting antirheumatic drugs in community rheumatology practice. *Journal of Rheumatology*, **19**, 704–8.

Morgan, S.L. *et al.* (1994). Supplementation with folic acid during methotrexate therapy for rheumatoid arthritis. A double-blind, placebo-controlled trial. *Annals of Internal Medicine*, **121**, 833–41.

Moritz, U. (1963). Electromyograpic studies in adult rheumatoid arthritis. (Thesis.) *Acta Rheumatologica Scandinavica*, Suppl. **5**.

Möttönen, T.T. (1988). Prediction of erosiveness and rate of development of new erosions in early rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **47**, 648–53.

Nelson, J.L., Hughes, K.A., Smith, A.G., Nisperos, B.B., Branchaud, A.M., and Hansen, J.A. (1993). Maternal–fetal disparity in HLA class II alloantigens and the pregnancy-induced amelioration of rheumatoid arthritis. *New England Journal of Medicine*, **329**, 466–71.

Nived, O., Sturfelt, G., and Wollheim, F. (1985). Systemic lupus erythematosus and infection. A controlled and prospective study including an epidemiological group. *Quarterly Journal of Medicine*, **55**, 271–87.

Norberg, R., Wollheim, F.A., and Gedda, P.O. (1980). Circulating protein complexes in D-penicillamine therapy of rheumatoid arthritis. Correlation between IgG and a_1-antitrypsin–IgA complexes and clinical response. *Acta Medica Scandinavica*, **208**, 393–6.

O'Dell, J. *et al.* (1994). Triple DMARD therapy for rheumatoid arthritis: efficacy. *Arthritis and Rheumatism*, **37**, s295.

Östensson, A. and Geborek, P. (1991). Septic arthritis as a nonsurgical complication in rheumatoid arthritis: relation to disease severity and therapy. *British Journal of Rheumatology*, **30**, 35–8.

O'Sullivan, J.B. and Cathcart, E.S. (1972). The prevalence of rheumatoid arthritis: follow-up evaluation of the effect of criteria on rates in Sudbury, Massachusetts. *Annals of Internal Medicine*, **76**, 573–7.

Owsianik, W.D.J. *et al.* (1980). Radiological involvement in the dominating hand in rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **39**, 508–10.

Paimela, L., Heiskanen, A., Kurki, P., Helve, T., and Leirisalo-Repo, M. (1991). Serum hyaluronate level as a predictor of radiologic progression in early rheumatoid arthritis. *Arthritis and Rheumatism*, **34**, 815–21.

Palmer, D.G., Hogg, N., Highton, J., Heissian, P.A., and Denholm, I. (1987). Macrophage migration and maturation within rheumatoid nodules. *Arthritis and Rheumatism*, **30**, 729–36.

Parke, A.L. (1990). Pregnancy and rheumatic diseases. *Ballière's Clinical Rheumatology*, **4**, 1–176.

Paulus, H.E. (1990). The use of combinations of disease-modifying antirheumatic agents in rheumatoid arthritis. *Arthritis and Rheumatism*, **33**, 113–20.

Pelletier, J.-P., Dibattista, J.A., Ranger, P., and Martel-Pelletier, J. (1994). The reduced expression of glucocorticoid receptors in synovial cells induced by nonsteroidal antiinflammatory drugs can be reversed by prostaglandin E_1 analog. *Journal of Rheumatology*, **21**, 1748–52.

Persellin, R.H. (1977). The effect of pregnancy on rheumatoid arthritis. *Bulletin of the Rheumatic Diseases*, **27**, 922–7.

Pincus, T. and Callahan, L.F. (1986). Taking mortality in rheumatoid arthritis seriously—predictive markers, socioeconomic status and comorbidity. (Editorial.) *Journal of Rheumatology*, **13**, 841–5.

Poole, R. (1994). Immunochemical markers of joint inflammation, skeletal damage and repair: where are we now? *Annals of the Rheumatic Diseases*, **53**, 3–5.

Porter, D.R. and Capell, H.A. (1990). The use of sulphasalazine as a disease modifying antirheumatic drug. *Baillière's Clinical Rheumatology*, **4**, 535–51.

Preevo, M.L.L. *et al.* (1993). Validity and reliability of joint counts. A longitudinal study in patients with recent onset rheumatoid arthritis. *British Journal of Rheumatology*, **32**, 589–94.

Preevo, M.L.L., van't Hof, M.A., Kuper, H.H., van Leeuwen, M.A., van de Putte, L.B.A., and van Riel, P.L.C.M. (1995). Modified disease activity scores that include twenty-eight joint counts. Development and validation. A prospective longitudinal study of patients with rheumatoid arthritis. *Arthritis and Rheumatism*, **38**, 44–8.

Raman, D. and Haslock, I. (1982). Trochanteritic bursitis—a frequent cause of 'hip' pain in rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **41**, 602–3.

Rasker, J.J. and Cosh, J.A. (1989). Course and prognosis of early rheumatoid arthritis. *Scandinavian Journal of Rheumatology*, Suppl. **79**, 45–56.

Romieu, I., Hernandez-Avila, M., and Lian, M.H. (1989). Oral contraceptives and the risk of rheumatoid arthritis: a metaanalysis of a conflicting literature. *British Journal of Rheumatology*, **28** (Suppl. I), 13–7.

Ropes, M.W., Bennett, G.A., Cobb, S., Jacox, R., and Jessar, R.A. (1958). 1958 revision of diagnostic criteria for rheumatoid arthritis. *Bulletin of the Rheumatic Diseases*, **9**, 175–6.

Saag, K.G. *et al.* (1994). Low dose long-term corticosteroid therapy in rheumatoid arthritis: an analysis of serious adverse events. *American Journal of Medicine*, **96**, 115–23.

Sack, U., Kinne, R.W., Marx, T., Heppt, P., Bender, S., and Emmrich, F. (1993). Interleukin-6 in synovial fluid is closely associated with chronic synovitis in rheumatoid arthritis. *Rheumatology International*, **13**, 45–51.

Sambrook, P.N., Browne, C.D., Champion, G.D., Day, R.O., Vallance, J.B., and Warwick, N. (1982). Terminations of treatment with gold sodium thiomalate in rheumatoid arthritis. *Journal of Rheumatology*, **9**, 932–4.

Sany, J. (1994). Intravenous immunoglobulin therapy for rheumatic diseases. *Current Opinion in Rheumatology*, **6**, 305–10.

Schwarzer, A.C. *et al.* (1990). The cycling of combination antirheumatic drug therapy in rheumatoid arthritis. *British Journal of Rheumatology*, **29**, 445–50.

Scott, D.L., Roden S., Marshall, T., and Kendall, M.J. (1982). Variations in response to non steroidal anti-inflammatory drugs. *British Journal of Clinical Pharmacology*, **14**, 691–4.

Scott, D.L. *et al.* (1989). Combination therapy with gold and hydroxychloroquine in rheumatoid arthritis: a prospective, placebo-controlled study. *British Journal of Rheumatology*, **28**, 128–33.

Sharp, J.T., Lidsky, M.D., Collins, L,C., and Moreland, J. (1971). Methods of scoring the progression of radiologic changes in rheumatoid arthritis: correlation of radiologic, clinical and laboratory abnormalities. *Arthritis and Rheumatism*, **14**, 706–20.

Sharp, J.T. *et al.* (1993). Radiological progression in rheumatoid arthritis: how many patients are required in a treatment trial to test disease modification? *Annals of the Rheumatic Diseases*, **52**, 332–7.

Sherrer, Y.S., Block D.A., Mitchell, D.M., Young, D.Y., and Fries, J.F. (1986). The development of disability in rheumatoid arthritis. *Arthritis and Rheumatism*, **29**, 494–500.

Short, C.L., Bauer, W., and Reynolds, W.S. (1957). *Rheumatoid arthritis.* Harvard University Press, Cambridge, Massachusetts.

Silman, A.J. (1989). Are there secular trends in the occurrence and severity of rheumatoid arthritis? *Scandinavian Journal of Rheumatology*, Suppl. **79**, 25–30.

Sjöblom, K.G., Saxne, T., Pettersson, H., and Wollheim, F.A. (1984). Factors related to the progression of joint destruction in rheumatoid arthritis. *Scandinavian Journal of Rheumatology*, **13**, 21–7.

Spector, T.D. (1989). Sex hormone measurements in rheumatoid arthritis. *British Journal of Rheumatology*, **28** (Suppl. I), 62–8.

Spector, T.D. (1990). Rheumatoid arthritis. In *Epidemiology of rheumatic diseases* (ed. M.C. Hochberg), pp. 513–37. Saunders, Philadelphia.

Spector, T.D. and Scott, D.L. (1988). What happens to patients with rheumatoid arthritis? The longterm outcome of treatment. *Clinical Rheumatology*, **7**, 315–30.

Spector, T.D., Roman, E., and Silman, A.J. (1990). The pill, parity and rheumatoid arthritis. *Arthritis and Rheumatism*, **33**, 782–9.

Sperber, K., Quraishi, H., Kalb, T.H., Panja, A., Stecher, V., and Mayer, L. (1993). Selective regulation of cytokine secretion by hydroxychloroquine: inhibition of interleukin 1 alpha (IL-1-a) and IL-6 in human monocytes and T cells. *Journal of Rheumatology*, **20**, 803–8.

Stenström, C.H. (1994). Home exercise in rheumatoid arthritis functional class II: goal setting versus pain attention. *Journal of Rheumatology*, **21**, 627–34.

Suzuki, A., *et al.* (1994). Cause of death in 81 autopsied patients with rheumatoid arthritis. *Journal of Rheumatology*, **21**, 33–6.

Svartz, N. (1942). Salazopyrin, a new sulfanilinamide preparation. *Acta Medica Scandinavica*, **110**, 577–98.

Symmons, D.P.M., Farr, M., Salmon, M., and Bacon, P.A. (1989). Lymphopenia in rheumatoid arthritis. *Journal of the Royal Society of Medicine*, **82**, 462–3.

Symmons, D.P., Barrett, E.M., Bankhead, C.R., Scott, D.G., and Silman, A.J. (1994). The incidence of rheumatoid arthritis in the United Kingdom: results from the Norfolk Arthritis Register. *British Journal of Rheumatology*, **33**, 735–9.

Taggart, A.J., Hill, J., Asbury, C., Dixon, J.S., Bird, H.A., and Wright, V. (1987). Sulphasalazine alone or in combination with D-penicillamine in rheumatoid arthritis. *British Journal of Rheumatology*, **26**, 32–6.

Tessler, H.H. (1985). The eye in rheumatic disease. *Bulletin of the Rheumatic Diseases*, **35**, 1–8.

Tett, S.E., Cutler, D., and Day, R.O. (1990). Antimalarials in rheumatic diseases. *Baillière's Clinical Rheumatology*, **4**, 467–89.

Thompson, M. and Bywaters, E.G.L. (1961). Unilateral rheumatoid arthritis following hemiplegia. *Annals of the Rheumatic Diseases*, **21**, 370–7.

Thompson, P.W. (1987). Laboratory markers of joint inflammation and damage. *British Journal of Rheumatology*, **26**, 83–5.

Thomson, W., Sanders, P.A., Davis, M., Davidson, J., Dyer, P.A., and Grennan, D.M. (1988). Complement C4B-null alleles in Felty's syndrome. *Arthritis and Rheumatism*, **31**, 984–9.

Tugwell, P. *et al.* (1990). Low-dose cyclosporin versus placebo in patients with rheumatoid arthritis. *Lancet*, **335**, 1051–5.

Vandenbroucke, J.P., Hazeroet, H.M., and Cats, A. (1984). Survival and cause of death in rheumatoid arthritis: a 25 year prospective follow-up. *Journal of Rheumatology*, **11**, 158–61.

van der Heide, D.M.F.M., van Riel, P.L.C.M., Nuver-Zwart, I.H., Gribnau, F.W.J., and van de Putte, L.B.A. (1989). Effects of hydroxychloroquine and sulphasalazine on progression of joint damage in rheumatoid arthritis. *Lancet*, **i**, 1036–8.

van der Heijde, D.M.F.M. (1991). Disease activity and outcome in rheumatoid arthritis. Thesis, University of Nijmegen, pp. 87–96.

van der Veen, M.J., van der Heide A., Kruize, A.A., and Bijlsma, J.W. (1994). Infection rate and use of antibiotics in patients with rheumatoid arthritis treated with methotrexate. *Annals of the Rheumatic Diseases*, **53**, 224–8.

van Leeuwen, M.A. *et al.* (1993). The acute-phase response in relation to radiographic progresion in early rheumatoid arthritis: a prospective study during the first three years of the disease. *British Journal of Rheumatology*, **32**, 3–13.

van Rijthoven, A.W.A.M. *et al.* (1986). Cyclosporine treatment for rheumatoid arthritis: a placebo-controlled double-blind, multicentre study. *Annals of the Rheumatic Diseases*, **45**, 726–31.

van Schaardenburg, D., Hazes, J.M.W., de Boer, A., Zwinderman, A.H., Meijers, K.A.E., and Breedveld, F.C. (1993). Outcome of rheumatoid arthritis in relation to age and rheumatoid factor at diagnosis. *Journal of Rheumatology*, **20**, 45–52.

van Zeben, D., Hazes, J.M.W., Vandenbroucke, J.P., Dijkmans, B.A.C., and Cats, A. (1990). Diminished incidence of severe rheumatoid arthritis associted with oral contraceptive use. *Arthritis and Rheumatism*, **33**, 1462–5.

van Zeben, D., Hazes, J.M.W., Zwinderman, A.H., Vandenbroucke, J.P., and Breedveld, F.C. (1994). The severity of rheumatoid arthritis: a 6-year followup study of younger women with symptoms of recent onset. *Journal of Rheumatology*, **21**, 1620–5.

Vischer, T.L. and van Eden, W. (1994). Oral desensitistaion in rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **53**, 708–10.

Vreugdenhil, G., Baltus, C.A.M., van Eijk, H.G, and Swaak, A.J.G. (1990). Anaemia of chronic disease: diagnostic significance of erythrocyte and serological parameters in iron deficient rheumatoid arthritis patients. *British Journal of Rheumatology*, **29**, 105–10.

Wadsworth, T.G. (1993). Prosthetic replacement of the arthritic elbow. *Current Opinion in Rheumatology*, **5**, 322–8.

Wegelius, O., Wafin, F., Falck, H.M., and Törnroth, T. (1980). Follow-up study of amyloidosis secondary to rheumatic disease. In *Amyloid and amyloidosis* (ed. G.G. Glenner, P.P. de Costa, and A.F. de Freitas), pp. 183–90. Excerpta Medica, Amsterdam.

Weiss, M.M. (1989). Corticosteroids in rheumatoid arthritis. *Seminars in Arthritis and Rheumatism*, **19**, 9–21.

Whalton, A., Stout, R.L., Spilman, P.S., and Klassen, D.K. (1990). Renal effects of ibuprofen, piroxicam and sulindac in patients with asymptomatic renal failure. *Annals of Internal Medicine*, **112**, 560–76.

Wiik, A. and Munthe, E. (1974). Complement fixing granulocyte-specific antinuclear factors in neutropenic cases of rheumatoid arthritis. *Immunology*, **26**, 1127–34.

Wilske, K.R. and Healey, L.A. (1990). Challenging the therapeutic pyramid: a new look at treatment strategies for rheumatoid arthritis. *Journal of Rheumatology*, Suppl. **25**, 4–7.

Wingrave, S.J. and Kay, S.R. (1978). Reduction in incidence of rheumatoid arthritis associated with oral contraceptives. *Lancet*, **i**, 569–71.

Wolfe, F. *et al.* (1994). The mortality of rheumatoid arthritis. *Arthritis and Rheumatism*, **4**, 481–94.

Wollheim, F.A. and Eberhardt, K.B. (1992). The search for laboratory measures of outcome in rheumatoid arthritis. *Baillière's Clinical Rheumatology*, **6**(1), 69–93.

Wollheim, F.A. and Saxne, T. (1992). Markers of cartilage destruction. In *Rheumatoid arthritis—recent research advances* (ed. J. Smolen). Springer-Verlag, Berlin.

Wollheim, F.A., Carlsson, J., Forsgren, A., and Pettersson, H. (1984). Rapidly progressing rheumatoid arthritis: an example of polycloncal B-cell activation? *Clinical Rheumatology*, **3**, 75–9.

Wollheim, F.A., Pettersson, H., Saxne, T., and Sjöblom, K.G. (1988). Radiographic assessment in relation to clinical and biochemical variables in rheumatoid arthritis. *Scandinavian Journal of Rheumatology*, **17**, 445–53.

Wollheim, F.A., Saxne, T., Eberhardt, K.B., and Johnson, U. (1995). The shared epitope and COMP as predictors of erosiveness in recent onset RA. (Abstract.) *15th Eruopean Workshop for Rheumatology Research*, Erlangen, 1995.

Woolf, A.D. (1991). Osteoporosis in rheumatoid arthritis. The clinical viewpoint. *British Journal of Rheumatology*, **30**, 82–4.

Zvaifler, N.J. and Pekin, T.J., Jr. (1963). Complement components in synovial fluids. *Clinical Research*, **11**, 180.

Zygmunt, S., Säveland, H., Brattström, H., Ljunggren, B., Larsson, E.M., and Wollheim, F. (1988). Reduction of rheumatoid periodontoid pannus following posterior occipito-cervical fusion visualised by magnetic resonance imaging. *British Journal of Neurosurgery*, **2**, 315–20.

5.4.3 Juvenile rheumatoid arthritis (rheumatoid factor positive polyarthritis)

Barbara M. Ansell

This subset of juvenile chronic arthritis is clinically and genetically indistinguishable from rheumatoid arthritis in adults. The overall incidence is unknown; in a prospective study of 148 patients with juvenile chronic arthritis seen within 3 months of onset, 6 per cent persistently carried IgM rheumatoid factor in the first 3 months and two further patients became seropositive by the 5-year follow-up (Ansell *et al*. 1987). There was a female preponderance of 2:1 and the disease has been seen in patients as young as 3 or 4 years, although more commonly the illness starts in those aged 10 years or more (Ansell 1983). The frequency in this selected referral group may well be an over-estimate as of 1328 patients on the British Paediatric Rheumatology Group register only 2.1 per cent (28) fall into this category. In the British Paediatric Rheumatology Group register the female preponderance was even more marked at 13:1 and the median age of onset was 11.9 years, with the age range 1.9 to 15 years (Symmons *et al*. 1995, personal communication).

As early as 1969 Hanson and his colleagues, in a prospective study of 110 children, recorded the value of the IgM rheumatoid factor as a marker for this variant (Hanson *et al* 1969); patients whose disease started between the ages of 12 and 16 years accounted for more than 80 per cent of those with positive tests for rheumatoid factor; seropositivity also correlated closely with the presence of nodules. Similarly, in a cross-sectional review of 110 patients whose arthritis commenced under 14 years of age, Cassidy and Valkenberg noted a significant correlation between the serological reaction, age of onset, presence of nodules, and bone erosion (Cassidy and Valkenberg 1967).

Mode of presentation

The disease usually presents as a polyarthritis involving the small joints of the hands and feet. A combination of soft tissue swelling of the wrists and carpi with involvement of metacarpophalangeal and proximal interphalangeal and metatarsophalangeal joints was closely associated with the persistent presence of IgM rheumatoid factor (Ansell and Wood 1976). The large joints can also be involved early, particularly knees and ankles, usually in association with small joint involvement. Very occasionally the onset was palindromic (two in 138 patients), while a few had an insidious onset as 'wrist or foot strain' for a few weeks before the polyarthritis became obvious (Ansell 1983). Those with an onset under the age of 10 years tended to have wrists, ankles and hindfeet, and knees affected early with the small joints becoming affected later. Approximately a quarter of our patients ultimately had a family history of seropositive rheumatoid arthritis (parent, grandparent, or sib), which could post-date the child's illness.

Clinical manifestations

Although lassitude, loss of weight, and general malaise were recorded in about half our patients, fever (even low-grade) was rare. Generalized lymphadenopathy was occasionally present, particularly in those with a very acute onset of polyarthritis. Other systemic manifestations were uncommon, although pleural effusion and pericarditis have been seen, and pulmonary fibrosis in a boy and a girl as the presenting feature.

The pattern is one of a progressive arthritis with severe generalized stiffness; shoulders are often badly affected early, while hip involvement can occur at any time from a few months to several years. Neck, elbows, and hindfeet are gradually involved so that by 5 years a deforming arthritis with weight loss is usual.

Dorsal sheath effusions and severe extensor tenosynovitis are not uncommon and can be associated subsequently with rupture of the extensor tendons of the fourth or fifth fingers and the thumb. Flexor tenosynovitis is common and tends to be more nodular than grossly prolific, causing triggering of fingers. As in adults, carpal tunnel syndrome may complicate the picture as can entrapment neuropathy at knee and elbow. Nodules with typical rheumatoid histology have been seen along the forearm in 30 per cent of patients in the first year of disease (Ansell 1983). Cutaneous nodules were not seen until later and vasculitis is very rare early in the course of the disease.

Diagnosis

As low titres of rheumatoid factor can occur in other autoimmune rheumatic disorders (e.g. systemic lupus erythematosus and some types of vasculitis as well as in hypergammaglobulinaemia, as is seen in sarcoidosis), such causes must be excluded. Because transiently positive rheumatoid factor tests can occur in infections (particularly those of viral original but also bacterial infections e.g. subacute bacterial endocarditis), the European League Against Rheumatism (EULAR) criteria suggest that three consecutive positive tests over a 3-month period are required to classify this type of arthritis.

Using standard tests, IgM rheumatoid factor can be present within a few weeks of the first symptoms and in the majority by 3 months; titres tend to rise as the disease becomes established. It is exceptionally rare for the rheumatoid factor test to become positive after more than 1 year of illness (Ansell 1987); once present, unless long-term slow-acting drugs are used early, it tends to remain positive. The importance of positive test results as a hall-mark of a subgroup of children with polyarthritis who generally have a poor outcome has been stressed (Cassidy *et al*. 1989).

Radiological changes tend to be early in all sites and particularly in the hands and feet. Periostitis along the shafts of the metacarpals and metatarsals and at the bases of the proximal phalanges extending on to the epiphyses of the joints have been seen radiologically as early as 6 weeks, although this is more usual between 3 and 6 months (Ansell and Kent 1977) (Fig. 1). Radiological changes tend to progress rapidly with the development of erosions.

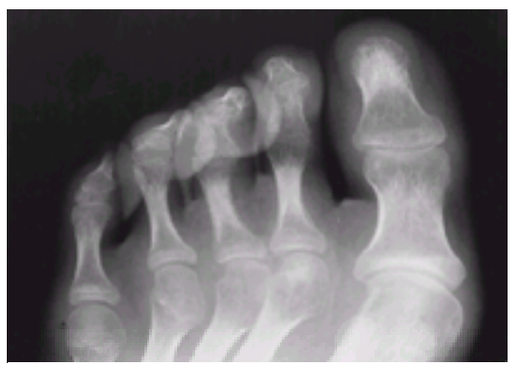


Fig. 1 Periostitis along the proximal phalanx of the second and third toe; note also the severity of the osteoporosis.

Course

In our early studies (Ansell and Wood 1976), the 15-year follow-up showed that a third of the patients had severe limitation of functional capacity (i.e. were unable to function independently) and the majority were still active. Slow-acting drugs, usually gold, had been introduced very late in treatment as previously the problems of this subgroup of seropositive juvenile arthritis had not been adequately recognized. Forty per cent had hip involvement, with destructive changes as early as 1 year; radiologically, protrusio, sometimes severe, had occurred within months from the first hip symptoms (Fig. 2).

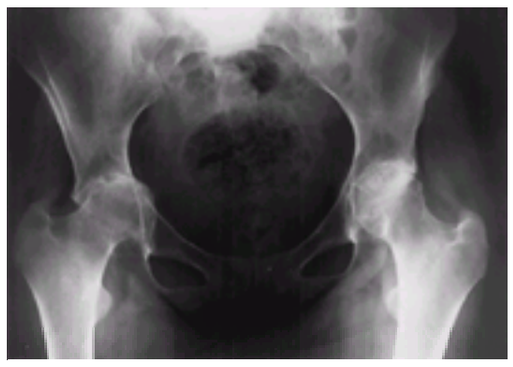


Fig. 2 Both hips were painful and limited in this 14-year-old girl with seropositive disease of 2 years duration. The right shows protrusio with marked narrowing of joint space and erosions; on the left the disease is less advanced.

A review of 138 patients, of whom two-thirds had been followed for at least 10 years, showed that death had occurred from renal failure due to amyloidosis in two, quadriplegia associated with both atlantoaxial and lower cervical subluxation in another, cardiac failure in a further patient, while infections of varying types caused death in the other three; non-fatal complications included vasculitis and peripheral neuropathy. No specific drug was noted to have been effective in suppressing disease activity, but the frequency of side-effects was a significant factor in preventing the maintenance of treatment schedules for long periods. In the minority where the disease was controlled, healing of erosions could occur (Fig. 3(a) and Fig. 3(b)). However, the overall review of 81 serial sets of radiographs (Williams and Ansell 1985), of whom two-thirds had received gold, penicillamine, or chloroquine and 13 per cent had received cytotoxic drugs, while confirming the association between wrist, carpus, and metacarpophalangeal joints in the hands and metatarsophalangeal joints in the feet, showed that lesions in these had progressed radiologically by 5 years (Fig. 4(a),Fig. 4(b) and Fig. 4(c)). In addition, a third of the patients showed erosive changes in large joints such as the hips, knees, or shoulders. Between 5 and 10 years from onset of disease, progression radiologically was evident in most patients with additional joints becoming involved in a further third. By this time, 17 per cent had had bilateral total hip replacement and 7 per cent knee replacement; atlantoaxial subluxation was common (Fig. 5). After 15 years or more, the radiological changes tended to stabilize, but various mechanical difficulties, often secondary to poor growth and degenerative change, as well as the primary destructive arthritis, were evident (Fig. 4(c)). The main differences between adults and children appeared to be a tendency to fusion of the carpal bones and distal interphalangeal joint erosion. Another problem in juveniles is alteration in growth in the presence of persistent disease activity.

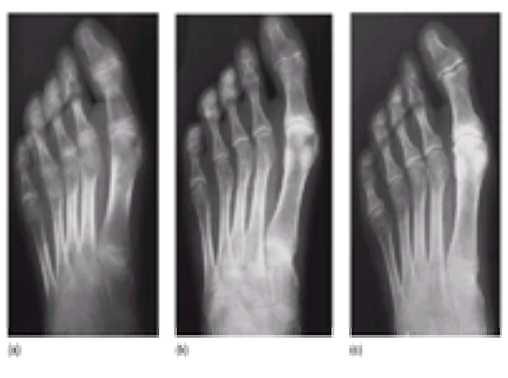


Fig. 3 (a) Radiograph of feet at presentation with disease duration of 3.5 months; note the erosions in the fifth metatarsals and periostitis along the fourth, third, and second proximal phalanges as well as irregularity in the shape of the first metatarsal head. Penicillamine therapy was commenced at this time and continued. (b) Eighteen months from the first picture there has been healing of the erosions in the fifth metatarsophalangeal joint and improvement in overall porosity with no new erosive changes. (c) This improvement has been maintained during the further 3-year treatment period.

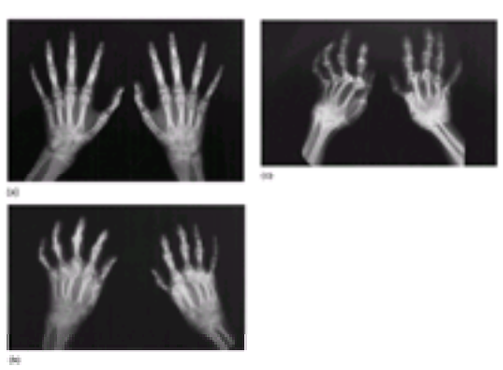


Fig. 4 (a) At presentation this girl (aged 13.5 years) had crowding of the carpus and changes between the distal row of the carpus; the bases of the metacarpi and particularly the head of the second metacarpals on both sides are thinning. (b) Despite a prolonged course of gold, 5 years from onset there has been gross destructive arthritis affecting the carpus which is fusing, all metacarpophalangeal joints, and proximal and distal interphalangeal joints. (c) Twenty years from onset destruction has occurred, particularly in the metacarpophalangeal joints and at the wrists.

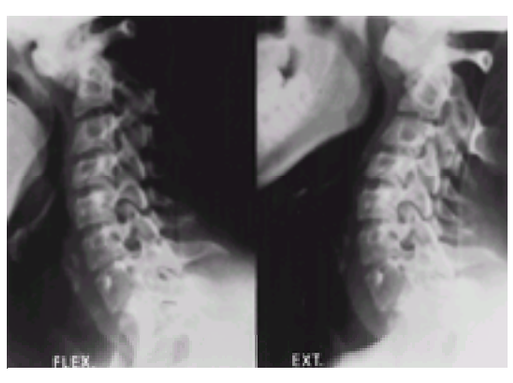


Fig. 5 Atlantoaxial subluxation causing compression on the cord but with relatively little change elsewhere in the cervical spine.

In addition to joint destruction, lone aortic regurgitation has been a serious complication. Reporting four cases in 1981, Leak *et al*. commented on the fact that this pursued a particularly aggressive course with sudden deterioration occurring in two patients, one of whom had an urgent aortic valve replacement, at which time pericarditis was also present, and the other died while awaiting assessment (Leak *et al*. 1981). The patient reported in full developed an aortic diastolic murmur, followed by bouts of chest pain associated with dyspnoea and a rapid deterioration in her clinical state. Histologically, the architecture of the aortic valve had been completely destroyed by multiple necrobiotic foci with the typical features of rheumatoid granulomata. There was active granulomatous tissue growing into the valve. These patients were young at onset (8–11 years) and aortic incompetence was seen to occur as early as 2 years from onset, but it has also been recorded as late as 15 years from onset. All had high titres of IgM rheumatoid factor and a severe destructive arthropathy requiring major joint replacement; three had subcutaneous nodules and one vasculitis. Further cases have been seen since this report; the valvular dysfunction can last only a short time before sudden deterioration (Fig. 6). In view of this, regular cardiac appraisal should be made part of the routine assessment in seropositive juvenile arthritis. Repeated non-invasive assessment can be made by echocardiography; measurement of the left ventricular diastolic dimension and fractional shortening are useful in predicting deterioration in function (O'Rourke *et al*. 1980). Development of the left ventricular dysfunction is an indication for urgent referral for consideration of surgery because the rapid deterioration with the development of angina in a young patient is associated with an increased risk of death (Reicheck *et al*. 1973). In addition to pericarditis and valvular lesions, a single case of endomyocardial fibrosis has been recorded in seropositive juvenile rheumatoid arthritis (Hughes *et al*. 1988). Whether this is indeed a complication of seropositive juvenile rheumatoid arthritis is not known; the patient had been on treatment with penicillamine at the time of her symptoms.

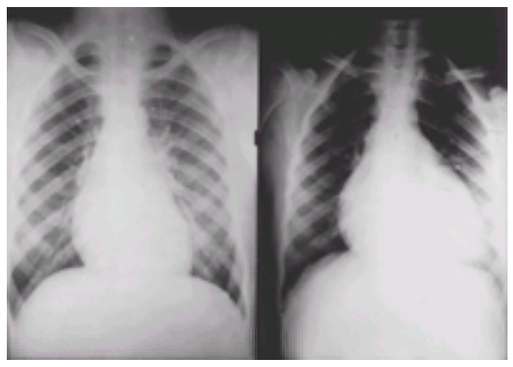


Fig. 6 Rapidly increasing cardiac silhouette over 4 months in a girl who had been noted to have a diastolic murmur 7 months before.

Pulmonary manifestations have been relatively uncommon in our long-term follow-up, although pleurisy with or without effusion has been seen as well as diffuse interstitial disease. Pulmonary arteritis has also been recorded (Gordon and Snyder 1964). The earliest account of rheumatoid lung disease in a child was that of Brinkman and Shaikoff who reported on a 13-year-old who had had progressive dyspnoea and cyanosis over 3 years (Brinkman and Shaikoff 1959). Nodules in the lungs appear excessively rarely, but pulmonary fibrosis is seen very occasionally (Fig. 7). Bronchiolitis obliterans has been reported in a 12-year-old girl with rheumatoid factor positive juvenile rheumatoid arthritis who was restarted on intramuscular gold about a month after mild side-effects had led to temporary withdrawal; this proved fatal despite intensive therapy (Pegg *et al*. 1994).

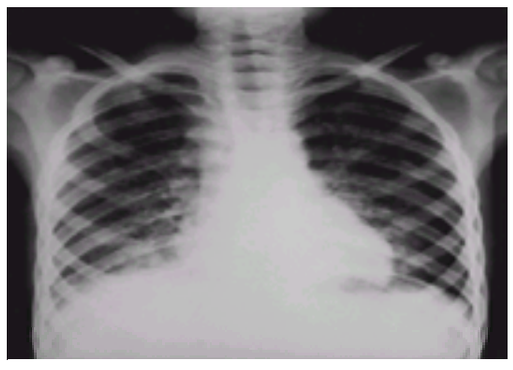


Fig. 7 This 15-year-old patient presented with increasing dyspnoea. Although seropositive, relatively mild juvenile rheumatoid arthritis of some 5-years duration was present affecting the hands and feet only. The duration of the chest symptoms could not be adequately assessed, but he had finger clubbing at presentation.

Genetic aspects

In a study of 52 Caucasian patients, all of whom had developed widespread erosions by 5 years from onset, 62 per cent were HLA DR4 (Clemens *et al*. 1983). Rheumatoid nodules were more frequent in the DR4 patients. There was no difference with respect to mean age of onset, family history of rheumatoid arthritis, or frequency of toxic reactions to long-acting drugs; the toxic reactions had occurred in 23 per cent and were more common in the DR4-negative patients. Although two patients in our group who were DR4 homozygous had required multiple arthroplasties (one had required replacement of the aortic valve within 5 years), the disease was of similar severity in both the DR4 heterozygotes and DR4-negative patients.

Nepom *et al*. found an association with HLA DW4 and HLA DW14 in seropositive polyarthritis juvenile rheumatoid arthritis subjects from the Pacific north-west United States (Nepom *et al*. 1984). Fraser *et al*. noted that late onset polyarticular juvenile rheumatoid arthritis, irrespective of seropositivity, was associated with one of the three common HLA DR4 extended haplotypes, notably HLA B44 SC30 DR4 (Fraser *et al*. 1990). The overall results are difficult to interpret, but there may be ethnic differences between the Caucasian subjects in different areas which would strongly influence the frequencies of the alleles and haplotypes under study.

Box 1 Summary

Seropositive juvenile rheumatoid arthritis is one of the rarer forms of arthritis of childhood. Affecting as it does older children and adolescents, it is important that as soon as the diagnosis is made, they come under a specialist clinic which has information and knowledge and at the same time is able to assess particular drug therapies. To date there has been no good prospective study and from our own personal experience it would seem that patients do better if they are introduced to a slow acting drug within 6 months of the disease, but only through long-term prospective studies, which should be co-ordinated through juvenile rheumatism centres, will this ever be established. In addition to the usual measures of joint protection, maintenance of function, and the use of non-steroidals, the question of local corticosteroid injections into a particularly troublesome joint will also need consideration. As genetically as well as clinically, they are similar to adult rheumatoid arthritis, it would seem that gold or methotrexate are probably the slow acting drugs of choice. There has been no work to date on cyclosporin. Should the onset of disease be under 10 years, it would appear that a proportion of these children will respond to hydrochloroquine; however, if after 6 months the joint count is increasing or the ESR is still very high, it is probably wise to switch to methotrexate or gold.

Management

Principles of management are the same as those for adults, namely maintenance of joint position and function and relief of pain by non-steroidal anti-inflammatory drugs. To date, those studied and accepted for children include ibuprofen, tolmetin, naprosyn, piroxicam, and diclofenac. These last three have been particularly valuable in appropriate dosage.

There have been no satisfactory prospective studies on the use of long-acting drugs, both because of the relative rarity of this subgroup of rheumatoid arthritis, and the problem of obtaining an overall picture. Children may attend paediatric clinics while adolescents become diverted to adult clinics; later the majority will attend adult clinics.

It is my impression that the introduction of a long acting drug within 6 months of the first symptoms will cause considerable improvement in some 70 per cent but nearly 30 per cent do not seem to respond satisfactorily to any present medication.

Chloroquine initially, and more recently hydroxychloroquine, in appropriate dosage has been used in some young patients as the sole long-acting drug. In a number, considerable improvement was seen in the first year of therapy (Hasson *et al*. 1993). Such patients were maintained on therapy for a further 2 years. If there was no benefit, another slow-acting drug was introduced.

At present, although there have been no studies of methotrexate in this subgroup, it would seem to be the slow-acting drug most frequently used. Its efficacy appears to be similar to that in adult rheumatoid arthritis and it is well tolerated. However, recurrence of activity has been seen occurring a year or two after control of the disease with no reduction in dosage. The effect on erosions is also not known, in that although no actual healing of erosions has as yet been demonstrated, certainly in some patients after 1 year no further erosions had developed. More work is required on this aspect.

A comparative study of gold and penicillamine early in the onset of disease was undertaken in 1980/81; this involved 24 patients (Ansell *et al*. 1981). At 1 year both gold and penicillamine were found to be effective; gold came into play more quickly. Three patients, two on gold and one on penicillamine, had stopped therapy by 1 year because of side-effects, but of those who were able to continue all had had a reduction in total active joint count and some fall in erythrocyte sedimentation rate. One in each group treated early in the disease became negative for IgM rheumatoid factor and both of these patients were in remission 3 years from commencement of therapy, with healing of erosions. However, even in this state, as in adults, relapses can occur on maintenance therapy. Overall, toxicity with gold and penicillamine is similar to that in adults but in an open study penicillamine appeared to have had the additional complication of myasthenia gravis, which occurred on several occasions, while one of our patients had myositis.

The role of sulphasalazine has not been adequately studied. Combination therapy in the form of gold and an antimalarial has been effective in a small group of patients; all of these had previously failed to respond adequately to gold or an antimalarial alone. In this group of patients with a relatively poor prognosis as regards persistence of disease activity and increasing erosions, multicentre studies are urgently needed.

For some patients, usually when they have become anaemic, have other complications such as weight loss and very severe arthritis, and at times poor growth, corticosteroids may be employed and whenever possible this should be on an alternate-day regimen. In general, cytotoxic therapy has been reserved for those patients whose disease is complicated by amyloidosis, but it has been used with apparent benefit for patients after aortic valve surgery who still had extremely active disease and two who had lung fibrosis and a further two with vasculitis.

In those few patients followed into adult life who have had children, a postpregnancy relapse of their disease has been seen in all but one, even though their disease was in a reasonable state at the commencement of pregnancy.

At the former Medical Research Coucil Rheumatism Unit at Taplow, some 10 per cent of all patients admitted underwent an orthopaedic surgical procedure. Approximately half of these were in patients suffering from seropositive juvenile rheumatoid arthritis, even though such patients accounted for only just over 10 per cent of admissions during that same period. Total replacement hip arthroplasty was one of the more common operations together with stabilization of the thumb and repair of extensor tendon ruptures; other surgery included fusion of the cervical spine and also of the hindfeet. Revision hip surgery has been required in many as the overall loosening rate has been 25 per cent with an average duration from initial surgery of 9.5 years (Witt *et al*. 1991). There has been a steady increase in the number of total replacement knee arthroplasties undertaken; long-term follow-up has not yet been reported. More recently, shoulder and elbow arthroplasty are under consideration in some special centres.

Chapter References

Ansell, B.M. (1983). Juvenile chronic arthritis with persistently positive tests for rheumatoid factor seropositive JRA. *Annale de Pediatre*, **30**, 545–50.

Ansell, B.M. (1987). Juvenile chronic arthritis. *Scandinavian Journal of Rheumatology Supplement*, **66**, 47–50.

Ansell, B.M. and Kent, P.A. (1977). Radiological changes in juvenile chronic polyarthritis. *Skeletal Radiology*, **1**, 129–44.

Ansell, B.M. and Wood, P.H.N. (1976). Prognosis in juvenile chronic polyarthritis. *Clinics of Rheumatic Diseases*, **2**, 397–412.

Ansell, B.M., Fink, C., and Wood, P.H.N. (1980). Juvenile arthritis in England: a long term follow-up. *Arthritis and Rheumatism*, **23**, 673.

Ansell, B.M., Hall, M.A., and Ruberio, S.A. (1981). A comparative study of gold and penicillamine in seropositive juvenile chronic arthritis (juvenile rheumatoid arthritis). *Annals of the Rheumatic Diseases*, **40**, 522–3.

Brinkman, E.L. and Shaikoff, L. (1959). Rheumatoid lung disease. Report of a case which developed in childhood. *American Review of Respiratory Disease*, **80**, 732–7.

Cassidy, J.T. and Valkenberg, H.A. (1967). A 5 year prospective study of rheumatoid factor test in juvenile rheumatoid arthritis. *Arthritis and Rheumatism*, **10**, 83–90.

Cassidy, J.T., Levinson, J.E., and Brewer, E.J. Jr. (1989). The development of classification criteria for children with juvenile rheumatoid arthritis. *Bulletin on the Rheumatic Diseases*, **38**, 1–7.

Clemens, L.E., Albert, E., and Ansell, B.M. (1983). HLA studies in IgM rheumatoid factor positive childhood arthritis. *Annals of the Rheumatic Diseases*, **42**, 431–4.

Fraser, P.A. *et al*. (1990). HLA extended haplotypes in childhood and adult onset HLA DR4 associated arthropathies. *Tissue Antigens*, **35**, 56–9.

Gordon, J.D. and Snyder, C.H. (1964). Rheumatoid disease of the lung and cor pulmonale. Observations in a child. *American Journal of Diseases of Children*, **108**, 174–80.

Hanson, V., Drexler, E., and Kornreich, M. (1969). The relationship of rheumatoid factor to age of onset in juvenile rheumatoid arthritis. *Arthritis and Rheumatism*, **12**, 82–6.

Hasson, N., Rooney, M., Ansell, B.M., and Woo, P. (1993). Chloroquine in rheumatoid factor positive polyarticular juvenile chronic arthritis. *Arthritis and Rheumatism*, **36**, Suppl 9, Abstract 123.

Hughes, L.O., Randle, S.M., and Raftery, E.G. (1988). Endomyocardial fibrosis in a white girl with seropositive juvenile chronic arthritis. *International Journal of Cardiology*, **18**, 101–5.

Leak, A.M., Miller-Craig, M.W., and Ansell, B.M. (1981). Aortic regurgitation in seropositive JRA. *Annals of the Rheumatic Diseases*, **40**, 229–34.

Nepom, B.S., Nepom, G.T., Michelson, E., Schaller, J.G., Antonelli, P., and Jansen, J.A. (1984). Specific HLA-DR4 associated histocompatability molecules characterised patients with seropositive juvenile rheumatoid arthritis. *Journal of Clinical Investigation*, **74**, 287–91.

O'Rourke, R.A. and Crawford, M.H. (1980). Editorial. Timing of valve replacement in patients with chronic aortic regurgitation. *Circulation*, **61**, 493–5.

Pegg, S.J., Lang, B.A., Mikhail, E.L., and Hughes, D.M. (1994). Fatal bronchiolitis obliterans in a patient with juvenile rheumatoid arthritis receiving chrysotherapy. *Journal of Rheumatology*, **21**, 549–51.

Reicheck, N., Shelburn, J.C., and Perlock, J.K. (1973). Clinical aspects in rheumatoid valvular disease. Aortic regurgitation. *Progress in Cardiovascular Disease*, **15**, 518–23.

Williams, R.A. and Ansell, B.M. (1985). Radiological findings in seropositive JCA with particular reference to prognosis. *Annals of the Rheumatic Diseases*, **44**, 685–93.

Witt, J.D., Swann, M., and Ansell, B.M. (1991). Total hip replacement in JCA. *Journal of Bone and Joint Surgery*, **73B**, 770–3.

5.5.1 Spondylarthropathy, undifferentiated spondylarthritis, and overlap

Andrei Calin

Historical review

As discussed by Calin, in 1974 Moll and colleagues introduced the term 'spondarthritis'(Moll *et al* 1974; Calin 1984). The concept was further developed in 1976 by Wright and Moll in their text entitled *Seronegative polyarthritis* (Wright and Moll 1976). As pointed out by Wright in a chapter entitled 'Relationships between ankylosing spondylitis and other spondarthritides' in Moll's text on ankylosing spondylitis (Wright 1980), we misquoted their term as 'spondylarthritis' in our 1978 monograph on the subject (Calin and Fries 1978). Since then, common usage has resulted in the widespread acceptance of the terms 'spondylarthritis', 'spondylarthropathy', and 'spondyloarthropathy'. We still favour (Calin 1997) the perhaps best known and most commonly applied term 'spondylarthritis', with respect and apologies to Moll, Wright, Khan, and other colleagues who may still prefer one of the other terms (Moll 1980; Wright 1980; Khan 1990). Regardless of preference, all agree that the spondylarthropathies include an exciting and intriguing group of disorders that ranges from asymptomatic sacroiliitis to symptomatic sacroiliitis, widespread multisystem ankylosing spondylitis, enteropathic arthropathies, certain subsets of juvenile onset arthritis, the reactive arthritides, and other entities. These are summarized in Table 1 and Fig. 1. The latter shows how the conditions may overlap with, or develop into, ankylosing spondylitis.

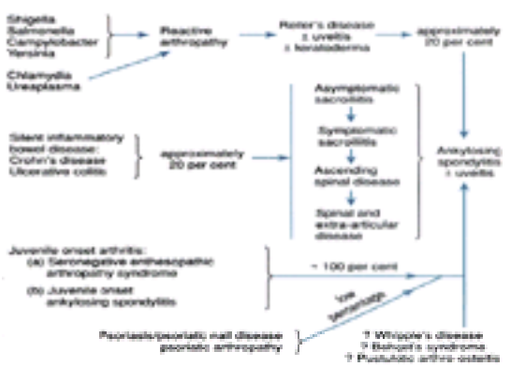


Fig. 1 The inter-related conditions making up the arthropathy group (well differentiated or undifferentiated); * = varying percentage.

Psoriatic arthropathy
Reiter's syndrome/reactive arthropathy (campylobacter, yersinia, shigella, salmonella)
Enteropathic spondylitis (Crohn's disease and ulcerative colitis)
Uveitis
Ankylosing spondylitis Juvenile ankylosing spondylitis Seronegative enthesopathic arthropathy syndrome
Pustulotic arthro-osteitis (considered by the Japanese to be part of spondylarthropathy spectrum (rare in USA and Europe))
Behçet's disease Whipple's disease (doubt exists as to whether these should be considered as part of the spectrum)
Undifferentiated spondylitis (i.e. subset of patients who have spondylarthropathic features but who fail to meet criteria for ankylosing spondylitis, Reiter's syndrome, or other condition, e.g. dactylitis, uveitis plus unilateral sacroiliitis)

Table 1 Individual conditions that overlap to form the spondylarthritides

Introduction

The seronegative spondylarthritides are characterized by involvement of the sacroiliac joints, by peripheral inflammatory arthropathy, insertional tendinitis (enthesopathy), and by the absence of rheumatoid factor (Calin 1989).

There are several other important features:

1. Pathological changes are concentrated at the site of insertion of ligaments or tendons into bone rather than the synovium. Further pathological changes may also develop in the eye, the aortic valve, lung parenchyma, and skin.
2. There is clinical evidence of overlap between the various spondylarthritides. Thus, a patient with psoriatic arthropathy may develop uveitis or sacroiliitis and a patient with inflammatory bowel disease may develop ankylosing spondylitis or mouth ulcers.
3. There is a tendency towards familial aggregation, with evidence that these disorders 'breed true' within families (Calin *et al*. 1984).
4. There is an association with HLA-B27, ranging from about 50 per cent (psoriatic and enteropathic spondylitis) to over 95 per cent (primary ankylosing spondylitis). The specific frequency depends on ethnic group and disease type (Table 2).

	HLA-B27 (%)
Healthy whites	6–14
(e.g. British)	8
(e.g. Northern Scandinavian)	14
Healthy blacks	~ 1–4
(e.g. African black)	~ 1
(e.g. USA black)	4
Indian Asians	2–6
Japanese	1–2
Chinese	5–10
Pakistanis	6–8
Haida Indians	50
Whites:	
Primary ankylosing spondylitis	95
Reactive arthropathy/Reiter's syndrome	80
Enteropathic spondylarthropathy	50
Uveitis	40–50
Psoriatic spondylitis	50–60
Blacks:	
Primary ankylosing spondylitis	50
Reactive arthropathy	50
Indonesian spondylarthropathy patients	8
Healthy Micronesians	9
Healthy Alaskan Eskimo	25–40

Table 2 Distribution of HLA-B27 among different healthy and diseased groups

Diagnostic criteria and classification

Classification or diagnostic criteria for several of the disorders belonging to the spondylarthropathy group already exist, for example the Rome (Kellgren *et al*. 1963), the New York (Bennet and Burch 1968), the Van der Linden *et al*. (Van der Linden *et al*. 1984), and other criteria for ankylosing spondylitis. Our group has long favoured the simple approach (i.e. that of symptomatic sacroiliitis) (Calin 1989a). Likewise, criteria exist for Reiter's syndrome (Willkens *et al*. 1981) and for psoriatic arthropathy (Vasey and Espinoza 1984).

There is a consensus that these criteria are too restricted, as there is a need to emphasize the existence of a much wider disease spectrum. For example radiographically detected sacroiliitis in the absence of symptoms would not be included in the existing classification. Moreover, patients with asymmetric sacroiliitis in addition to, for example, a dactylitis or uveitis, would be excluded from classification and yet clearly are part of the spondylarthropathy spectrum. Furthermore, patients with such limited or atypical forms of disease would be excluded from the typical clinical or epidemiological study. For this reason the European Spondylarthropathy Study Group (ESSG) has proposed criteria for the entire spondylarthropathy group of patients, which would encompass those with clearly defined entities such as Reiter's syndrome or ankylosing spondylitis on the one hand and those with an undifferentiated spondylarthropathy on the other (Dougados *et al*. 1991). In essence, patients with inflammatory spinal pain or asymmetric synovitis predominantly of the lower limb, together with at least one of the following: positive family history, psoriasis, inflammatory bowel disease, enthesopathic lesions, or asymmetric sacroiliitis, have 'undifferentiated spondylarthropathy', with an acceptable sensitivity and specificity. The proposed classification criteria known as the European Seronegative Study Group (ESSG) criteria for spondylarthropathy, will help broaden our acceptance and understanding of the entire spondylarthropathic disorders (Table 3(a)). Parallel to the ESSG criteria, Amor has developed an excellent point-scale that has good sensitivity and specificity in the assessment of patients with spondylarthritis (Table 3(b)). The two sets are compared in Table 3(c) (Amor *et al*. 1991). Reactive arthritis and enteropathic arthropathy are readily defined with Amor's criteria.

Table 3 Diagnostic criteria and classification (a) European Spondylarthropathy Study Group (ESSG) criteria (b) Criteria for diagnosing spondylarthropathies (Amor *et al*. 1991) (c) Comparison of the 12 items criteria (Amor *et al*. 1991) and ESSG criteria

Parallel to the awareness that patients may have a limited or atypical form of spondylarthropathy, there is an appreciation that first degree relatives of HLA-B27-positive probands with classical disease frequently have an inflammatory process that appears to be related in terms of pathology or clinical type to the probands disease and yet would not satisfy any of the above criteria. For this reason several of us have used the term 'undifferentiated spondylarthropathy' (Khan and Van der Linden 1990; Burns and Calin 1984) to describe such individuals. With the proposed new classification (Table 3(a)) all such individuals would be part of the diagnostic group.

Clinical subsets

The interrelated group of conditions constituting the spondylarthropathies have a variety of signs and symptoms (Fig. 2, Table 4). Edmunds *et al*. compared primary ankylosing spondylitis and psoriatic and enteropathic disease in a large controlled study (Edmunds *et al*. 1991). They are categorized according to the specific articular or extra-articular pattern. Depending on the associated clinical features (i.e. urethritis, eye disease, skin involvement) and the way the disease progresses (i.e. remission, relapse) a specific diagnostic label is given. However, as discussed above, clearly defined criteria are frequently absent and one often meets patients who have a spondylarthropathy but in whom the symptoms and signs are such that one is left with an undifferentiated picture. Family and epidemiological studies confirm this clinical finding. For example a patient may appear with unilateral sacroiliitis and little else, or chest wall symptoms due to intercostal muscle insertional tendinitis, and, for example, uveitis. Clearly, the specific phenotypic expression is the end product of a variety of interrelating genetic and environmental factors. Finally, the link between the skin and arthropathy should be stressed (Rosner *et al*. 1993).

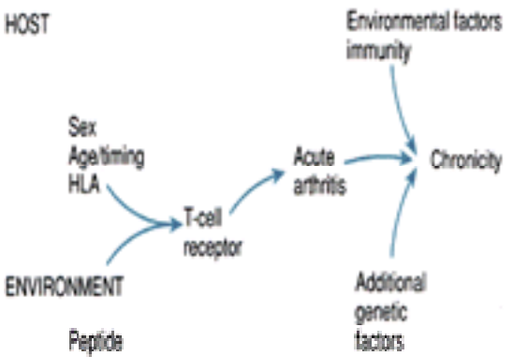


Fig. 2 The relationship between host, environmental, and other factors in determination of phenotypic expression.

	Ankylosing spondylitis	Reiter's syndrome	Psoriatic arthropathy	Enteropathic spondylitis	Juvenile arthropathy (juvenile ankylosing spondylitis subset)	Reactive arthropathy	Undifferentiated spondylo arthropathy
Sex	Male > female	Male=female	Female > male	Female=male	Male > female	Male=female	Male=female
Age of onset (years)	< 20	Any age	Any age	Any age	< 16	Any age	Any age
Uveitis	—	++	–	+	+	+	±
Conjunctivitis	–	+	–	–	–	+	+
Peripheral joints	Lower > upper	Lower usually	Upper > lower	Lower > upper	Lower > upper	Lower > upper	Lower > upper	Again lower
Sacroiliitis	Always	Often	Often	Often	Often	Often	±
HLA-B27(%)	90	80	20 (30 with sacroiliitis)	50	60	80	+
Enthesopathy	+	+	–	+	+	+	+
Aortic regurgitation	+	+	T_6	?	?	+	±
Familial aggregation	+	+	+	+	+	+	+
Risk for HLA-B27 positive individual (%)	1–20	20	?	?	?	20	?
Onset	Gradual	Sudden	Variable	Gradual	Variable	Sudden	Variable
Urethritis	–	+	–	–	–	?	?
Skin involvement	–	+	+/–	+	+	+	?
Mucous membrane involvement	–	+	–	+	–	–	±
Spinal involvement (spinal)	+	+	+	+	+	–	+
Self-limiting	–	?	?	–	?	?	?
Remission / relapses	–	?	?	–	?	?	?

Table 4 Comparison of seronegative spondylarthropathies

Pathogenesis

Even in the reactive arthropathies where the infective trigger is recognized (e.g. yersinia, shigella, salmonella, campylobacter, chlamydia) and the genetic background (HLA-B27) clearly defined, the precise pathogenesis is not well understood (Schur 1994; Inman and Scofield 1994; De Castro 1994; Khan 1996; Lopez-Larrea *et al*. 1996; Scofield 1996). The various steps may include:

1. low-grade inflammatory change in the bowel;
2. the absorption of micro-organisms or parts thereof;
3. endocytosis of the inciting fragments by antigen-presenting cells;
4. degrading of the material followed by linking of the putative antigenic peptide with the HLA molecule;
5. the expression on the cell surface of the combined HLA–peptide complex as a binary product;
6. finally, this HLA–antigenic peptide composite interacts with the T-cell receptor determinant, the three forming a tertiary product.

An acute arthropathy results and, perhaps, following further unknown environmental factors and poorly defined genetic characteristics (Fig. 2 and Fig. 3), the chronic disease state may develop (McClean *et al*. 1993; Scofield *et al*. 1993; Stieglitz and Lipsky 1993; Khan 1993; Rojo *et al*. 1993; Whelan and Archer 1993; De Vries *et al*. 1992; Hermann *et al*. 1993; Taurog *et al*. 1993a; Breban *et al*. 1993a; Skurnik *et al*. 1993; Madden *et al*. 1992). Why the sacroiliac joints are preferentially involved remains unknown (Braun and Sieper 1996).

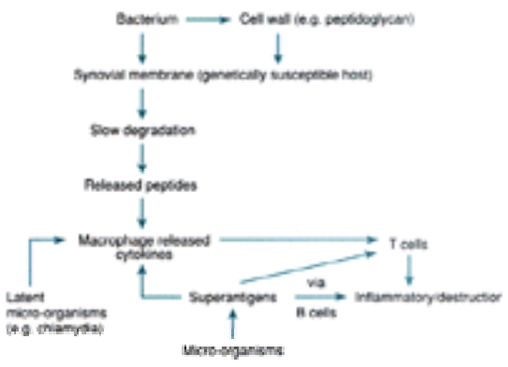


Fig. 3 The putative pathway between the infective trigger and disease pathogenesis.

HLA-B27

The HLA-B27 molecule is a two-chain structure that consists of a polymorphic glycoprotein, termed the heavy or a chain. The a chain is encoded by the HLA complex and is noncovalently linked to a non-HLA-encoded, non-polymorphic protein, b $_2$-microglobulin. The entire molecule is anchored to the cell membrane by the heavy chain. The extracellular portion of the heavy chain is divided into three domains, termed a $_1$, a_2, and a_3, each of which contains approximately 90 amino acids. The a_1 and a_2 domains are distal to the cell membrane and it is these that demonstrate the greatest HLA polymorphism. The a $_1$ and a_2 domains each consist of four b strands and an a helix. The eight b strands of these two domains form a b-pleated sheet or platform, which in turn supports the two a helices. These a helices create a groove which serves as the antigen-binding site for the putative peptide fragment which has been processed from the larger arthritogenic antigen. The two a helices with the bound antigen fragment comprise a ligand which is recognized by the T-cell receptor on a class 1-restricted, CD8-positive T cell.

There are several hypotheses explaining the HLA-B27 association with spondylarthropathy. In contrast to the situation with Reiter's disease, where clearly defined organisms are known to be operative (e.g. chlamydia, shigella, salmonella, campylobacter), the trigger for ankylosing spondylitis remains unclear and could either be one of the above organisms or viruses, or indeed other environmental phenomena (Ebringer 1991).

The first hypothesis suggests that a particular B27 molecule can act as a receptor for the aetiological agent. Support for such an hypothesis comes from observations that other cell surface molecules can act as receptors for viruses, such as CD4 for the human immunodeficiency virus. Individuals with B27 who contact the putative trigger will develop disease.

The second hypothesis suggests that the antigen-binding groove of only certain HLA molecules can accept the processed antigenic fragment that is ultimately responsible for causing disease. Thus, when the ankylosing spondylitis-causing organism enters the cell the antigens are degraded to peptides and only certain HLA molecules accept the antigen. The HLA molecule–antigenic peptide is then presented to the T cell and disease takes place.

The third hypothesis postulates that the T-cell antigen receptor which recognizes the HLA molecule-peptide complex is responsible for disease but because the T-cell recognition is restricted by an HLA molecule there is an association between the disease and B27.

Finally, there is an hypothesis related to the 'molecular mimicry' phenomenon. Here the peptide derived from the organism that causes disease is immunologically identical to HLA-B27. Therefore, the peptide is not recognized as foreign—no immune response is mounted and the disease process develops. Alternatively, the peptide is recognized as foreign and a vigorous immune response is mounted against the organism but the response cross reacts with self tissue causing disease. (For reviews see Khan 1993; Khan 1996; Schur 1994; Inman and Scofield 1994; De Castro 1994).

We still do not know which of these many possibilities—if any—is correct. Interestingly, certain strains of shigella which contain a plasmid (an extrachromosomal piece of DNA) are known to be arthritogenic, causing Reiter's syndrome. Recent data suggest that B27-positive individuals share structural and immunological homology with a sequence of five amino acids present in the plasmid (Stieglitz and Lipsky 1993).

At least 11 variants of the HLA-B27 molecule exist and until recently it was thought that disease predisposition did not appear to be restricted to a particular allele. However, Hill *et al*. have now described a rapid method of HLA Class 1 typing using the polymerase chain reaction and oligonucleotide hybridization that eliminates requirements for viable lymphocytes and, in addition, allows subtypes to be clearly defined (Hill *et al*. 1991). The authors studied black subjects in the Gambia, West Africa, and showed that the predominant subtype was HLA-B27.03. This is particularly rare or absent in other racial groups and interestingly the subtype is not recognized by cytotoxic T cells—perhaps explaining why spondylarthritis is rare in black Africa and when it does occur, may not be associated with HLA-B27.

HLA-B27.03 differs from the other common HLA-B27 subtypes by a single amino acid substitution of histidine for tyrosine at position 59 of the a $_1$ domain. This has been predicted by the 'arthritogenic peptide' model of HLA-B27 disease (Benjamin and Parham 1990). This model evokes a central role for cytotoxic T lymphocytes, which react with an HLA-B27-specific peptide carried by a foreign pathogen. It is proposed that HLA-B27-restricted lymphocytes cause disease when they recognize a similar or identical peptide normally expressed in joint tissue. This is in contrast to the 'altered self' and 'molecular mimicry' models where either the cysteine residue at position 67 or neighbouring residues are thought to be of paramount importance. (HLA-B27.03 shares this cysteine residue at position 67 with the other HLA-B27 subtypes.) Using monoclonal anti-HLA-B27 antibodies, several bacterial components have been recognized. The relevance of this cross-reaction remains unclear. Through a computerized search, a klebsiella protein has been identified that carries a stretch of six amino acids identical to residue 72 to 77 of two of the HLA-B27 variants. Moreover, a synthetic peptide carrying the six amino acids of the HLA-B27 protein is reactive with serum antibodies in some patients with disease. Another amino acid residue of interest is that of position 45, which has provided particular interest relating to the '45 pocket' hypothesis (Benjamin and Parham 1990). The effect of this pocket in peptide presentation is not known but it could have some functional role. The distribution of different HLA-B27 subtypes is given in Table 5.

Population	HLA-B27 subtype
Caucasoids	HLA-B*2704 HLA-B*2702
Native North Americans	HLA-B*2705
Siberian Chukchis	HLA-B*2704
Chinese	HLA-B*2705
Asian Indians	HLA-B*2705 HLA-B*2704 HLA-B*2702
West Africans	HLA-B*2703 HLA-B*2705

Table 5 Distribution of B27 subtypes in various population

The immune response to various triggers has been studied at length. For example during yersinia reactive arthropathy, high and persistent IgA anti-yersinia antibodies have been detected (Granfors and Toivanen 1986). This prolonged antibody response to yersinia, as well as the higher chlamydia antibody titres in synovial fluid as compared to that in serum in chlamydial reactive arthropathy (Hughes and Keat 1992), could indicate an antigenic persistence and impaired antigen elimination from the joint. In addition, proliferative responses specific to the organism which causes the infection preceding the arthropathy have been described (Gaston *et al*. 1989; Hermann *et al*. 1989). A response to a 65-kDa mycobacterial heat shock protein has been detected in patients with reactive arthropathy but its specificity is only minimal. The persistence of intrasynovial antigenic material is now recognized (Nikkari *et al*. 1992; Toivanen and Toivanen 1996).

Initial interest was created by Geczy *et al*. and Ebringer *et al*. with regard to klebsiella and ankylosing spondylitis (Geczy *et al*. 1985; Ebringer *et al*. 1985). The former group reported a specific modification of HLA-B27-positive lymphocytes in patients with ankylosing spondylitis—a phenomenon that has received little support elsewhere. The Ebringer data supported molecular mimicry with cross-reactivity between klebsiella and HLA-B27, a finding supported by other studies with different arthritogenic triggers (van Bohemen *et al*. 1984; Raybourne *et al*. 1988). Meanwhile, Schwimmbeck *et al*. described homology between the 72 to 77 amino acid sequence from HLA-B27.05 subtype and residues 188 to 193 of *Klebsiella pneumoniae* nitrogenase (Schwimmbeck *et al*. 1987). Likewise, Toivanen *et al*. have described a similar sequence between HLA-B27 and Yap 1, an outer membrane protein of yersinia (Toivanen *et al*. 1990). Relevance of the mimicry phenomenon to pathogenesis is complicated by the publication of several studies that fail to agree. Regardless of the putative klebsiella link, there is still no certainty that this organism is arthritogenic in ankylosing spondylitis. It may well be that numerous bacteria (and indeed viruses) can induce disease in genetically susceptible individuals (for review see Kingsley and Sieper 1993; Schoen 1996).

Transgenic rats have now been created that express HLA-B27 and this model will help clarify some of the above issues (Hammer *et al*. 1990). The HLA-B27-carrying rats spontaneously develop a disease characterized by involvement of the gastrointestinal tract, peripheral and axial joints, male genital tract, skin, nails, and heart. In Table 6 this disease is compared with the human spondylarthritides and adjuvant-induced arthritis in the rat (for review see Taurog 1997).

	Adjuvant disease	Transgenic rat	Ankylosing spondylitis	Reiter's syndrome/ reactive arthritis
Peripheral joints	+	+ (M>F)	+	+
Axial joints	+	+ (M>F)	++	±
Gastrointestinal tract	?	++ (M=F)	±	±
Male genital tract	+	+	−	±
Skin lesions	+	+	−	±
Nails	?	+	−	±
Heart	?	+	±	±
Eye lesions	+	− (?)	+	±
Enthesitis	+	+	+	+
Synovitis	+	+	+	+

Table 6 Comparison of adjuvant rat model, transgenic (HLA-B27+) rat, ankylosing spondylyitis, and reactive arthritis (Reiter's syndrome)

Disease susceptibility relates to the transgene product copy number in lymphoid cells, with the high expressing lines prone to disease. Differences in expression of the HLA-B27 gene product occur by disease onset, but differences in thymic and splenic messenger RNA levels are found *in utero*. HLA-B27 transgenic rats have been studied in cell transfer experiments (Breban *et al*. 1993b). The data reveal that disease can be transferred by engraftment of adult bone marrow cells from the diseased animals into irradiated non-transgenic rats. Bone marrow precursor cells are responsible for passage of disease. Moreover, the spondylarthropathy-like illness of rats has been found to be T-cell dependent (Breban *et al*. 1993b). These data argue for HLA-B27, itself, being involved in the pathogenesis of disease rather than there being a gene in linkage disequilibrium with HLA-B27 that is operative. When HLA-B27 positive b $_2$-microglobulin transgenic disease rats are raised in a germ-free environment, disease occurs but in a manner distinct from the regular situation. Specifically, the animals do not develop gut inflammation or arthropathy but they do develop more nail and skin changes as well as genital inflammation. These results suggest that gut flora are involved in some way in the development of spondylarthropathy (Taurog *et al*. 1993b).

Reactive arthropathy or intra-articular 'infection'?

Reactive arthropathy (Toivanen and Toivanen 1996) is defined as an inflammatory joint disease that relates to an infective organism that is distant in both time and place from the arthropathy. For example chlamydia induced urethritis may be followed 2 weeks later by uveitis, arthropathy, and other stigmata of Reiter's disease. A similar situation can follow dysenteric infection with campylobacter, yersinia, or other organism. However, there have been studies demonstrating chlamydia, yersinia, and most recently salmonella (Granfors *et al*. 1990) antigenic components within the synovial tissues. Whether the finding of intrasynovial micro-organism-related antigen represents an epiphenomenon or whether the material is central to disease pathogenesis remains unclear but at least at the moment it is not considered that viable organisms are found in the joint space (Highton and Poole 1993; Nikkari *et al*. 1992; Hughes and Keat 1992). Moreover, we have described postsalmonella vaccination arthropathy—a rheumatic condition related to the inoculation of dead organisms (Calin *et al*. 1987).

Interestingly, Lyme disease, which is known to be a multisystem disorder related to *Borrelia burgdorferi*, may produce a 'reactive' phenomenon in addition to the direct infective process (Weyand and Goronzy 1989).

Epidemiology and the spondylarthritides (see van der Linden and van der Heijde 1990)

Epidemiological studies tell us that reactive arthropathy is more common in the epidemic situation than where the organism is endemic. Adults are more at risk than are children in contradistinction to the situation with acute rheumatic fever. In postdysenteric arthropathy, patients frequently suffer only minimal gastrointestinal symptoms. As mentioned, postvaccination (salmonella) Reiter's disease can occur. Ankylosing spondylitis is seen in both the developing and developed world, with an higher age of onset in the latter. Other epidemiological studies demonstrate that ankylosing spondylitis occurs predominantly in the HLA-B27-positive Haida and Pima Indians and the HLA-B27-positive relatives of white patients with ankylosing spondylitis, while Reiter's syndrome occurs among the Navaho Indians and Inupiat Eskimo in addition to white family members of HLA-B27-positive probands with Reiter's disease. Presumably HLA-B27 is not sufficient in itself and, as stated above, an additional gene or genes modify the phenotypic expression (i.e. HLA-B27 plus appropriate environmental trigger, plus the ankylosing spondylitis gene results in ankylosing spondylitis while HLA-B27 plus environmental trigger plus the Reiter's gene leads to Reiter's syndrome) (Fig. 4).

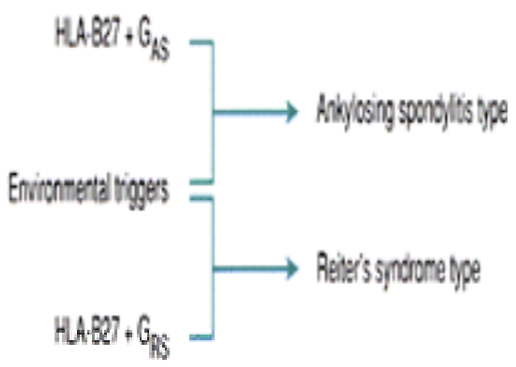


Fig. 4 Epidemiology and the spondylarthritides. G_{AS} = genes determining phenotypic expression of ankylosing spondylitis type. G_{RS} = genes determining phenotypic expression of Reiter's syndrome type.

The relationship between sex, phenotypic expression, mode of inheritance, and age at onset in the spondylarthropathies has recently been studied by Kennedy and colleagues (Kennedy *et al*. 1993b). Epidemiological studies have focused on the presence of spondylarthritis in Alaskan Eskimos (Boyer *et al*. 1994), the native population of Chukotka in Russia (Alexeeva *et al*. 1994), the Indonesian Chinese and native Indonesians (Nasution *et al*. 1993), and in Togo (Mijiyawa 1993). The varied prevalence rate is of great interest.

The risk for the HLA-B27-positive individual depends on the nature of that individual. For example if related to an HLA-B27-positive patient with ankylosing spondylitis, the chance of developing that same disease is approximately 1 in 3. HLA-B27-positive relatives of healthy HLA-B27-positive subjects are at much less risk of developing spondylarthropathy. Interplay between different chromosomes almost certainly occurs. Chromosome 6 is clearly important, with HLA-B27, -B60, and -CW6 all relevant. In addition, recent interest focuses on the tumour necrosis factor and class III genes (complement) on the same sixth chromosome. Other possible genetic factors relate to the T-cell receptor b gene (at a locus on the seventh chromosome) and the T-cell receptor a gene, the a_1-antitrypsin gene, and the IgG heavy-chain gene products related to loci on the fourteenth chromosome. Additional gene loci and products thereof include the IgG light g chain (chromosome 22). Studies using restriction fragment length polymorphisms have failed to confirm the presence of additional genetic material of relevance to disease pathogenesis. For example, no differences have been found between the frequencies of such polymorphisms for tumour necrosis factor between patients and controls (Verjans *et al*. 1991).

Intact antigen is taken up by the cell and partly digested by the so-called large multifunctional protease system, the product from a gene on the short arm of the 6th chromosome. This results in a nonapeptide (i.e. a nine amino acid peptide) reaching the endoplasmic reticulum via the transporter associated with peptide gene product. (The transporter associated with peptide genes are also found in the major histocompatibility complex.) The nonapeptide then becomes associated with the HLA-B27 Class 1 antigen and extrudes on the cell surface as a bimolecular product to face the T-cell receptor gene product. The resulting HLA-B27–antigen–T-cell receptor trimolecular complex releases cytokines and inflammatory mediators. This has been the focus for study (Maksymowych *et al*. 1994).

Reactive arthropathy appears to be mediated by HLA Class 1-restricted, CD8-positive T cells, given that both psoriatic arthropathy and Reiter's disease thrive in the presence of human immunodeficiency virus infection. This is in contrast to the CD4-positive, T-cell-maintained arthropathy of rheumatoid arthritis, which appears to remit in the presence of human immunodeficiency virus infection. Ankylosing spondylitis, itself, appears unaffected by the virus but few data exist.

The changing epidemiology of rheumatic diseases Calin *et al*. 1988; Will *et al*. 1990; Will *et al*. 1992)

The spondylarthropathies should perhaps be included in any discussion of changing epidemiological features. There are many reasons why the pattern (perceived or real) of a disease may change. Increased interest and recognition by the medical profession and greater concern and pressure from patients may lead to more emphasis on a particular disease. For example osteoporosis, until recently ignored by rheumatologists, has become a major focus of interest.

Chronic tophaceous gout, once the scourge of medical clinics, is an example of a rheumatic disorder which has all but disappeared, in part due to changing dietary habits and in part because of hypouricaemic agents. By contrast, a recent rise in prevalence of gout in older females relates to the greater use of diuretics.

Rheumatic fever, at least until recently, had become rare in the more affluent communities. The decline in fatal rheumatic carditis accelerated after 1945, perhaps due to an alteration in the pathogenicity of the streptococcal group A M antigen induced by penicillin. Altered streptococcal antigenicity may also be responsible for the recent recognition of a poststreptococcal arthropathy.

Recent studies also suggest that rheumatoid arthritis may be declining in frequency in the developed world, perhaps due to the advent of the contraceptive pill, but it is becoming more severe in the developing world. For example recent published data have demonstrated that as rural black Africans migrate to an urban environment, rheumatoid arthritis increases in frequency and becomes a more destructive disease. Hypothetical explanations for this phenomenon may include either exposure to new antigens present in an urban environment but uncommon in a less crowded rural setting or conversely a reduction in the antigenic load due to fewer parasitic infestations, which in turn could result in less immunosuppression.

There is increasing evidence accumulating from developing countries to suggest that the pattern of ankylosing spondylitis may be changing. This relates to both the age of onset of the disease and the pattern of joint involvement. Ankylosing spondylitis develops at an earlier age in countries with poor living conditions and as these improve the age of onset increases. We have recently suggested that the age of onset of the disease may also be increasing over time in Britain. The influence of potential left and right censoring biases on the United Kingdom data has been emphasized elsewhere, though the importance of this phenomenon is difficult to quantify. Can the conclusion of an increasing age of onset of ankylosing spondylitis in Britain be substantiated by data from other communities? The putative change may be due to a later age of exposure to the presumed 'infective trigger(s)' or altered pathogenicity of the trigger(s), perhaps due to a modifying factor such as the widespread use of antibiotics.

Amor *et al*. in France noted that the age of onset of ankylosing spondylitis was influenced by the geographic background of their patients (Amor *et al*. 1991). They observed that 25 per cent of patients from North Africa develop disease before the age of 15 years whereas only 10 per cent of patients in France did so. Moreover, 47 second generation North Africans with ankylosing spondylitis born and living in France (but whose parents were born in North Africa) were identified from a survey conducted by the French Society of Rheumatology in 1983. None of these 'Beur', as they are known, developed disease before the age of 15. Other studies have also noted a lower frequency of juvenile onset ankylosing spondylitis (at under age 16) in white Caucasian populations as compared to subjects in the developing world. In spite of the inevitable ascertainment biases, there is now a series of studies consistently demonstrating a greater frequency of patients with a lower age of disease onset from developing countries.

An entirely separate epidemiological route has been taken and the changing pattern of new patient referral to the London Hospital has been studied (Will *et al*. 1992). Of interest, patients with non-specific mechanical back pain have become progressively younger over the last three decades whereas new patients with ankylosing

spondylitis have become progressively older. These data, together with the French data of Amor, add further support for the changing pattern.

Family studies of sibling pairs with ankylosing spondylitis also suggest that the date of onset in each sib is similar while the age at onset is discordant suggesting exposure to the trigger at about the same time. This emphasises the importance of the environmental trigger determining the time of onset of disease in sibling pairs who have a similar genetic predisposition (Calin and Elswood 1989a).

The age of disease onset also influences the need for total hip replacement. Calin and Elswood have shown that 16 per cent of a juvenile cohort (10-15 years), 10 per cent of an early onset (18-20 years), and 1 per cent of a late onset cohort (30-40 years) had total hip replacements performed (Calin and Elswood 1989b). Amor *et al*. also observed that the frequency of total hip replacement in a spondylitic population correlates well with the mean age of disease onset (Amor *et al*. 1991). A study of total hip replacements performed in patients from four French hospitals between 1977 and 1983 was undertaken. Twenty-two of a total of 71 patients (30 per cent) who had total hip replacements were born in North Africa as compared with only 7.9 per cent of non-surgical patients treated in French hospitals during this same period. The explanation for more severe hip joint involvement in a juvenile cohort is unknown. The developing hip joint may be at greater risk of damage. Conversely, a more marked inflammatory response may result in greater joint destruction in juveniles compared with patients who develop the disease at an older age.

It is of interest that a changing pattern in inflammatory bowel disease is now also recognized and in many studies Crohn's disease, which became progressively more common until the mid-1970s, has now begun to be less prevalent. The intriguing interrelationship between the bowel and ankylosing spondylitis is, of course, well known (Mielants *et al*. 1993; Leirisalo-Repo *et al*. 1994; Mielants *et al*. 1995).

In conclusion, chronic rheumatic diseases should not be considered immutable processes. A changing pattern of disease is to be expected given the interaction of genes and the environment. Ankylosing spondylitis, and perhaps the other spondylarthropathies, should now be considered as another example of a rheumatic disorder whose characteristics may be altering as a result of a changing environment. We might expect a diminished need for total hip replacement among patients as a consequence of an increasing age of onset of patients in the West and perhaps in developing countries as the environment changes. Physicians will need to be increasingly aware that patients may present with symptoms of disease in their thirties or later. The epidemiological pattern among the other spondylarthropathies is less well defined.

The bowel and spondylarthritis (Mielants *et al*. 1993; Leirisalo-Repo 1994; Mielants *et al*. 1995)

The relationship between the bowel and arthritis is an intimate one. For example an active arthropathy may follow enteropathic infections such as those related to shigella, salmonella, yersinia, campylobacter, and other organisms. Moreover, patients with reactive arthropathy or Reiter's disease related to chlamydia may also develop enteric symptoms as a manifestation of the disease process.

It is well recognized that patients with inflammatory bowel disease (ulcerative colitis and Crohn's disease) may develop both peripheral and axial arthropathy. In general, the peripheral arthritis appears to be a complication of severe bowel involvement with a close relationship between the two components. By contrast, axial disease (sacroiliitis and ascending spinal disease) should be considered a manifestation of the same genetic background, in as much as there is no obvious link between the timing or disease severity of the two entities.

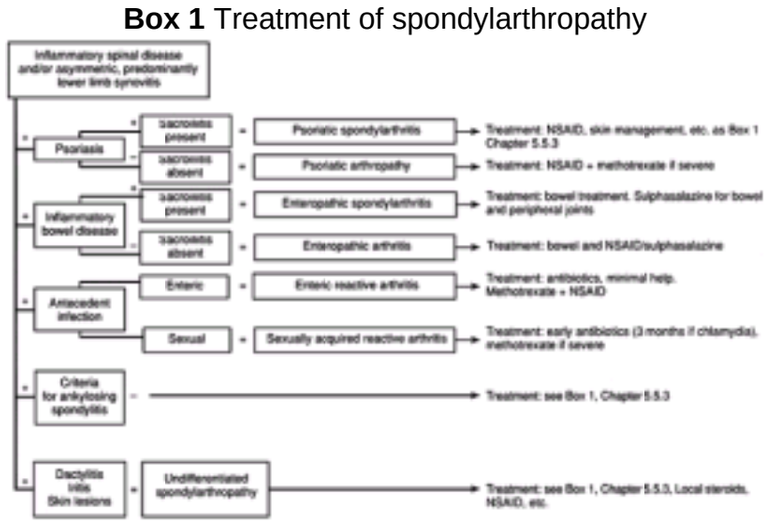
The story has been taken further by Mielants and Veys who performed ileocolonoscopy with biopsies of the gut and described clinical involvement of the terminal ileum as a common finding in the spondylarthropathy patients as a whole (Mielants and Veys 1990; Mielants *et al*. 1993; Mielants *et al*. 1995). Interestingly, the low grade bowel involvement that mimics mild Crohn's disease was often present as a silent phenomenon and appeared unrelated to the HLA-B27 marker. The authors concluded, from multiple studies, that a number of patients with undifferentiated spondylarthropathy could in fact be suffering from a form of subclinical Crohn's disease, of which arthritis is the only clinical manifestation (Leirisalo-Repo *et al*. 1994).

The putative relationship between the bowel, enteric micro-organisms, and arthropathy has been further addressed by way of therapy. Sulphasalazine has been used for years as a treatment for ulcerative colitis, colonic Crohn's disease, and, more recently, for rheumatoid arthropathy. It is metabolized to 5-aminosalicylic acid (active in inflammatory bowel disease) and sulphapyridine (active in rheumatoid disease). Work by Feratz *et al*. suggests efficacy also in ankylosing spondylitis and perhaps the other spondylarthropathies (Feratz *et al*. 1990). Whether this drug works by way of altering the bowel microflora or another mechanism remains unclear.

Titres of serum and secretory IgA (presumed to be of bowel mucosal origin) are raised in patients with ankylosing spondylitis and in patients with yersinia arthropathy, again arguing for a close link between events in the bowel wall and arthritis.

Finally, it is of note that the transgenic rat expressing human HLA-B27 develops a spontaneous arthropathy that mimics spondylarthritis. Strikingly, the main organ involvement is the bowel with marked diarrhoea (Hammer *et al*. 1990).

Box 1 Treatment of spondylarthropathy



Treatment (Box 1)

The management of the various components of spondylarthropathy are discussed in the separate chapters that follow. However, there are some general points that should be stressed. Drug therapy is often disappointing. For example for ankylosing spondylitis, itself, the major thrust of treatment relates to an exercise programme. In contrast to the situation in rheumatoid disease where rest is good for the joints, the patient with ankylosing spondylitis deteriorates with rest and improves with exercise. For those with reactive arthritis, the role of antibiotics has been considered but even the proponents of this approach would agree that outcome is only marginally improved (Leirisalo-Repo 1993; Bardin *et al*. 1992; Lauhio *et al*. 1991; Lauhio *et al*. 1992; Toivanen *et al*. 1993). One difficulty relates to the interpretation of the effect, if any, of an antibiotic. Does this relate to the antimicrobial action or, for example, the anticollagenolytic potential of the drug? There are theoretical reasons why sulphasalazine should be efficacious in the spondylarthropathies given the relationship between bowel inflammation and arthropathy. However, again, improvement is marginal and in large part beneficial only for peripheral joints rather than axial disease (Youssef *et al*. 1992; Kirwan *et al*. 1993; Job-Duslandre *et al*. 1993; Dougados *et al*. 1995).

Natural history and prognosis

Outcome in spondylarthropathy is notoriously difficult to define.

Since our early studies in the 1970s when we showed that Reiter's syndrome is typically not a self-limiting disorder (Calin and Fries 1976), there have been numerous studies showing that for many patients the outcome can be relentlessly progressive. For ankylosing spondylitis, itself, the natural history is varied although most believe that an intensive exercise programme has an excellent effect on disease progression. No longer is it believed that the disease 'burns out' for the majority (Kennedy *et al*. 1993a). Amor and colleagues have attempted to define predictive factors for the long-term outcome of spondylarthropathies and have suggested that those individuals with hip involvement, an erythrocyte sedimentation rate over 30, poor response to non-steroidal anti-inflammatory drugs, a decreased Schober's test, dactylitis, oligoarthropathy, and young age at onset have the worse prognosis (Amor *et al*. 1994). Although this study was not a formal prospective investigation, few

would disagree that these criteria relate to a less good outcome. Recently, we have defined outcome in terms of metrology, disease activity, function, global status, severity, and radiology (Calin 1995a; Calin 1995b; Mackay *et al*. 1996).

Finally, the influence of sex in arthritis and its relationship to cartilage damage cannot be ignored (Da Silva and Willoughby 1994).

Conclusion

During the next years we will understand why and how HLA-B27, other genes, the bowel, sex, and environmental triggers interact, resulting in the development of spondylarthropathy in its many guises. Once we understand this intricate, interrelated network we will know how to manage and perhaps cure our patients.

Chapter References

Alexeeva, L., Krylov, M., Vturin, V., Mylov, N., Erdesz, S., and Benevolenskaya, L. (1994). Prevalence of spondylarthropathies and HLA-B27 in the native population of Chukotka, Russia. *Journal of Rheumatology*, **22**, 2298–300.

Amor, B. *et al*. (1991). Evaluation des criteres de spondylarthropathies d'Amor et de Liesseg: une etude transversale de 2228 patients. *Annales de Medicine Interne*, **142**, 85–9.

Amor, B., Santos, R.S., Nahal, R., Listrat, V., and Dougados, M. (1994). Predictive factors for the longterm outcome of spondylarthropathies. *Journal of Rheumatology*, **21**, 1883–7.

Bardin T., Enel, C., Comelis, F., Saiski, C., Jorgensen, C., Ward, R., and Lanthrop, M. (1992). Antibiotic treatment of venereal disease and Reiter's syndrome in a Greenland population. *Arthritis and Rheumatism*, **35**, 190–4.

Benjamin, R. and Parham, P. (1990). Guilt by association: HLA-B27 and ankylosing spondylitis. *Immunology Today*, **11**, 137–43.

Bennett, P.H. and Burch, T.A. (1968). *Population studies of the rheumatic diseases*, pp. 456–7. Excerpta Medica Foundation, Amsterdam.

Boyer, G.S. *et al*. (1994). Prevalence of spondylarthropathies in Alaskan Eskimos. *Journal of Rheumatology*, **21**, 2292–7.

Braun, J. and Sieper, J. (1996). The sacroiliac joint in the spondyloarthropathies. *Current Opinion in Rheumatology*, **8**, 275–87.

Breban, M., Hammer, R.E., Richardson, J.A., and Taurog, J.D. (1993*a*). Transfer of the inflammatory disease of HLA-B27 transgenic rats by bone marrow engraftment. *Journal of Experimental Medicine*, **178**, 1607–16.

Breban, M. *et al*. (1993*b*). T cells but not thymic exposure to B27 are required for the inflammatory disease of HLA-B27 transgenic rats. *Arthritis and Rheumatism*, **36**, S73.

Burns, T.M. and Calin, A. (1984). Undifferentiated spondylarthopathy. In *Spondylarthropathies* (ed A. Calin), pp. 253–64. Grune and Stratton, Orlando, FL.

Calin, A. (1984). *Spondylarthropathies*. Grune and Stratton, Orlando, FL.

Calin, A. (1989). Ankylosing spondylitis. *Textbook of rheumatology*. WB Saunders, Philadelphia, PA.

Calin, A. (1995*a*). The Dunlop Dotteridge Lecture. Ankylosing spondylitis: defining disease status and the relationship between radiology, metrology, disease activity, function and outcome. *Journal of Rheumatology*, **22**, 740–44.

Calin, A. (1995*b*). The Margaret Holroyd Essay. The individual with ankylosing spondylitis: defining disease status and the impact of illness. *British Journal of Rheumatology*, **34**, 663–72.

Calin, A. and Elswood, J. (1989*a*). Relative role of genetic and environmental factors in disease expression: sib-pair analysis in ankylosing spondylitis. *Arthritis and Rheumatism*, **32**, 77–81.

Calin, A. and Elswood, J. (1989*b*). The outcome of 138 total hip replacements and 12 revisions in ankylosing spondylitis. *Journal of Rheumatology*, **16**, 955–8.

Calin, A., Elswood, J., Rigg, S., and Skevington, S.M. (1988). Ankylosing spondylitis – an analytical review of 1500 patients: the changing pattern of disease. *Journal of Rheumatology*, **15**, 1234–8.

Calin, A. and Fries J.F. (1976). An 'experimental' epidemic of Reiter's syndrome revisited. *Annals of Internal Medicine*, **84**, 564–6.

Calin, A. and Fries, J.F. (1978). *Ankylosing spondylitis: discussions in patient management*.. Medical Examination Publishing Company, New York.

Calin, A., Goulding, D., and Brewerton, D. (1987). Post salmonella vaccination reactive arthropathy. *Arthritis and Rheumatism*, **30**, 969–70.

Calin, A., Marder, A., Marks, S., and Burns, T. (1984). Familial aggregation of Reiter's syndrome and ankylosing spondylitis: a comparative study. *Journal of Rheumatology*, **11**, 672–77.

Da Silva, J.A.P. and Willoughby, D.A. (1994). The influence of sex in arthritis: is cartilage an overlooked factor. *Journal of Rheumatology*, **21**, 791–6.

De Castro, J.A.L. (1994). Structure, function, and disease association of HLA-B27. *Current Opinion in Rheumatology*, **6**, 371–7.

De Vries, D.D., Dekker-Saeys, A.J., Gyodi, E., Bohn, U., and Ivanyi, P. (1992). Absence of autoantibodies to peptides shared by HLA-B27.5 and klebsiella pneumoniae nitrogenase in serum samples from HLA-B27 positive patients with ankylosing spondylitis and Reiter's syndrome. *Annals of the Rheumatic Diseases*, **51**, 783–9.

Dougados, M. *et al*. (1991). Preliminary criteria for the classification of spondylarthropathy. *Arthritis and Rheumatism*, **34**, 1218–27.

Dougados, M., van der Linden, S., Amor, B., Calin, A., *et al*. (1995). Sulphasalazine in the treatment of spondylarthropathy: a randomized multi-centre double blind placebo controlled study. *Arthritis and Rheumatism*, **38**, 618–27.

Ebringer, A. (1991). Ankylosing spondylitis and klebsiella—the debate continues. Editorial. *Journal of Rheumatology*, **18**, 312–13.

Ebringer, A., Baines, M., Childerstone, M., Ghuloom, M., and Ptaszynska, T. (1985). Etiopathogenesis of ankylosing spondylitis and the cross-tolerance hypothesis. In *Advances in inflammation research*. *The spondylarthropathies*, Vol. 9 (ed. M. Ziff and S.B. Cohen), pp. 101–28. Raven Press, New York.

Edmunds, L., Elswood, J., Kennedy, L.G., and Calin, A. (1991). Primary ankylosing spondylitis, psoriatic and enteropathic spondylarthropathy: a controlled analysis. *Journal of Rheumatology*, **18**, 696–8.

Ferraz, M.B., Tugwell, P., Goldsmith, C.H., and Atra, E. (1990). Metaanalysis of sulfasalazine in ankylosing spondylitis. *Journal of Rheumatology*, **17**, 11.

Gaston, J.S.H. *et al*. (1989). Synovial T lymphocyte recognition of organisms that trigger reactive arthritis. *Clinical and Experimental Immunology*, **76**, 348–53.

Geczy, A.F., Prendergast, J.K., Sullivan, J.S., Upfold, L.I., Edmonds, J.P., and Bashi, H.V. (1985). Possible role of enteric organisms in the pathogenesis of seronegative arthropathies. In *Advances in inflammation research*. *The spondylarthropathies*, Vol. 9. (ed. M. Ziff and S.B. Cohen), pp. 129–37. Raven Press, New York.

Granfors, K. and Toivanen, A. (1986). IgA anti-Yersinia antibodies in Yersinia triggering reactive arthritis. *Annals of the Rheumatic Diseases*, **45**, 401–12.

Granfors, K. *et al*. (1990). Salmonella lipopolysaccharide in synovial cells from patients with reactive arthritis. *Lancet*, **335**, 685–8.

Hammer, R.E., Maika, S.D., Richardson, J.A., Tang, J-P., and Taurogi, J.D. (1990). Spontaneous inflammatory disease in transgenic rats expression HLA-B27 and Human bm: an animal model of HLA-B27-associated human disorders. *Cell*, **63**, 1099–112.

Hermann, E., Fleischer, B., Mayet, W.J., Poralla, T., and zum Buschenfelde, K.H.M. (1989). Response of synovial fluid T cell clones to Yersinia enterocolitica antigens in patients with reactive Yersinia arthritis. *Clinical and Experimental Immunology*, **75**, 365–70.

Hermann, E., Yu, D.T.Y., zum Buschenfelde, K.H.M., and Fleischer, B. (1993). HLA-B27-restricted CD8 T cells derived from synovial fluids of patients with reactive arthritis and ankylosing spondylitis. *Lancet*, **342**, 646–50.

Highton, J. and Poole, E. (1993). Sexually acquired reactive arthritis: inflammation or sepsis? *British Journal of Rheumatology*, **32**, 649–52.

Hill, A.V.S, Allsopp, C.E.M., Kwiatkowski, D., Anstey, N.M., Greenwood, B.M., and McMichael, A.J. (1991). HLA class 1 typing by PCR: HLA-B27 and an African B27 subtype. *Lancet*, **337**, 640–2.

Hughes, R.A., Treharne, J.D., and Keat, A.C. (1989). Chlamydial antibody titres in serum and synovial fluid (abstract). *British Journal of Rheumatology*, **28**, 4.

Hughes, R. and Keat, A. (1992). Reactive arthritis: the role of bacterial antigens in inflammatory arthritis. *Baillières Clinical Rheumatology*, **6**, 285–308.

Inman, R.D. and Scofield, R.H. (1994). Etiopathogenesis of ankylosing spondylitis and reactive arthritis. *Current Opinion in Rheumatology*, **6**, 360–70.

Job-Duslandre, C. and Menkes, C.J. (1993). Sulphasalazine treatment for juvenile arthropathy. *Revue du Rheumatisme*, **60**, 489–93.

Kellgren, J.H., Jeffrey, M.R., and Ball, J. (1963). *The epidemiology of chronic rheumatism*, Vol. 1, pp. 326–7. Blackwell Scientific Publications, Oxford.

Kennedy, L.G., Edmunds, L., and Calin, A. (1993 *a*). Does ankylosing spondylitis burn out? A prospective 2-year study. *Journal of Rheumatology*, **20**, 688–92.

Kennedy, L.G., Will, R., and Calin, A. (1993 *b*). Sex ratio in the spondylarthropathies and its relationship to phenotypic expression, mode of inheritance and age at onset. *Journal of Rheumatology*, **20**, 1900–4.

Khan, M.A. (ed) (1990). *Ankylosing spondylitis and related spondyloarthropathies in spine*, Vol. 4, No. 3. Hanley & Belfas, Philadelphia.

Khan, M.A. (1993). Pathogenesis of ankylosing spondylitis: recent advances. *Journal of Rheumatology*, **20**, 1273–7.

Khan, M.A. (1996). Spondylarthropathies (Editorial overview). *Current Opinion in Rheumatology*, **8**, 267–8.

Khan, M.A. and van der Linden, S.M. (1990). Undifferentiated spondyloarthropathies. In *Ankylosing spondylitis and related spondyloarthropathies in spine*, Vol. 3, No. 3 (ed. M.A. Khan), pp. 657–64. Hanley & Belfas, Philadelphia.

Kingsley, G. and Sieper, J. (1993). Current prospectives in reactive arthritis. *Immunology Today*, **14**, 387–91.

Kirwan, J., Edwards, A., Huitfeldt, B., Thompson, P., and Currey, H. (1993). The course of established ankylosing spondylitis and the effects of Sulphasalazine over 3 years. *British Journal of Rheumatology*, **32**, 729–33.

Lauhio, A., Leirisalo-Repo, M., Lahdevirta, J., Saikku, P., and Repo, H. (1991). Double-blind, placebo-controlled study of three-month treatment with lymecycline in reactive arthritis, with special reference to chlamydia arthritis. *Arthritis and Rheumatism*, **34**, 6–14.

Lauhio, A. *et al*. (1992). The anticoliagenolytic potential of lymecycline in the long-term treatment of reactive arthritis. *Arthritis and Rheumatism*, **35**, 195–8.

Leirisalo-Repo, M. (1993). Areantibiotics of any use in reactivearthritis? *Acta Pathologica, Microbiologica, et Immunologica Scandinavica*, **101**, 575–81.

Leirisalo-Repo, M. (1994). Enteropathic arthritis, Whipple's disease, juvenile spondyloarthropathy, and uveitis. *Current Opinion in Rheumatology*, **6**, 385–90.

Leirisalo-Repo, M., Turunen, U., Stenman, S., Helenius, P., and Seppala, K. (1994). High frequency of silent inflammatory bowel disease in spondylarthropathy. *Arthritis and Rheumatism*, **37**, 23–31.

Lopez-Larrea, C. *et al*. (1996). HLA B27. Structure, function, and disease association. *Current Opinion in Rheumatology*, **8**, 296–308.

Mackay, K., Mack, C.S., Pande, I., Rousson, E., and Calin, A. (1996). The Bath Ankylosing Spondylitis Radiology Index (BASRI). *Arthritis and Rheumatism*, **39**, 206.

Madden, D.R., Gorga, J.C., Strominger, J.L., and Wiley, D.C. (1992). The three-dimensional structure of HLA-B27 at 2.1: a resolution suggests a general mechanism for tight peptide binding to MHC. *Cell*, **70**, 1035–48.

Maksymowych, W.P. *et al*. (1994). Polymorphism in an HLA linked proteasome gene influences phenotypic expression of disease in HLA-B27 positive individuals. *Journal of Rheumatology*, **21**, 665–9.

McClean, I.L. *et al*. (1993). HLA-B27 subtypes in the spondyloarthropathies. *Clinical and Experimenal Immunology*, **91**, 214–19.

Mielants, H. *et al*. (1993). Gut inflammation in children with late onset of pauciarticular juvenile chronic arthritis and evolution to adult spondyloarthropathy: a prospective study. *Journal of Rheumatology*, **20**, 1567–72.

Mielants, H. *et al*. (1995). The evolution of spondyloarthropathies in relation to gut histology. *Journal of Rheumatology*, **22**, 2266–72.

Mijiyawa, M. (1993). Spondylarthropathies in patients attending the rheumatology unit of Lome Hospital. *Journal of Rheumatology*, **20**, 1167–9.

Moll, J.M.H. (ed) (1980). *Ankylosing spondylitis*. Churchill Livingstone, Edinburgh.

Moll, J.M.H. *et al*. (1974). Associations between ankylosing spondylitis, psoriatic arthritis, Reiter's disease, the intestinal arthropathies and Behçet's syndrome. *Medicine*, **53**, 343.

Nasution, A.R., Mardjuadi, A., Suryadhana, N.G., Daud, R., and Muslichan, S. (1993). Higher relative risk of spondylarthropathies among B27 positive Indonesian Chinese than native Indonesians. *Journal of Rheumatology*, **20**, 988–90.

Nikkari, S. *et al*. (1992). Yersinia-triggered reactive arthritis: use of polymerase chain reaction and immuno-cytochemical staining in the detection of bacterial components from synovial specimens. *Arthritis and Rheumatism*, **35**, 682–7.

Raybourne, R.B., Bunning, V.K., and Williams, K.M. (1988). Reaction of anti-HLA-B monoclonal antibodies with envelope proteins of Shigella species. Evidence for molecular mimicry in the spondylarthropathies. *Journal of Immunology*, **140**, 3489–95.

Rojo, S., Garcia, F., Villadangos, J.A., and Lopez de Castro, J.A. (1993). Changes in the repertoire of peptides bound to HLA-B27 subtypes and to site specific mutants inside and outside Pocket B. *Journal of Experimental Medicine*, **177**, 613–20.

Rosner, I.A., Burg, C.G., Wisnieski, J.J., Schacter, B.Z., and Richter, D.E. (1993). The clinical spectrum of the arthropathy associated with hidradenitis suppurativa and acne conglobata. *Journal of Rheumatology*, **20**, 4.

Schoen, R.T. (1996). Infectious arthritis and immune dysfunction. *Current Opinion in Rheumatology*, **8**, 317–21.

Schur, P.H. (1994). Arthritis and autoimmunity. *Arthritis and Rheumatism*, **37**, 1818–25.

Schwartz, B.D. (1990). Infectious agents, immunity and rheumatic diseases. *Arthritis and Rheumatism*, **33**, 457–65.

Schwimmbeck, P.L., Yu, D.T.Y., and Oldstone, M.B.A. (1987). Autoantibodies to HLA-B27 in the sera of HLA-B27 patients with ankylosing spondylitis and Reiter's syndrome. Molecular mimicry with Klebsiella pneumonias as potential mechanism of autoimmune disease. *Journal of Experimental Medicine*, **166**, 173–81.

Scofield, R.H. (1996). Etiopathogenesis and biochemical and immunologic evaluation of spondyloarthropathies. *Current Opinion in Rheumatology*, **8**, 309–15.

Scofield, R.H., Warren, W.L., Koelsch, G., and Harley, J.B. (1993). A hypothesis for the immune dysregulation in spondyloarthropathies: contributions from enteric organisms, B27 structure. peptides bound by B27 and convergent evolution. *Proceedings of the National Academy of Science*, **90**, 9330–4.

Skurnik, M., Batsford, S., Mertz, A., and Tolvanen, P. (1993). The putative arthrogenic cationic 19-kilodalton antigen of yersinia enterocolitica is a urease b-subunit. *Infectious Immunology*, **61**, 2498–504.

Stieglitz, H. and Lipsky, P. (1993). Association between reactive arthritis and antecedent infection with shigella flexneri carrying a 2-Md plasmid and encoding an HLA-B27 mimetic epitope. *Arthritis and Rheumatism*, **36**, 1387–91.

Taurog, J.D. (1997). Animal models of the spondylarthropathies. In *Spondylarthropathies* (ed. A. Calin and J. Taurog). Oxford University Press (in press).

Taurog, J.D., Maika, S.D., Simmons, W.A., Breban, M., and Hammer, R.E. (1993 *a*) Susceptibility to inflammatory disease in HLA-B27 transgenic rat lines correlates with level of B27 expression. *Journal of Immunology*, **150**, 4168–72.

Taurog, J.D. *et al*. (1993*b*) Effect of the germfree state on the inflammatory disease of HLA-B27 transgenic rats: a split result. *Arthritis and Rheumatism*, **36**, S46.

Toivanen, P. *et al*. (1990). Tetrapeptide shared by HLA-B27 and Yersinia outer membrane protein YOP1 is not decisive for antigenicity of YOP1 (abstract). *Clinical Rheumatology*, **9**, 133.

Toivanen, A., Yli-Kerttula, T., Luukkainen, R., Meilahti-Palo, R., Granfors, K., and Seppala, J. (1993). Effect of antimicrobial treatment on chronic reactive arthritis. *Clinical and Experimental Rheumatology*, **11**, 301–7.

Toivanen, A. and Toivanen, P. (1996). Reactive arthritides. I>Current Opinion in Rheumatology, **8**, 334–40.

van Bohemen, C.G., Grumet, F.C., and Zanen, H.C. (1984). Identification of HLA-B27 M1 and M2 cross-reactive antigens in Klebsiella, Shigella and Yersinia. *Immunology*, **52**, 605–9.

van der Linden, S. and van der Heijde, D.M. (1996). Clinical and epidemiologic aspects of ankylosing spondylitis and spondyloarthropathies. *Current Opinion in Rheumatology*, **8**, 269–74.

van der Linden, S., Valkenburg, H.A., and Cats, A. (1984). Evaluation of diagnostic criteria for ankylosing spondylitis. *Arthritis and Rheumatism*, **27**, 361–8.

Vasey, F. and Espinoza, L.R. (1984). Psoriatic arthropathy. In *Spondylarthropathies* (ed. A, Calin), pp. 151–85. Grune and Stratton, Orlando, FL.

Verjans, G.M.G.M., van der Linden, S.M., van Eys, G.J.J.M., de Waal, L.P., and Kijistra, A. (1991). Restriction fragment length polymorphism of the tumor necrosis factor region in patients with ankylosing spondylitis. *Arthritis and Rheumatism*, **34**, 486–9.

Weyand, C.M. and Goronzy, J.J. (1989). Immune responses to borrelia-burgdorferi in patients with reactive arthritis. *Arthritis and Rheumatism*, **32**, 1057–64.

Whelan, M.A. and Archer, J.R. (1993). Chemical reactivity of an HLA-B27 thiol group. *European Journal of Immunology*, **23**, 3278–85.

Will, R., Amor, B., and Calin, A. (1990). The changing epidemiology of rheumatic diseases: should ankylosing spondylitis now be included? *British Journal of Rheumatology*, **29**, 299–300.

Will, R., Calin, A., and Kirwan, J. (1992). Increasing age at presentation with ankylosing spondylitis and rheumatic diseases. *Annals of the Rheumatic Diseases*, **52**, 340–2.

Willkens, R.F. *et al*. (1981). Reiter's syndrome. Evaluation of preliminary criteria for definite disease. *Arthritis and Rheumatism*, **24**, 844–9.

Wright, V. (1980). Relationships between ankylosing spondylitis and other spondarthritides. In *Ankylosing spondylitis* (ed. J.M.H. Moll). Churchill Livingstone, Edinburgh.

Wright, V. and Moll, J.M.H. (1976). *Seronegative polyarthritis.* North Holland, Amsterdam.

Youssef, P.P., Berlouch, J.V., and Jones, P.D. (1992). Successful treatment of human immunodeficiency virus-associated Reiter's syndrome with Sulfasalazine. *Arthritis and Rheumatism*, **35**, 723–4.

5.5.2 Spondylarthropathies in childhood

Taunton R. Southwood and Murray H. Passo

Synopsis

Juvenile spondylarthropathy is an umbrella term covering a relatively homogeneous group of diseases in children, usually beginning during the preadolescent or adolescent years. The diagnosis of spondylarthropathy is often more difficult in childhood than during the adult years because symptoms of back pain are uncommon in juvenile spondylarthropathy, particularly at the onset of the illness. Additionally, the radiographic features of spondylitis and sacroiliitis are often absent during childhood. The diagnosis usually rests on a combination of arthritis affecting large joints of the lower limbs and extra-articular inflammatory features including enthesitis, bowel pathology, acute uveitis, and psoriasis. The disease is reported more frequently in boys than girls, and more than 50 per cent of affected children eventually develop ankylosing spondylitis (Burgos-Vargas and Clark 1989; Cabral *et al*. 1992).

Because of the potentially lifelong implications of juvenile spondylarthropathy, every effort should be made to increase patients' and their families' knowledge about the disease and compliance with treatment. However, no published studies confirm that children with spondyarthropathy should be treated any differently from those with other forms of juvenile chronic arthritis. The principles of treatment are summarized as follows, but for further details, reference should be made to the chapters on juvenile chronic arthritis.

Education about the disease for the parents and the patient, presented in an age-appropriate form, is important. Physiotherapy, appropriate limb splinting, and hydrotherapy are aimed at maintaining back posture, flexibility and strength, as well as correcting any reduced range of peripheral joint movement. Orthotics, such as custom-moulded insoles, often alleviate the symptoms of plantar enthesitis. Pharmacotherapy is usually begun with a non-steroidal anti-inflammatory drug. Tolmetin sodium (30 mg/kg per day in three divided doses) and indomethacin (1.5–2.5 mg/kg per day in three divided doses, or as a daily dose if an appropriate slow-release preparation is available) appear to be the most effective. Intra-articular depot corticosteroids (e.g. triamcinolone hexacetonide: 1 mg/kg per joint for large joints, 0.5 mg/kg per joint for smaller joints) are used to control the peripheral arthritis of this condition, and should be considered early in the treatment programme. Sulphasalazine (commencing at a dose of 12.5 mg/kg per day and increasing weekly over a month to 50 mg/kg per day in divided doses) is the slow-acting antirheumatic drug of choice for uncontrolled spondylarthropathy, especially for significant axial involvement, polyarthritis, or severe, persistent oligoarthritis. Regular blood monitoring is required for bone marrow depression and hepatotoxicity, but skin rash and abdominal pain are the most common adverse effects. Methotrexate also may have a role in the long-term control of spondylarthropathy.

It is not uncommon for juvenile spondylarthropathy to run a fluctuating course, and occasionally the doses of non-steroidal anti-inflammatory drugs can be adjusted by the patients or their families to reflect these changes. In those children fortunate enough to experience prolonged remission, pharmacotherapy should be discontinued with caution, as relapse may occur after many symptom-free months or even years.

Aetiology

In considering the aetiology of juvenile spondylarthropathy, both inherited and environmental predisposing factors must be taken into account. Juvenile spondylarthropathy has a strong familial association with ankylosing spondylitis, and other inflammatory diseases such as acute uveitis, psoriasis or inflammatory bowel disease occur with increased frequency in family members (Petty 1990). The genetic predisposition to juvenile spondylarthropathy was first associated with the *HLA-B27* gene over 25 years ago (Ansell *et al*. 1969). Carriage of the gene confers an increased risk of developing ankylosing spondylitis; HLA-B27-positive children of a parent who has both HLA-B27 and ankylosing spondylitis are at least five times more likely to develop the disease than their HLA-B27-negative siblings (Van der Linden *et al*. 1984).

With the advent of molecular techniques, spondylarthropathies have been linked to several of the HLA-B27 subtypes B*2702, 04 and 05 (MacLean 1992), but only B*2705 has been associated with childhood disease. Compelling evidence in support of an aetiological role for HLA-B27 is found in animal models of the disease. Transgenic Lewis rats expressing a high cell-surface density of human B*2705 and b $_2$-microglobulin develop a spondylarthropathy during the first 6 weeks of life, characterized by peripheral arthritis, spinal inflammation, and inflammatory bowel disease. These features appear to be specific for the *HLA-B27* gene, as they are not seen in the rats expressing the *B*2705* gene in low copy number, or in rats transgenic for *HLA-B7* (Taurog *et al*. 1994). Genetic factors other than HLA-B27 are also likely to play an important part in humans. Jarvinen (1995) has demonstrated a disease concordance rate of 50 per cent for HLA-B27-positive monozygotic twins and only 20 per cent for HLA-B27-positive dizygotic twins.

Environmental factors, including enteric bacteria, are linked to the aetiology of juvenile spondylarthropathy by several intriguing strands of evidence. Enteric infections with salmonella, shigella, campylobacter or yersinia during childhood may be accompanied by peripheral arthritis, and occasionally by typical Reiter's syndrome (arthritis, urethritis, and conjunctivitis). Ninety per cent of children with Reiter's syndrome are HLA-B27 positive, and 80 per cent have had a preceding dysentery. Although the majority of these cases remit spontaneously, at least 1 in 5 develops radiographic evidence of sacroiliac arthritis (Cuttica *et al*. 1992) and there are anecdotal reports of progression to clinical spondylarthropathy (Southwood and Gaston 1993). The role of enteric bacteria in the arthritis associated with inflammatory bowel disease is less clear. There is no doubt that many children with inflammatory bowel disease develop arthritis (7–21 per cent). Successful treatment of the bowel inflammation with sulphasalazine may be accompanied by remission of peripheral arthritis, the benefit of remission appearing to depend on the antibiotic (sulpha-) component, which may alter enteric bacterial colonization. Patients with inflammatory bowel disease who are HLA-B27 positive are more likely to develop ankylosing spondylitis (Lindsley and Schaller 1974; Passo *et al*. 1986).

There are reports that subacute bowel inflammation (histological evidence of acute or chronic ileocolonic inflammation) is found in up to 80 per cent of patients with juvenile spondylarthropathy, and a similar proportion of patients with late-onset pauciarticular juvenile chronic arthritis (Mielants *et al*. 1987; Veys *et al*. 1992; Mielants *et al*. 1993). The finding of chronic mucosal inflammation appeared to predict evolution to ankylosing spondylitis in these patients. There is also evidence that immune responses to enteric bacteria are found in the synovial compartment of children with juvenile spondylarthropathy. Approximately 90 per cent of children who have HLA-B27 positive chronic arthritis have lymphocytes in synovial fluid responsive to enteric bacteria, compared with only 25 per cent of children with HLA-B27-negative disease (Life *et al*. 1993). However, the stimulus to accumulation of such lymphocytes within the joint space remains unclear. There have been no consistent reports of bacteria or foreign antigenic fragments in the synovial membrane or fluid. Potentially cross-reactive immune responses directed against both bacterial and human heat-shock proteins have been proposed, but there are several arguments against such a directly pathogenetic role. Cells recognizing specific pathogens account for only a tiny minority (1:500) of the synovial cell population (Kingsley, personal communication) and it is possible that the inflammatory milieu of the arthritic joint favours the relatively non-antigen-specific accumulation of memory T lymphocytes.

The importance of enteric colonization has been shown in the *HLA-B27* transgenic rat model. A group of *HLA-B27* transgenic rats maintained in a germ-free environment did not develop the spondylarthropathy or bowel inflammation seen in their littermates. The responsible enteric pathogen or pathogens have yet to be isolated (Taurog 1994). If this model is a true reflection of human disease, the *HLA-B27* gene may predispose an individual to spondylarthropathy by reducing immune resistance to enteric bacteria at the level of the antigen-presenting cell (Feltkamp *et al*. 1996).

Clinical features and outcome

It is frequently difficult to distinguish early or undifferentiated juvenile spondylarthropathy from juvenile chronic arthritis. The spondylarthropathies are dynamic diseases that may continue to evolve over several years before reaching full expression. A child with arthritis in a peripheral joint(s) may not be suspected of having a spondylarthropathy until classical signs of axial involvement develop during the adult years. This section will deal firstly with the clinical features of the undifferentiated spondylarthropathies, and then with recognizable spondylarthropathies of ankylosing spondylitis, inflammatory bowel disease, Reiter's syndrome, and psoriasis in childhood (Table 1).

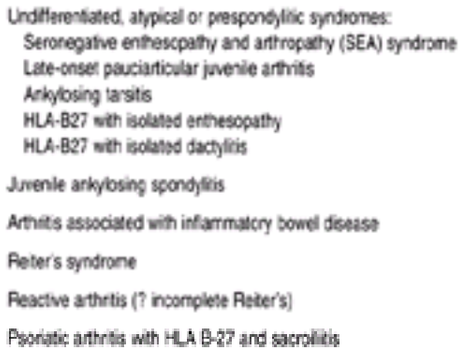


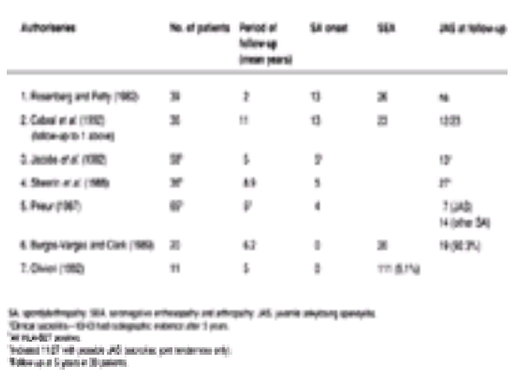
Table 1 Classification of spondylarthropathies in children

Undifferentiated or atypical spondylarthropathy

In a population-based study from Sweden, Andersson Gare *et al*. (1987) reported that 5 per cent of children with chronic arthritis were classified with a spondylarthropathy, but it has been estimated that between 10 and 15 per cent of children diagnosed as having juvenile chronic arthritis may eventually develop a spondylarthropathy. This poses a diagnostic challenge: to identify characteristic clinical features that predate the onset of back and sacroiliac symptoms, and therefore allow the prediction of outcome with some accuracy in this group of children with arthritis. Several descriptive, retrospective clinical studies of children who eventually developed an identifiable spondylarthropathy have provided useful information for this purpose (Schaller *et al*. 1976; Jacobs *et al*. 1982; Schaller 1983), and can be summarized as follows.

1. A family history of spondylarthropathy (defined as inflammatory low back disease) was described in up to 60 per cent of patients.
2. There was a strong male preponderance in patients; male:female ratios up to 10:1 have been reported.
3. Antecedent insults included febrile illnesses and musculoskeletal trauma.
4. The arthritis was of 'late onset'; only 40 per cent of the children were symptomatic by their ninth year. The earliest reported onset was at 12 months of age.
5. The arthritis was typically pauciarticular and asymmetrical, predominantly involving the large, weight-bearing joints of the lower limbs.
6. Extra-articular manifestations occurred in 42 per cent of patients, including urethritis, iritis, conjunctivitis, and keratoderma blennorrhagicum.
7. Enthesopathy was prominent in up to 75 per cent of patients.

Several sets of criteria have been proposed to differentiate childhood spondylarthropathy from other forms of juvenile arthritis. Most recently, the term 'late-onset pauci-articular juvenile chronic arthritis' has become fashionable, although there are no studies to indicate the validity of this designation (Table 2).



SA, spondyloarthropathy; SEA, seronegative enthesopathy and arthropathy; JAS, juvenile ankylosing spondylitis.
*Clinical sacroiliitis—10/43 had radiographic evidence after 5 years.
†All HLA-B27 positive.
‡Included 11/27 with juvenile JAS (sacroiliac joint tenderness only).
§Follow-up in 5 years in 39 patients.

Table 2 Unclassified or atypical spondylarthropathies in childhood

The SEA syndrome

Clinical features

Rosenberg and Petty (1982) described 39 children (35 boys and 4 girls) with a syndrome of seronegative enthesopathy and arthropathy (**SEA** syndrome). The mean age of onset of the first musculoskeletal symptom was 9.8 years (range 1–16 years). This group included 13 patients who fulfilled diagnostic criteria for the diagnosis of juvenile ankylosing spondylitis, inflammatory bowel disease, reactive arthritis, or Reiter's syndrome. The remaining 26 did not have one of these identifiable diseases, but did have the combination of enthesitis and arthritis or arthralgia. Enthesitis was demonstrated by discrete, localized tenderness at the bony insertions of ligaments, tendons or fascias. Principle sites of enthesitis included the calcaneal insertions of the plantar fascia and Achilles tendons, the metatarsal heads and base of the fifth metatarsal, the ischial tuberosities, iliac crests, and patellar tendon insertions. Low back pain and stiffness were present in only nine patients, but many had abnormal flattening of the lumbar curve on forward flexion. HLA typing of 32 children demonstrated that 23 (72 per cent), including the eight patients with juvenile ankylosing spondylitis and one with Reiter's syndrome, were HLA-B27 positive.

Juvenile ankylosing tarsitis has been described in up to 80 per cent of patients with SEA syndrome and 87 per cent of those with juvenile ankylosing spondylitis (Levi *et al*. 1990; Burgos-Vargas 1991). These patients develop inflammation of synovial sheaths and bursas, tendons, entheses and joints of the feet, leading to radiographic or MRI evidence of ankylosis. Additionally, non-traumatic atlantoaxial subluxation has been reported in two HLA-B27-positive children with SEA syndrome (Foster *et al*. 1995).

Outcome

There have been several longer-term follow-up studies of patients with SEA syndrome. Cabral *et al*. (1992) reported on 36 of the original 39 patients described by Rosenberg and Petty (1982), 2.5 to 23.5 years (mean 11 years) after the onset of their symptoms. Assessment of outcome in the patients who did not have ankylosing spondylitis, inflammatory bowel disease, reactive arthritis, or Reiter's syndrome in the original study revealed that the disease had progressed to definite or probable ankylosing spondylitis in half (12 patients). The presence of arthralgia was less specific than definite arthritis for predicting a spondylitic outcome. Burgos-Vargas and Clark (1989) reported 20 Mexican patients with SEA syndrome, and compared their outcome after at least 5 years of follow-up with 25 patients with polyarticular-onset juvenile chronic arthritis and 28 patients with definite ankylosing spondylitis of juvenile onset. Radiographic evidence of sacroiliitis of the ankylosing spondylitis type showed in four patients with SEA syndrome before the third year of follow-up. From the third to the fifth year of follow-up, back complaints and radiographically confirmed sacroiliitis fulfilling the diagnostic criteria for ankylosing spondylitis developed in an increasing proportion of the patients (47.1–75 per cent) and ultimately

affected 92.3 per cent. Other than the absence of back problems at the initial presentation, no significant differences in outcome were seen between the group with SEA syndrome and the group with juvenile ankylosing spondylitis.

Olivieri *et al*. (1992) reported that only 1 of 11 Caucasian HLA-B27-positive children (9.1 per cent) with SEA syndrome developed bilateral sacroiliitis after 5 years of disease. Ethnic and environmental factors were thought to have contributed to the discrepancy between these findings and those of the Mexican population reported by Burgos-Vargas and Clark (1989).

Atypical spondylarthropathy

Clinical features

Hussein *et al*. (1989) have proposed a set of diagnostic criteria for atypical spondylarthritis in a study of 26 children. Cases were classified according to criteria that were shown to be highly sensitive in differentiating atypical spondylarthropathies from other forms of juvenile chronic arthritis. The major criteria were (i) a family history of spondylarthropathy or oligoarthritis, (ii) enthesopathy, (iii) arthritis of digital joints including big toes, (iv) sacroiliitis, (v) presence of the *HLA-B27* gene, and (vi) recurrent arthralgia or arthritis. Minor criteria were (i) age of onset after 10 years of age, (ii) male sex, (iii) only the lower extremities affected, (iv) acute iritis or conjunctivitis, (v) arthritis of hips, and (vi) onset following an idiopathic enteritis. When four of the six major criteria were present, 96.1 per cent of the patients were correctly classified as having atypical spondylarthropathy, with a sensitivity of 84.6 per cent and a specificity of 100 per cent. The same diagnostic accuracy was achieved when three major and three minor criteria were present. Long-term follow-up studies of patients with atypical spondylarthropathy defined by the above criteria have not been reported.

HLA-B27-positive arthritis

Clinical features

The usefulness of HLA-B27 as a diagnostic marker of the juvenile spondylarthropathies has been discussed by many investigators, but interpretation of a positive HLA-B27 result is complicated by the occurrence of the gene in at least 8 per cent of the normal, non-arthritic childhood population. In addition, HLA-B27 has been associated with a number of different musculoskeletal diseases in children and it is unclear if these all form part of the spondylarthropathic spectrum. For example, there have been reports of children with isolated dactylitis (Siegel and Baum 1988), isolated peripheral enthesitis (Olivieri *et al*. 1990; Olivieri and Pasero 1992), and isolated hip flexion contractures (Bowyer 1995).

Outcome

A number of investigators have reported the outcome of children with arthritis who are HLA-B27 positive. Jacobs *et al*. (1982) described 58 such patients who were followed for a mean period of 5 years. The arthritis at the initiation of the illness was often transient and recurrent, pauciarticular and asymmetrical in distribution, and occurred primarily in the knees or ankles. Hip signs were noted in only seven of the patients at onset, but the hips were ultimately affected in 36 per cent. Radiographs of the sacroiliac joints were obtained for 43 of the patients, 10 of whom showed signs of sacroiliitis after a mean symptomatic period of 5 years (range 1–12 years). Other radiographic findings included periostitis, severe osteopenia, calcaneal erosions, or spurs. Rapid destruction of a single joint occurred in 3 of 58 patients.

Prieur (1987) reported a study of 65 children with HLA-B27-associated arthritis. There were 45 boys and 20 girls, with a mean age at onset of symptoms of 10 years (range 2.5–16 years). Just over a quarter of the patients (27 per cent) had arthritis in an upper limb during the first 6 months of disease. Enthesitis was present in 10 patients and 'sausage' digits in one-third. After 5 years' follow-up, 32 per cent had fulfilled criteria for the diagnosis of a defined spondylarthropathy: ankylosing spondylitis (7 patients), Reiter's syndrome (2 patients), psoriatic arthritis (9 patients), and Crohn's disease (3 patients).

Thirty-six HLA-B27-positive children with arthritis who had been followed for a mean period of 8.9 years were described by Sheerin *et al*. (1988). Five patients had an initial diagnosis of juvenile ankylosing spondylitis (14 per cent), and 24 (67 per cent) initially had peripheral arthritis without axial involvement. The most frequently involved joint was the knee (15 patients), followed by the ankle (8 patients), and the foot (5 patients). During the follow-up period, 22 patients had symptoms consistent with enthesitis, although it was demonstrable in only 16 patients (44 per cent). Eight patients (22 per cent) had extra-articular manifestations: acute iritis (4 patients), inflammatory bowel disease (1), psoriasis (2), and localized scleroderma (1). One of the patients with psoriasis had recurrent episodes of urethritis suggestive of incomplete Reiter's syndrome. Of the 28 patients in whom radiographs of sacroiliac region had been taken, 13 had normal results. A clinical course consistent with juvenile ankylosing spondylitis ultimately developed in 27 of the 36 patients.

Juvenile ankylosing spondylitis

The Rome and New York criteria for the diagnosis of ankylosing spondylitis (Kellgren *et al*. 1963; Bennett and Burch 1967) were derived for use in adults, and have not been validated in the paediatric population. However, there is a small proportion of children with arthritis who do fulfil the criteria for the diagnosis of the disease.

Epidemiology

Historical averages in populations of children with arthritis attending paediatric rheumatology clinics have suggested that 5 to 8 per cent have ankylosing spondylitis compared with 75 to 83 per cent with juvenile chronic arthritis (Ladd *et al*. 1971). The prevalence of juvenile ankylosing spondylitis can also be extrapolated from the numbers of adult patients who report that onset occurred before they were 16 years old. Lawrence *et al*. (1989) reported an estimated prevalence of ankylosing spondylitis in populations from the United States ranging from 129/100 000 (Carter *et al*. 1979) to 222/100 000 (Mikkelsen *et al*. 1967). Generally between 10 and 19 per cent of cases begin before the age of 16 years (Hart and Maclagan 1955). The prevalence depends on the population under study; for example, Burgos-Vargas *et al*. (1989), describing a Mexican Mestizo population in which ankylosing spondylitis was diagnosed between 1980 and 1987, reported that 54 per cent had onset of symptoms before the age of 16 years. Over 90 per cent of patients in published series of juvenile ankylosing spondylitis were HLA-B27 positive as is true in the adult population (Edmonds *et al*. 1974; Sturrock *et al*. 1974; Veys *et al*. 1976; Hafner 1987; Cassidy and Petty 1995). The sum of these data suggests that the prevalence of juvenile ankylosing spondylitis is approx. 12 to 18/100 000.

Clinical characteristics

Juvenile ankylosing spondylitis in children commonly presents as an arthritis affecting the large joints of the lower limbs, particularly the knees and ankles (Schaller *et al*. 1969; Ladd *et al*. 1971; Bywaters 1976; Kleinman *et al*. 1977; Schaller 1977; Ansell 1980; Garcia-Morteo *et al*. 1983; Hafner 1987; Burgos-Vargas *et al*. 1989; Burgos-Vargas and Vazquez-Mellado 1995). Although it is rare for the hips to be involved at presentation, there is a small group of children in whom recurrent, transient hip symptoms eventually develop into ankylosing spondylitis. The early course of the disease is frequently episodic, and other features of an undifferentiated spondylarthropathy may be present. Inflammation of the entheses may be prominent early in the disease course. The most common sites include Achilles tendon insertions, peripatellar insertions including the quadriceps and patellar tendons, and the insertions of the plantar fascia into calcaneum and metatarsal heads. Subtalar and mid-tarsal arthritis is not uncommon. Tarsometatarsal arthritis is sometimes associated with hindfoot involvement and results in a supination deformity. Indeed, spontaneous fusion and complete obliteration of the tarsometatarsal joints may occur (Levi *et al*. 1990; Burgos-Vargas 1991) (Fig. 1). Persistent hip involvement leading eventually to total hip arthroplasty is one of the most common peripheral articular complications of juvenile ankylosing spondylitis (Calin and Elswood 1988). Burgos-Vargas and Vazquez-Mellado (1995) assessed early clinical features that help to differentiate juvenile chronic arthritis from juvenile ankylosing spondylitis, using retrospective analysis of a group of patients at 6 months, 12 months, and 10 years after onset. Enthesitis and tarsal disease in children who have arthritis of the lower but not of the upper extremities appear to differentiate juvenile ankylosing spondylitis from juvenile chronic arthritis during the first year of symptoms.

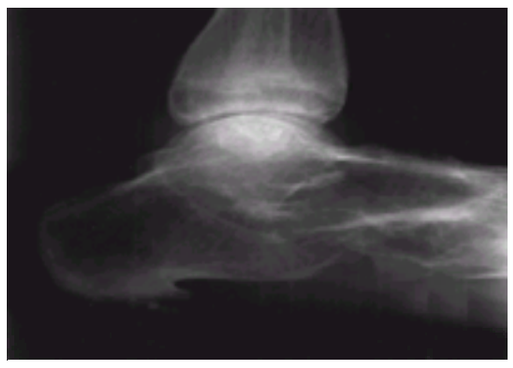


Fig. 1 Radiograph of ankle and hindfoot of the patient also illustrated in Fig. 2. The subtalar and talonavicular articulations are essentially ankylosed. A calcaneal spur is present from long-standing plantar fasciitis.

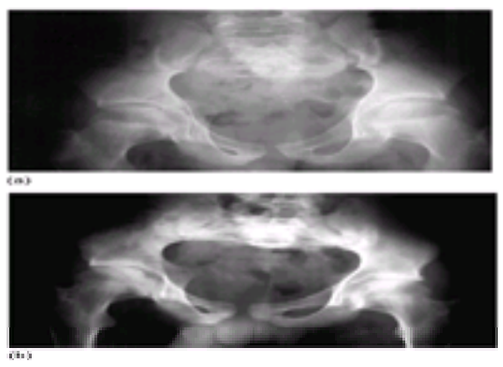


Fig. 2 (a) Radiograph of pelvis in a 14-year-old boy demonstrating normal sacroiliac joints. He was initially diagnosed as juvenile chronic arthritis. (b) Radiograph of the same patient after 4 years of disease activity. Note the sclerosis and erosion of the sacroiliac joints. The patient did not specifically complain of back pain but has limited lumbar flexion on Schober testing.

The joints of the arms are much less commonly affected, but occasionally the shoulders are involved. In contrast to adults with ankylosing spondylitis, most affected children do not report pain in the axial skeleton initially. Only 12.8 to 24 per cent have pain, stiffness or limitation of motion of the lumbosacral spine or sacroiliac within the first year of disease (Burgos-Vargas and Petty 1992). This may be due in part to the rather poorly localized symptoms of lumbosacral or sacroiliac joint disease, which may cause pain in the buttock, groin, or thigh. Evidence of sacroiliac involvement can be elicited by pressure over the sacroiliac joints or by compression of the joints using manoeuvres such as Patrick's test or Gaenslen's sign (Hoppenfeld 1976). Serial follow-up with Schober's measurements and measurement of chest expansion can reveal progressive restriction of the range of motion of the axial skeleton. Most patients develop axial involvement between 5 and 10 years after the onset of symptoms. Definite involvement of the spine and sacroiliitis in patients with juvenile ankylosing spondylitis occurred after a mean of 7.3 ± 2 years in a series of 35 patients (Burgos-Vargas and Vazquez-Mellado 1995). Serial plain radiographs of the sacroiliac joints may be helpful in monitoring the course of the disease (Fig. 2).

The systemic complications of juvenile ankylosing spondylitis include acute iritis (Ladd *et al*. 1971; Schaller 1977; Ansell 1980; Hafner 1987), aortic valve insufficiency (Stewart *et al*. 1978; Reid *et al*. 1979; Ansell 1980; Gore *et al*. 1981; Kean *et al*. 1980; Gerster *et al*. 1987; Stamato *et al*. 1995), C1–C2 subluxations (Reid and Hill 1978), and amyloidosis (Ansell 1980). The frequency of aortic valve insufficiency detected by imaging with colour Doppler ultrasound may be as high as 30 per cent of patients with a juvenile spondylarthropathy (Gerster *et al*. 1987), but others have reported figures of 8 per cent for aortic regurgitation and 5 per cent for mild mitral regurgitation (Stamato *et al*. 1995).

Several studies have compared juvenile with adult ankylosing spondylitis (Marks *et al*. 1982; Burgos-Vargas *et al*. 1989; Garcia-Morteo *et al*. 1983). Significant differences noted in the juvenile-onset type are the peripheral articular onset, precocious hip destruction, and the insidious, late development of axial involvement. Systemic complications are similar, apart from apical pulmonary disease, IgA nephropathy, and cauda equina syndrome, which have not been reported in children.

Outcome

The long-term outcome of juvenile ankylosing spondylitis is variable. Garcia-Morteo *et al*. (1983) found no juvenile-onset patients who retained full functional capacity after a mean disease duration of 15 years. Twelve patients (50 per cent) were in class II, seven in class III, and five in class IV using the Steinbocker classification. In contrast, the functional outcome of patients with the disease appears to be reasonably good, with a large proportion in employment (73 per cent) or continuing education (Marks *et al*. 1982). Most patients (95 per cent) require long-term anti-inflammatory therapy and have persistent peripheral arthritis. Calin and Elswood (1988) followed 135 patients with juvenile ankylosing spondylitis and reported that 74 per cent remained in full-time employment, a higher proportion than a comparable number of patients with adult-onset disease (54 per cent).

Inflammatory bowel disease-associated arthritis

Epidemiology

Peripheral arthritis and spondylitis are the most common extra-intestinal manifestations of chronic inflammatory bowel disease. Inflammatory bowel disease of unknown aetiology is divided into two major clinicopathological subtypes, ulcerative colitis and Crohn's disease; the common articular complications of each subtype are the same. In most series of adult patients, the prevalence of joint involvement is 4 to 22 per cent in ulcerative colitis and 2 to 22 per cent in Crohn's disease (Gavallese and Kantrowitz 1988). Two major reviews of arthritis associated with inflammatory bowel disease in children revealed similar prevalence: 9 to 21 per cent of patients with ulcerative colitis and 10 to 16 per cent of patients with Crohn's disease (Lindsley and Schaller 1974; Passo *et al*. 1986).

Clinical manifestations

The peripheral arthritis is pauciarticular, usually affecting the weight-bearing joints of the lower extremities, but occasionally involving joints of the upper extremities. The duration of attacks ranges from 2 days to 12 weeks, and averages about 1 month. These attacks occasionally precede the onset of bowel symptoms but usually coincide with or occur after the onset of bowel disease. After the initial attack, episodes of peripheral arthritis frequently occur during exacerbations or active periods of bowel disease. Passo *et al*. (1986) noted eight patients with Crohn's disease who had episodes of arthritis at times when they were free of bowel symptoms. These patients had anaemia and hypoalbuminaemia, however, suggesting that subclinical bowel inflammation may have been present. Importantly, most flares of bowel disease were not accompanied by arthritic manifestations. There was no correlation between the severity of the bowel inflammation and the occurrence of arthritis. Treatment of the bowel inflammation was usually associated with resolution of the peripheral arthritis.

Lindsley and Schaller (1974) describe five children with inflammatory bowel disease, all of whom had classical signs of ankylosing spondylitis (pain, loss of motion in the low back, and radiographic evidence of arthritis in the sacroiliac joints and lumbar spine). The symptoms were persistent and progressive, resulting in permanent impairment. Destructive hip disease was seen in four of these patients and all five suffered peripheral arthritis, including hip inflammation, at some time during follow-up. Joint symptoms preceded bowel symptoms in two of the five patients by 3 months and by 8 years; one had concurrent onset of bowel and joint symptoms. The activity of spondylitis and concurrent peripheral arthritis did not correlate with that of the bowel disease nor was treatment of the bowel disease effective in treatment of the spondylitis.

There are few reproducible data on HLA typing in children with spondylitis and inflammatory bowel disease. In adult patients, the prevalence of HLA-B27 positivity is 53 to 75 per cent, considerably lower than the 90 to 95 per cent prevalence seen in idiopathic ankylosing spondylitis (Gavallese and Kantrowitz 1988).

Reiter's syndrome

Prevalence

There are no estimates of the prevalence of Reiter's syndrome in the child population. Although fewer than 100 childhood cases have been described, reactive arthritis often goes unreported and probably many more cases have been encountered. There are several reviews of the literature and summaries of the data (Lockie and Hunder 1971; Iverson *et al*. 1975; Singsen *et al*. 1977; Rosenberg and Petty 1979; Cuttica *et al*. 1992). There is undoubtedly a considerable overlap between Reiter's syndrome and reactive arthritis.

Clinical manifestations

The clinical syndrome in children is similar to that described in adult patients. The individual features of arthritis, urethritis, conjunctivitis, keratoderma blennorrhagica, circinate balanitis, and oral mucosal ulcers may occur sequentially over a period of several weeks. Singsen *et al*. (1977) reported seven boys in a 2-year period who developed the triad of urethritis, arthritis, and conjunctivitis over a range of 4 to 24 days (mean 11 days). The evolution of signs and symptoms occurs asynchronously in many children. Whereas adult Reiter's syndrome often follows venereal infection, most cases of childhood Reiter's syndrome are preceded by enteric infection with species of salmonella, shigella, yersinia, or campylobacter. There are also individual reports of *Clostridium difficile* and parasitic infestation provoking Reiter's syndrome. A 14-year-old female was reported with Reiter's syndrome following an urinary-tract infection with *Escherichia coli* (Thomas and Roberton 1994). In 1992, Cuttica *et al*. (1992) reported the largest series of children with Reiter's syndrome and confirmed that diarrhoea antedated the onset of arthritis in 18/26 (69 per cent) patients; no patient reported venereal disease. The majority of patients (69 per cent) had pauciarticular involvement of lower extremities. Spine involvement was seen in six patients (23 per cent). The full triad of arthritis, conjunctivitis, and urethritis was seen in 9/16 (35 per cent). Reiter's syndrome in most children is a short-lived condition, characterized by complete resolution of the signs and symptoms; however, a protracted course and recurrent disease have been described (Iverson *et al*. 1975; Singsen *et al*. 1977). In Cuttica's series, mean duration of follow-up was 28.5 months (range 2–13.5 years); 15/26 (58 per cent) were in complete remission when last seen, seven (27 per cent) had a sustained course, three (11.5 per cent) had a fluctuating course, and only one child had a relapsing course. Functional outcome was grade I to II in 96 per cent. Radiographic erosions were noted in 4/26 (15 per cent) peripheral joints and sacroiliitis in 5/26 (21 per cent) (Cuttica *et al*. 1992). The long-term outcome of Reiter's syndrome in children is not well documented, although a proportion of cases undoubtedly progress to ankylosing spondylitis.

Juvenile psoriatic arthritis

Juvenile psoriatic arthritis has been traditionally grouped with the spondylarthropathies, but there is emerging evidence that the clinical and laboratory features of this disease have greater similarity to pauciarticular-onset juvenile chronic arthritis with an asymmetrical polyarticular course. Important dissimilarities to the spondylarthropathies are the lack of association with HLA-B27, and the relatively rare outcome of ankylosing spondylitis.

Prevalence

Most estimates of the prevalence of psoriasis in the general population range from 1 to 2 per cent (Baker 1966). In children, its estimated prevalence is approx. 0.3 per cent (Farber and Carsen 1966). The onset of psoriasis is frequently between 5 and 15 years of age, with one-third of all cases starting by the age of 15 years. Based on a population study by Hellgren (1969), Lawrence *et al*. (1989) calculated that approx 4.5 per cent of those with psoriasis have psoriatic arthritis. They further suggested that the prevalence of psoriatic arthritis in the United States is 0.67 per cent; accordingly, about 160 000 persons in the United States have psoriatic arthritis. Oriente *et al*. (1994) reported a similar prevalence of juvenile psoriatic arthritis of 1 per cent in a review of 425 patients with the onset of psoriasis before the age of 31 years. Cassidy and Petty (1995) estimate the prevalence of juvenile psoriatic arthritis at between 10 to 15 cases per 100 000 based on population studies reported by Espinoza (1985), making it approximately one-tenth as common as juvenile chronic arthritis, although fewer than 200 cases have been reported (Lambert *et al*. 1976; Calabro 1977; Sills 1980; Shore and Ansell 1982; Hamilton *et al*. 1990; Truckenbrodt and Hafner 1990).

Overall, juvenile psoriatic arthritis accounts for 2 to 8 per cent of chronic arthritic conditions in childhood (Lambert *et al*. 1976; Ansell 1977; Sills 1980). However, on using more liberal criteria for the diagnosis, Kuster and Quoss (1983) reported that it may account for as much as 40 per cent of juvenile chronic arthritis.

Clinical characteristics

Juvenile psoriatic arthropathy is a heterogeneous group of arthritic conditions that pose classificational problems. The majority of affected children resemble cases of juvenile chronic arthritis and a smaller fraction juvenile spondylarthropathy. In addition, there are probably children who have coincidental psoriasis and juvenile chronic arthritis (Petty 1994). Juvenile psoriatic arthritis has been defined by Lambert *et al*. (1976) as 'an inflammatory arthritis commencing prior to the age of 16 years, associated with psoriasis either preceding the onset of arthritis or occurring within the subsequent 15 years and usually with an absence of rheumatoid factor in the serum'. The psoriasis (nail and cutaneous manifestations) may lag behind the onset of arthritis by several months to years. In an effort to include patients who lack the typical cutaneous stigmata, Southwood *et al*. (1989) proposed a different set of criteria (the Vancouver criteria) for the definition of juvenile psoriatic arthritis (Table 3). Patients with conventionally defined psoriatic arthritis constituted 6 per cent of Southwood's patient population; indeed, according to the revised criteria, psoriatic arthritis constitutes approx. 19 per cent of the chronic arthritides of childhood. Shore and Ansell (1982) also described children without skin lesions but with probable juvenile psoriatic arthritis. Their 12 patients had asymmetrical arthritis of both upper and lower limbs, and a positive family history of psoriasis; 8 of the 12 had significant nail pitting. They suggested that such patients should be regarded as having 'probable juvenile psoriatic arthritis'. In the 5 years of follow-up, half of these patients developed overt psoriasis.

Definite psoriatic arthritis
Arthritis with typical psoriatic rash
Arthritis with three of the four following minor criteria:
Dactylitis
Nail pitting or onycholysis
Psoriasis-like rash
Family history of psoriasis (1st or 2nd degree relatives)
Probable psoriatic arthritis
Arthritis with 2 of the 4 minor criteria listed

[a]The clinical manifestations need not be present simultaneously. After Southwood *et al*. (1989) and Cassidy and Petty (1995).

Table 3 Vancouver criteria for the diagnosis of psoriatic arthritis in children [a]

Juvenile psoriatic arthritis differs from the other spondylarthropathies in that females are affected more often than males (1.5:1), but with marked variability among series (range 2:8 to 1:1.3, female:male) (Lambert *et al*. 1976; Shore and Ansell 1982; Southwood *et al*. 1989; Hamilton *et al*. 1990; Truckenbrodt and Hafner 1990; Koo *et al*. 1991). Simultaneous onset of arthritis and psoriasis is relatively uncommon. Most studies listed above show that the onset of arthritis is oligoarticular, but an asymmetrical polyarthritis evolves over the ensuing years. Sacroiliitis occurs in a significant number of cases (11–47 per cent), but is not inevitable (Southwood *et al*. 1989; Hamilton *et al*. 1990). The knee is the most commonly affected joint; dactylitis, tendinitis, and involvement of the distal interphalangeal joints are also extremely common. Dactylitis, distal interphalangeal inflammation, and the associated classical psoriatic nail changes are sometimes obvious (Fig. 3 and Fig. 4). The course of the disease tends to be more severe than that of pauciarticular juvenile arthritis; it is less severe, however, than the polyarticular course of juvenile chronic arthritis, especially in patients who are positive for rheumatoid factor. Overall, functional class I or II outcomes are common (Lambert *et al*. 1976; Calabro 1977). Fewer than 10 to 15 per cent of patients evolve into functional classes III or IV (Shore and Ansell 1982).

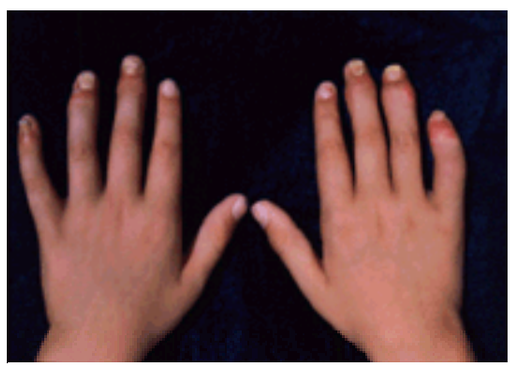


Fig. 3 The hands of an 8-year-old boy with severe psoriatic nail changes. They show arthritis of distal interphalangeal joints, dactylitis of the right fifth digit, and diffuse metacarpophalangeal involvement.

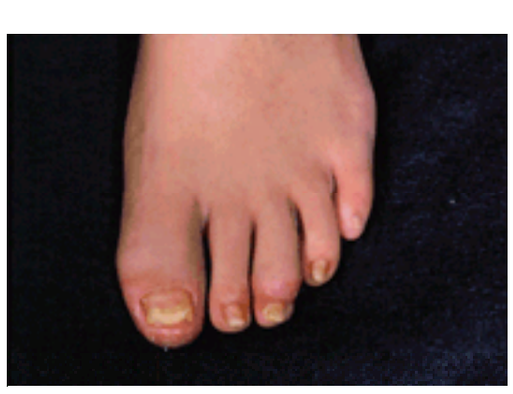


Fig. 4 The foot of the boy whose hands are depicted in Fig. 3, showing arthritis of the interphalangeal joint of the great toe and dactylitis of the third toe, as well as classical psoriatic nail changes.

Chronic iridocyclitis similar to that seen in juvenile arthritis developed in 8 to 17 per cent of reported patients with psoriatic arthritis (Sills 1980; Shore and Ansell 1982, Southwood *et al*. 1989; Truckenbrodt and Hafner 1990). Most patients with chronic iridocyclitis have also been positive for antinuclear antibody, which draws a closer association with juvenile chronic arthritis than the typical spondylarthropathies that manifest acute iritis.

No strong HLA association with juvenile psoriatic arthritis has been demonstrated. Hamilton *et al*. (1990) showed increased frequency of HLA-A2 and -B17, a possible increase in HLA-DR7, and no increase in HLA-B27. Southwood *et al*. (1989) found increased HLA-DR8 and decreased frequencies of HLA-DR4; HLA-B27 was found in eight patients (23 per cent), five of whom had clinical features of spondylarthropathy (back pain, limited range of motion in the back, and sacroiliitis). Ansell *et al*. (1993) confirmed the heterogeneity of 70 patients: HLA-B27 was increased in frequency, particularly in boys with older onset. HLA-A2, HLA-DR5, and HLA-DRw8 were found to be increased in early-onset cases. No correlation was found with the characteristic dactylitis with resultant 'sausage' toe or finger. It is likely that improvement in the classificational specificity of juvenile psoriatic arthritis as well as the spondylarthropathies in general will result in greater understanding of the aetiology, pathogenesis, and treatment of these diseases in the future.

Chapter References

Andersson Gare, B. *et al*. (1987). Incidence and prevalence of juvenile chronic arthritis: a population survey. *Annals of the Rheumatic Diseases*, **46**, 277–81.

Ansell, B.M. Bywaters, E.G.L. Lawrence, J.S. (1969). Familial aggregation and twin studies in Stills Disease, juvenile chronic polyarthritis. *Rheumatology*, **2:**37.

Ansell, B. M. (1977). Juvenile chronic polyarthritis. *Arthritis and Rheumatism*, **20** (Suppl.), 176–8.

Ansell, B. M. (1980). Juvenile spondylitis and related disorders. In *Ankylosing spondylitis* (ed. J. M. H. Moll), pp. 120–36. Churchill Livingstone, London.

Ansell, B., Beeson, M., Hall, P., Bedford, P., and Woo, P. (1993). Paediatric rheumatology: short report HLA and juvenile psoriatic arthritis. *British Journal of Rheumatology*, **32**, 836–7.

Baker, H. (1966). Epidemiological aspects of psoriasis and arthritis. *British Journal of Dermatology*, **78**, 249–61.

Bennett, P. H. and Burch, T. A. (1967). New York symposium on population studies in the rheumatic diseases: new diagnostic criteria. *Bulletin of Rheumatic Diseases*, **17**, 453–8.

Bowyer, S. (1995). Hip contracture as the presenting sign in children with HLA B-27 arthritis. *Journal of Rheumatology*, **22**, 165–7.

Burgos-Vargas, R. (1991). Ankylosing tarsitis: clinical features of a unique form of tarsal disease in the juvenile-onset spondylarthropathies. *Arthritis and Rheumatism*, Suppl. S196.

Burgos-Vargas, R. and Clark, P. (1989). Axial involvement in the seronegative enthesopathy and arthropathy syndrome and its progression to ankylosing spondylitis. *Journal of Rheumatology*, **16**, 192–7.

Burgos-Vargas, R. and Petty, R. E. (1992). Juvenile ankylosing spondylitis. *Rheumatic Diseases Clinics of North America*, **18**, 123–42.

Burgos-Vargas, R. and Vazquez-Mellado, J. (1995). The early clinical recognition of juvenile-onset ankylosing spondylitis and its differentiation from juvenile arthritis. *Arthritis and Rheumatism*, **38**, 835–44.

Burgos-Vargas, R., Naranjo, A., Castillo, J., and Katona, G. (1989). Ankylosing spondylitis in the Mexican Mestizo: patterns of disease according to age at onset. *Journal of Rheumatology*, **16**, 186–91.

Bywaters, E. G. L. (1976). Ankylosing spondylitis in childhood. *Clinical Rheumatic Diseases*, **2**, 387–95.

Cabral, D. A., Oen. K., and Petty, R. E. (1992). A follow-up study of children with seronegative enthesopathy arthropathy (SEA) syndrome (Abstr.). *Journal of Rheumatology*.

Calabro, J. J. (1977). Psoriatic arthritis in children. *Arthritis and Rheumatism*, **20** (Suppl.), 415–17.

Calin, A. and Elswood, J. (1988). The natural history of juvenile-onset ankylosing spondylitis: a 24-year retrospective case-control study. *British Journal of Rheumatology*, **27**, 91–3.

Carter, E. T., McKenna, C. H., and Brian, D. D. (1979). Epidemiology of ankylosing spondylitis in Rochester, Minnesota 1935–1973. *Arthritis and Rheumatism*, **22**, 365–70.

Cassidy, J. T. and Petty, R. (1995). Spondyloarthropathy. In *The textbook of pediatric rheumatology*, (3rd edn) (ed. J. T. Cassidy and R. E. Petty), pp. 224–59. Saunders, Philadelphia.

Cuttica, R. J., Scheines, E. J., Garay, S. M., Romanelli, M. D., and Cocco, J. A. M. (1992). Juvenile onset Reiter's syndrome. A retrospective study of 26 patients. *Clinical and Experimental Rheumatology*, **10**, 285–8.

Dougados, M. *et al*. (1991). The European study group preliminary criteria for the classification of spondyloarthropathy. *Arthritis and Rheumatism*, **34**, 1218–30.

Edmonds B. H. *et al*. (1974). Follow-up study of juvenile chronic polyarthritis with particular reference to histocompatibility antigen W27. *Annals of the Rheumatic Diseases*, **33**, 289–92.

Espinoza, L. R. (1985). Psoriatic arthritis: further epidemiologic and genetic considerations. In *Psoriatic arthritis* (ed. L. H. Gerber and L. R. Espinoza), pp. 9–32. Grune and Stratton, Orlando, FL.

Farber, E. M. and Carsen, R. A. (1966). Psoriasis in childhood. *California Medicine*, **105**, 415–20.

Feltkamp, T.E., Khan, M.A., and Lopez de Castro, J.A. (1996). The pathogenetic role of HLA-B27. *Immunology Today*, **17**, 5–7.

Foster, H. E. *et al*. (1995). Atlantoaxial subluxation in children with seronegative enthesopathy and arthropathy syndrome: 2 case reports and a review of the literature. *Journal of Rheumatology*, **22**, 548–51.

Garcia-Morteo, O., Maldonado-Cocco, A., Suarez-Almazor, M. E., and Garay, E. (1983). Ankylosing spondylitis of juvenile onset: comparison with adult onset disease. *Scandinavian Journal of Rheumatology*, **12**, 246–8.

Gavallese, E. M. and Kantrowitz, F. G. (1988). Arthritis manifestations of inflammatory bowel disease. *American Journal of Gastroenterology*, **84**, 703–9.

Gerster, J. C., Payot, M., and Piccinin, P. (1987) Clinical and echocardiographic assessment in juvenile-onset. *British Journal of Rheumatology*, **26**, 155–6.

Gore, J. E., Vizcarrondo, L. E., and Rieffel, C. N. (1981). Juvenile ankylosing spondylitis and aortic regurgitation: a case presentation. *Pediatrics*, **68**, 423–6.

Hafner, R. (1987). Die juvenile spondarthritis. Retrospektive untersuchung on 91 patienten. *Monatsschrift Kinderheilkund*, **135**, 41–6.

Hamilton, M. L., Gladman, D. D., Shore, A., Laxer, R. M., and Silverman, E. D. (1990). Juvenile psoriatic arthritis and HLA antigens. *Annals of the Rheumatic Diseases*, **49**, 694–7.

Hart, F. D. and Maclagan, N. F. (1955). Ankylosing spondylitis. A review of 184 cases. *Annals of the Rheumatic Diseases*, **14**, 77.

Hellgren, L. (1969). Association between arthritis and psoriasis in total populations. *Acta Rheumatologica*, **15**, 316–26.

Hoppenfeld, S. (1976). Physical examination of the lumbar spine In *Physical examination of the spine and extremities* (ed. S. Hoppenfeld), pp. 261–2. Appleton-Century-Croft, New York.

Hussein, A., Abdul-Khaliq, H., and von der Hardt, H. (1989). Atypical spondyloarthropathies in children: proposed diagnostic criteria. *European Journal of Pediatrics*, **148**, 513–17.

Iverson, J. M. I., Nanda, B. S., Hancock, J. A. H., Pownall, P. J., and Wright, V. (1975). Reiter's disease in three boys. *Annals of the Rheumatic Diseases*, **34**, 364–8.

Jacobs, J. C., Berdon, W. E., and Johnson, A. D. (1982). HLA-B27-associated spondylarthritis and enthesopathy in childhood: clinical, pathologic, and radiographic observations in 58 patients. *Pediatrics*, **100**, 521–8.

Jarvinen, P. (1995). Occurrence of ankylosing spondylitis in a nationwide series of twins. *Arthritis and Rheumatism*, **38**, 381–3.

Kean, W. F., Anastassiades, T. P., and Ford, P. M. (1980). Aortic incompetence in HLA B27-positive juvenile arthritis. *Annals of the Rheumatic Diseases*, **39**, 294–5.

Kellgren, J. H., Jeffrey, M. R., and Ball, J. (1963). *The epidemiology of chronic rheumatism*, Vol. 1, p. 326. Blackwell Scientific, Oxford.

Kleinman, P., Riverlis, M., Schneider, R., and Kaye, J. J. (1977). Juvenile ankylosing spondylitis. *Pediatric Radiology*, **125**, 775–80.

Koo, E., Balogh, Z. S., and Gomor B. (1991). Juvenile psoriatic arthritis. *Clinical Rheumatology*, **10**, 245–9.

Kuster, R. M. and Quoss, I. (1983). Juvenile psoriatic arthritis and spondylitis with or without psoriasis: probable and definite courses (Abstr.). *Zeitschrift für Rheumatologie*, **42**, 310.

Ladd, J. R., Cassidy, J. T., and Martel, W. (1971). Juvenile ankylosing spondylitis. *Arthritis and Rheumatism*, **14**, 579–90.

Lambert, J. R., Ansell, B. M., Stephenson, E., and Wright, V. (1976). Psoriatic arthritis in childhood. *Clinical Rheumatic Diseases*, **2**, 339–52.

Levi, S., Ansell, B. M., and Klenerman, L. (1990). Tarsometatarsal involvement in juvenile. *Foot and Ankle Journal*, **11**, 90–2.

Life, P., Hassel, A., Williams, K., *et al*. (1993). Responses to Gram-negative bacterial antigens by synovial T cells from patients with juvenile chronic arthritis: recognition of heat shock protein HSP60. *Journal of Rheumatology*, **20**, 1388–96.

Lindsley, C. B. and Schaller, J. G. (1974). Arthritis associated with inflammatory bowel disease in children. *Journal of Pediatrics*, **84**, 16–20.

Lockie, G. N. and Hunder, G. G. (1971). Reiter's syndrome in children: a case report and review. *Arthritis and Rheumatism*, **14**, 767–72.

MacLean, L. (1992). HLA-B27 subtypes: implications for the spondyloarthropathies. *Annals of the Rheumatic Diseases*, **51**, 929–31.

Marks, S. H., Barnett, M., and Calin, A. (1982). A case-control study of juvenile- and adult-onset ankylosing spondylitis. *Journal of Rheumatology*, **9**, 739–41.

Mielants, H., Veys, E. M., Joos, R., Covelier, C., De Vos, M., and Proot, F. (1987). Late onset pauciarticular juvenile chronic arthritis: relation to gut inflammation. *Journal of Rheumatology*, **14**, 459–65.

Mielants, H. *et al*. (1993). Gut inflammation in children with late onset pauciarticular juvenile chronic arthritis and evolution to adult—a prospective study. *Journal of Rheumatology*, **20**, 1567–72.

Mikkelsen, W. M. *et al*. (1967). Estimates of the prevalence of rheumatic diseases in the population of Tecumseh, Michigan, 1959–60. *Journal of Chronic Diseases*, **20**, 351–69.

Olivieri, I. and Pasero, G. (1992). Long-standing isolated juvenile onset HLA B-27 associated peripheral enthesitis. *Journal of Rheumatology*, **19**, 164–5.

Olivieri, I., Barbieri, P., Gemignani, G., and Pasero, G. (1990). Isolated juvenile onset HLA-B27 associated peripheral enthesitis. *Journal of Rheumatology*, **17**, 567–8.

Olivieri, I., Foto, M., Ruju, G. P., Gemignani, G., Giustarini, S., and Pasero, G. (1992). Low frequency of axial involvement in Caucasian pediatric patients with seronegative enthesopathy and arthropathy syndrome after 5 years of disease. *Journal of Rheumatology*, **19**, 469–75.

Oriente, C. B., Scarpa, R., and Oriente, P. (1994). Prevalence and clinical features of juvenile psoriatic arthritis in 425 psoriatic patients. *Acta Dermatologica Venereologica (Stockholm)*, Suppl. 186, 109–10.

Passo, M. H., Fitzgerald, J. F., and Brandt, K. D. (1986). Arthritis associated with inflammatory bowel disease in children: relationship of joint disease to activity and severity of bowel lesion. *Digestive Disease Science*, **31**, 492–7.

Petty, R. E. (1990). HLA-B27 and rheumatic diseases of childhood. *Journal of Rheumatology*, **17**, 7–10.

Petty, R. E. (1994). Juvenile psoriatic arthritis, or juvenile arthritis with psoriasis? *Clinical and Experimental Rheumatology*, **12** (Suppl. 10), S55–8.

Prieur, A. M. (1987). HLA B27 associated chronic arthritis in children: review of 65 cases. *Scandinavian Journal of Rheumatology*, **66** (Suppl.), 51–6.

Reid, G. D. and Hill, R. H. (1978). Atlanta-axial subluxation in juvenile ankylosing spondylitis. *Journal of Pediatrics*, **93**, 531–2.

Reid, G. D., Patterson, M. W. H., Patterson, A. C., and Cooperberg, P. L. (1979). Aortic insufficiency in association with juvenile ankylosing spondylitis. *Journal of Pediatrics*, **95**, 78–80.

Rosenberg, A. M. and Petty, R. E. (1979). Reiter's disease in children. *American Journal of Diseases of Children*, **133**, 394–8.

Rosenberg, A. M. and Petty, R. E. (1982). A syndrome of seronegative enthesopathy and arthropathy in children. *Arthritis and Rheumatism*, **25**, 1041–7.

Schaller, J. G. (1977). Ankylosing spondylitis of childhood onset. *Arthritis and Rheumatism*, **20**, 398–401.

Schaller, J. G. (1983). Pauciarticular arthritis of childhood (pauciarticular juvenile arthritis). *Annales de Pediatrie*, **30**, 557–63.

Schaller, J. G. (1985). The spondyloarthropathies (variant diseases) In *Practice of pediatrics*, Vol. 8, (ed. V. E. Kelly), pp. 1–10. Harper and Row, Philadelphia.

Schaller, J. G., Bitnum, S., and Wedgwood, R. J. (1969). Ankylosing spondylitis with childhood onset. *Journal of Pediatrics*, **74**, 505–16.

Schaller, J. G., Ochs, H. D., Thomas, E. D., Nisperos, B., Feigl, P., and Wedgewood, R. J. A. (1976). Histocompatibility antigens in childhood onset arthritis. *Journal of Pediatrics*, **88**, 926–30.

Sheerin, K. A., Giannini, E. H., Brewer, E. J., Jr, and Barron, K. S. (1988). HLA-B27-associated arthropathy in childhood: long-term clinical and diagnostic outcome. *Arthritis and Rheumatism*, **31**, 1165–70.

Shore, A. and Ansell, B. M. (1982). Juvenile psoriatic arthritis—an analysis of 60 cases. *Journal of Pediatrics*, **4**, 529–35.

Siegel, D. M. and Baum, J. (1988). HLA-B27 associated dactylitis in children. *Journal of Rheumatology*, **15**, 976–7.

Sills, E. M. (1980). Psoriatic arthritis in childhood. *Johns Hopkins Medical Journal*, **146**, 49–53.

Singsen, B. H., Bernstein, B. H., Koster-King, K. G., Glovsky, M. M., and Hanson, V. (1977). Reiter's syndrome in childhood. *Arthritis and Rheumatism*, **20** (Suppl.), 402–7.

Southwood, T.R. and Gaston, J.S.H. (1993). Evolution of synovial fluid mononuclear cell responses in a HLA-B27-positive patient with *Yersinia*-associated juvenile arthritis. *British Journal of Rheumatology*, **32**, 845–8.

Southwood, T. R. *et al*. (1989). Psoriatic arthritis in children. *Arthritis and Rheumatism*, **32**, 1007–13.

Stamato, T. *et al*. (1995). Prevalence of cardiac manifestations of juvenile ankylosing spondylitis. *American Journal of Cardiology*, **75**, 744–6.

Stewart, S. R., Robbins, D. L., and Castles, J. J. (1978). Acute fulminant aortic and mitral insufficiency in ankylosing spondylitis. *New England Journal of Medicine*, **299**, 1448–9.

Sturrock, R. D., Dick, H. M., Henderson, N., Goel, G. K., Lee, W. C., and Buchanan, W. W. (1974). Association of HL-A27 and AJ in juvenile arthritis and ankylosing spondylitis. *Journal of Rheumatology*, **1**, 269–73.

Taurog, J.D., Richardson, J.A., Croft, J.T., *et al*. (1994). The germ free state prevents the development of gut and joint inflammatory disease in HLA-B27 transgenic rats. *Journal of Experimental Medicine*, **180**, 2359–64.

Thomas, D. G. and Roberton, D. M. (1994). Reiter's syndrome in an adolescent girl. *Acta Paediatrica*, **83**, 339–40.

Truckenbrodt, H. and Hafner, R. (1990). Die psoriasisarthritis intramuscular kindesalter. *Z Rheumatol*, **49**, 88–94.

Van der Linden, S.M., Valkenburg, H.A., de Jongh, B.M., and Cats, A. (1984). The risk of developing ankylosing spondylitis in HLA-B27 positive individuals. A comparison of relatives of spondylitis patients with the general population. *Arthritis and Rheumatism*, **27**, 241–9.

Veys, E. M., Coigne, E., Mielants, H., and Vergruggen, A. (1976). HLA and juvenile chronic polyarthritis. *Tissue Antigens*, **8**, 61–5.

Veys, E. M., Mielants, H., Joos, R., and Clercq, L. D. (1992) Juvenile spondyloarthropathies in 1992. *Journal of Rheumatology*, **20** (Suppl. 37), 19–25.

5.5.3 Ankylosing spondylitis

Andrei Calin

Historical review

As summarized elsewhere (Calin 1988) Raymond, in 1912, provided convincing illustrations of mummies from Egypt who appeared to have ankylosing spondylitis. The paleopathologist, Sir Mark Ruffer, described a gentleman called Nefermaat who lived some 3000 years before Christ. His mummified remains demonstrated a rigid block of bone from the mid-cervical region to the sacrum. Short described 18 further instances of ankylosing spondylitis from Egyptian sources over the next 3 millennia and other examples exist from the Danish and French neolithic periods. More recent (AD 1200) remains of New Mexican Indians with ankylosing spondylitis have been described. There has, over the last decades, been some debate as to whether these individuals had diffuse idiopathic skeletal hyperostosis or ankylosing spondylitis; although there is no absolute certainty, the fused sacroiliac joints in many instances attest to such patients having ankylosing spondylitis.

In the nineteenth century Bechterew, Strumpell, and Marie promoted the general recognition of ankylosing spondylitis.

Diagnostic criteria

Criteria for the classification and diagnosis of different rheumatic diseases have been developed to help different clinicians in different locations use the same diagnostic labels for similar groups of patients. Originally, criteria were formulated in Rome in 1961 and later revised in New York in 1966 (Bennett and Burch 1968) (Table 1). The most frequent means of satisfying these is a radiographic demonstration of grade 3 or 4 bilateral sacroiliitis (Fig. 1), together with a history of back pain. In fact, few clinicians rely on these criteria and most would consider a patient as having ankylosing spondylitis if symptomatic and with sacroiliitis. Naturally, as discussed in Chapter 5.5.1, there are often exceptions to any list of criteria and, although of major concern to the epidemiologist, the clinician is often satisfied to use 'common sense' rather than rigid criteria. For example, we do recognize ankylosing spondylitis *sine* sacroiliitis on the one hand or atypical forms of disease on the other.

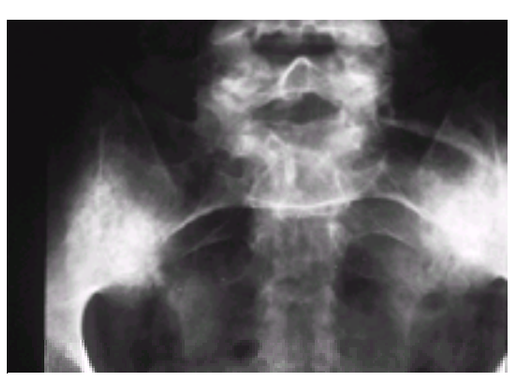


Fig. 1 Pelvic radiograph (anteroposterior view) revealing grade IV sacroiliitis. Note marked juxta-articular sclerosis and destruction of joint with fusion.

Rome criteria
Clinical criteria
1. Low back pain > 3 months, relieved by rest
2. Thoracic pain and stiffness
3. History of iritis
4. Limited motion of lumbar spine
5. Limited chest expansion
Radiological criteria
1. Bilateral sacroiliitis
Ankylosing spondylitis is diagnosed if bilateral sacroiliitis plus one clinical criterion, or four out of five clinical criteria, are present.

New York criteria
Clinical criteria
1. Limited movement of lumbar spine in three planes
2. Pain in lumbar spine or at dorsolumbar junction
3. Chest expansion = 2.5 cm
Radiological criteria
1. Bilateral sacroiliitis, grade 3–4
2. Unilateral sacroiliitis, grade 3–4, or bilateral sacroiliitis, grade 2

Table 1 Criteria for ankylosing spondylitis

The disease is considered primary if no other rheumatological disorder is present, or secondary if the sacroiliitis is related to psoriatic arthropathy, inflammatory bowel disease or Reiter's syndrome. One great difficulty lies in the prevalence of patients with clinical symptoms and radiological signs of disease but in whom the radiologist has failed to recognize sacroiliitis. Such individuals are often misdiagnosed as having mechanical back pain or other inappropriate labels. In essence, a single anteroposterior film of the pelvis is sufficient to define the radiological entity. More sophisticated investigations such as the use of magnetic resonance imaging are always expensive and may be inappropriate and unhelpful (Braun *et al*. 1994).

A review of the Rome Criteria reveals an obvious flaw: the limitation of spinal movement is not defined. The New York Criteria do define limitation of chest expansion as 2.5 cm or less but there is inevitably a major subjective component to chest expansion, and decreased movement typically occurs only late in the course of the disease. Pain, as such, is too sensitive and too non-specific to be used as a criterion and we have developed a simple analysis of the qualitative nature of pain in patients with inflammatory spinal disease (i.e. ankylosing spondylitis). Specifically, (Calin and Fries 1977) (i) patients are typically below 40 years of age at onset; (ii) the onset is insidious; (iii) duration has been at least 3 months at first attendance; (iv) there is an association with morning stiffness; and (v) improvement occurs with exercise. For those with three or more of these features a pelvic radiograph should elucidate whether there is evidence of sacroiliitis. HLA B27 typing has led to immense strides in the understanding of the spondylarthropathy group but should not be considered a diagnostic test or necessary for the diagnosis of ankylosing spondylitis. A patient with symptomatic sacroiliitis lacking HLA B27 still has ankylosing spondylitis and, moreover, the test is frequently negative in secondary forms of ankylosing spondylitis, where only some 50 per cent of those with enteropathic spondylitis or psoriatic spondylitis carry the antigen. Occasionally sacroiliitis is seen as a chance finding in the absence of pain. Presumably the precipitating trigger has been pulled in the susceptible individual, but for reasons not clearly understood, symptoms have never reached a threshold noticed by patient or physician. A review of Bayesian theory reminds us that HLA B27 typing can only be helpful when we are 50 per cent certain of the diagnosis.

By using Bayesian analysis it has been pointed out that if the clinician is 50 per cent certain of the diagnosis, then HLA B27 typing is helpful, whereas at the extremes of a priori probability such testing is of no use. Clearly, the physician is unlikely to know when he or she is 50 per cent confident! The clinician perhaps should simply treat with a non-steroidal anti-inflammatory drug if a diagnosis of ankylosing spondylitis appears possible, rather than relying on further testing such as computed tomography and other investigations (Calin 1980; Calin 1982).

In epidemiological and familial studies patients are sometimes found with unilateral sacroiliitis, dactylitis, syndesmophytes, and other stigmata of 'undifferentiated spondylarthropathy'(see Chapter 5.5.1).

Until the specific environmental trigger(s), gene(s), and biological mechanism(s) leading to disease pathogenesis are fully elucidated, the obsessional use of criteria for classification may be inappropriate.

Epidemiology

During the 1950s it was recognized that ankylosing spondylitis occurred in twins, brothers, fathers, mothers and other relatives of affected individuals. Indeed, Strecher (1957) found the disease to be 30 times more prevalent among relatives of spondylitics than among controls. The author suggested that the disease was inherited as a single autosomal dominant factor with '70 per cent penetrance in men and 10 per cent penetrance in women'. In 1967 Emery and Lawrence studied 188 available first-degree relatives of 76 probands with ankylosing spondylitis and appropriately matched controls. Sixteen per cent of the first-degree relatives of patients had sacroiliitis, with 20 per cent of the males affected and 8 per cent of the females—a ratio higher than the often quoted 10:1 (male:female) ratio claimed at that time.

The search for an explanation for this increased heritability took a dramatic step forward in 1973 with two reports of the association between ankylosing spondylitis and HLA B27. Indeed, the link between HLA B27 and ankylosing spondylitis could elucidate several observations:

1. The family clustering as discussed above. HLA B27 is inherited as an autosomal codominant characteristic, 50 per cent of first-degree relatives of probands with HLA B27 possessing the antigen.
2. Uveitis is a common accompaniment of ankylosing spondylitis. HLA B27 is found in some 40 per cent of individuals with acute unilateral self-limiting uveitis, even in the absence of underlying rheumatological disease.
3. Many patients with Reiter's disease develop sacroiliitis. Overall, some 80 per cent of patients with reactive arthropathy are HLA B27 positive, those developing sacroiliitis and ascending spinal disease being more closely linked to HLA B27. Whether the sacroiliitis in such patients should be considered a complication of Reiter's syndrome or a further manifestation of the HLA B27 status remains unclear.
4. Juvenile chronic arthropathy, psoriatic arthropathy, and inflammatory bowel disease can all be associated with sacroiliitis and ankylosing spondylitis and it is known that HLA B27 is increased in all three groups who have a spondylarthropathy picture.

For a further discussion of the link between HLA B27 and the spondylarthropathies, the reader is directed to Chapter 5.5.1 and elsewhere. Herein some further clinical aspects may be summarized:

1. Some 5 to 10 per cent of HLA B27 positive individuals develop ankylosing spondylitis after an unknown environmental event and 20 per cent of subjects with HLA B27 develop reactive arthropathy after contact with an arthritogenic agent (chlamydia, salmonella, etc.).
2. Up to 5 per cent of Caucasian patients with ankylosing spondylitis are not HLA B27 positive.
3. Only 50 per cent of those with psoriatic or enteropathic spondylitis are B27 positive.
4. The association between ankylosing spondylitis and HLA B27 in non-Caucasians (around 50 per cent) is much less than that seen in Caucasians.
5. Relatives of probands with both sacroiliitis and HLA B27, even when carrying an identical HLA haplotype, frequently remain disease free.
6. The Pima and Haida Indians, both with a high frequency of HLA B27, develop ankylosing spondylitis frequently but Reiter's syndrome rarely, if ever. By contrast, the Navaho Indians and Alaskan Inupiat Eskimos develop Reiter's syndrome more frequently than ankylosing spondylitis.
7. Ankylosing spondylitis and Reiter's syndrome tend to 'breed true' within families (Calin *et al*. 1984).
8. HLA B27 relatives of HLA B27-positive patients are 20 times more likely to develop ankylosing spondylitis than are HLA B27-positive relatives of healthy HLA B27 subjects. The distribution of HLA B27 among different healthy and disease groups is summarized in Chapter 5.5.2, Table 2, and discussed more fully elsewhere (Calin 1989; Calin and Elswood 1989a; De Castro 1994; Inman and Schfield 1994; Khan 1996; Toivanen and Toivanen 1996; Lopez-Larrea *et al*. 1996; Scofield 1996; Braun and Sieper 1996).

	Males	Females
Family history	+	++
Association with psoriasis	++	+
Association with inflammatory bowel disease	++	+++
HLA B27	>90 per cent	>90 per cent
Disease activity	++	+++
Function (severity)	++	+++
Peripheral joint disease:		
Initial	+	++
Subsequent	+	+++
Spinal ankylosis*	++	+
Cervical spine symptoms	+	++
Ostelitis pubis	+	+++

*Skipping thoracic and lumbar spine in females

Table 2 Major differences between ankylosing spondylitis in men and in women

9. Some 11 subtypes of HLA-B27 are now recognized.

Prevalence (see van der Linden and van der Heijde 1996)

With the increased awareness and interest in the disease, many patients who previously were thought to have mechanical back pain, 'seronegative rheumatoid arthritis', and other disorders are now recognized as having ankylosing spondylitis. The true prevalence of ankylosing spondylitis appears to be in the region of 0.25 to 1 per cent with a peak of 2 per cent in northern Norway. The figures contrast sharply with older data reporting a ratio of ankylosing spondylitis to rheumatoid disease of about 1:15.

Several studies of blood donor populations suggest that up to 20 per cent of HLA-B27-positive individuals develop symptomatic ankylosing spondylitis, the majority of whom did not carry a diagnosis. Of interest, some 20 per cent of HLA-B27-positive individuals develop reactive arthropathy/Reiter's syndrome following infection with an arthritogenic trigger. Other studies have suggested that some 10 per cent of individuals with HLA-B27 develop ankylosing spondylitis (Dawkins *et al*. 1981), while up to 20 per cent of Pima and Haida men with HLA-B27 have the condition.

The prevalence and nature of spondylarthropathy varies in different ethnic groups. For example, Boyer *et al*. (1994) identified 104 cases of spondylarthropathy in an Eskimo population, a prevalence of 2.5 per cent in adults aged 20 years and over. They found undifferentiated spondylarthropathy and reactive arthropathy to be more common than ankylosing spondylitis *per se*. Strikingly, they found the prevalence of spondylarthropathy to have an equal sex distribution. Of note, although the prevalence of HLA-B27 is in the range of 20 to 40 per cent, the number of cases of spondylarthropathy was not as high as that found among the Canadian Haida Indians (6 to 10 per cent) (Braun *et al*. 1994). Elsewhere, Alexeeva *et al*. (1994) studied the prevalence of spondylarthritis amongst the native population of Chukotaa in Russia. Among these circumpolar subjects, they found the prevalence of spondylarthropathy to be 2.5 per cent, with 1 per cent having ankylosing spondylitis. HLA-B27 occurred in 34 per cent of the population. Nasution *et al*. (1993) found intriguing data in Indonesia. Specifically, HLA-B27 was found in 62 per cent of the Chinese patients with ankylosing spondylitis compared with under 3 per cent among the healthy controls. In contrast, only 8 per cent of the native Indonesians with disease were HLA-B27 positive compared with 9 per cent of the healthy controls, indicating the lack of association between HLA-B27 and disease in native Indonesians. Finally, Mijiyawa (1993) found spondylarthritis in 31 of 2000 patients in Lomé, Togo. In this population eight patients were HIV carriers. They therefore concluded that spondylarthropathy would appear to be less rare in black Africans than hitherto considered. Moreover, the frequency is likely to rise with the increase in HIV prevalence.

Delay in diagnosis

Until recently delays of between 5 and 10 years were recorded between the onset of symptoms and the diagnosis at last being made. There are now preliminary data suggesting that this delay is decreasing (Calin *et al*. 1988).

Sex distribution

That ankylosing spondylitis may occur much more frequently in females than hitherto suggested has been confirmed by several studies. In our large series in Britain, the sex ratio is in the region of 2.5 to 1 in favour of men, a figure confirmed by other studies. Distinguishing features between men and women with disease are discussed below (and see Table 2) (Will *et al*. 1990a; Kennedy *et al*. 1993).

Racial distribution (see Khan 1990; Khan 1996)

It has long been recognized that ankylosing spondylitis occurs more frequently in Caucasian populations. In fact it appears that ankylosing spondylitis roughly follows the distribution of HLA-B27. For example, in the American Indian where HLA-B27 prevalences have been reported ranging from 18 to 50 per cent, ankylosing spondylitis is particularly frequent, whereas the condition is less common in the black American, where HLA-B27 has a prevalence of 3 to 4 per cent, and correspondingly rarer in black Africans, where HLA-B27 occurs in under 1 per cent. Strikingly, the association with HLA-B27 is less dominant in those races where this phenotype is less frequent (e.g. only 60 to 70 per cent of Japanese patients are HLA-B27 positive; the frequency of which in the general population is correspondingly low (1 to 2 per cent)).

Age distribution (see Kennedy et al. 1994)

Sacroiliac and spinal disease usually develop in the late teenage years or in the early twenties in primary ankylosing spondylitis, whereas in those with secondary forms of disease, older ages at onset are seen. Interestingly, in the developing world, ankylosing spondylitis occurs more frequently at a younger age, teenagers with onset of disease being frequently found. By contrast, in the developed world onset in the mid- to late twenties is relatively more common. As discussed in Chapter 5.5.1, there may well be a changing pattern in disease. Our studies (Will *et al*. 1990b; Will *et al*. 1992) have suggested that within France and Britain the age at onset of ankylosing spondylitis is increasing. The fact that different studies using different epidemiological techniques have produced these findings is of particular interest.

Pathological features

The unique pathology of ankylosing spondylitis and the spondylarthritides was clearly defined in 1971 by Ball and developed further by Bywaters (1984). In contrast to the situation in rheumatoid disease the primary pathological site is the enthesis (insertion of ligaments and capsules into bone) rather than the synovium. In addition, the enthesopathic change is characterized by fibrosis and ossification rather than joint destruction and instability. In the spine, enthesopathic changes at the site of insertion of the outer fibres of the anulus fibrosus result in squaring of vertebral bodies (Fig. 2), vertebral end-plate destruction, and syndesmophyte formation (Fig. 3). The enthesis is a metabolically active site, perhaps explaining why early changes in ankylosing spondylitis occur during growth in the teenage years. In spite of our understanding about HLA-B27 and the pathology we still do not understand why the enthesis preferentially becomes affected.

Fig. 2 Squaring of lumbar vertebral body in ankylosing spondylitis. Note sclerosis of bone at site of enthesopathic change at insertion of anterior fibres of the annulus fibrosis.

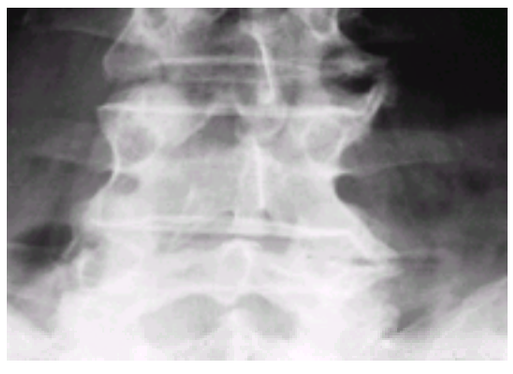


Fig. 3 Note syndesmorphytes between vertebral bodies. These are vertically directed new bone lesions associated with normal joint space.

In summary, ossification occurs in the region of the discs, the epiphyseal and sacroiliac joints and extraspinal sites, initiated by lesions at the site of ligamentous insertion. Synovitis itself does occur in peripheral joints and a proliferative synovitis can mimic that seen in rheumatoid disease. For reasons that we do not yet understand, soluble HLA Class II antigens are found in high levels in the synovial fluid in both rheumatoid arthritis and degenerative arthritis but in very low levels in spondylarthritis (Armas *et al*. 1990).

Clinical features

Ankylosing spondylitis will only be diagnosed when there is a high index of suspicion in a patient presenting with back pain of an inflammatory nature. In any such individual the differential diagnosis relates to mechanical dysfunction. The two conditions are contrasted in Table 3.

	Mechanical	Inflammatory
Pain history	+	++
Family history	–	+
Onset	Acute	Insidious
Age (years)	15–90	< 40
Sleep disturbance	+	++
Morning stiffness	+	+++
Involvement of other systems	–	+
Effect of exercise	Worse	Better
Effect of rest	Better	Worse
Radiation of pain	Anatomic (S1, L5)	Diffuse (thoracic, buttock)
Sensory symptoms	+	–
Motor symptoms	+	–
Scoliosis	+	–
Range of movement decreased	Asymmetrically	Symmetrically
Local tenderness	Local	Diffuse
Muscle spasm	Local	Diffuse
Straight-leg raising	Decreased	Normal
Sciatic nerve stretch	Positive	Absent
Hip involvement	–	+
Neurological	+	–
Other systems	–	+

Table 3 Differential findings in mechanical and inflammatory back pain

We have demonstrated a striking loss of bone mineral content in early disease (Will *et al*. 1990c). Juxta-articular and generalized osteoporosis are well recognized in rheumatoid disease but the relative role of hormonal factors, immobility, drug therapy, and the disease process is unclear. As discussed, ankylosing spondylitis is characterized by inflammation at the site of the enthesis which can lead to local bone erosion and juxtainsertional osteoporosis, followed by new bone formation. Although it has long been recognized that patients with severe ankylosing spondylitis may develop a dorsal kyphosis with some anterior wedging of the vertebrae, early osteoporosis has not until now been recognized. Whether the pattern of bone loss relates to tumour necrosis factor or other mediators remains unclear. HLA-B27-positive brothers of HLA-B27-positive patients have normal bone density. To what extent this early osteoporosis in patients who have normal spine mobility should be considered an early pathological marker, with changes at the enthesis being secondary in nature, is unknown.

Spinal symptoms

Late spinal complications

Few patients progress relentlessly to the classical late 'bamboo spine' (Fig. 4 and Fig. 5). The fused spine may fracture (Fig. 6) following trivial or even unrecognized injury and microfractures and clinical fractures are relatively common in severe ankylosing spondylitis. The spinal deformity may make it difficult to see a fracture site, particularly in the low neck, and special views may be required. The fracture may be clinically silent or a dramatic event which can be fatal in outcome. A sudden exacerbation of back pain, spontaneously or following mild trauma, may relate to a fracture or a localized defect of the vertebral end plate (destruction of the disc–bone border). Spondylodiscitis is the term given to this lesion and may require rest and analgesia, with pain decreasing over 2 or 3 weeks. This contrasts with the usual exercise programme required for patients. The nature of spondylodiscitis and its prevalence has recently been defined in a cross-sectional study of over 100 patients, some 12 per cent of whom had radiological evidence of discitis, though often asymptomatic in nature (Kabasakal *et al*. 1994).

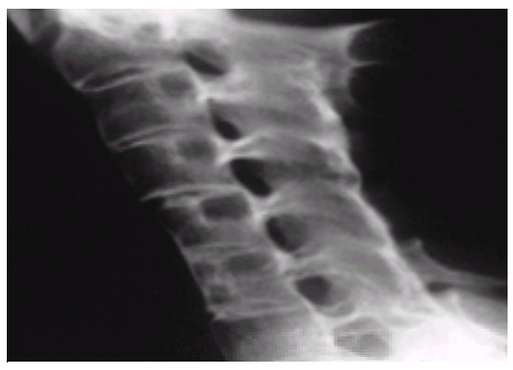


Fig. 4 Cervical spine in severe ankylosing spondylitis. Note fusion of facetal joints and anterior fusion of bodies with squaring of vertebrae.

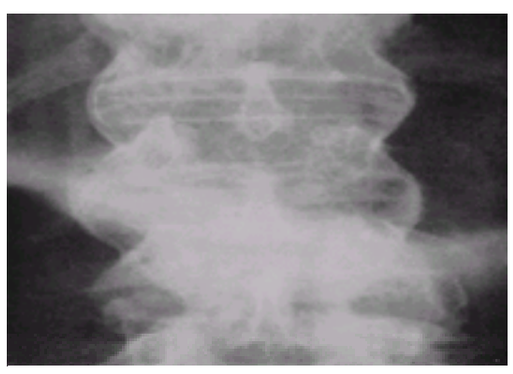


Fig. 5 Fused lumbar spine vertebrae with syndesmorphytes linking vertebrae.

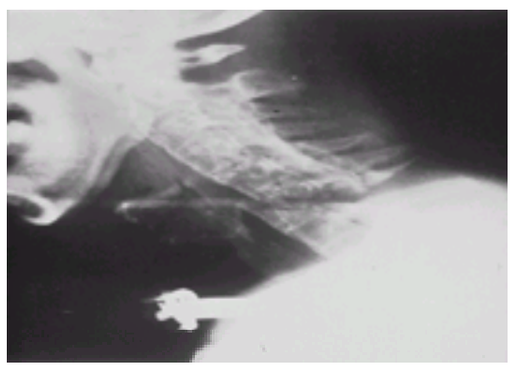


Fig. 6 Fracture of fused cervical spine in ankylosing spondylitis following minimal trauma—note defect in C6.

Extraspinal joint disease

Some 20 to 40 per cent of patients have peripheral joint disease at some stage during their illness. This may be asymmetric and often affects the lower limbs predominantly. Hip (Fig. 7) ankylosis and shoulder disease may provide major disability, with temporomandibular joint dysfunction occurring in up to 10 per cent of patients. HLA-DR4 may be associated with peripheral joint disease.

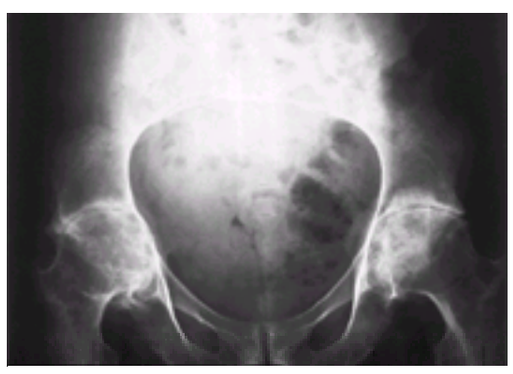


Fig. 7 Bilateral hip arthropathy in a female with ankylosing spondylitis. Note joint space reduction and erosive change in heads of femora.

We have recently shown that there is a striking inverse correlation between the age at onset of disease and hip involvement; the vast majority of individuals requiring a total hip replacement having onset of disease during the teenage years. Hip involvement in those with onset in the twenties or later is vanishingly rare. For those requiring hip replacement, bilateral surgery is frequent. The long-term outcome for those with total hip replacements is excellent, the majority doing well some 10 years after replacement (Calin and Elswood 1989b).

Enthesopathic lesions

In view of the pathological disorder it is not surprising that patients have insertional tendinitis at any site typified by involvement of the Achilles tendon, intercostal muscle insertions, plantar fasciitis, and dactylitis.

Apart from low back pain some patients have a 'pleuritic' type of chest pain that may cause sleep disturbance and anxiety. This pain is worse on inspiration and relates to an insertional tendinitis of the small costosternal and costovertebral muscles.

Extra-articular disease

Until recently ankylosing spondylitis was predominantly considered to be a spinal disease with little constitutional systemic involvement. It is now recognized that the disorder may affect all body systems and indeed may not be immunologically silent.

General symptoms

Constitutional features include fatigue, weight loss, low grade fever, hypochromic or normochromic anaemia, and increased erythrocyte sedimentation rate. For many patients, fatigue is the major component. We have shown (Calin *et al*. 1993a) that in a comparison of pain, stiffness, and fatigue, the three major features of ankylosing spondylitis, the last of these has been a major component for a large minority. The main difficulty relates to management. Exercise and anti-inflammatory drugs are good for both the pain and stiffness but, to date, fatigue remains a frustrating symptom to treat. Certain features such as an arthropathy and uveitis can occur at any time in the course of the disease; other problems are predominantly associated with severe chronic involvement. Examples of the latter include aortic regurgitation, cord compression, upper lobe pulmonary fibrosis, and amyloid deposition.

Eye disease

Iritis (Fig. 8) occurs in up to 40 per cent of patients with ankylosing spondylitis and, as mentioned, has little correlation with the disease activity in the spine. Although the visual episodes are often self-limiting, local steroid drops or systemic therapy may be required. In a recent study (Edmunds and Calin 1991) we have shown that there is no relationship between the inflammation of the uveal tract and spine. In addition we failed to define any obvious environmental trigger even in those with recurrent disease. For example, there is no seasonal pattern.

Fig. 8 An example of severe recurrent iritis with both active change and evidence of old damage.

Pulmonary involvement

Chronic infiltrative and fibrotic changes in the upper lobe of the lungs may occur. This upper lobe pulmonary fibrosis is now well recognized. Cough, sputum, and dyspnoea may develop with the sputum becoming profuse. Radiographs reveal usually bilateral upper lobe pulmonary fibrosis, sometimes with cyst formation and parenchial destruction. The lesions can be invaded by aspergillus with changes mimicking tuberculosis. Histology reveals patchy pneumonia with round cell and fibroblast infiltration progressing to interalveolar fibrosis. Dense pleural and pulmonary fibrosis can occur. Treatment is of no avail and death may follow massive haemoptysis.

A rigid chest wall may result from fusion of the thoracic joints but pulmonary ventilation is usually well maintained by the diaphragm. In an analysis of deaths in patients with ankylosing spondylitis Court-Brown and Doll (1965) noted deaths from respiratory causes to be some three times higher than in a control population.

Cardiovascular disease

Cardiovascular involvement is well recognized. The stated prevalence of this complication varies from 3.5 per cent of cases within 15 years to 10 per cent with up to 30 years duration. Although heart disease occurs more frequently in those with more severe spondylitis, cardiac conduction defects and other abnormalities may occur in those with minimal disease.

Aortic incompetence, cardiomegaly, and persistent defects in cardiac conduction are the most common findings with complete atrioventricular block occasionally occurring. Pericarditis has been described and cardiac involvement can range from being clinically silent to dominating the picture.

Aortic incompetence is the best studied complication. Up to 20 per cent of patients with ankylosing spondylitis may have anatomic evidence of involvement of the aortic valve but few of these have clinically detectable valvular dysfunction. Scar tissue and intimal fibrous proliferation may affect the aortic valve cusps and aorta behind and above the sinuses of Valsalva. Scar tissue may extend below the base of the aortic valve producing a subaortic fibrous ridge. Occasionally the aortitis is evident before any evidence of ankylosing spondylitis and the condition can thus be considered a *forme fruste*, similar to the occurrence of uveitis in a HLA-B27-positive individual without underlying rheumatological disease.

Amyloidosis

Amyloid is an occasional complication of ankylosing spondylitis. In one study, 3 of 35 patients were found to have amyloid on routine rectal biopsy. Although relatively common the event is rarely of clinical significance.

Renal disease

There appears to be no impairment of renal glomerular function in ankylosing spondylitis, in spite of the recognized pathological changes. In a study of 38 consecutive patients with severe ankylosing spondylitis, investigation of glomerular function failed to show any marked abnormality (Calin 1975). Nevertheless, immunoglobulin A (IgA) nephropathy is well recognized. However this appears to be of relatively little clinical significance.

Neurological syndromes

Involvement of the cauda equina may occur in the later stages of the disease. The syndrome presents with insidious onset of leg and buttock pain with sensory and motor impairment in association with bowel and bladder dysfunction. Lumbar diverticulae are found on myelographic examination. Unfortunately, treatment appears to be of no help, but a single case report suggests the value of a peritoneal shunt in one of our patients in whom the deterioration has been arrested.

Bowel disease (see Mielants et al. 1993a; Mielants et al. 1993b)

As discussed above and in Chapter 5.5.2 there is an intimate relationship between the bowel and ankylosing spondylitis. Low-grade bowel inflammation has been described (on ileocolonoscopy) in the absence of symptoms. The relevance of these findings remains unclear, although we know that the HLA-B27-positive transgenic rat develops a picture typical of spondylarthropathy with a major degree of bowel involvement. There is also a suggestion that there is increased gut permeability in patients and relatives, perhaps relating to disease pathogenesis (Martinez-Gonzalez *et al*. 1994).

Physical examination

Spinal mobility is symmetrically decreased but may still be normal in the earlier stages of disease. A variety of measurements have been described, although the modified Schober is the most useful. The distraction of a line drawn from the midpoint between the posterior iliac spines to an arbitrary site 10 cm above this point is measured and the distraction noted on forward flexion. In a normal individual this 10 cm line increases by some 50 to 100 per cent whereas a patient with active disease may only have a distraction of some 20 per cent or less. Lateral spinal flexion may be measured by noticing the distraction on contralateral flexion of a line drawn in the midaxillary plane. As mentioned above, chest expansion can be reduced, although this may only occur late in the disease. Moreover, in the female—for obvious reasons—chest expansion is difficult to measure precisely. Intermalleolar straddle on abducting the legs is a useful measurement of non-specific pelvic inflammation, and neck mobility can be measured. A formal approach to defining mobility in disease has now been validated (the Bath Ankylosing Spondylitis Radiology Index, BASRI) (Jenkinson *et al*. 1994).

There may be muscle spasm with loss of the normal lumbar lordosis, while some individuals may present with pain but no physical abnormality. Peripheral joints may be normal or grossly involved— particularly in women, who tend to have more peripheral joint disease than men.

Radiological evaluation

The five grades of sacroiliitis introduced by the New York Criteria range from 0 to 4. These are summarized in Table 4. Grade 0 refers to normal joints with clear sacroiliac margins and uniform joint space. There is no juxta-articular sclerosis. Grade 1 signifies suspicious change but no definite abnormality, while grades 2 and 3 relate to an increasing degree of sacroiliitis, as defined by blurring of the joint margin, juxta-articular sclerosis, decreased joint width, and erosive change. Grade 4 describes complete fusion or ankylosis of the joints with or without residual sclerosis.

Grade	
0 Normal	Grade depends on degree
1 Suspicious	of blurring of joint margins,
2 Minimal sacroiliitis	juxta-articular sclerosis,
3 Moderate sacroiliitis	erosive change, and
4 Ankylosis	narrowing

[a]According to New York criteria

Table 4 Sacroiliac changes in ankylosing spondylitis [a]

A single anteroposterior view radiograph of the pelvis is sufficient. Sacroiliitis does not occur as an age-related phenomenon although degenerative change of the sacroiliac joints may occur. Osteitis condensans ilii, a sclerotic condition of unknown cause producing dense bone in one or both iliacs but not the sacrum, may confuse the unwary. However, typically this iliac change allows the joint margin to be defined more clearly and therefore there should be no confusion between the two conditions. It is not unknown for an individual patient to have both sacroiliitis and osteitis condensans ilii.

Patients with bilateral, grade-4 sacroiliitis and no extrapelvic disease are frequently seen. Why some patients do not have progressive disease above the pelvic rim remains unclear.

In view of the enthesopathic nature of the disease it is not surprising that radiological changes occur elsewhere. Thus, vertebral squaring may occur, vertebral-end plate collapse may be seen, and varying degrees of ossification develop. The radiological changes may be contrasted with those seen in diffuse idiopathic hyperostosis (Table 6). Likewise syndesmophytes must be differentiated from osteophytes; the latter associated with disc-space narrowing. In essence, syndesmophytes move in a vertical direction whereas osteophytes are typically horizontal. McEwen *et al*. (1971) have defined the radiological differences between primary ankylosing spondylitis and the secondary spondylarthropathies, as summarized in Table 5. The BASRI allows for a simple and rapid assessment of entire spine and hips (Pande *et al*. 1995).

	Primary ankylosing spondylitis — and ankylosing spondylitis associated with inflammatory bowel disease	Ankylosing spondylitis associated with Reiter's syndrome and psoriatic arthropathy
Sacroiliac changes	Symmetrical	Asymmetrical
Osteitis pubis	++	+
Facetal joint involvement	+++	+
Squaring of vertebrae	++	+
Syndesmophytes	+++	+
Ossification	++	+
Spread	Ascending	Random

Table 5 Radiographic changes in primary and secondary forms of ankylosing spondylitis

	DISH	Ankylosing spondylitis
Usual age of onset (years)	> 50	< 40
Thoracolumbar kyphosis	±	++
Limitation of spinal mobility	±	++
Pain	±	++
Limitation of chest expansion	±	++
Radiography:		
Hyperostosis	+	+
SI joint erosion	–	++
SI joint synovial obliteration	±	++
SI joint ligamentous obliteration	+	++
Apophyseal joint obliteration	–	++
ALL ossification	++	±
PLL ossification	+	0
Osteoporosis/lysis	–	++
Enthesopathies (whiskering) with erosions	–	++
Enthesopathies (whiskering) without erosions	++	+
HLA-B27 (white patients) (%)	6	95
HLA-B27 (black patients) (%)	2	50

[illegible] ALL, anterior longitudinal ligament; PLL, posterior longitudinal ligament.

Table 6 Differentiating features of diffuse idiopathic skeletal hyperostosis (DISH) and ankylosing spondylitis. Of note osteoporosis is an early and dramatic finding in ankylosing spondylitis (Will *et al*. 1989)

Radionuclide scanning may be more sensitive than radiography but the change is non-specific and is likely to confuse rather than help. Likewise computed tomography and MRI is expensive and usually unnecessary.

Ankylosing spondylitis and diffuse idiopathic skeletal hyperostosis are compared in Table 6. Osteoporosis is now well recognized as is the consequence (i.e. compression fractures) (Donnelly *et al*. 1994; Mullaji *et al*. 1994).

Laboratory abnormalities

Apart from the presence of HLA-B27 in over 95 per cent of white patients, the majority have no evidence of change in laboratory tests. In fact there is little correlation between erythrocyte sedimentation rate and other so-called mediators of inflammatory disease and disease activity as defined by clinical symptoms and findings. Serum alkaline phosphatase may be elevated and we and others have described elevated serum creatine phosphokinase levels. The explanation for this remains unclear. Circulating immune complexes may be found and one study suggested that psoriasis-associated retrovirus-like particles may be present. Likewise antibodies to proteoglycan occur and antibodies to heat shock protein are described. Interleukin 6 and acute phase reactants have recently been compared (Tutuncu *et al*. 1994). Interleukin 6 is more frequently raised in disease, compared to the situation with erythrocyte sedimentation rate, plasma viscosity or C-reactive protein (approximately 80 per cent compared with 50 per cent).

Ankylosing spondylitis in women

The disease is frequently missed in women unless a high degree of suspicion exists. Moreover certain features of the disease in women make definition of the disease more difficult. The major differences between the disease in men and women are summarized in Table 2. Women tend to have more peripheral joint involvement and the spinal disease progresses less rapidly. However, dramatic spinal involvement in women as in men is sometimes seen. Historically, the delay in diagnosis is greater in women than men but hopefully this difference is now decreasing. Interestingly, the sex distribution of primary ankylosing spondylitis favouring men in the ratio of 2.5 to 1 is even greater in psoriatic spondylitis (4 to 1), but for those with enteropathic disease the ratio is equal. Age is also a factor (Kennedy *et al*. 1993; Kennedy *et al*. 1994).

The sex ratio is said to favour males (about 6:1 in the early teenage years) but this ratio decreases with increasing age. By the late twenties the ratio is down to 2.3 to 1 with intermediate ratios in the ages between. However, our recent data suggest that even in teenagers, the ratio is approximately 3:1 in favour of boys.

Aetiology (see Khan 1993; Khan 1996; Rubin *et al*. 1994; Reveille *et al*. 1994 and further discussion in Chapter 5.5.1)

The cause of ankylosing spondylitis remains unknown. For the last decade interest has focused on a specific klebsiella serotype and the putative relationship between this organism and HLA-B27-positive lymphocytes of patients. However, it appears more likely that many different organisms can precipitate disease in susceptible individuals. The relationship between HLA-B27, other genes (Burney *et al*. 1994), and the environment is under intense scrutiny. As discussed more fully in Chapter 5.5.1 there are several recognized facts that must be taken into consideration:

1. Approximately 5 to 10 per cent of HLA-B27-positive individuals develop ankylosing spondylitis—a similar percentage to those who develop reactive arthropathy following infection with an arthritogenic organism.
2. Relatives of healthy HLA-B27-positive subjects only rarely develop disease while those of HLA-B27-positive patients are more likely to do so.
3. Both ankylosing spondylitis and Reiter's syndrome breed true within families.
4. The prevalence of ankylosing spondylitis follows the distribution of HLA-B27—where the latter is rare the association is less close.
5. Some 5 per cent of individuals with ankylosing spondylitis are HLA-B27 negative.
6. Certain ethnic groups such as the Haida and Pima are likely to develop ankylosing spondylitis while others such as the Navaho and Inupiat Eskimo develop Reiter's syndrome.
7. HLA-B27 is heterogeneic with some 11 subtypes recognized. To date few differences have been found between HLA-B27-positive patients and HLA-B27-positive controls. HLA*2706 and 09 may protect against disease in the Thai and Indonesians. and in Sardinia, respectively (Lopez-Larrea *et al*. 1996).
8. Serum IgA and secretory IgA are raised in spondylitics.
9. Iliocolonoscopic biopsy evaluation of the distal small bowel reveals some clinical inflammation in both HLA-B27-positive and -negative individuals, regardless of bowel symptoms.
10. Diet may modify arthritis in animals and perhaps in humans.
11. Silent carriage of certain microorganisms can precipitate disease in reactive arthritis.
12. There may be influence from genetic loci on chromosomes 2, 14, 19, and perhaps others.
13. There is apparent sharing of homologous amino-acid sequence by HLA-B27 (host antigen residue 72 to 77) and *Klebsiella pneumoniae* (residues 188 to 193). Moreover, there is cross-reactivity between plasmid and HLA-B27 and between HLA-B27 and heat shock protein.
14. The link between psoriasis, inflammatory bowel disease, and ankylosing spondylitis must not be forgotten.
15. Sulphasalazine, which is efficacious in the management of inflammatory bowel disease, has at least some effect in modifying the disease process in ankylosing spondylitis.
16. That ankylosing spondylitis is not immunologically silent is now recognized. Immunoglobulin allotypes are similar to those of controls.
17. Animal models are of major value. The adjuvant rat and the HLA-B27-positive transgenic rat and mice models will certainly further elucidate the situation. The latter suggests that HLA-B27 itself is of paramount importance.
18. Calin *et al*. (1993b) compared familial ankylosing spondylitis with sporadic diseases and determined that familial disease tended to be milder than the sporadic version, suggesting that the latter had fewer susceptibility but more severity genes in contrast to the former.

Prognosis

The outcome for an individual patient is difficult to define. Some patients have disease limited to the pelvis while others have progressive intractable inflammation of spine and elsewhere. HLA-B27 homogeneity is not associated with more severe disease. In an attempt to define entry variables leading to poor outcome we have recently shown that some 15 per cent of patients younger than 15 years of age at onset require one or more total hip replacements within 18 years compared with only 1 per cent of those in their late twenties. Clearly the developing hip is at particular risk.

Sib pairs concordant for disease (Calin and Elswood 1989a) have a similar disease process from the radiological standpoint but for the fact that the age at onset appears to be determined by environmental rather than genetic factors (i.e. the sib pairs share the same calendar date at onset rather than a similar age at onset). The younger sib is more likely to develop hip involvement as discussed above. By contrast, offspring concordant for disease with their parent develop a disorder that relates only little to the prior generation. Lehtinen (1993) has studied the nature of mortality and causes of death in 398 patients with ankylosing spondylitis. There are no major surprises. Patients died from cardiovascular causes, amyloidosis, and sometimes the effect of treatment.

In an attempt to define predictive factors in spondylarthropathy, Amor *et al*. (1994) suggested that early hip involvement, an erythrocyte sedimentation rate over 30, poor initial response to non-steroidal anti-inflammatory drug treatment, a decreased Schober, the presence of dactylitis or oligoarticular disease, and age at onset below 16 years were all associated with a worse prognosis. We would agree but also add lower social-educational background. In addition, sporadic disease (Calin *et al*. 1993b) would appear to have a worse prognosis than the familial disorder.

The story is further complicated because we have to develop an understanding not only of the natural history of ankylosing spondylitis but also of the effect of treatment. As recently discussed (Calin 1994), the situation in ankylosing spondylitis is infinitely more difficult than that for rheumatoid disease or systemic lupus erythematosus. In the study of outcome, in spondylitis, we do not have the advantage of valuable laboratory tests. We therefore need to rely much more on what the patient tells us. Happily, we know from data provided by Hidding *et al*. (1994) that there is an excellent correlation between the self-report of symptoms and observed status in patients with ankylosing spondylitis, in contrast to the situation in those with fibromyalgia and, to a lesser extent, rheumatoid arthritis.

Defining disease status in ankylosing spondylitis

We have recently developed and validated self-administered instruments that define disease status in ankylosing spondylitis (Garrett *et al*. 1994) and functional ability (Calin *et al*. 1994). The Bath Ankylosing Spondylitis Disease Activity Index and the Bath Ankylosing Spondylitis Functional Index, together, define with clarity and simplicity the clinical status of patients with this condition. These two self-administered instruments, in addition to an objective measurement (the Bath Ankylosing Spondylitis Metrology Index) will allow many more studies in terms of natural history and response to management. The Metrology Index (Jenkinson *et al*. 1994) consists of five simple measurements of cervical rotation, tragus to wall distance, lateral spinal flexion, modified Schober, and intermalleolar distance. Finally, we have produced the Bath Ankylosing Spondylitis Global Status, which allows the definition of the different components of clinical well-being, or otherwise, to be defined and compared (Pande *et al*. 1995), and the Bath Radiology Index (BASRI).

Management: general considerations (see Cuellar and Espinoza 1996)

The majority of patients with ankylosing spondylitis have good prognosis for a successful life pattern, despite chronic discomfort over many years. The disease progresses to severe and total ankylosis in relatively few patients. A summary of treatment is given in Box 1.

Box 1 Treatement of ankylosing spondylitis

The patients should stop smoking. The combination of chest and pulmonary disease is a tragic occurrence. Swimming is the best routine sport to pursue. For those few individuals with relentless disease, admission to an active rehabilitation unit is advisable. Hydrotherapy and aggressive remedial exercises provide benefit over the short term and our ongoing studies are addressing the long-term outcome in a controlled study.

Postural exercise

The patient must realize that the aim of therapy is to maintain normal posture and physical activity. A hard bed at night is more useful than a soft mattress. One pillow only should be used. Extension exercises should remind the patient that the natural tendency of the disease is towards flexion and loss of height. Thus, extension exercises must be performed at least twice daily. A hot shower provides decreased stiffness and allows the exercise regimen to be followed. During the day, adequate attention must be given to the position of work, the style of chairs, and the chance for mobility.

One major difficulty for the patient with ankylosing spondylitis is that of fatigue. Our recent study suggests that low-dose amitriptyline at night may ameliorate this phenomenon (Koh *et al*. 1996).

Therapy

There are few well-controlled studies comparing different agents. Phenylbutazone was for a long time the drug of choice but this agent has fallen into disrepute, although we still favour its use in those with severe disease who do not respond to indomethacin or other agents. Indomethacin is usually considered the drug of choice and we use the slow-release preparation of 75 mg given once at night or twice daily. Patients are advised to titrate the dosage downwards, some requiring one tablet every other night or simply one tablet once or twice weekly. Indeed, it is likely that although indomethacin has been available for some 30 years, there is no

newer non-steroidal anti-inflammatory drug with greater efficacy. For the majority of patients indomethacin is still the favoured agent (Calin and Elswood 1990).

Sulphasalazine has been shown in a meta-analysis to be efficacious when compared with placebo. The improvement is perhaps not dramatic and needs to be weighed against the relative cost and the inevitable need for blood testing. In a European-wide study of sulphasalazine in patients with ankylosing spondylitis, reactive arthropathy, and psoriatic spondylarthropathy, we have shown that sulphasalazine has a small role to play, particularly in those with peripheral stigmata (Dougados *et al*. 1995) We have recently focused on patients' perception of drug risks and preparedness to take this risk (O'Brien *et al*. 1990a; O'Brien *et al*. 1990b). Finally, we await with interest, outcome studies relating to a controlled evaluation of inpatient compared with outpatient management.

Other drug therapy includes auranofin, which appears to be only minimally efficacious if at all, and possibly azathioprine. Surgical treatment is indicated, particularly for hip replacements, but also vertebral wedge osteotomies may be done in order to correct deformity.

Radiotherapy

Radiotherapy may be of limited value but it has fallen into disrepute because of the risk of leukaemia. A retrospective case–controlled study comparing those who had received radiotherapy and those who had not (Calin and Elswood 1989c) showed that radiotherapy may have had little more than minimal effect in altering the course of the disease.

Genetic counselling

Family members of patients with ankylosing spondylitis and other HLA-B27-related arthropathies are advised against routine typing. It is important for the family and physician to have knowledge regarding the disease in order that this can be recognized early on in family members.

The National Ankylosing Spondylitis Society (NASS)

In Britain there is a flourishing patient association known as the National Ankylosing Spondylitis Society and this has now been followed in many other countries around the world. All patients should be directed to their national association where membership provides excellent educational material, newsletters, and advice about everyday and professional activities. For information regarding NASS or the international body contact: The Director, Mr Fergus Rogers, National Ankylosing Spondylitis Society, 5 Grosvenor Crescent, London, SW1X 7ER (Tel. 0171 235 9585; Fax 0171 235 5827).

Summary

Ankylosing spondylitis is a common disease. There are still many aspects of the disorder that remain poorly understood, both in terms of aetiology and natural history. For example, an epidemiological study suggested that even in the absence of specific and recognized complications, the death rate in patients with ankylosing spondylitis is greater than in matched controls (Radford *et al*. 1977). Nevertheless the physician and patient can usually remain optimistic about the long-term outlook. The vast majority of patients are following a satisfying professional, personal, and family life.

Chapter References

Alexeeva, L., Krylov, M., Vturin, V., Mylov, N., Erdesz, S., and Benevolenskaya, L. (1994). Prevalence of spondylarthropathies and HLA-B27 in the native population of Chukotka, Russia. *Journal of Rheumatology*, **21**, 2298–300.

Amor, B., Silva Santos, R., Nahal, R., Listrat, V., and Dougados, M. (1994). Predictive factors for the longterm outcome of spondylarthropathies. *Journal of Rheumatology*, **21**, 1883–87.

Armas, J., Stevenson, F.K., Spellerberg, M.B., and Calin, A. (1990). Release of cell surface MHC class II antigen into synovial fluid: correlation with disease type and severity. *British Journal of Rheumatology*, **29** (Suppl. l), 23.

Ball, J. (1971). Enthesopathy of rheumatoid and ankylosing spondylitis. *Annals of the Rheumatic Diseases*, **30**, 213.

Bennett, P.H. and Burch, T.A. (1968). *Population studies of rheumatic diseases*, pp. 456–7. Excerpta Medica Foundation, Amsterdam.

Boyer, G.S. *et al*. (1994). Prevalence of spondylarthropathies in Alaskan Eskimos. *Journal of Rheumatology*, **21**, 2292–7.

Braun, J. and Sieper, J. (1996) The sacroiliac joint in the spondyloarthropathies. *Current Opinion in Rheumatology*, **8**, 275–87.

Braun, J., Bollow, M., Eggens, U., Konig, H., Distler, A., and Sieper, J. (1994). Use of dynamic magnetic resonance imaging with fast imaging in the detection of early and advanced sacroiliitis in spondylarthropathy patients. *Arthritis and Rheumatism*, **37**, 1039–45.

Burney, R.O. *et al*. (1994). Analysis of the MCH Class II encoded components of the HLA Class I antigen processing pathways in ankylosing spondylitis. *Annals of the Rheumatic Diseases*, **53**, 58–60.

Bywaters, E.G.L. (1984). Pathology of the spondylarthropathies. In *Spondylarthropathies* (ed. A. Calin), pp. 43–68. Grune and Stratton, Orlando, FLA.

Calin, A. (1975). Renal glomerular function in ankylosing spondylitis. *Scandinavian Journal of Rheumatology*, **4**, 241–3.

Calin, A.(1980). HLA B27—to type or not to type. (Editorial). *Annals of Internal Medicine*, **92**, 208–11.

Calin, A. (1982). HLA B27 in 1982. Reappraisal of a clinical test. *Annals of Internal Medicine*, **96**, 114–15.

Calin, A. (1988). The history of the seronegative spondylarthropathies. In: *The antiquity of the erosive arthropathies*. ARC Conference Proceedings No. 5.

Calin, A. (1989). Ankylosing spondylitis. In: *Textbook of rheumatology*, (3rd edn) (ed. W.N. Kelley, E.D. Hards, S. Ruddy, and C.G. Sledge), pp. 1021–37. W.B. Saunders, Philadelphia.

Calin, A. (1994). Can we define the outcome of ankylosing spondylitis and the effect of physiotherapy management? (Editorial). *Journal of Rheumatology*, **21**, 184–5.

Calin, A. and Elswood, J. (1989*a*). Relative role of genetic and environmental factors in disease expression: sib-pair analysis in ankylosing spondylitis. *Arthritis and Rheumatism*, **32**, 77–81.

Calin, A. and Elswood, J. (1989*b*). The outcome of 138 total hip replacements and 12 revisions in ankylosing spondylitis: high success rate after a mean follow-up of 7.5 years. *Journal of Rheumatology*, **16**, 955–8.

Calin, A. and Elswood, J. (1989*c*). Retrospective case-control analysis of 376 irradiated patients with ankylosing spondylitis. *Journal of Rheumatology*, **16**, 1443–5.

Calin, A. and Elswood, J. (1990). A prospective nation-wide cross-sectional study of NSAID useage over 2 years and 1321 patients with ankylosing spondylitis. *Journal of Rheumatology*, **17**, 81–3.

Calin, A. and Fries, J.F. (1977). The clinical history as a screening test in ankylosing spondylitis. *Journal of the American Medical Association*, **237**, 2613–14.

Calin, A., Marder, A., Marks, S., and Burns, T. (1984). Familial aggregation of Reiter's syndrome and ankylosing spondylitis: a comparative study. *Journal of Rheumatology*, **11**, 672–7.

Calin, A., Elswood, J., Rigg, S., and Skevington, S.M. (1988). Ankylosing spondylitis—an analytical review of 1500 patients: the changing pattern of disease. *Journal of Rheumatology*, **15**, 1234–8.

Calin, A., Edmunds, L., and Kennedy, L.G. (1993*a*). Fatigue in ankylosing spondylitis—why is it ignored? *Journal of Rheumatology*, **20**, 991–5.

Calin, A., Kennedy, L.G., Edmunds, L., and Will, R. (1993 *b*). Familial vesus sporadic ankylosing spondylitis: two differing diseases. *Arthritis and Rheumatism*, **36**, 676–81.

Calin, A. *et al*. (1994). A new approach to defining functional ability in ankylosing spondylitis: the development of the Bath Ankylosing Spondylitis Functional Index. *Journal of Rheumatology*, **21**, 2281–5.

Courtbrown, W.M. and Doll, R. (1965). Mortality from cancer and other causes after radiotherapy for ankylosing spondylitis. *British Journal of Medicine*, **2**, 1327.

Cuellar, M.L. and Espinoza, L.R. (1996). Management of spondyloarthropathies. *Current Opinion in Rheumatology*, **8**, 288–95.

Dawkins, R.C. *et al*. (1981). Prevalence of ankylosing spondylitis and radiological abnormalities of the sacroiliac joints in HLA B27 positive individuals. *Journal of Rheumatology*, **8**, 1025.

Donnelly, S. *et al*. (1994). Bone mineral density and vertebral compression fracture rates in ankylosing spondylitis. *Annals of the Rheumatic Diseases*, **53**, 117–21.

Dougados, M., van der Linden, S., Amov, B., Calin, A., *et al*. (1995). Sulphasalazine in the treatment of spondylarthropathy: a randomized, multicentre, double-blind, placebo-controlled study. *Arthritis and Rheumatism*, **38**, 618–27.

Edmunds, L. and Calin, A. (1991). New light on uveitis in ankylosing spondylitis. *Journal of Rheumatology*, **18**, 50–52.

Emery, A.E. and Lawrence, J.S. (1967). Genetics of ankylosing spondylitis. *Journal of Medical Genetics*, **4**, 239.

Garrett, S., Jenkinson, T., Kennedy, L.G., Whitelock, H., Gaisford, P., and Calin, A. (1994). A new approach to defining disease status in ankylosing spondylitis: The Bath Ankylosing Spondylitis Disease Activity Index. *Journal of Rheumatology*, **21**, 2286–91.

Hidding, A. *et al*. (1994). Comparison between self-report measures and clinical observations of functional disability in ankylosing spondylitis, rheumatoid arthritis and fibromyalgia. *Journal of Rheumatology*, **21**, 818–23.

Inman, R.D. and Schfield, R.H. (1994). Etiopathogenesis of ankylosing spondylitis and reactive arthritis. *Current Opinion*, **6**, 360–70.

Jenkinson, T.R., Mallorie, P.A., Whitelock, H.C., Kennedy, L.G., Garrett, S.L., and Calin, A. (1994). Defining spinal mobility in ankylosing spondylitis (AS). The Bath AS Metrology Index. *Journal of Rheumatology*, **21**, 1694–8.

Kabasakal, Y., Garrett, S.L., and Calin, A. (1994). The epidemiology of spondylodiscitis in ankylosing spondylitis—a controlled study. *British Journal of Rheumatology*, **33** (51), 122.

Kabasakal, Y., Garrett, S.L., and Calin, A. (1995). Outcome of total hip replacement (THR) in ankylosing spondylitis (AS): a thirteen year follow-up study. *British Journal of Rheumatology*, **34** (51), 114.

Kabasakal, Y., Garrett, S.L., and Calin, A. (1996). The epidemiology of spondylodiscitis in ankylosing spondylitis—a controlled study. *British Journal of Rheumatology*, **35**, 660–3.

Kennedy, L.G. and Calin, A. (1994). Sex ratios and age at onset in probands and secondary cases in the spondylarthropathies. *Journal of Rheumatology*, **20**, 1900–4.

Kennedy, L.G., Will, R., and Calin, A. (1993). Sex ratio in the spondylarthopathies and its relationship to phenotypic expression, mode of inheritance and age at onset. *Journal of Rheumatology*, **20**, 1062–3.

Khan, M.A. (ed.) (1990). *Spine. Ankylosing spondylitis and related spondylarthropathies*, 4, No. 3. Hanley and Belfus, Philadelphia.

Khan, M.A. (1993). Pathogenesis of ankylosing spondylitis: recent advances. (Editorial) *Journal of Rheumatology*, **20**, 1273–7.

Khan, M.A. (1996). Spondylarthropathies. Editoral overview. *Current Opinion*, **8**, 267–8.

Koh, W.H., Pande, I., Jones, S., Samuels, A., and Calin, A. (1996). A placebo-controlled study of low dose amitriptyline in ankylosing spondylitis. *British Journal of Rheumatology*, **35** (Suppl. 1), 288 (Abstr.)

Lehtinen K. (1993). Mortality and causes of death in 398 patients admitted to hospital with ankylosing spondylitis. *Annals of the Rheumatic Diseases*, **52**, 174–7.

Lopez-Larrea, C., Gonzalez-Roces, S., and Alvarez, V. (1996). HLA-B27 structure, function, and disease association. *Current Opinion in Rheumatology*, **8**, 296–308.

McEwen, C. *et al*. (1971). Ankylosing spondylitis and spondylitis accompanying ulcerative colitis, regional enteritis, psoriasis, and Reiter's disease. *Arthritis and Rheumatism*, **14**, 291.

Martinez-Gonzalez, O., Cantero-Hinojosa, J., Paule-Sastre, P., Gomez-Magan, J.C., and Salvatierra-Rios D. (1994). Intestinal permeability in patients with ankylosing spondylitis and their healthy relatives. *British Journal of Rheumatology*, **33**, 644–7.

Mielants, H., Veys, E.M., Goemaere, S., Covelier, C., and DeVos, M. (1993 *a*). A prospective study of patients with spondylarthropathy with special reference to HLA-B27 and to gut histology. *Journal of Rheumatology*, **20**, 1353–8.

Mielants, H. *et al*. (1993*b*). Gut inflammation in children with late onset of pauciarticular juvenile chronic arthritis and evolution to adult spondyloarthropathy: a prospective study. *Journal of Rheumatology*, **20**, 1567–72.

Mijiyawa, M. (1993). Spondylarthropathies in patients attending the Rheumatology Unit of Lomé Hospital. *Journal of Rheumatology*, **20**, 1167–9.

Mullaji, A.B., Upadhay, S.S., and Ho, E.K.W. (1994). Bone mineral density in ankylosing spondylitis. *Journal of Bone and Joint Surgery*, **76B**, 660–72.

Nasution, A.R., Mardjuadi, A., Suryadhana, N.G., Daud, R., and Muslichan, S. (1993). Higher relative risk of spondyloarthropathies among B27 positive Indonesian Chinese than native Indonesians. *Journal of Rheumatology*, **20**, 988–90.

O'Brien, B.J., Elswood, J., and Calin, A. (1990 *a*). Perception of prescription drug risks: a survey of patients with ankylosing spondylitis. *Journal of Rheumatology*, **17**, 503–57.

O'Brien, B., Elswood, J., and Calin, A. (1990 *b*). Willingness to accept risk in the treatment of rheumatic disease. *Journal of Epidemiology and Community Health*, **44**, 249–52.

Pande, I., Mackay, K., Chatfield, K., and Calin, A. (1995). The Bath Ankylosing Spondylitis Radiology Index (BASRI): a new validated approach to disease assessment. *British Journal of Rheumatology*, **34** (Suppl. 2), 37 (Abstr.).

Radford, E.P., Doll, R., and Smith, P.G. (1977). Mortality among patients with ankylosing spondylitis not given X-ray therapy. *New England Journal of Medicine*, **297**, 572.

Reveille, J.D. *et al*. (1994). HLA in ankylosing spondylitis—is HLA B27 the only MHC gene involved in disease pathogenesis? *Seminars in Arthritis and Rheumatism*, **23**, 295–305.

Rubin, L.A. *et al*. (1994). Investigating the genetic basis for ankylosing spondylitis—linkage studies with the major histocompatibility complex region. *Arthritis and Rheumatism*, **37**, 1212–16.

Scofield, R.H. (1996). Etiopathogenesis and biochemical and immunologic evaluation of spondyloarthropathies. *Current Opinion in Rheumatology*, **8**, 309–15.

Strecher R,M. (1957). Heredity in joint disease. *Documenta Rheumatica* (Geigy), **12**.

Toivanen, A. and Toivanen, P. (1996). Reactive arthritis. *Current Opinion*, **8**, 334–40.

Tutuncu, Z.N., Bilgic, A., Kennedy, L.G., and Calin, A. (1994). Interleukin-6. Acute phase reactants in clinical status and ankylosing spondylitis. *Annals of the Rheumatic Diseases*, **53**, 425–6.

van der Linden, S. and van der Heijde, M.F.M. (1996). Current and epidemiologic aspects of ankkylosing spondylitis and spondyloarthropathies. *Current Opinion in Rheumatology*, **8**, 269–74.

Will, R., Palmer, R., Bhalla, A.K., Ring, F., and Calin A. (1989). Marked osteoporosis is present in early ankylosing spondylitis and may be a primary pathological event. *Lancet*, **ii**, 1483–4.

Will, R. *et al*. (1990*a*). Is there sexual inequality in ankylosing spondylitis? The study of 498 women and 1201 men. *Journal of Rheumatology*, **17**, 1649–52.

Will, R., Amor, B., and Calin, A. (1990 *b*). The changing epidemiology of rheumatic diseases: should ankylosing spondylitis now be included? *British Journal of Rheumatology*, **29**, 299–300.

Will, R., Palmer, R., Bhalla, A.K., Ring, F., and Calin, A. (1990 *c*). Marked osteoporosis is present in early ankylosing spondylitis and progresses in late disease. *Proceedings 2nd Bath Conference on Osteoporosis and Bone Mineral Measurement*, 25–27 June.

Will, R., Calin, A., and Kirwan, J. (1992). Increasing age at presentation for patients with ankylosing spondylitis. *Annals of the Rheumatic Diseases*, **52**, 340–42.

5.5.4 Psoriatic arthritis

Dafna D. Gladman

Introduction

Psoriatic arthritis is an inflammatory arthritis, associated with psoriasis (Wright and Moll 1976). Its original definition as seronegative for rheumatoid factor, has been replaced by 'usually seronegative' since as many as 15 per cent of the general population, particularly over age 60, may have a positive rheumatoid factor, and rheumatoid factor may be present in more than 10 per cent of patients with psoriasis who do not have arthritis (Gladman *et al*. 1986). The majority of patients with psoriatic arthritis run a benign course. However, in about a fifth of the patients a chronic, progressive, deforming arthritis may develop, resulting in significant joint destruction and limitation of daily activities.

Epidemiology

Psoriasis is a chronic skin condition which affects 1 to 3 per cent of the population (Farber and Scott 1979). The association between psoriasis and arthritis might be fortuitous. Since psoriasis is a common condition, and arthritis, particularly osteoarthritis, is quite prevalent, it is conceivable that psoriasis and some unrelated form of arthritis may occur in the same patient. Indeed, some patients with psoriasis do present with a coincidental rheumatoid arthritis, or osteoarthritis. Cats (1990) has argued that psoriasis is just a measure of disease expression in certain patients with peripheral arthritis and spondylarthropathy. However, epidemiological evidence described below supports the notion that psoriatic arthritis is a distinct form of arthritis associated with psoriasis.

Although the first description of arthritis associated with psoriasis was provided by Aliberti (Eccles and Wright 1985; O'Neill and Silman 1994), psoriatic arthritis was considered to be a variant of rheumatoid arthritis until forty years ago. Epidemiological studies over the past four decades have confirmed the association between psoriasis and arthritis. These studies have shown an increased frequency of arthritis among patients with psoriasis and an increased prevalence of psoriasis among patients with arthritis. Thus, 6 to 42 per cent of patients with psoriasis may have psoriatic arthritis (Table 1), while the prevalence of arthritis in the general population is about 3 per cent. Likewise, the prevalence of psoriasis among patients with seronegative arthritis is reported to be 20 per cent, while arthritis occurs in only 2 to 3 per cent of the general population (Eccles and Wright 1985; O'Neill and Silman 1994). Lawrence *et al*. (1989) recently estimated the prevalence of psoriatic arthritis in the United States to be 0.67 per cent, whereas the overall prevalence for rheumatoid arthritis was estimated at 1.2 per cent.

Authors (year)	Centre	Number of patients studied	Percentage with arthritis
Leczinsky (1948)	Sweden	534	7
Vilanova (1951)	Barcelona	214	25
Little *et al*. (1975)	Toronto	100	32
Leonard *et al*. (1978)	Rochester	77	39
Green *et al*. (1981)	Cape Town	61	42
Scarpa *et al*. (1984)	Naples	180	34
Stern (1985)	Boston	1285	20
Zanelli and Wilde (1992)	Winston-Salem	469	17
Falk and Vandbakk (1993)	Kautokeino	35	17
Barišić-Druško *et al*. (1994)	Osijek region	553	10

Table 1 The prevalence of psoriatic arthritis among patients with psoriasis

The discovery of the rheumatoid factor and its association with rheumatoid arthritis helped separate psoriatic arthritis as a distinct entity, since patients with arthritis and psoriasis tended to be seronegative. Radiographical features in psoriatic arthritis were found to be different from those of rheumatoid arthritis (Avila *et al*. 1960). A female preponderance was found in rheumatoid arthritis, whereas the gender ratio among patients with psoriatic arthritis was almost equal (Wright and Moll 1976; Eccles and Wright 1985; O'Neill and Silman 1994). Unlike patients with rheumatoid arthritis, patients with psoriatic arthritis may present with a spondylarthropathy. Psoriatic arthritis is therefore classified with the seronegative spondylarthropathies. The studies by Wright and Moll (1976) are notable for presenting the unique features of psoriatic arthritis, and paving the way for other clinical descriptions.

The frequency of psoriatic arthritis has been reported in 6 to 42 per cent of patients with psoriasis (Leczinsky 1948; Little *et al*. 1975; Leonard *et al*. 1978; Green *et al*. 1981; Scarpa *et al*. 1984; Stern 1985; Zanelli and Wilde 1992; Falk and Vandbakk 1993; Barii-Druko *et al*. 1994). The most quoted prevalence of 6.8 per cent (Leczinsky 1948), is based on a study of the prevalence of rheumatoid-type polyarthritis in a population of inpatients with psoriasis. This frequency does not include other patterns of psoriatic arthritis. More recently, Scarpa *et al*. (1984) and Stern (1985) identified 34 per cent and 20 per cent of their psoriatic patients, respectively, to have psoriatic arthritis, while Green *et al*. (1981) reported arthritis in 42 per cent of their outpatients with psoriasis. A recent study of the prevalence of psoriasis in the Lapp population (Falk and Vandakk 1993) identified 17 per cent of patients with psoriasis who had psoriatic arthritis, while a study from Croatia (Barii-Druko *et al*. 1994) found that 9.8 per cent of patients with psoriasis had psoriatic arthritis. Since psoriasis may affect 1 to 3 per cent of the population, and as many as 30 per cent of psoriatic patients may develop psoriatic arthritis, almost 1 per cent of the population may suffer from psoriatic arthritis, which is the expected prevalence of rheumatoid arthritis. This is indeed close to the estimated prevalence of 0.67 per cent reported for psoriatic arthritis in the United States (Lawrence *et al*. 1989). Little *et al*. (1975) and Leonard *et al*. (1978) suggested that psoriatic arthritis was more common in patients with severe psoriasis. Both groups of investigators reported frequencies of 30 per cent of psoriatic arthritis among patients whose psoriasis required admission to hospital. However, psoriatic arthritis may precede the diagnosis of psoriasis in about 15 per cent of the patients (Wright and Moll 1976; Kammer *et al*. 1979; Gladman *et al*. 1987; Jones *et al*. 1994) (Table 2). Moreover, the highest prevalence of psoriatic arthritis was recorded among patients attending an outpatient dermatology clinic in Cape Town (Green *et al*. 1981).

Features	Roberts *et al.* (1976)	Kammer *et al.* (1979)	Scarpa *et al.* (1984)	Gladman *et al.* (1987)	Veale (1994)	Jones (1994)
	(N=166)*	(N=100)	(N=42)	(N=200)	(N=100)	(N=136)
Male/female ratio	67:101	47:53	29:33	104:116	59:41	43:57
Age of onset (years)	30–45	33–45	43–60	37	34	37.6
Asymmetrical oligoarthritis (%)	33	64†	16	51†	43	28†
Symmetrical polyarthritis (%)	70	23†	39.0	19†	33	53
Distal (%)	17	?	7.5	10	16	1
Back (%)	5	21	21.0	2†	4	6
Mutilans (%)	5	?	2.5	16	2	4
Sacroiliitis (%)	?	?	16.0	27	15	6
Joints before skin (%)	10	30	?	57	?	18

*Number of patients in the series.
†Includes patients with only distal joints involved.
‡1–4 including symmetrical oligoarthritis.
§5–10 including asymmetrical polyarthritis.
¶20 including peripheral joint plus back involvement.
? data not/unclear.
‡ unspecified.

Table 2 Clinical features of psoriatic arthritis in large series

Clinical features

Psoriatic arthritis affects women and men almost equally, usually in their third or fourth decade (Wright 1956; Kammer *et al*. 1979; Green *et al*. 1981; Scarpa *et al*. 1984; Gladman *et al*. 1987). Nail lesions proved to be the only clinical feature which may identify patients with psoriasis destined to develop arthritis (Gladman *et al*. 1986). These lesions occur in close to 90 per cent of patients with psoriatic arthritis (Little *et al*. 1975; Wright and Moll 1976; Kammer *et al*. 1979; Green *et al*. 1981; Gladman *et al*. 1987) and in 46 per cent of patients with psoriasis uncomplicated by arthritis (Gladman *et al*. 1986).

The arthritis is inflammatory in nature. It may affect any peripheral joint, as well as the axial skeleton and the sacroiliac joints. Patients usually present with pain, associated with stiffness, which is more marked in the morning and improves with activity. More than half the patients complain of morning stiffness of more than 30-min duration (Gladman *et al*. 1987). Evidence of inflammation may be detected clinically by the presence of stress pain or tenderness, as well as effusions (Gladman *et al*. 1990a), although these signs may not be as easily detectable as they are in rheumatoid arthritis, since patients with psoriatic arthritis are less tender than patients with rheumatoid arthritis (Buskila *et al*. 1992). The inflamed joints in patients with psoriatic arthritis may have a purplish-red discoloration, a feature which is not often seen in rheumatoid arthritis. The effusions in psoriatic arthritis joints tend to be tense, and are often difficult to detect. There is no predilection to particular joints, with the exception of the distal interphalangeal joints. Roberts *et al*. (1976) suggested that the knees were more commonly involved, as well as the proximal interphalangeal and metacarpophalangeal joints in the three groups they described. Jones *et al*. (1994) found that women tended to have small joints and upper limb involvement, whereas men tended to have axial involvement. Features which appear to differentiate psoriatic arthritis from rheumatoid arthritis clinically are shown in Table 3.

	Psoriatic arthritis	Rheumatoid arthritis
Female preponderance	Uncommon	Common
Distal interphalangeal involvement	Common	Uncommon
Symmetry	Less common	Common
Erythema over affected joint	Common	Uncommon
Back involvement	Common	Uncommon
Enthesopathy	Common	Uncommon
Skin lesions	Common	Uncommon
Nail lesions	Common	Uncommon
Rheumatoid factor	Uncommon	Common
Osteopenia	Uncommon	Common
Osteolysis	Common	Uncommon
Ankylosis	Common	Uncommon

Table 3 Comparison between psoriatic arthritis and rheumatoid arthritis

The spondylarthropathy may present with an inflammatory type of back pain, which is associated with stiffness and improves with activity. Clinical evidence of sacroiliitis may be obtained by specific tests, including the Gaenzlen's manoeuvre, the FABER (flexion, abduction, external rotation of the hip) test, and direct pressure over the sacroiliac joints (Gladman *et al*. 1987; Hanly *et al*. 1988). In some patients restricted range of back movements may be documented, by a reduction of flexion–extension as well as lateral flexion and rotation (Gladman *et al*. 1987; Hanly *et al*. 1988). Unlike ankylosing spondylitis, many of the patients with psoriatic spondylarthropathy are asymptomatic, and demonstrate a full range of back movement (Gladman *et al*. 1987; Hanly *et al*. 1988; Gladman *et al*. 1992b; Gladman *et al*. 1993). In these patients the diagnosis of the spondylarthropathy is made radiographically.

Clinical spectrum of psoriatic arthritis

Wright and Moll (1976) presented the seminal work on the clinical patterns of psoriatic arthritis. They proposed that these include: distal arthritis, involving the distal interphalangeal joints (Fig. 1); an asymmetric oligoarthritis involving small or medium-sized joints in an asymmetric distribution (Fig. 2); a symmetric polyarthritis, indistinguishable from rheumatoid arthritis; arthritis mutilans, which is a deforming, destructive, and disabling form of arthritis (Fig. 3 and Fig. 4); and a spondylarthropathy (Fig. 5 and Fig. 6). Similar descriptions have been reported by others (Kammer *et al*. 1979; Scarpa *et al*. 1984; Gladman *et al*. 1987; Helliwell *et al*. 1991; Jones *et al*. 1994; Veale *et al*. 1994b). The frequency of the various patterns has varied in the literature (Table 2). Although in their initial description of the psoriatic arthritis patterns Wright and Moll (1976) suggested that the most common pattern was asymmetric oligoarthritis, more recent studies confirm that polyarthritis is most commonly seen among patients with psoriatic arthritis. Psoriatic arthritis is asymmetric in distribution in about half the cases, even when polyarticular distribution is noted. Thus, in comparison with rheumatoid arthritis, psoriatic arthritis tends to be characterized as an asymmetric form of arthritis. The exact prevalence of each of the psoriatic arthritis patterns has been difficult to establish, since investigators have not used the exact same definitions, particularly with regards to symmetry. Helliwell *et al*. (1991) suggested a method for defining symmetrical involvement in patients with psoriatic arthritis, such that for each level of joints if the ratio of the number of matched pairs to the total number of joints was more than 0.5 then the distribution was considered symmetrical. Jones *et al*. (1994) showed that using this method more patients were found to have a symmetrical arthritis. Although the distal pattern has been described as typical for psoriatic arthritis, its frequency has varied widely, and some investigators have not been able to identify patients with isolated distal joint involvement. It has also been recognized that the patterns themselves may change with time in individual patients (Gladman 1992). A patient may present initially with an oligoarthritis which later becomes polyarticular, or develop an initial polyarthritis, which persists in only a few joints. Indeed, Jones *et al*. (1994) recently documented these changes in pattern over time in over 60 per cent of their patients with psoriatic arthritis. Moreover, unless radiographs are performed on all patients, joints which had been previously involved may not be identified, and the spondylarthropathy may be missed (Little *et al*. 1975; Gladman *et al*. 1987; Hanly *et al*. 1988). While the patterns of psoriatic arthritis may facilitate the diagnosis, it is not clear whether they have a prognostic significance.

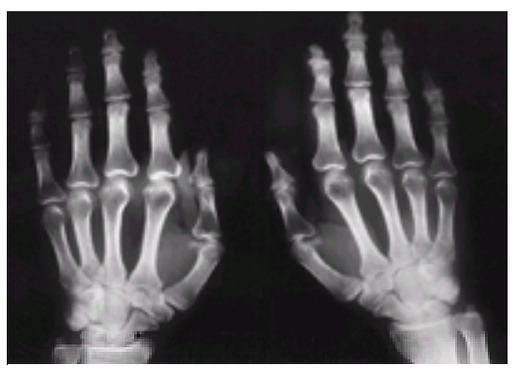


Fig. 1 Distal arthritis, involving the distal interphalangeal joints, with erosions and joint space narrowing.

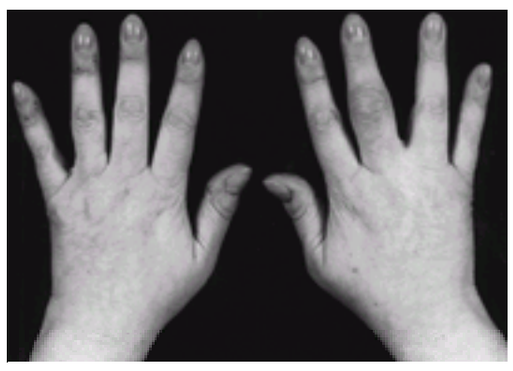


Fig. 2 An asymmetric oligoarthritis involving the third proximal interphalangeal joint on the right. Note the psoriatic lesions in the periungual areas of the left fourth and fifth fingers.

Fig. 3 Arthritis mutilans, which is a deforming, destructive, and disabling form of arthritis, showing the inability to use the hands fully.

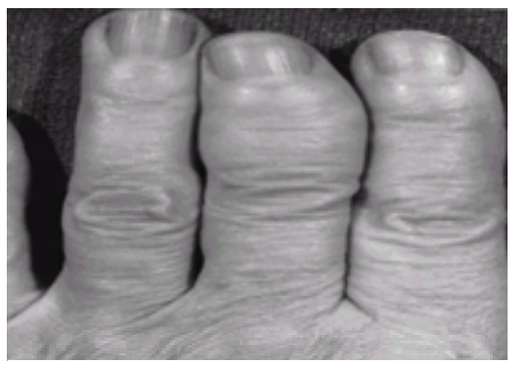


Fig. 4 Telescoping of the third distal interphalangeal joint, seen in patients with psoriatic arthritis, which may be part of arthritis mutilans.

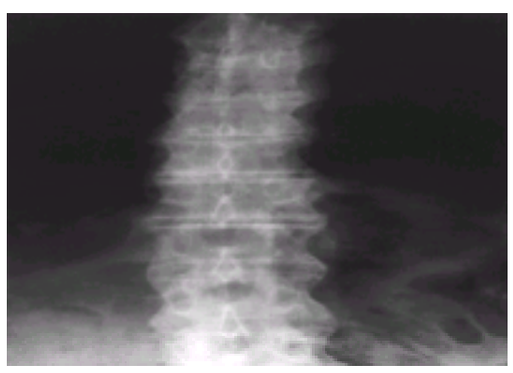


Fig. 5 Thoracolumbar spine in a patient with psoriatic spondylarthropathy, demonstrating syndesmophytes.

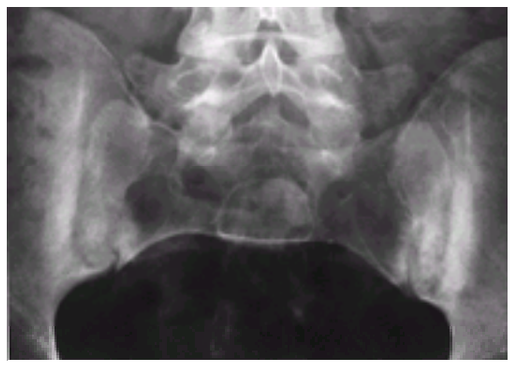


Fig. 6 Bilateral sacroiliitis in a patient with psoriatic spondylarthropathy.

A typical feature of psoriatic arthritis is the development of dactylitis, which presents as a swelling of a whole digit, with inflammation involving distal and proximal interphalangeal, and occasionally the metacarpophalangeal joints (Fig. 7). Dactylitis occurs in over a third of the patients (Gladman *et al*. 1987). The exact pathogenesis of the dactylitis is unclear. It may be related either to extensive inflammation and effusion in all the joints of a particular digit, with an associated tenosynovitis, or to soft tissue inflammation in the whole digit. The use of more advanced imaging techniques may help delineate the pathogenesis of 'sausage digits'.

Indeed, a recent scintigraphy study using human immunoglobulin labelled with technetium-99m demonstrates the inflammation in the digit which is missed by a bone scan (Stoeger *et al*. 1994). Tenosynovitis by itself is also a feature of psoriatic arthritis. Although inflammation of the extensor carpi ulnaris has been considered typical for rheumatoid arthritis, we have seen it quite often among the patients attending the psoriatic arthritis clinic. As in the other spondylarthropathies, such as Reiter's syndrome and ankylosing spondylitis, Achilles tendinitis, heel pain, and plantar fascitis are common among patients with psoriatic arthritis. Enthesitis, or inflammation at sites of tendon insertion, is frequent, particularly at the Achilles tendon, the insertion of the plantar fascia, and ligamentous insertions around the pelvic bones. These are commonly diagnosed radiographically as spurs (Fig. 8). A recent study from Finland (Lehtinen *et al*. 1994) offers an improved method, using ultrasound, to identify the presence of enthesitis in patients with spondylarthropathy.

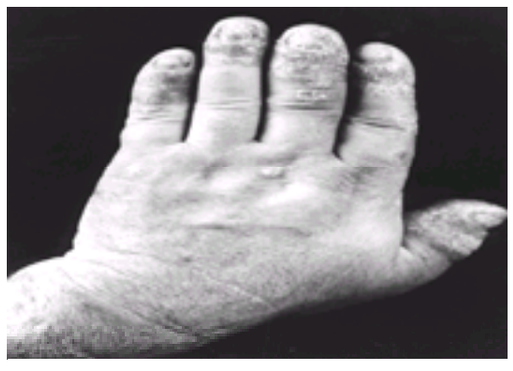


Fig. 7 Dactylitis, which presents as swelling of a whole digit, with inflammation of the distal and proximal interphalangeal, and occasionally the metacarpophalangeal joints, involving the thumb and third finger. Note the psoriatic skin and nail lesions.

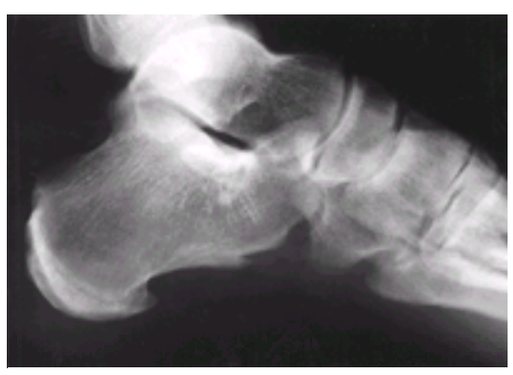


Fig. 8 Spur formation at the insertion of the plantar fascia, representing enthesitis.

The spondylarthropathy of psoriatic arthritis

The frequency of spinal involvement in psoriatic arthritis has varied from 2 per cent, as isolated back disease, to as high as 40 per cent, when associated with peripheral arthritis (Wright and Moll 1976; Lambert and Wright 1977; Kammer *et al*. 1979; Scarpa *et al*. 1984; Gladman *et al*. 1986; Gladman *et al*. 1987; Hanly *et al*. 1988; Moll 1994). Lambert and Wright (1977), found that 40 per cent of 130 patients with psoriatic arthritis had back involvement, based on back pain and reduced spinal mobility. Gladman *et al*. (1987), documented spinal involvement, based on both clinical and radiographical evidence, in 35 per cent of their patients at their first visit to the psoriatic arthritis clinic. This number increased to 51 per cent at follow-up (Gladman *et al*. 1992b). In both studies patients with spinal involvement tended to be male and older than patients without back involvement. Hanly *et al*. (1988) described 52 patients with psoriatic spondylarthropathy who had been followed for a minimum of 30 months, and for a mean of almost 5 years. Despite clinical and radiographical evidence of progression of sacroiliitis and spondylitis, patients' symptoms tended to improve with time, and their spinal mobility was generally good. They suggested that patients with psoriatic arthritis who have spinal involvement are not as symptomatic from their back disease as are patients with ankylosing spondylitis, and do not demonstrate the same degree of limitation of back movement that is noted in patients with ankylosing spondylitis (Hanly *et al*. 1988; Scarpa *et al*. 1988). The spondylarthropathy of patients with psoriatic arthritis was indeed found to be less severe than that seen in ankylosing spondylitis. This was evidenced by the lower frequency of symptomatic neck and back disease, as well as less limitation of movement and grade 4 sacroiliitis in patients with psoriatic arthritis compared with those with ankylosing spondylitis (Gladman *et al*. 1993). Moreover, among patients with psoriatic spondylarthropathy, there are gender-related differences in disease expression, with more advanced spondylarthropathy noted among men (Gladman *et al*. 1992b). The cervical spine in psoriatic arthritis received special attention in two recent studies. Salvarani *et al*. (1992a) studied 57 patients with psoriatic arthritis, of whom 70 per cent had radiographical evidence of cervical spine disease. They identified a high prevalence (23 per cent) of atlantoaxial subluxation. Jenkinson *et al*. (1994) detected cervical spine disease in 57 per cent of their patients with psoriatic arthritis, of whom only 3 had atlantoaxial subluxation. The majority of their patients had spondylitic type changes with apophysial joint narrowing or fusion and syndesmophytes. In both these studies neck involvement was related to prolonged disease duration. At the University of Toronto Psoriatic Arthritis Clinic, only 4 patients with atlantoaxial subluxation have been identified of a total of 450 patients followed over the past 17 years.

Extra-articular features

Skin psoriasis

The skin lesions of psoriasis consist of an erythematous scaly area that varies from a localized plaque on the elbows and knees to an incapacitating, generalized skin involvement with significant effect on the cardiovascular and heat regulating mechanism (Farber and Scott 1979; Goodfield 1994). The skin lesions are classified as: typical psoriasis vulgaris, with major involvement of the extensor surfaces; inverse psoriasis, affecting the flexural areas; pustular psoriasis, which may be localized to the palms and soles, or may be of the more generalized serious form called Von Zambush, and which may pose a threat to life; and the erythrodermic generalized group. The majority of patients with psoriatic arthritis demonstrate the classic psoriasis vulgaris pattern (Wright *et al*. 1979; Gladman *et al*. 1986; Gladman *et al*. 1987). Only 35 per cent of the patients describe a relationship between their skin and joint manifestations (Gladman *et al*. 1987). All areas of the skin may be affected, including the mucosa and the nails. Nail lesions include pitting, ridging, and onycholysis (Wright and Moll 1976). Two or all of these features in the same patient are in favour of a psoriatic origin for the nail dystrophy (Eastmond and Wright 1979). As already mentioned, nail lesions are particularly common among patients with psoriatic arthritis. Occasionally, distal interphalangeal joint disease may follow the development of onycholysis in a particular finger nail, but overall, there has not been a direct correlation between the presence of distal interphalangeal joint disease and nail lesions, since the former occurs in 35 per cent of the patients, whereas the latter occur in almost 90 per cent of the patients with psoriatic arthritis.

Other extra-articular features

The extradermal extra-articular features of psoriatic arthritis are similar to the features described in other seronegative spondylarthropathies and include iritis, which may occur in 7 per cent of the patients (Gladman *et al*. 1987), mouth ulcers, urethritis, colitis, and aortic valve disease (Wright and Moll 1976). A case of a patient with psoriatic arthritis with pyoderma gangrenosum was recently described (Smith and White 1994), and we have seen a case in our psoriatic arthritis clinic. The development of lymphoedema of the upper limb in patients with psoriatic arthritis was recently described in the literature (Mulherin *et al*. 1993). Similar cases have been seen at the University of Toronto Psoriatic Arthritis Clinic. The mechanism of this presentation is not entirely clear, but may represent a similar change to that occurring in dactylitis, with inflammation in both joints and tendons, leading to the clinical picture of 'lymphoedema'. Alternatively, in some of these cases it is possible that a ruptured capsule, as seen in a Baker's cyst in the knee, occurs in the wrist or elbow joint, leading to the soft tissue swelling noted in the distal portion of the upper limb.

Laboratory investigations in psoriatic arthritis

There are no specific laboratory tests which are diagnostic for psoriatic arthritis. Anaemia occurred in 14 per cent of the patients presenting to the psoriatic arthritis clinic (Gladman *et al*. 1987), perhaps reflecting the chronic disease, or an iron deficiency anaemia, secondary to therapy with non-steroidal anti-inflammatory drugs. Indeed, the much higher frequency of anaemia at follow-up (Gladman *et al*. 1990b), was thought to represent the untoward effect of these drugs. Elevated white-cell counts and other acute phase reactants may also be present (Gladman *et al*. 1987). This may reflect the acute phase inflammatory reaction, and may be seen in patients with psoriasis without arthritis. Elevated rythrocyte sedimentation rates may be seen in more than 40 per cent of patients with psoriatic arthritis, and probably reflect both joint and skin inflammation (Gladman *et al*. 1986; Gladman *et al*. 1987). Hyperuricaemia is not uncommon among patients with psoriasis and arthritis (Gladman *et al*. 1986; Gladman *et al*. 1987), and is probably related to the high turnover of skin cells. However, since both psoriasis and gout may occur in young males, one must rule out the possibility that the arthritis is crystal induced, before making the diagnosis of psoriatic arthritis in a patient with psoriasis. On the other hand, the presence of an acute monoarthritis, even in the first metatarsophalangeal joint in the presence of psoriasis does not mean the patient has gout. In both these situations a careful search for negatively birefringent, uric acid crystals should be carried out on the fluid obtained by joint aspiration.

Patients with psoriatic arthritis are usually seronegative for rheumatoid factor. However, in each series of patients with psoriatic arthritis there are about 10 to 15 per cent of the patients who have a positive rheumatoid factor, albeit in a low titre. It should be noted that patients with psoriasis uncomplicated by arthritis demonstrated the same frequency of a positive rheumatoid factor, despite the fact that they were younger on average than the patients with psoriatic arthritis (Gladman *et al*. 1986). Antinuclear factor has also been demonstrated in the sera of patients with uncomplicated psoriasis and patients with psoriatic arthritis, in the same frequency (Gladman *et al*. 1986). Whether this antinuclear antibody reflects the presence of antibodies to stratum corneum antigens is unclear (Gladman 1985). The demonstration of hypergammaglobulinaemia in patients with psoriatic arthritis provides further evidence for immunological abnormalities in patients with this condition (Gladman 1985).

Radiographical features of psoriatic arthritis

Radiographical abnormalities may be seen in both peripheral joints and the axial skeleton in patients with psoriatic arthritis (Resnick and Niwayama 1981). The features commonly associated with psoriatic arthritis and which help differentiate it from rheumatoid arthritis include: absence of juxta-articular osteoporosis; the predilection for distal interphalangeal joints; 'whittling' (lysis) of terminal phalanges (Fig. 9); lack of symmetry; gross destruction of isolated joint; 'pencil-in-cup' appearance (Fig. 10); ankylosis (Fig. 10); fluffy periostitis (Fig. 11); and both classical and atypical spondylitis (Wright and Moll 1971; Resnick and Niwayama 1981; Fig. 5a, Fig. 6 and Fig. 7).

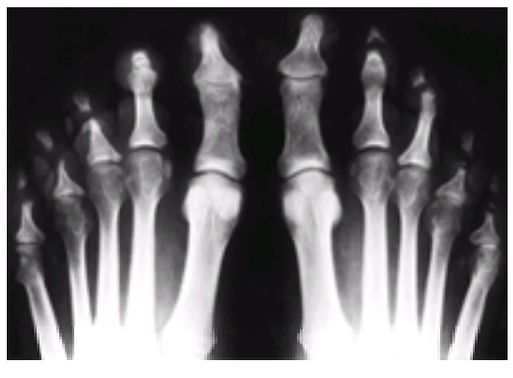


Fig. 9 'Whittling' (lysis) of terminal phalanges of the first and second toes bilaterally.

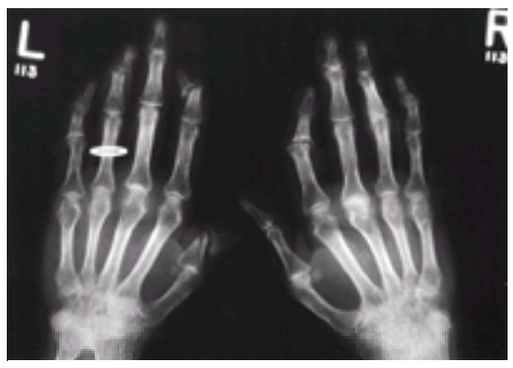


Fig. 10 'Pencil-in-cup' appearance seen in its early phase in the second right proximal, the fifth right distal, and the left fifth distal interphalangeal joints. Fully developed changes are seen in the left index distal and the left thumb interphalangeal joints. In addition, ankylosis is seen in the right distal interphalangeal joint

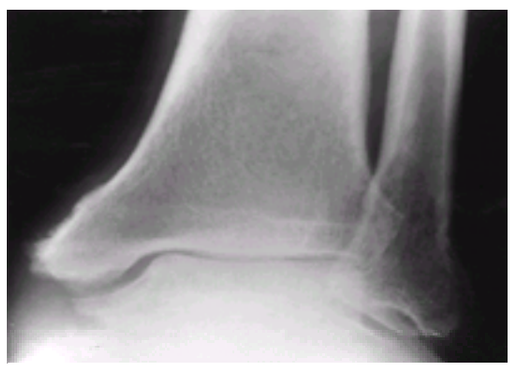


Fig. 11 Fluffy periostitis in the distal end of the tibia.

Diagnosis of psoriatic arthritis

There are no available diagnostic criteria for psoriatic arthritis (Gladman 1995). None the less, the diagnosis of psoriatic arthritis is generally based on the definition of the disease: an inflammatory arthritis in the presence of psoriatic skin lesions, usually seronegative for rheumatoid factor. In a patient with psoriasis, the development of an inflammatory arthritis makes the diagnosis easier. The clinical and radiographical features described above help identify the patient with psoriatic arthritis who had not previously demonstrated skin lesions. Thus, a patient who presents with an asymmetric oligoarthritis, or an inflammatory polyarthritis which includes distal interphalangeal joints, or peripheral arthritis with a spondylarthropathy, should be investigated for the presence of psoriasis, and psoriatic arthritis should clearly be considered in the differential diagnosis. The presence of dactylitis is certainly helpful, as is the presence of enthesitis. It should be noted that the skin lesions may be minimal, and indeed 'hidden'. One must therefore search for these lesions, particularly in the umbilical area, the anal cleft, the scalp, and the ears. Nail lesions are not always recognized by the patient, and should be looked for carefully. The common occurrence of distal joint involvement means that psoriatic arthritis needs to be differentiated from osteoarthritis. The distal interphalangeal lesions in patients with psoriatic arthritis are inflammatory in nature, and tend to be swollen, such that for

the most part they can be differentiated clinically as softer than the hard bony enlargement of Heberden's nodes. The presence of more proximal joint involvement, particularly the wrist and metacarpophalangeal joints, also helps distinguish psoriatic arthritis from osteoarthritis. However, osteoarthritis is a common condition, particularly with advancing age, and a patient may have Heberden's nodes complicating pre-existing psoriatic arthritis. Reiter's disease occasionally presents a diagnostic difficulty. The skin lesions in pustular psoriasis may be indistinguishable both clinically and pathologically from those of Reiter's syndrome, and the clinical features of the arthritis and the spondylarthropathy are similar. Psoriatic arthritis tends to be polyarticular, which may help. Iritis and mucous membrane lesions may be more common in Reiter's disease.

Pathogenesis of psoriatic arthritis

Understanding the pathogenetic mechanisms of disease is crucial for both the development of appropriate therapeutic approaches and the ultimate cure of a disease. It is likely that similar mechanisms would be operating in the development of both skin and joint disease in psoriatic arthritis. Although the exact pathogenetic mechanisms in psoriatic arthritis remain to be elucidated, factors thought to be important include genetic, immunological, and environmental (Gladman 1992; Abu-Shakra and Gladman 1994).

Genetic factors

More than 40 per cent of patients with psoriasis or psoriatic arthritis have a family history of the skin or joint disease in first-degree family members (Gladman *et al*. 1986; Gladman *et al*. 1987). Further support for genetic factors as possible aetiological mechanisms for psoriatic arthritis has come from the observations of high concordance for psoriasis in monozygotic twins and from clustering of both psoriasis and psoriatic arthritis within families (Espinoza 1985; Eastmond 1994). The discovery of the HLA system on chromosome 6 of humans, and the ability to detect HLA alloantigens of both class 1 and class 2, has allowed further elaboration of genetic mechanisms in psoriasis and psoriatic arthritis. Population studies in psoriasis revealed an increased frequency of HLA antigens B13, B17, B37, Cw6 and DR7 (Espinoza 1985; Gladman *et al*. 1986; Eastmond 1994). In psoriatic arthritis, increased frequencies of HLA-B13, -B17, -B27, -B38, -B39, -DR4, and -DR7 have been reported (Espinoza 1985; Gladman *et al*. 1986; Sakkas *et al*. 1990). Gladman *et al*. (1986) compared 158 patients with psoriatic arthritis to 101 patients with uncomplicated psoriasis. They found that the HLA-B7 or -B27 antigens were more common among patients with psoriatic arthritis, whereas B17, Cw6, and DR7 were more common among patients with uncomplicated psoriasis. HLA-B27 has clearly been associated with back disease in psoriatic arthritis, thus lending further credence to its grouping with the HLA-B27-associated spondylarthropathy. HLA-DR4 appears to be associated with the peripheral articular pattern of psoriatic arthritis (Gladman *et al*. 1986). A search for other genetic markers in psoriatic arthritis revealed that there were no specific T-cell receptor genes unique to the disease (Sakkas *et al*. 1990). However, Southern blot analysis using DNA probes for the immunoglobulin heavy chain gene (IgH) on chromosome 14q32 suggests that the gene may confer susceptibility to arthritis in patients with psoriasis (Sakkas *et al*. 1991). A recent identification of a gene for familial psoriasis on chromosome 17 is intriguing, but its relationship to psoriatic arthritis is unclear (Tomfohrde *et al*. 1994).

Immunological mechanisms

The clinical and pathological features of both psoriasis and psoriatic arthritis support the role of immunological factors in the pathogenesis of these conditions. The inflammatory nature of the disease, the cellular infiltrates seen both in skin and joint lesions, and the deposition of immunoglobulins in the epidermis as well as the synovial membrane, all support an immune mechanism (Gladman 1985; Panayi 1994). Autoantibodies, such as antinuclear antibodies, rheumatoid factor, and antibodies against skin antigens, as well as immune complexes, have been found in the sera of patients with psoriasis and psoriatic arthritis, supporting a hyperactive humoral immune mechanism (Gladman 1985). It has recently been hypothesized that antibodies against psoriasis-specific non-histone proteins found in patients with psoriasis, may facilitate displacement of the non-histone complex from DNA which may contain the psoriasis gene (Cormane and Asghar 1987), thus allowing for the development of the disease. This hypothesis is of particular interest since the identification of a psoriasis gene on chromosome 17.

On the cellular side, an imbalance of T-cell activity has been shown, which may result from either lack of T-cell suppression or excess of helper cell activity. Indeed, T-lymphocyte hyporeactivity to mitogens, impaired suppressor cell function, and a decrease in certain T-cell subpopulations have been demonstrated in psoriatic arthritis (Gladman 1985). The presence of activated T cells in psoriatic skin lesions was suggested by the demonstration of T lymphocytes bearing HLA-DR molecules as well as receptors for interleukin 2 in the lesions (Gottlieb 1988). In addition, keratinocytes bearing HLA-DR molecules were found in association with an increased frequency of psoriatic arthritis in patients with psoriasis. This prompted the proposal that these cells serve to present antigen to T cells, leading to mediator release and an inflammatory response (Gottlieb 1988). These T cells have indeed been found to express HLA-DR molecules, receptors for IL-2, and a variety of adhesion molecules, and to secrete proinflammatory cytokines, in particular IL-6 (Abu-Shakra and Gladman 1993). Fibroblasts from the skin and synovium of patients with psoriatic arthritis have an increased proliferative activity and the capability of secretion of increased amounts of IL-1b, IL-6 and platelet-derived growth factors. Activated T cells have been noted in the affected tissues (both skin and joints) in psoriatic arthritis by most investigators (Gladman 1993; Veale *et al*. 1993; Panayi 1994; Veale *et al*. 1994a). The results of several studies suggest that cytokines secreted from activated T cells and other mononuclear proinflammatory cells induce proliferation and activation of synovial and epidermal fibroblasts. It is of interest that both psoriasis and psoriatic arthritis, but not rheumatoid arthritis, have been reported to flare in the presence of the acquired immune deficiency virus (Buskila and Gladman 1990), suggesting that helper T (CD4) cells are not required for the disease process in psoriatic arthritis (Vasey *et al*. 1989). Indeed a decreased helper/suppressor T-cell ratio has been shown in patients with severe psoriasis (Rubins and Merson 1987). Moreover, deficient helper T-cell function in psoriatic patients was demonstrated using an antibody-specific induction system (Ventura *et al*. 1989). Panayi (1994) recently reviewed the immunological abnormalities seen in psoriatic arthritis. These include the presence of activated T lymphocytes in synovial membranes, an increase in macrophage numbers, and the presence of B cells. Although the presence of activated T cells in the affected tissues (both skin and joints) in psoriatic arthritis has been noted by most investigators (Gladman 1993; Veale *et al*. 1993; Panayi 1994; Veale *et al*. 1994a), the presence of an increased number of macrophages has not been uniformly described. This may be related to the stage in the disease at which the observations were made (Gladman 1993). It is still unclear whether the activated T cells are the cause of the arthritis, or the result of as yet unidentified factor.

Over the past 10 years the role of metabolites of arachidonic acid, such as prostaglandins and particularly leukotrienes, in the pathogenesis of both psoriasis and psoriatic arthritis has been evaluated (Voorhees 1983). Levels of leukotriene B4 have been shown to be increased in the psoriatic skin lesions, and injections of this compound has caused intraepidermal microabscesses. Drugs which lower the levels of leukotriene B4 are effective in controlling the skin lesions. However, the evidence is not conclusive. Greaves and Camp (1988) recently proposed an integrated approach to inflammation of human skin, considering the lipoxygenase system, platelet activating factor, and cytokines. This is an attractive proposal which allows for the integration of all the immunological abnormalities described. However, in a recent study, Veale *et al*. (1994c) demonstrated that Efamol marine was able to alter prostaglandin metabolism but did not produce a clinical improvement in either skin or joint manifestations in patients with psoriatic arthritis.

Environmental factors

Infection

Some investigators believe that guttate psoriasis is initiated by an infectious agent (Vasey 1985). Support for the role of bacterial antigens in the pathogenesis of psoriasis and psoriatic arthritis comes from indirect observations of enhanced humoral and cellular immunity to Gram-positive bacteria typically found in the psoriatic plaques. However, psoriatic plaques often become secondarily infected, thus the cause–effect relationship of bacteria and psoriasis is complicated. Moreover, Grinlinton *et al*. (1993) demonstrated that the response to streptococcal antigens by gd+ T cells from patients with psoriatic arthritis was also noted by cells from patients with rheumatoid arthritis, suggesting that this reaction may be a feature of inflammatory arthritis.

The exacerbation of psoriasis and psoriatic arthritis seen in the context of acquired immune deficiency virus infection is intriguing (Vasey *et al*. 1989). The possibility that psoriatic arthritis might be virus induced has recently been proposed by Luxembourg *et al*. (1987). However, an investigation of antigens related to the major internal protein p27 of a psoriasis-associated retrovirus-like particle failed to reveal specificity for psoriatic patients (Rødahl and Iversen 1985). Thus an exact viral aetiology for psoriasis or psoriatic arthritis has yet to be described.

Trauma

In almost all accounts of psoriatic arthritis there are reports of patients whose arthritis developed after trauma. However, the majority of these are anecdotal case reports (Langevitz *et al*. 1990). Scarpa *et al*. (1992) reviewed the records of 138 patients with psoriatic arthritis and compared the frequency of a traumatic event prior to the development of the arthritis to that recorded in 138 patients with rheumatoid arthritis. They found that trauma of some type preceded the diagnosis of psoriatic arthritis in 9 per cent of the cases, whereas in rheumatoid arthritis it was found in only 1 per cent of the patients, suggesting a role for trauma in some patients with

psoriatic arthritis.

Treatment of psoriatic arthritis

The treatment modalities employed in psoriatic arthritis are in part based on the pathogenetic mechanisms discussed above. They concentrate on control of inflammation, and an attempt to modify the immunological mechanisms thought to be operating in this disease (Gladman 1992).

The treatment of psoriasis

The treatment of psoriatic arthritis includes treatment for the skin condition as well as treatment for the joint disease. The skin lesions are treated by topical medications, aimed at controlling the inflammation and skin proliferation, including tar, anthralin, and corticosteroids (Marks 1980). In refractory cases, systemic medications such as methotrexate (Roenigk *et al*. 1969), **PUVA** (psoralen and ultraviolet A light) (Parish *et al*. 1974), retinoic acid derivatives (Klinkhoff *et al*. 1989), and more recently cyclosporin (Ellis *et al*. 1991) are used. The possible pathogenetic role of leukotrienes in the development of psoriatic arthritis has resulted in fish oil recently being introduced as a treatment for psoriasis (Gupta *et al*. 1989a).

The treatment of psoriatic arthritis

Non-steroidal anti-inflammatory therapy

The initial treatment for psoriatic arthritis consists of non-steroidal anti-inflammatory drugs (**NSAIDs**), including enteric-coated acetylsalicylic acid (**ECASA**), ibuprofen, naproxen, indomethacin, tolmetin, piroxicam, diclofenac sodium, and others (Abu-Shakra and Gladman 1993). In patients whose primary problem is that of peripheral arthritis, medications such as ECASA, ibuprofen, naproxen, or diclofenac sodium might be preferred. However, if there is evidence of a spondylarthropathy, it seems that indomethacin or tolmetin would be more appropriate. The latter two drugs would also be appropriate if morning stiffness is prolonged, and may be used in conjunction with other NSAIDs. Several of the NSAIDs have been incriminated in exacerbating the psoriasis (Abel *et al*. 1986), perhaps through the prostaglandin mechanism. It may therefore be necessary to change medications if an exacerbation of psoriasis occurs.

Disease-modifying drugs for psoriatic arthritis

If the arthritis persists despite the use of non-steroidal anti-inflammatory medications, then the next level of medications is embarked upon, that is, disease-modifying antirheumatic drugs.

Gold

Gold has been studied in a controlled fashion in psoriatic arthritis, using either intramuscular (Dowart *et al*. 1978) or oral (Carrett and Calin 1989) preparations. It has recently been shown that the intramuscular preparation is superior to the oral gold in patients with this condition (Palit *et al*. 1990). Moreover, although gold may control the inflammatory process in patients with psoriatic arthritis, it has not prevented progression of erosive disease over a 2-year period (Mader *et al*. 1995).

Penicillamine

Penicillamine has also been used successfully in psoriatic arthritis (Roux *et al*. 1979). Both gold and penicillamine are quite slow acting, however, requiring at least 6 months for a therapeutic effect. Therefore, other medications whose onset of action is faster have been tried.

Antimalarials

Although physicians have been reluctant to use antimalarials because of anecdotal reports of flares of psoriasis, and despite the lack of controlled trials, both chloroquine phosphate and hydroxychloroquine have been used (Kammer *et al*. 1979). Indeed, chloroquine has been shown to reduce disease activity in patients with psoriatic arthritis over a period of 6 months, and the frequency of psoriatic flares was no greater than that observed in the control group (Gladman *et al*. 1992a).

Methotrexate

Methotrexate, which has been found to be effective in controlling the skin psoriasis, has been used in psoriatic arthritis since 1964, when Black *et al*. (1964) performed a double-blind study of 21 patients using parenteral methotrexate. There have been two controlled trials of the use of low-dose weekly methotrexate in psoriatic arthritis. Willkens *et al*. (1984) demonstrated improvement in grip strength, morning stiffness, and joint count in patients with psoriatic arthritis, and physician global assessment was improved in the methotrexate group at 3 months, but only physician global assessment was significantly higher in the methotrexate group compared with the placebo. Zacharia and Zacharia (1987) found significant improvement in pain and functional scores as well as a decrease in the erthrocyte sedimentation rate during treatment with low-dose weekly methotrexate for psoriatic arthritis. Espinoza *et al*. (1992), in a retrospective uncontrolled study of 40 patients with psoriatic arthritis treated with a mean of dose of 11.2 mg/week of methotrexate during a mean period of 34 months, found that 37 per cent of the patients had an excellent response (no evidence of active synovitis) while 58 per cent had a good response (no more than four active joints and a decrease of at least 50 per cent in the number of those previously involved). Only two patients had discontinued the drug because of toxicity; one with leucopenia and the other stomatitis. Eleven patients developed liver test abnormalities. However, cirrhosis related to methotrexate was not noted. Abu-Shakra *et al*. (1995) found that while methotrexate reduced the actively inflamed joint count in patients with psoriatic arthritis, it did not prevent disease progression in these patients over a period of 2 years of treatment. None the less, methotrexate is used regularly for the treatment of psoriatic arthritis, particularly in the face of severe psoriasis. Methotrexate has an advantage over gold and penicillamine since it is effective within a few weeks. In addition, because it is given as an intermittent dose, once a week, patients prefer to take it rather than take other medications which are required daily, and often in repeated doses. Concerns about severe liver disease from methotrexate therapy, which resulted from reports in the late 1960s and early 1970s, seem to have been alleviated since more judicious use of intermittent dose has become commonplace.

Sulphasalazine

Sulphasalazine has been shown in two double-blind, placebo-controlled trials to be effective in psoriatic arthritis (Farr *et al*. 1990; Fraser *et al*. 1993). Rahman and Gladman (1995) recently demonstrated that sulphasalazine may not be tolerated by all patients with psoriatic arthritis, but in patients who are able to continue therapy the clinical response is impressive, and there is also improvement in the skin psoriasis.

Azathioprine

Azathioprine has also been used in psoriatic arthritis (Levy *et al*. 1972), but since its effect on psoriasis is not well recognized, it has not been as useful as it is in rheumatoid arthritis. We have recently used azathioprine in patients with psoriatic arthritis who had not responded to or were unable to tolerate methotrexate, with encouraging results.

Retinoids

Retinoids have only been studied in uncontrolled fashion, since it is difficult to blind both patients and observers to their side-effects (Klinkhoff *et al*. 1989). While these drugs may be effective against both skin and joint manifestations, their toxicity appears high. There is a new retinoid being tested at present which may prove less toxic.

Cyclosporin

Cyclosporin has recently been studied as a therapeutic option for both psoriasis and psoriatic arthritis (Gupta *et al*. 1989b; Salvarani *et al*. 1992b). Although it has been suggested that it is effective and safe, NSAIDs cannot be used concomitantly. Moreover, its adverse effects, particularly on the kidney, preclude its widespread use.

Steroids

Oral steroids are usually avoided in psoriatic arthritis, since on dose reduction they can cause significant flares of the skin psoriasis. However, intra-articular steroids may be used at any time, especially when there is a joint which is particularly inflamed. We tend to avoid injecting joints which are surrounded by psoriatic plaques because of fear of causing infections.

Dietary modification

Based on the pathogenetic mechanisms proposed for both psoriasis and psoriatic arthritis, there may be a role for fish oil preparations or specific immunomodulators in the treatment of both conditions. A pilot study of MaxEpa in psoriatic arthritis was promising (Peloso and Gladman 1992), but a more recent placebo-controlled trial of Efamol marine demonstrated no efficacy in psoriatic arthritis (Veale *et al*. 1994c). The use of oral vitamin D_3 for the treatment of psoriatic arthritis has recently been proposed (Huckins *et al*. 1990)

Other medications

Psoralen and ultraviolet A light has been used in psoriatic arthritis with some success (Perlman *et al*. 1979). More recently, the use of extracorporeal photochemotherapy has been proposed (de Misa *et al*. 1994). Buskila *et al*. (1991) described a woman with psoriatic arthritis, who experienced a remarkable improvement of her arthritis while she was taking bromocriptine for primary infertility due to hyperprolactinaemia. Others have also reported similar results (Abu-Shakra and Gladman 1993). Peptide T has also been advocated for the treatment of psoriatic arthritis, but information is currently available only from case reports (Abu-Shakra and Gladman 1993).

Physiotherapy and occupational therapeutic modalities should be used both for symptomatic relief and to avoid development of deformities. Patients may require splints and need to be instructed as to energy conservation and joint protection. In patients with spondylarthropathy, specific back exercises may be necessary.

The approach to the management of a patient with psoriatic arthritis should be to control the inflammatory process, in both the skin and the joints. If conservative measures such as topical medications and NSAIDs are not helpful, disease-suppressive medications should be used. In cases where both skin and joint disease are active, medications such as methotrexate and retinoids should be considered early. Sulphasalazine and azathioprine may also work in these situations. Where the arthritis is a problem, and skin lesions are well controlled with topical medications, then the other medications, including gold, chloroquine, penicillamine, and azathioprine might be used first. Newer modalities, such as cyclosporin, should be reserved for patients with particularly aggressive disease.

Monitoring disease activity during courses of NSAIDs and disease-modifying drugs should be based on the actively inflamed joint count and effusion count, which have been found to be reliable measures of disease activity (Gladman *et al*. 1990a) despite the fact that patients with psoriatic arthritis do not demonstrate the same degree of tenderness as do patients with rheumatoid arthritis (Buskila *et al*. 1992). It is important to emphasize that the pain scale has not been shown to correlate with measures of disease severity (Duffy *et al*. 1992; Blackmore *et al*. 1995).

Surgery

Surgery is reserved for patients whose joints have become deformed and damaged. There are no systematic studies of surgery in psoriatic arthritis. There is a concern that the joints do not recover function very well, and that there may be bone formation similar to what has been described in ankylosing spondylitis (Dwosh *et al*. 1976). Our own experience with knee and hip surgery has been good. However, surgical procedures performed on several of our patients in an attempt to correct flexion deformities of the fingers have not been rewarding. Patients developed recurrence or worsening of their deformities, and ended up with a less functional hand than prior to surgery.

Recently an approach to the treatment of temporomandibular joint disease in psoriatic arthritis has been described (Peterson and Shepherd 1992). Hicken *et al*. (1994) recently reviewed their experience with foot and ankle surgery in 17 patients with psoriatic arthritis collected over a 15-year period. There were 27 operations which included forefoot arthroplasty or arthrodesis and toe arthroplasty or arthrodesis. The operations were considered successful in 89 per cent of the cases. Complications occurred in a patient who had a local infection associated with a local flare of psoriasis and required corrective surgery, another patient who required an additional procedure, and a third patient who had delayed union.

The course and prognosis of psoriatic arthritis

The course of psoriatic arthritis is variable. There are patients who have few episodes and who recover completely, but in many the disease is persistent. A third of the patients relate the course of their skin disease and joint disease (Gladman *et al*. 1987). The lack of a systematic approach to disease assessment in psoriatic arthritis has hindered the performance of follow-up studies. Clinical measures of inflammatory activity as well as damage have only recently been validated in psoriatic arthritis (Gladman *et al*. 1990a). Roberts *et al*. (1976) concluded in their follow-up study that 'apart from the deforming group the arthritis was not notably progressive'. However, there was time lost from work in at least 60 per cent of all patients, and radiographical progression was recorded in about 15 per cent. Stern (1985) noted that more than 50 per cent of the patients with psoriatic arthritis had some limitation on their daily activities, and they were twice as likely to be unemployed as patients with other joint disease. Hanly *et al*. (1988) and Gladman *et al*. (1990b), suggested that the disease is progressive, based on the increased number of deformed and damaged joints observed in their patients, who were followed according to a standard protocol in the University of Toronto Psoriatic Arthritis Clinic. Coulton *et al*. (1989) reported on the outcome of 40 patients admitted to hospital for psoriatic arthritis who were followed for a mean of 8 years. At the end of that period, none of the patients died. However, 35 per cent of their patients were in Steinbrocker's classes III and IV, supporting the notion that a proportion of patients with psoriatic arthritis become disabled. Clear evidence of progression of deformities was demonstrable when patients were compared at presentation and at follow-up, based on duration of follow-up at the psoriatic arthritis clinic (Gladman 1994). None the less, the Health Assessment Questionnaire in patients with psoriatic arthritis does not give the same high scores seen in rheumatoid arthritis (Jones *et al*. 1994; Blackmore *et al*. 1995). This may very well be due to the fact that the questionnaire correlates highly with pain, which is less likely to be an issue for patients with psoriatic arthritis than for those with rheumatoid arthritis (Buskila *et al*. 1992).

It would be of benefit to be able to identify those patients who are destined to develop the more severe disease and treat them appropriately. In an attempt to identify such prognostic factors, Gladman *et al*. (1995) studied 305 patients who entered the psoriatic arthritis clinic with less than ten deformed joints, and identified clinical indicators for progression through four stages: no deformities, one to four deformities, five to nine deformities and ten or more deformities. Patients who had five or more effused joints at presentation to the clinic were more likely to progress, as were patients treated with disease-modifying drugs, while patients who had a low erythrocyte edimentation rate were less likely to develop more deformities during follow-up. There was no correlation with disease duration. Moreover, Gladman *et al*. (1995) recently demonstrated that the HLA antigens B27, B39, and DQw3 were more important than the clinical features in predicting such progression. Thus, prognostic factors for severe disease in psoriatic arthritis are identifiable and may serve as markers for specific therapeutic modalities.

Approach to the management of psoriatic arthritis

My overall approach to a patient with psoriatic arthritis includes confirming the diagnosis, assessing the extent of disease activity in terms of both joint and skin disease, and assessing the degree of damage that has occurred. This involves a careful history, physical examination, laboratory assessment, and radiographical evaluation. The goal of treatment is control of inflammation which will hopefully lead to control of symptoms, and prevention of deformities and damage, so that the patient may continue to lead an active and productive life.

Once the diagnosis has been made and the patient assessed, my approach to management begins with patient education. I explain to the patient that he/she suffers from an inflammatory arthritis, that the treatment is aimed at controlling inflammation, and therefore medications need to be taken regularly. The role of daily stresses on exacerbations of both skin and joint disease is reviewed. The need to treat both skin and joint manifestations of the disease is also discussed. These topics often need to be repeated during follow-up.

The actual therapeutic approach is then tailored to the individual patient (see Box 1). A patient who has mild disease, with minimal skin lesions and mild arthritis without deformities, is treated with topical ointments for the skin, and non-steroidal anti-inflammatory drugs for the joints. I tend to use enteric-coated acetylsalicylic

acid, ibuprofen, naproxen, or diclofenac for polyarticular disease, and those that complain of pain. In patients with oligoarticular disease, and those with spondylarthropathies, I tend to use indomethacin or tolmetin. While there is no scientific proof to support this approach, it seems to be empirically correct. However, since patients vary both in their response to and tolerance of different NSAIDs, this sequence is often changed. More for individuals who clearly express aversion to taking pills, I tend to choose those NSAIDs which can be given once daily.

I also use intra-articular corticosteroid injections for individual joints. I find that in psoriatic arthritis the inflammation is intense and deformity can ensue rapidly. My patients are educated to call immediately when a red, hot joint appears, and to present themselves for joint injections. Although some people feel that intra-articular steroids are not as effective in seronegative disease as they are in rheumatoid arthritis, this has not been my experience. We have been able to control severe inflammation and prevent damage in joints which we injected early (as judged by what happened to other joints in the same individual). This applies to both large and small joints in this disease. I use joint injections as an adjunct to systemic therapy as well, at any point in the disease, provided the joint is not completely destroyed.

In a patient with mild psoriasis but more active and severe arthritis, I tend to use second-line medications early. I use antimalarials, sulphasalazine, methotrexate, gold, and azathioprine, for this type of patient. If the patient demonstrates a spondylarthropathy, I would tend to choose sulphasalazine first, since it seems to control spinal disease as well as peripheral disease, whereas the other medications have not been shown to be as effective for spinal involvement.

In a patient who has severe psoriasis, even if the arthritis is not that severe, I tend to start with methrotrexate, since it has been shown to be effective for both components of the disease. If the methotrexate is not tolerated, then I switch to either sulphasalazine (unless there is sulpha allergy) or azathioprine. If the methotrexate is tolerated, but not completely effective, I add either an antimalarial or sulphasalazine. In patients with severe psoriasis and mild arthritis I have used PUVA, which works well for the skin and has worked well for the joints.

Box 1 Therapeutic approach to the management of psoriatic arthritis

Type of presentation	NSAIDs	Second line	Intra-articular
Mild skin Mild joint	Yes	No	As required
Mild skin Moderate joint	Yes	Sulphasalazine, methotrexate, antimalarials, gold, indomethacin	As required
Severe skin Mild joint	Yes	Methotrexate, sulphasalazine, cyclosporin, PUVA, retinoids	As required
Severe skin Severe joint	Yes	Methotrexate, sulphasalazine, indomethacin, PUVA, cyclosporin, retinoids	As required

I reserve the use of cyclosporin A and retinoids for patients with severe psoriasis and arthritis who either refuse to take methotrexate, or are unable to tolerate it. These drugs are more toxic than the others and need to be used with caution.

Chapter References

Abel, E.A., DiCicco, L.M., Orenberg, E.K., Fraki, J.E., and Farber, E.M. (1986). Drugs in exacerbation of psoriasis. *Journal of the American Academy of Dermatology*, **15**, 1007–22.

Abu-Shakra, M. and Gladman, D.D. (1993). Management of refractory psoriatic arthritis. *Rheumatology Reviews*, **2**, 201–6.

Abu-Shakra, M. and Gladman, D.D. (1994). Aetiopathogenesis of psoriatic arthritis. *Rheumatology Reviews*, **3**, 1–7.

Abu-Shakra, M., Gladman, D.D., Thorne, J.C., Long, J., Gough, J., and Farewell, V.T. (1995). Longterm methotrexate therapy in psoriatic arthritis: clinical and radiologic outcome. *Journal of Rheumatology*, 22, 241–5.

Avila, R., Pugh, D.G., Slocumb, C.H. *et al*. (1960). Psoriatic arthritis. A roentgenologic study. *Radiology*, **75**, 691–702.

Barii-Druko, V., Dobri, I., Pai, A, *et al*. (1994). Frequency of psoriatic arthritis in general population and among psoriatics in department of dermatology. *Acta Dermatologica et Venereologica*, **74** (Suppl. 186), 107–8.

Blackmore, M., Gladman D.D., Husted J., Long J., and Farewell V.T. (1995). Measuring health status in psoriatic arthritis. *Journal of Rheumatology*, 22, 886–92.

Black, R.L., O'Brien, W.M., Van Scott, E.J., Auerbach, R., Eisen, A.Z., and Bunim, J.J. (1964). Methotrexate therapy in psoriatic arthritis. Double blind study on 21 patients. *Journal of the American Medical Association*, **189**, 743–7.

Buskila, D., Sukenik, S., Holcberg, G., *et al*. (1991). Improvement of psoriatic arthritis in a patient treated with bromocriptine for hyperprolactinemia. *Journal of Rheumatology*, **18**, 611–12.

Buskila, D., Langevitz, P., Gladman, D.D., *et al*. (1992). Patients with rheumatoid arthritis are more tender than those with psoriatic arthritis. *Journal of Rheumatology*, **19**, 1115–19.

Carrett, S. and Calin, A. (1989). Evaluation of auranofin in psoriatic arthritis: a double blind placebo controlled trial. *Arthritis and Rheumatism*, **32**, 158–65.

Cats, A. (1990). Psoriasis and arthritis. *Cutis*, **46**, 323–9.

Cormane, R.D. and Asghar, S.S. (1987). Psoriasis specific chromosomal proteins, antibodies against them and disease activity. *Medical Hypotheses*, **22**, 369–72.

Coulton, B.L., Thomson, K., Symmons, D.P.M., and Popert, A.J. (1989). Outcome in patients hospitalised for psoriatic arthritis. *Clinical Rheumatology*, **2**, 261–5.

de Misa, R.F., Azafia, J.M., Harto, A, *et al*. (1994). Psoriatic arthritis: one year of treatment with extracorporeal photochemotherapy. *Journal of the American Academy of Dermatology*, **30**, 1037–8.

Dowart, B.B., Gall, E.P., Schumacher, H.R., and Krauser, R.E. (1978). Chrysotherapy in psoriatic arthritis: efficacy and toxicity compared to rheumatoid arthritis. *Arthritis and Rheumatism*, **21**, 513–15.

Duffy, C.M., Watanabe Duffy, K.N., Gladman, D.D., *et al*. (1992). Utilization of the Arthritis Impact Measurement Scales (AIMS) for patients with psoriatic arthritis. *Journal of Rheumatology*, **19**, 1727–32.

Dwosh, I.L., Resnick, D., and Becker, M.A. (1976). Hip involvement in ankylosing spondylitis. *Arthritis and Rheumatism*, **19**, 683–92.

Eastmond, C.J. (1994). Genetics and HLA antigens. *Baillière's Clinical Rheumatology*, **8**, 263–76.

Eastmond, C.J. and Wright, V. (1979). The nail dystrophy of psoriatic arthritis. *Annals of the Rheumatic Diseases*, **38**, 226–8.

Eccles, J.T. and Wright, V. (1985). The history and epidemiologic definition of psoriatic arthritis as a distinct entity. In *Psoriatic arthritis*, (ed. L.H. Gerber and L.R. Espinoza), pp. 1–8. Grune & Stratton, Orlando, FL.

Ellis, C.N., Fradin, M.S., Messana, J.M., *et al*., (1991). Cyclosporin for plaque-type psoriasis. Results of a multidose, double-blind trial. *New England Journal of Medicine*, **324**, 277–84.

Espinoza, L.R. (1985). Psoriatic arthritis: further epidemiologic and genetic considerations. In *Psoriatic arthritis*, (ed. L.H. Gerber and L.R. Espinoza), pp. 9–32. Grune & Stratton, Orlando, FL.

Espinoza, L.R., Zakraoni, L., Espinoza, C.G., *et al*. (1992). Psoriatic arthritis: clinical response and side effects of methotrexate therapy. *Journal of Rheumatology*, **19**, 872–7.

Falk, E.S. and Vandbakk, Ø. (1993). Prevalence of psoriasis in a Norwegian Lapp population. *Acta Dermatological et Venereologica*, **73** (Suppl.), 6–9.

Farber, E.M. and Scott, E.V. (1979). Psoriasis. In *Dermatology in general medicine*, (2nd edn) (ed. T.B. Fitzpatrick *et al*.), pp. 233–47, McGraw-Hill, New York.

Farr, M., Kitas, G.D., Waterhouse, L., Jubb, R., Felix-Davies, D., and Bacon, P.A. (1990). Sulphasalazine in psoriatic arthritis: a double-blind placebo-controlled study. *British Journal of Rheumatology*, **29**, 46–9.

Fraser, S.M., Hopkins, R., Hunter, J.A., *et al*. (1993). Sulphasalazine in the management of psoriatic arthritis. *British Journal of Rheumatology*, **32**, 923–5.

Gladman, D.D. (1985). Immunologic factors in the pathogenesis of psoriatic arthritis. In *Psoriatic arthritis*, (ed. L.H. Gerber and L.R. Espinoza), pp. 33–44. Grune & Stratton, Orlando, FL.

Gladman, D.D. (1992). Psoriatic arthritis: recent advances in pathogenesis and treatment. *Rheumatic Disease Clinics of North America*, **18**, 247–56.

Gladman, D.D. (1993). Toward unravelling the mystery of psoriatic arthritis. *Arthritis and Rheumatism*, **36**, 881–4.

Gladman, D.D. (1994). Natural history of psoriatic arthritis. *Baillière's Clinical Rheumatology*, **8**, 379–94.

Gladman, D.D. (1995) Classification criteria for psoriatic arthritis. *Baillièr*es Clinical Rheumatology 9, 319–29.

Gladman, D.D. and Farewell, V.T. (1995). The role of HLA antigens as indicators of disease progression in psoriatic arthritis (PSA): multivariate relative risk model. *Arthritis and Rheumatism*, 38, 845–50.

Gladman, D.D., Anhorn, K.A.B., Schachter, R.K., and Mervart, H. (1986). HLA antigens in psoriatic arthritis. *Journal of Rheumatology*, **13**, 586–92.

Gladman, D.D., Shuckett, R., Russell, M.L., Thorne, J.C., and Schachter, R.K. (1987). Psoriatic arthritis (PSA)—an analysis of 220 patients. *Quarterly Journal of Medicine*, **62**, 127–41.

Gladman, D.D., Farewell, V., Buskila, D., *et al*.(1990*a*). Reliability of measurements of active and damaged joints in psoriatic arthritis. *Journal of Rheumatology*, **17**, 62–4.

Gladman, D.D., Stafford-Brady, F., Chang, C.H., Lewandowski, K., and Russell, M.L. (1990 *b*). Clinical and radiological progression in psoriatic arthritis. *Journal of Rheumatology*, **17**, 809–12.

Gladman, D.D., Blake, R., Brubacher, B., and Farewell, V.T. (1992 *a*). Chloroquine therapy in psoriatic arthritis. *Journal of Rheumatology*, **19**, 1724–6.

Gladman, D.D., Brubacher, B., Buskila, D., Langevitz, P., and Farewell, V.T. (1992 *b*). Psoriatic spondyloarthropathy in men and women: a clinical, radiographic, and HLA study. *Clinical And Investigation Medicine*, **15**, 371–5.

Gladman, D.D., Brubacher, B., Buskila, D., Langevitz, P., and Farewell, V.T. (1993). Differences in the expression of spondyloarthropathy: a comparison between ankylosing spondylitis and psoriatic arthritis. Genetic and gender effects. *Clinical And Investigation Medicine*, **16**, 1–7.

Gladman, D.D., Farewell, V.T., and Nadeau, C. (1995). Clinical indicators of progression in psoriatic arthritis (PSA): multivariate relative risk model. *Journal of Rheumatology*, 22, 675–9.

Goodfield, M. (1994). Skin lesions in psoriasis. *Baillière's Clinical Rheumatology*, **8**, 295–316.

Gottlieb, A.B. (1988). Immunologic mechanisms in psoriasis. *Journal of the American Academy of Dermatology*, **18**, 1276–380.

Greaves, M.W. and Camp, R.D.R. (1988). Prostaglandins, leukotrienes, phospholipase, platelet activating factor, and cytokines: an integrated approach to inflammation of human skin. *Archives of Dermatology*, **280** (Suppl.), S33–S41.

Green, L., Meyers, O.L., Gordon, W., and Briggs, B. (1981). Arthritis in psoriasis. *Annals of the Rheumatic Diseases*, **40**, 366–9.

Grilington, F.M., Skinner, M.A., Birchall, N.M., and Tan, P.L.I. (1993). gd+ T cells from patients with psoriatic and rheumatoid arthritis respond to streptococcal antigen. *Journal of Rheumatology*, **20**, 983–7.

Gupta, A.K., Ellis, C.N., Telliner, D.C., Anderson, T.F., and Voorhees, J.J. (1989 *a*). Double-blind, placebo-controlled study to evaluate the efficacy of fish oil and low-dose UVB in the treatment of psoriasis. *British Journal of Dermatology*, **120**, 801–7.

Gupta, A.K., Matteson, E.I., Ellis, C.N., *et al*. (1989*b*). Cyclosporin in the treatment of psoriatic arthritis. *Archives of Dermatology*, **125**, 507–10.

Hanly, J.G., Russell, M.L., and Gladman, D.D. (1988). Psoriatic spondyloarthropathy: a long term prospective study. *Annals of the Rheumatic Diseases*, **47**, 386–93.

Helliwell, P., Marchesoni, A., Peters, M., Barker, M., and Wright, V. (1991). A re-evaluation of the osteoarticular manifestations of psoriasis. *British Journal of Rheumatology*, **30**, 339–45.

Hicken, G.J., Kitaoka, H.B., and Valente R.M. (1994). Foot and ankle surgery in patients with psoriasis. *Clinical Orthopedics and Related Research*, **300**, 204–6.

Huckins, D., Felson, D.T., and Holick, M. (1990). Treatment of psoriatic arthritis with oral 1,25-dihydroxyvitamin D_3: a pilot study. *Arthritis and Rheumatism*, **33**, 1723–7.

Jenkinson, T., Armas, J., Evison, G., *et al*. (1994). The cervical spine in psoriatic arthritis: a clinical and radiological study. *British Journal of Rheumatology*, **33**, 255–9.

Jones, S.M., Armas, J.B, Cohen, M.G., Lovell, C.R., Evison, G., and McHugh, N.J. (1994). Psoriatic arthritis: outcome of disease subsets and relationship of joint disease to nail and skin disease. *British Journal of Rheumatology*, **33**, 834–9.

Kammer, G.M., Soter, N.A., Gibson, D.J., and Schur, P.H. (1979). Psoriatic arthritis. A clinical, immunologic and HLA study of 100 patients. *Seminars in Arthritis and Rheumatism*, **9**, 75–97.

Klinkhoff, A.V., Gertner, E., Chalmers, A., *et al*. (1989). Pilot study of etretinate in psoriatic arthritis. *Journal of Rheumatology*, **16**, 789–91.

Lambert, J.B. and Wright, V. (1977). Psoriatic spondylitis: a clinical and radiological description of the spine in psoriatic arthritis. *Quarterly Journal of Medicine*, **46**, 411–25.

Langevitz P., Buskila D., and Gladman D. (1990). Psoriatic arthritis precipitated by physical trauma. *Journal of Rheumatology*, 17, 695–7.

Lawrence, R.C., Hochberg, M.C., Kelsey, J.L., *et al*. (1989). Estimates of the prevalence of selected arthritis and musculoskeletal diseases in the United States. *Journal of Rheumatology*, **16**, 427–41.

Leczinsky, C.G. (1948). The incidence of arthropathy in a ten-year series of psoriasis cases. *Acta Dermatologica et Venereologica*, **28**, 483–7.

Lehtinen, A., Traavisainen, M., and Leirisalo-Repo, M. (1994). Sonographic analysis of enthesopathy in the lower extremities of patients with spondyloarthropathy. *Clinical and Experimental Rheumatology*, **12**, 143–8.

Leonard, D.G., O'Duffy, J.D., and Rogers, R.S. (1978). Prospective analysis of psoriatic arthritis in patients hospitalized for psoriasis. *Mayo Clinic Proceedings*, **53**, 511–18.

Levy, J.J., Paulus, H.E., Barnett, E.V., *et al*. (1972). A double-blind controlled evaluation of azathioprine treatment in rheumatoid arthritis and psoriatic arthritis. *Arthritis and Rheumatism*, **15**, 116–17.

Little, H., Harvie, J.N., and Lester, R.S. (1975). Psoriatic arthritis in severe psoriasis. *Canadian Medical Association Journal*, **112**, 317–19.

Luxembourg, A., Cailla, H., Roux, H., and Roudier, J. (1987). Do viruses play an etiologic role in ankylosing spondylitis and psoriatic arthritis? *Clinical Immunology and Immunopathology*, **45**, 292–5.

Mader, R., Gladman, D.D., Long, J., *et al*. (1995). Injectable gold for the treatment of psoriatic arthritis—long term follow-up. *Clinical And Investigation Medicine*, **18**, 139–43.

Marks, J.M. (1980). Psoriasis: utilizing the treatment options. *Drugs*, **19**, 429–36.

Moll, J.M. (1994). The place of psoriatic arthritis in the spondarthritides. *Baillière's Clinical Rheumatology*, **8**, 395–417.

Mulherin, D.M., FitzGerald, O., and Bresnihan, B. (1993). Lymphedema of the upper limb in patients with psoriatic arthritis. *Seminars in Arthritis and Rheumatism*, 22, 350–6.

O'Neill, T. and Silman, A.J. (1994). Historical background and epidemiology. *Baillière's Clinical Rheumatology*, **8**, 245–61.

Palit, J., Hill, J., Capell, H.A., *et al*. (1990). A multicentre double-blind comparison of auranofin, intramuscular gold thiomalate and placebo in patients with psoriatic arthritis. *British Journal of Rheumatology*, **29**, 280–3.

Panayi, G. (1994). Immunology of psoriasis and psoriatic arthritis. *Baillière's Clinical Rheumatology*, **8**, 419–27.

Parish, J.A., Fitzpatrick, T.B., Tanenbaum, L., and Pathak, M.A. (1974). Photochemotherapy of psoriasis with oral methoxsalen and longwave ultraviolet light. *New England Journal of Medicine*, **291**, 1207–11.

Peloso, P. and Gladman, D.D. (1992). Fish oils in the treatment of psoriatic arthritis: an open study. *Arthritis and Rheumatism*, **35** (Suppl. 9), S225 (Abstract).

Perlman, S.G., Gerber, L.H., Roberts, M., *et al*. (1979). Photochemotherapy and psoriatic arthritis. A prospective study. *Annals of Internal Medicine*, **91**, 717–22.

Peterson, A.W. and Shepherd, J.P. (1992). Facia lata interpositional arthroplasty in the treatment of temporomandibular joint ankylosis caused by psoriatic arthritis. *International Journal of Oral and Maxillofacial Surgery*, **21**, 137–9.

Rahman, P.A. and Gladman, D.D. (1995). Sulphasalazine in psoriatic arthritis [Abstract]. *Journal of Rheumatology*, 22, 1601.

Resnick, D. and Niwayama, G. (1981). Psoriatic arthritis. In *Diagnosis of bone and joint disorders*, pp. 1103–29. W.B. Saunders, Philadelphia.

Roberts, M.E.T., Wright, V., Hill, A.G.S., and Mehra, A.C. (1976). Psoriatic arthritis: a follow-up study. *Annals of the Rheumatic Diseases*, **35**, 206–19.

Rødahl, E. and Iversen, O.J. (1985). Antigens related to the major internal protein, p27, of a psoriasis associated retrovirus-like particle are expressed in patients with chronic arthritis. *Annals of the Rheumatic Diseases*, **44**, 761–5.

Roenigk, H.H., Fowler-Bergfeld, W., and Curtis, G.H. (1969). Methotrexate for psoriasis in weekly oral doses. *Archives of Dermatology*, **99**, 86–93.

Roux, H., Schiano, A., Maestracci, D., and Serratrice, G. (1979). Notre experience du traitement du rheumatisme psoriasique par la D-penicillamine. *Revue du Rheumatisme*, **46**, 631–3.

Rubins, A.Y. and Merson, A.G. (1987). Subpopulations of T lymphocytes on psoriasis patients and their changes during immunotherapy. *Journal of the American Academy of Dermatology*, **17**, 972–7.

Sakkas, L.I., Loqueman, N., Bird, H., Vaughan, R.W., Welsh, K.I., and Panayi, G.S. (1990). HLA class II and T cell receptor gene polymorphisms in psoriatic arthritis and psoriasis. *Journal of Rheumatology*, **17**, 1487–90.

Sakkas, L.I., Jachesoni, A., Kerr, L.A., *et al*. (1991). Immunoglobulin heavy chain gene polymorphism in Italian patients with psoriasis and psoriatic arthritis. *British Journal of Rheumatology*, **30**, 449–50.

Salvarani, C., Macchioni, P., Cremonesi, T., *et al*. (1992*a*). The cervical spine in patients with psoriatic arthritis: a clinical, radiological and immunogenetic study. *Annals of the Rheumatic Diseases*, **51**, 73–7.

Salvarani, C., Macchioni, P., Boiardi, L., *et al*. (1992*b*). Low dose cyclosporin A: relation between soluble interleukin-2 receptors and response to therapy. *Journal of Rheumatology*, **19**, 74–9.

Scarpa, R., Oriente, P., Pulino, A., *et al*. (1984). Psoriatic arthritis in psoriatic patients. *British Journal of Rheumatology*, **23**, 246–50.

Scarpa, R., Oriente, P., Pucino, A., *et al*. (1988). The clinical spectrum of psoriatic spondylitis. *British Journal of Rheumatology*, **27**, 123–37.

Scarpa, R., Del Puente, A., di Girolamo, C., *et al*. (1992). Interplay between environmental factors, articular involvement, and HLA B27 in patients with psoriatic arthritis. *Annals of the Rheumatic Diseases*, **51**, 78–9.

Stern, R.S. (1985). The epidemiology of joint complaints in patients with psoriasis. *Journal of Rheumatology*, **12**, 315–20.

Stoeger, A., Mur, E., Penz-Schneeweiss, D., *et al*. (1994). Technitium-99m human immunoglobulin scintigraphy in psoriatic arthropathy: first results. *European Journal of Nuclear Medicine*, **21**, 342–4.

Tomfohrde, J., Silverman, A., Barnes, R., *et al*. (1994). Gene for familial psoriasis susceptibility mapped to the distal end of human chromosome 17q. *Science*, **264**, 1141–5.

Torre-Alonso, J.C., Rodriguez-Perez, A., Arribas-Castrillo, J.M., *et al*. (1991). Psoriatic arthritis (PA): a clinical, immunological and radiological study of 180 patients. *British Journal of Rheumatology*, **30**, 245–50.

Vasey, F.B. (1985). Etiology and pathogenesis of psoriatic arthritis. In *Psoriatic arthritis*, (ed. L.H. Gerber and L.R. Espinoza), pp. 45–57. Grune & Stratton, Orlando, FL.

Vasey, F.B., Seleznick, M.J., Fenske, N.A., and Espinoza, L.R. (1989). New signposts on the road to understanding psoriatic arthritis. *Journal of Rheumatology*, **16**, 1405–7.

Veale, D., Yanni, G., Rogers, S., Barnes L, Bresnihan, B., and Fitzgerald, O. (1993). Reduced synovial membrane macrophage numbers, ELAM-1 expression, and lining layer hyperplasia in psoriatic arthritis as compared with rheumatoid arthritis. *Arthritis and Rheumatism*, **36**, 893–900.

Veale, D.J., Barnes, L., Rogers, S., and FitzGerald, O. (1994*a*). Immunohistochemical markers for arthritis in psoriasis. *Annals of the Rheumatic Diseases*, **53**, 450–4.

Veale, D., Rogers, S., and Fitzgerald, O. (1994*b*). Classification of clinical subsets in psoriatic arthritis. *British Journal of Rheumatology*, **33**, 133–8.

Veale, D.J., Torley, H.I., Richards, I.M., *et al*. (1994*c*). A double-blind placebo controlled trial of efamol marine on skin and joint symptoms of psoriatic arthritis. *British Journal of Rheumatology*, **33**, 954–8.

Ventura, M., Colizzi, M., Ottolenghi, A., *et al*. (1989). Cell-mediated immune response in psoriasis and psoriatic arthritis. *Recenti Progressi in Medicina*, **80**, 449–54.

Voorhees, J.J. (1983). Leukotrienes and other lipoxygenase products in the pathogenesis and therapy of psoriasis and other dermatoses. *Archives of Dermatology*, **119**, 541–7.

Willkens, R.F., Williams, H.J., Ward, J.R., *et al*. (1984). Randomized, double-blind, placebo controlled trial of low-dose pulse methotrexate in psoriatic arthritis. *Arthritis and Rheumatism*, 27, 376–81.

Wright, V. (1956). Psoriasis and arthritis. *Annals of Rheumatic Diseases*, 15, 348–53.

Wright, V. and Moll, J.M.H. (1971). Psoriatic arthritis. *Bulletin of Rheumatic Diseases*, 21, 627–32.

Wright, V. and Moll, J.M.H. (ed.) (1976). Psoriatic arthritis. In *Seronegative polyarthritis*, pp. 169–223. North Holland Publishing Co., Amsterdam.

Wright, V., Roberts, M.C., and Hill, A.G.S. (1979). Dermatological manifestations in psoriatic arthritis: a follow-up study. *Acta Dermatologica et Venereologica*, **59**, 235–40.

Zacharias, H. and Zacharias, E. (1987). Methotrexate treatment of psoriatic arthritis. *Acta Dermatologica et Venereologica*, **67**, 270–3.

Zanelli, M.D. and Wilde, J.S. (1992). Joint complaints in psoriasis patients. *International Journal of Dermatology*, **31**, 488–91.

5.5.5 Reactive arthropathy, Reiter's syndrome, and enteric arthropathy in adults

Bernard P. Amor and Antoine A. Toubert

Historical considerations

A brief historical survey may help in understanding the different names for what is now called reactive arthritis. The occurrence of arthritis following acute diarrhoea or a urethral discharge was intriguing enough to be mentioned occasionally in the ancient medical literature. Hippocrates noted that 'a youth does not suffer from gout until after sexual intercourse', gout at that time meaning acute arthritis. He also mentioned that diarrhoea which stops suddenly may create deposits in the chest and the joints. The observations of Martiniere in 1664, Stoll in 1776, and Brodie in 1818 are also often cited, among many others (see Gounelle and Marche 1941). Clinical descriptions improved when microbiological techniques allowed differentiation of septic from aseptic arthritis. The first description of an identifiable infectious agent triggering aseptic arthritis was made during the First World War by Fiessinger and Leroy (1916), who reported an outbreak of *Shigella dysenteriae* and described four of the dysenteric patients as suffering from what they called an 'oculo-urethro-synovial' syndrome. Hans Reiter (1916) described the same oculo–urethro–synovial triad in a young officer and discussed the role of a treponema called 'treponema arthritidis'. During the Second World War and the period of confusion that followed, epidemics of dysentery (Bauer and Engelman 1942; Marche 1946; Paronen 1948; Masbernard 1959; Noer 1966) were associated with new cases of what was now called 'Reiter's syndrome'. Further observation of these cases showed an unsuspected frequency of relapses and chronic evolution, often unrelated to new infectious episodes (Csonka 1958).

During the same period, the curing of gonococcal infections with penicillin confirmed the identification of non-gonococcal urethritis and disrupted the old and mixed concept of chronic gonococcal arthritis. Among the different aetiological agents of non-gonococcal urethritis, *Chlamydia trachomatis* (Harkness 1950; Ford and Rasmussen 1964; Dunlop *et al*. 1965; Schachter *et al*. 1969; Keat *et al*. 1980; Amor 1985) has progressively become recognized as the one most closely associated with a triad similar to that described by Fiessinger, Leroy, and Reiter after diarrhoea. 'Post-dysenteric or epidemic Reiter' and 'postvenereal or endemic Reiter's syndrome' were defined, and a first set of criteria proposed (Willkens *et al*. 1981). Concurrently, the clinical overlap and genetic links between complete or incomplete Reiter's syndrome, ankylosing spondylitis, and psoriasis appeared more and more impressive as observation periods lengthened (Marche 1954; Khan and Hall 1965; Sairanen 1969).

The association between HLA-B27 antigen, ankylosing spondylitis (Brewerton *et al*. 1973a; Schlosstein *et al*. 1973), and Reiter's syndrome or reactive arthritis (Amor *et al*. 1974; Brewerton *et al*. 1973b) provided strong support to the clinical data.

In 1973 and 1974, Aho and colleagues reported aseptic arthritis following gut infection with *Yersinia enterocolitica*, mostly in HLA-B27-positive individuals, and proposed a new term, 'reactive arthritis', which was so successful that it has progressively replaced all the previous terms (Aho *et al*. 1973; Aho *et al*. 1974).

During the same period, the clinical overlap and genetic links between reactive arthritis, ankylosing spondylitis, and other apparent entities were growing stronger, and the introduction of the concept and the term 'spond(ylo)arthropathies' was welcomed (Moll *et al*. 1974).

Definition

Reactive arthritides are often defined as aseptic arthritis triggered by an infectious agent located outside the joint. This definition is unsatisfactory because in some cases non-viable agents can be identified in the joints, and, moreover, in many patients no triggering agent may be found despite identical clinical features.

Reactive arthritides are now considered as belonging to the spondylarthropathies. Therefore, as a first step, they can be defined as spondylarthropathy according to the criteria shown in Table 1 if the sum of the scores for the 12 items is 6 or more (Amor *et al*. 1990). They can then be defined as reactive arthritis if item 7 (urethritis or cervicitis less than 1 month before an arthritic episode) and/or item 8 (acute diarrhoea less than 1 month before an arthritic episode) are present. The validity of these criteria has been established in a study including 140 spondylarthropathies and 1829 control rheumatic patients (see Table 3(b) of Chapter 5.5.1). The serial recording of the disease pattern of one patient throughout its evolution (Fig. 1) explains this definition.

Criteria	Score
A. Clinical symptoms or past history of:	
1. Lumbar or dorsal pain during the night or morning stiffness of the lumbar or dorsal spine	1
2. Asymmetrical oligoarthritis	2
3. Buttock pain—if alternating alternately the right or the left buttock	1 or 2
4. Sausage-like toe or digit	2
5. Heel pain	2
6. Iritis	2
7. Non-gonococcal urethritis or cervicitus accompanying, or within 1 month before, the onset of arthritis	1
8. Acute diarrhoea accompanying, or within 1 month before, the onset of arthritis	1
9. Presence or history of psoriasis and/or balanitis and/or of inflammatory bowel disease (ulcerative colitis, Crohn's disease)	2
B. Radiological finding:	
10. Sacroiliitis (grade ≥ 2 if bilateral, grade ≥ 3 if unilateral)	3
C. Genetic background:	
11. Presence of HLA-B27 and/or familial history of ankylosing spondylitis, Reiter's syndrome, uveitis, psoriasis or chronic enteropathies	2
D. Response to treatment:	
12. Close-out improvement of rheumatic complaints with non-steroidal anti-inflammatory drugs (NSAIDs) in less than 48 h or relapse of the pain in less than 48 h if NSAIDs discontinued	2

Table 1 Criteria for diagnosing spondylarthropathies (Amor *et al*. 1990)

Factors observed during first 2 years of disease	Mild (n=80)	Severe (n=20)	Risk for severity		Proposed weighting
			Ratio	95% CI	
Hip arthritis	0	7	22.65	4.43–115.8	4
ESR > 30	10	16	7.88	4.84–9.15	3
Poor efficacy of NSAIDs	5	11	6.33	2.56–27.11	3
Limitation of lumbar axis	2	7	7.88	2–25	3
Sausage-like finger or toe	2	5	6.48	1.48–48	2
Oligarthritis	12	12		1.38–13.1	1
Onset < 16 years	12	10	3.47	1.059–12.75	1

A sum of 2 or more is predictive of a benign outcome (sensitivity 92.5%, specificity 78%). A sum of 7 or more indicates a risk of severe outcome (sensitivity 50%, specificity 97.5%).
ESR erythrocyte sedimentation rate, NSAID non-steroidal anti-inflammatory drug.

Table 3 Seven predictive factors of long-term outcome in spondylarthropathy (Amor *et al*. 1994)

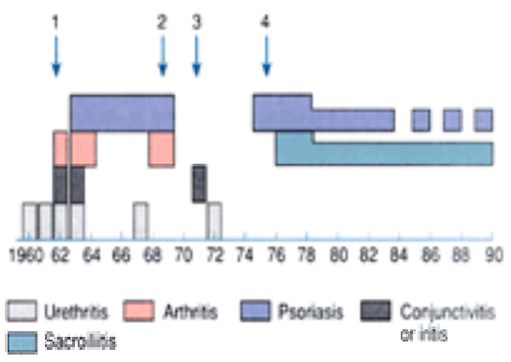


Fig. 1 Synoptic recording of disease pattern across 30 years of follow-up in a typical patient with Reiter's syndrome. 1. The diagnosis is sexually acquired reactive arthritis. 2. Psoriasis: the diagnosis could be psoriatic arthritis. 3. Isolated iritis. 4. Buttock and back pain/sacroiliitis: the diagnosis could be ankylosing spondylitis. It is clear that this patient does not have four different diseases but only one, a spondylarthropathy, which has taken the clinical aspect of reactive arthritis at onset. He is B27-positive and has enough items (see Table 1) to be defined as spondylarthropathy at all stages of his disease.

It may still be useful in clinical practice or for research purposes to subdivide reactive arthritis according to the triggering agent, when it can be identified.

Clinical features

Reactive arthritis, like other spondylarthropathies, combines four syndromes: a peripheral arthritis syndrome, an enthesopathic syndrome, a pelviaxial syndrome, and an extramusculoskeletal syndrome. The combination of these four may vary from one spondylarthropathy to another, in the subgroup of reactive arthritis from one patient to another, and in a given patient during the course of the disease. Nevertheless, the diagnosis of spondylarthropathy will remain unchanged.

General conditions

The onset of reactive arthritis is sometimes acute, with fever as high as 39°C, severe weight loss, and diffuse, polyarticular involvement. However, in most cases in which synovitis is limited to a few joints, low-grade fever or no fever at all is the rule.

The peripheral arthritis syndrome

The onset is acute and within a few days two to four joints (oligoarthritis), mainly knees and ankles, are painful and swollen with an asymmetrical distribution (item 1). Mono- and polyarthritis may be observed. They are not helpful for diagnosis, but they do not exclude it if the other items of the criteria of spondylarthropathies are present.

Diffuse swelling of an entire finger(s) or toe(s), commonly described as sausage digit or toe, is very specific (item 4).

The enthesopathic syndrome

Localized painful parts of bones at the level of the enthesis, in isolation or associated with arthritis, are described spontaneously or after questioning in 42 per cent of patients. Heel pain is the most frequently recognized enthesopathic pain because it is easier to differentiate from joint pain than enthesopathic pains localized around the knee or the hip.

The pelviaxial syndrome

Highly distinctive features are inflammatory dorsal or low back pain (item 1), or buttock pain (item 3), observed in 50 per cent of patients.

The extramusculoskeletal syndrome

Intestinal and genitourinary symptoms are not only of diagnostic but also of pathophysiological and therapeutic importance; they will be described and discussed in more detail at the end of this chapter.

Mucocutaneous lesions (Montgomery et al. 1959)

Balanitis circinata is a painless, erythematous lesion of the glans penis. The size of the lesion may vary but its boundaries with normal mucosa are always clearly

defined (Fig. 2). Similar painless, erythematous, well-defined lesions may be found on the oral mucosa (hard and soft palate, gingiva, tongue, cheeks) (Fig. 3).

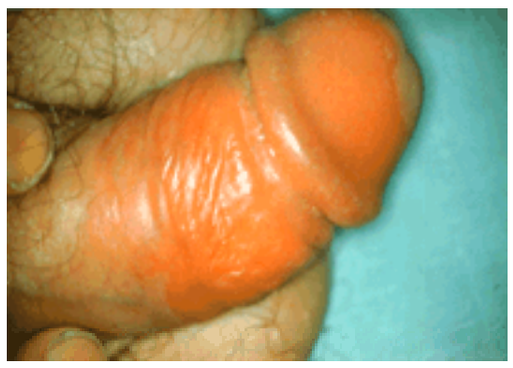


Fig. 2 Balanitis.

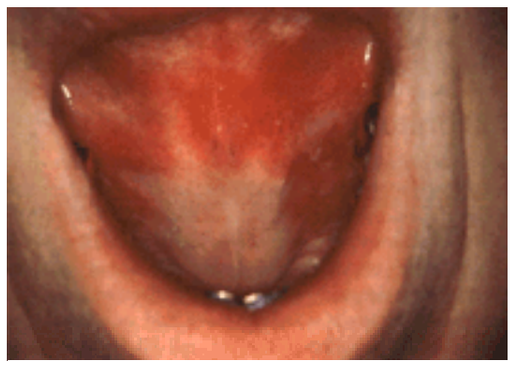


Fig. 3 An erythematous, painless, palatal mucosal lesion.

Mucosal lesions are not frequent (9–40 per cent) but are very specific. Being painless, they may be ignored by the patient.

Lesions of pustular psoriasis, when localized on the palms and/or the soles of the feet, are called keratoderma blennorrhagica (Fig. 4) and are diagnostically very suggestive, as are hyperkeratosis and parakeratosis of the nails. These lesions are infrequent but are often associated with a severe outcome (Temine *et al*. 1972). Ordinary psoriatic lesions may also be seen.

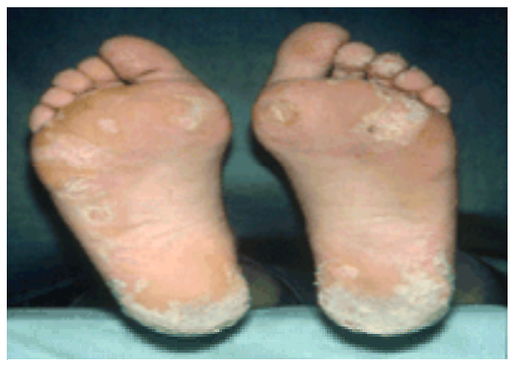


Fig. 4 Keratoderma blennorrhagica.

Visceral involvement, very similar to that observed in other spondylarthropathies, is infrequent (1 per cent), and includes cardiac conduction abnormalities (prolongation of the PR interval, complete heart block) in early disease and aortic insufficiency in late disease. An association with transient neurological dysfunction such as peripheral or cranial nerve palsy, Parsonage–Turner syndrome, and hemiplegia has been described (Oates and Hancock 1959).

'Pleuritic' chest wall or pseudonephritic pains are probably unrecognized enthesopathic pains of the chest or due to involvement of the costovertebral joints (Benhamou *et al*. 1987).

Ocular lesions

Conjunctivitis is a part of the classical triad and is observed very early, before or at the onset of arthritis; the discharge is sterile and subsides in 1 to 4 weeks. Uveitis is less frequent in early disease but occurs in 15 per cent of patients with recurring disease, often as an incident separate from arthritis.

Intestinal symptoms

Acute diarrhoea may precede the musculoskeletal symptoms by 1 month or less (Table 1; item 8). This diarrhoea, resulting from an outbreak of enterobacterial infection, may affect many individuals, but only some of them will go on to suffer from what has been called epidemic reactive arthritis. Dysenteric symptoms are sometimes very mild in these patients and even milder than in non-arthritic patients. If no medical advice or treatment was required, the symptoms may be ignored if patients are not closely questioned. The provocative agent has commonly disappeared from the gut when the joint symptoms arise, and during an epidemic it is identified in diarrhoeal but not in arthritic patients. The isolation of a triggering agent from the stool of an arthritic patient is a rare event, except for salmonellae (Hannu and Leirisalo-Repo 1988). An agent is called arthritogenic if, during an outbreak, some cases of reactive arthritis are observed among a number with only dysentery. *Shigella flexneri* type 1 and 2, *Salmonella enteritidis* and *S. typhimurium, S. abony, S. blocley, S. schwarzengrund, S. heidelberg, S. haifa, S. manila, S. newport, Clostridium difficile* (Puterman and Rubinow 1993), *Vibrio parahaemoliticus* (Tamura *et al*. 1993), *Yersinia pseudotuberculosis*, and *Y. enterocolitica* fall within this definition. *Shigella sonnei, S. typhi*, and *Escherichia coli* dysenteric syndrome, in contrast, have so far not been complicated by reactive arthritis. *Yersinia enterocolitica* is often mentioned as the triggering agent in Northern Europe. Infection by *Y. enterocolitica* in Belgium, the country with the highest incidence of this infection, is strongly associated with eating raw pork (Tauxe *et al*. 1987).

Chronic diarrhoea or other gut symptoms are not associated with reactive arthritis; nevertheless, overlap with inflammatory bowel disease-associated arthropathies and the histological finding of silent inflammatory gut lesions in some cases of reactive arthritis make the demarcation blurred and the diagnosis of undifferentiated spondylarthropathies very helpful (see 'Enteric arthropathies in adults' below).

Urogenital symptoms

Non-specific urethritis is generally limited to a mild, painless, and non-purulent urethral discharge. It may be asymptomatic and detected by examination of the first

morning urine specimen. Occasionally, however, it is severe and accompanied by prostatitis. In women, cervicitis is usually marked by vaginal discharge but may be asymptomatic.

Urethritis can also be a postdysenteric phenomenon, apparently not sexually acquired, raising the possibility that inflammation may be due to mechanisms other than direct infection.

Non-gonococcal, sexually acquired urethritis was followed by a clinical picture of reactive arthritis in 16 out of 531 men followed prospectively (Keat *et al*. 1987). Agents associated with non-gonococcal urethritis, and particularly *Ch. trachomatis*, which is the most frequently isolated, have therefore been considered as candidates for triggering sexually acquired reactive arthritis. *Chlamydia pneumonia* infection and antibodies against *Ch. pneumoniae* are highly prevalent in the adult population These antibodies make the genus/species tests unhelpful as a means of screening for the serological diagnosis of *Ch. tracomatis* infections (Freidank *et al*. 1993). But is *Ch. pneumoniae* itself able to trigger a reactive arthritis (Braun *et al*. 1994)?

Mycoplasmas and even ureaplasmas that are pathogenic for the human genital tract still require rigorous investigation. It is now clear that sexually acquired reactive arthritis is not attributable to a single micro-organism. Recurrent or repeated infections do not always lead to recurrence of arthritis and may occur in the absence of further sexual intercourse.

In women, reports of reactive arthritis following acute urogenital symptoms are scarce. Of women reviewed 2 to 4 years after the onset of salpingitis, 72 per cent had radiological or scintiscan evidence of sacroiliitis and a substantial proportion of these also had low back pain (Szanto and Hagenfeldt 1979). However, these findings have not, to date, been correlated with any specific microbiological data and raise difficulties of interpretation that remain unresolved.

Genetic background

Family studies of first- and second-degree relatives may be helpful for diagnosis. Questions should not be limited to the history of reactive arthritis, even if multiple cases have been reported in some families (Wright 1978), but should also concern other rheumatological aspects of spondylarthropathies (ankylosing spondylitis, episodes of oligoarthritis) and their most common extra-articular manifestations or associations (iritis, psoriatic skin lesions, or inflammatory bowel disease). A well-defined family history (Table 1; item 11) may provide a reliable diagnosis in the absence of specific clinical symptoms when testing for HLA-B27 antigen is not desirable for technical or economical reasons, and in HLA-B27-negative individuals (Dougados *et al*. 1991).

Response to treatment

Response to treatment is not a common tool for diagnosis, except when it is dramatic or specific, such as with colchicine in gout, penicillin in gonococcal arthritis, and corticosteroids in polymyalgia rheumatica. Pain at night and morning stiffness in spondylarthropathies are very sensitive to non-steroidal anti-inflammatory drugs. Item 12 in the diagnostic criteria (Table 1) takes advantage of this response and defines its limits (dramatic improvement within 48 h with these drugs or relapse of pain in less than 48 h after their discontinuation).

Laboratory features

The inflammatory nature of the disease is biologically confirmed by an elevated erythrocyte sedimentation rate and an increased concentration of C-reactive protein. Analysis of synovial fluid shows more than 2000 cells/mm^3, with a majority of polymorphonuclear leucocytes; synovial biopsy shows inflammatory changes including vascular congestion and perivascular cell infiltration, mainly of neutrophils. Cultures of synovial tissue or fluid are negative. The synovial complement concentration is normal. HLA-B27 is present in 50 to 80 per cent of cases. In the absence of, or in addition to, a family history, the presence of the HLA-B27 antigen constitutes item 11 in the criteria.

For therapeutic decisions in some cases, but mainly for epidemiological or for research purposes, many biological tests can be performed: these include stool cultures, which might be positive during an outbreak of gastroenteritis but are often negative when arthritic symptoms occur. *Chlamydia trachomatis* may be detected by microimmunofluorescence with specific antisera on urethral or cervical scrapings. Tests for detecting specific antibodies or the stimulation of blood mononuclear cells by bacterial antigen have no significant predictive value for reactive arthritis in the general population or even in the rheumatology clinic (Sieper *et al*. 1993a); the proliferation of mononuclear cells from synovial fluid appears more specific and could give some information on the triggering agent when the preceding infection is asymptomatic (Ford *et al*. 1985; Sieper *et al*. 1993b). The search for the products of triggering agents using specific monoclonal antibodies, or DNA or RNA hybridization, is limited to research purposes (Granfors *et al*. 1989).

In the early stages, no radiographic signs are found except for changes in the sacroiliac joint in 4 per cent of patients, probably present before the onset of reactive arthritis.

Evolution

In the short term

The first oligoarticular episode subsides in 3 to 6 months, during which time symptomatic treatment is generally required. The first episode may be longer in severe polyarticular involvement, or very short and often wrongly attributed to trauma when only one joint is affected. Two sites, the metatarsophalangeal joints and the heels, may remain painful for months and in some cases for 1 to 4 years. Extramusculoskeletal symptoms also subside, but balanitis and psoriatic lesions may persist for longer. When all symptoms are taken into account, 75 per cent of patients are in complete remission at the end of the second year after onset (Fig. 5). One per cent of patients with reactive arthritis, particularly those with keratoderma blennorrhagica, may have a very severe and even fatal outcome (Temine *et al*. 1972).

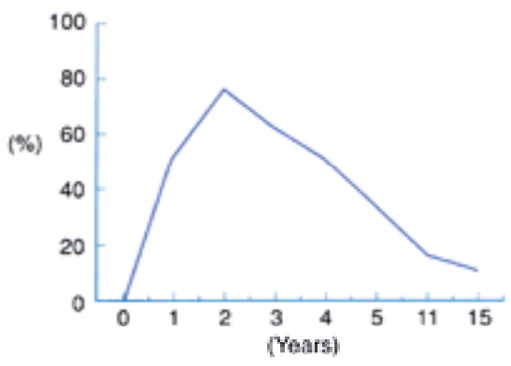


Fig. 5 Complete remission: incidence across 15 years.

In the long term and prognosis (Amor 1979; Amor et al. 1994)

Figure 6 is based on synoptic recording of the disease pattern, as shown in Fig. 1, applied to 140 patients with reactive arthritis across 15 years. It shows the percentage of relapses and clinical involvement in these patients over time. Relapses begin 3 to 4 years after the first episode and can consist of recurrence of peripheral arthritis or enthesopathic pain (Fig. 6a), of pelviaxial symptoms (Fig. 6b), or of iritis or other extra-articular symptoms (Fig. 6c); these symptoms can be isolated or associated. Radiographic changes may now be observed. Narrowing of joint spaces and erosions are very uncommon, except when reactive arthritis is associated with psoriatic lesions.

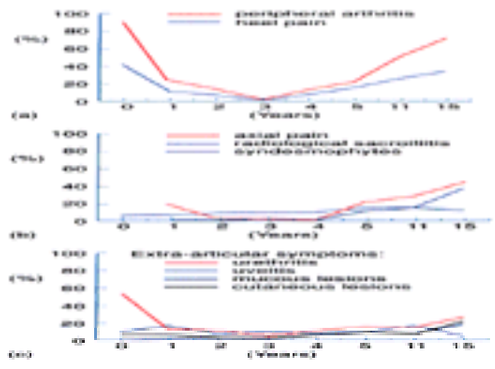


Fig. 6 Synoptic recording, as in Fig. 1, of the disease pattern in 140 patients with reactive arthritis across 15 years. (a) Recurrence of peripheral arthritis or enthesopathic pain; (b) recurrence of pelviaxial symptoms; (c) recurrence of iritis or other extra-articular symptoms.

Enthesopathic lesions can be visualized at an early stage by focal uptake of [$^{99}Tc^{m}$] methylene diphosphonate (**MDP**) (Fig. 7). Later radiographic changes, such as bone erosions (Fig. 8) and large and fluffy bone spurs, may develop along the metatarsals, the phalanges of the feet, the calcaneum (Fig. 9), and the pelvis.

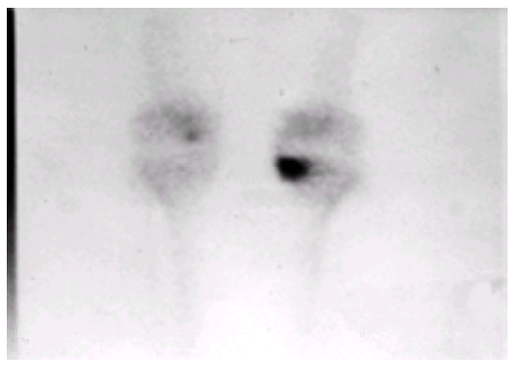


Fig. 7 Enthesopathic lesion visualized by focal uptake of [$^{99}Tc^{m}$]MDP.

Fig. 8 Advanced changes in enthesopathic lesions: erosion and calcaneal spur.

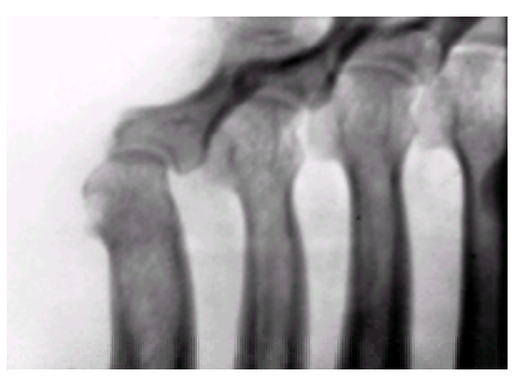


Fig. 9 Enthesopathic erosive lesion of the fifth metatarsal head.

Sacroiliitis, indistinguishable from that seen in ankylosing spondylitis, is observed in 4 per cent of patients at onset. Its frequency increases with time: it is observed in 37 per cent of patients followed for 15 years or more (see Fig. 6b) and is associated with axial lesions of ankylosing spondylitis in 15 per cent of cases. Over 25 per cent of patients were forced to change occupation or were unemployable. The percentage of remissions and the clinical nature of relapses over time are similar in HLA-B27-positive and -negative patients (Table 2). Seven predictive factors of long-term outcome have been selected and weighted (Table 3). If the sum of these factors at entry was 3 or less, a benign outcome could be predicted with a sensitivity of 92 per cent and a specificity of 78 per cent (Amor *et al*. 1994).

Duration of follow-up (years)	No. of patients	HLA-B27 (percentage remissions)		
		B27	B27+	B27 –
1	123	75	50	55
5–10	69	68	36	38.4
11–15	40	65	17.6	20

Table 2 Percentage of complete remissions according to HLA-B27 antigen status over time

Reactive arthritis according to sex, age, and triggering agents

Better knowledge of the disease and more accurate diagnostic criteria have allowed recognition of reactive arthritis in women and children. Nevertheless, the sex ratio (M:F) is still 3:1 and three-quarters of affected patients are young adults under 40 years of age.

As an example, more than 1000 cases of postdysenteric reactive arthritis were observed in Algeria between 1954 and 1960 among young French adults during their military service, whereas the local population, whatever its ethnic origin, remained unaffected. The local population has a frequency of HLA-B27 antigen of 6 per cent and ankylosing spondylitis is as common as in other caucasoid countries but more severe (Will *et al*. 1990). This observation suggests that some agents, especially those that are water-borne, do not trigger reactive arthritis in very young children.

In patients with yersinial arthritis who were HLA-B27-positive, inflammatory symptoms in the urinary tract, ocular inflammation, and back pain during acute disease were significantly more common than in HLA-B27-negative patients (Leirisalo *et al*. 1982). In contrast, erythema nodosum was more frequent in HLA-B27-negative patients with yersinial arthritis.

Differential diagnosis outside the spondylarthropathy group

Other sexually acquired arthritis

Gonococcal arthritis must be suspected in any cases of acute or subacute arthritis, usually in women. A search for *Neisseria gonorrhoea* in urethral and cervical discharges, and by culture of blood and synovial fluid, must be made concurrently with a search for agents that trigger reactive arthritis. During isolation procedures, if a gonococcal infection is suspected clinically, a therapeutic test using antibiotics exclusively is justified because arthritis will subside dramatically within 3 days and completely within 10 days.

Human immunodeficiency virus (HIV)-associated arthritis

Acute but transitory joint pains are described during the early stage of the disease; they differ from the synovitis of reactive arthritis, but reactive arthritis, like other spondylarthropathies, may be worsened by HIV (Winchester *et al*. 1987; Rowe and Keat 1989). The search for HIV infection is often recommended in even the typical case of reactive arthritis.

Behçet's syndrome

In Behçet's syndrome the painful aphthous ulcerations of oral or genital mucosa are different from the painless mucosal lesions of reactive arthritis. The total or posterior uveitis of Behçet's syndrome may be distinguished from the iritis or anterior uveitis of spondylarthropathies. Strictly oral but recurrent aphthous lesions are sometimes described by spondylarthropathic patients, leading to some confusion in their classification, which may explain some of the discrepancies concerning the frequency of sacroiliitis in Behçet's syndrome.

Parvovirus arthropathies

In parvovirus arthropathies, an acute arthritis associated with lumbar pain may mimic the onset of reactive arthritis. However, the transient erythematous skin lesions and the spontaneous recovery may help in the clinical identification of this disease (White *et al*. 1985).

Lyme disease

Oligoarthritis in Lyme disease (Steere *et al*. 1983) in young people may be confused with reactive arthritis if patients are not asked about a past history of erythema chronicum migrans or a stay in an endemic area. Enthesopathic pains are sometimes described in Lyme disease. When serological indications of Lyme disease are unclear, treatment with high doses of penicillin alone (or other antibiotics), which promptly relieves symptoms in Lyme disease but is ineffective in reactive arthritis, may be helpful.

Differential diagnosis inside the spondylarthropathy group: limits of this concept

Among patients identified as having reactive arthritis in the absence of the typical oculo–urethro–synovial triad, differential diagnosis from the other spondylarthropathies is often disputed because overlaps are numerous. From a practical point of view, these disputes are probably unimportant. However, questions as to the limits of the spondylarthropathy group do arise; they concern psoriatic arthropathies and **SAPHO** (see below).

According to the criteria for spondylarthropathies, patients with psoriatic arthritis are classified into those that meet these criteria and those that do not and in whom psoriasis appears to be associated with other rheumatic disorders such as rheumatoid arthritis.

SAPHO (synovitis, acne, palmoplantar pustulosis, hyperostosis, aseptic osteomyelitis) (Chamot et al. 1987)

The use of this grouping may help to classify patients with different clinical presentations, such as acute pseudoseptic arthritis associated with palmoplantar pustular lesions and multifocal aseptic osteomyelitis whether associated or not with the same skin lesions. The boundaries between palmoplantar pustulosis and psoriasis, the occurrence of bone lesions around the sacroiliac joints, and the presence of HLA-B27 antigen in some patients may make their classification difficult.

The boundaries of reactive arthritis

Postyersinial erythema nodosum may be used to introduce the issue of the boundaries of reactive arthritis. The highly statistical difference in the occurrence of erythema nodosum between HLA-B27-positive and -negative individuals may indicate that postyersinial erythema nodosum is not reactive arthritis but the result of the interaction of the same infectious agent on a different genetic background. Erythema nodosum is not otherwise described among the cutaneous manifestations observed in reactive arthritis triggered by infectious agents other than *Yersinia* spp.

The definition of reactive arthritis given earlier clarifies this issue because postyersinial erythema nodosum does not fit the diagnostic criteria for the spondylarthropathies.

The same can be said for diseases that have sometimes incorrectly been called reactive arthritis, such as rheumatic fever or parasitic arthritis. No relation exists at any level (clinical, pathological, genetic, or therapeutic) between rheumatic fever and reactive arthritis. It is nevertheless suggested that b-haemolytic streptococci may trigger reactive arthritis as well as rheumatic fever in genetically predisposed patients (Valtonen *et al*. 1993). The treatment of carcinoma of the bladder by intravesical injection of bacillus Calmette–Guérin may induce an arthritis which in some cases fits the diagnostic criteria for the spondylarthropathies. Anecdotal reports of arthritis associated with brucellae, *Mycobacterium phlei*, or parasitic infections cannot be classified as reactive arthritis because they are cured by a specific treatment, which is not the case in true reactive arthritis.

Pathogenesis of reactive arthritis

The precise disease mechanism for reactive arthritis is still unknown and research in this field has focused mainly on immunogenetics because of the association between this arthritis and HLA-B27, and on microbiology in view of the clinical data concerning the triggering of the disease by micro-organisms (Enterobacteriaciae and *Ch. trachomatis*). The interplay between these two areas of research has generated hypotheses that apply not only to reactive arthritis but also to the other forms of spondylarthropathies and even more generally to an HLA–disease association (for a full review of possible pathogenetic mechanisms, see Chapter 5.4.1.). Several controversial aspects of the links between HLA-B27 and reactive arthritis-triggering bacteria have been unravelled by some recent findings. They all point to an important role of the *in situ* antibacterial immune response. Concerning *Ch. trachomatis*, analysis by polymerase chain reaction has shown that the organism could be

present in the synovium of reactive arthritis (Rahman *et al*. 1992), albeit in a form lacking the expression of the lipopolysaccharide and major outer-membrane surface antigen (Beutler *et al*. 1994), so reflecting a state of latency and an inability of the host to clear the infection. This phenomenon could be caused by the local production of interferon-g, consistent with the predominant Th1 phenotype of the T-helper cells in synovial fluid (Simon *et al*. 1993). Another major advance has been the isolation from lymphocytes in the synovial fluid of reactive arthritis of HLA-B27-restricted, CD8 + and bacteria-specific, cytotoxic T lymphocytes (Herman *et al*. 1993). Finally, the absence of arthritis in HLA-B27 transgenic rats (Hammer *et al*. 1990) raised in a germ-free environment (Taurog *et al*. 1994) provides a strong argument for the direct effects of both HLA-B27 and bacterial antigens in the pathogenesis of reactive arthritis and provides a precious tool for future experimental studies.

Management (Box 1)

There is no cure for reactive arthritis, but symptomatic treatments may considerably reduce the discomfort resulting from its musculoskeletal and extra-articular features.

Non-steroidal anti-inflammatory drugs promptly relieve pelviaxial pain but do not always have such clear-cut efficacy for arthritis and enthesopathies. Local injections of corticosteroids are in this instance helpful while awaiting the usual remission, 3 to 6 months later. Arthritic metatarsophalangeal joints should receive local injections very early in the disease course, and should then be passively and actively mobilized to avoid permanent deformity from fibrous retraction. In a double-blind, placebo-controlled study of 3 months, lymecycline significantly decreased the duration of illness in patients with *Ch. trachomatis*-triggered reactive arthritis but not in those with other reactive arthritides (Lauhio *et al*. 1991). A 6-month follow-up of an outbreak of *S. enteritidis* enterocolitis found arthritic symptoms in 17 of 108 patients (15 per cent) that were not prevented by early antibiotic treatment; neither did antibiotics affect the duration of reactive arthritis (Locht *et al*. 1993).

Urethritis following dysentery subsides spontaneously in a few days. Postvenereal or apparently spontaneous urethritis may be aseptic and does not respond to antibiotics. Its chronicity or frequent relapses may disturb the patient and the family. A feeling of guilt and anxiety about sexual misconduct should be allayed and the patient enlightened. Intraurethral instillation of corticosteroid may stop the chronic urethral discharge.

Early conjunctivitis subsides spontaneously. Inflammation of the uveal tract requires steroid eye-drops or subconjunctival injections. However, systemic steroids have no proven, as opposed to anecdotal, role in the management of reactive arthritis. For those patients with relentless progression of the disease, second-line drugs may be given. The efficacy of sulphasalazine (2–3 g/24 h) in ankylosing spondylitis (Dougados *et al*. 1986) has been shown in a 6-month controlled study against placebo. The clinical benefit of this drug is more clear-cut in spondylarthropathies with peripheral involvement (Amor *et al*. 1984) such as reactive arthritis. Uncontrolled data suggest that methotrexate (7.5–15 mg/week) and gold salts given in a schedule similar to that used in rheumatoid arthritis may be suitable. In a placebo-controlled study (Calin 1986), azathioprine (1–2 mg/kg body wt daily) was shown to be helpful. However, immunosuppressive therapy should be avoided in young adults, unless other treatment has failed, and must be used as a last resort.

Prevention of reactive arthritis and of relapses

The diversity of agents that may trigger reactive arthritis and the significant number of cases in which no infectious agent at all is identified render preventive treatment, akin to the use of penicillin in rheumatic fever, impracticable.

Nevertheless, the number of patients with sexually acquired reactive arthritis has notably decreased in recent years and this may be due to early treatment of non-gonococcal urethritis. A study in Greenland, where epidemiological conditions are favourable (Bardin *et al*. 1990), seems to indicate that the treatment of non-gonococcal urethritis with tetracycline significantly reduces relapses of the arthritis.

Many workers have observed that in epidemic dysentery, patients with reactive arthritis have not required treatment during the dysenteric episode. This observation may also indicate that the genetic background of reactive arthritis has protects from severe diarrhoea.

Enteric arthropathies in adults

Connections between gut diseases and arthritis have been described for reactive arthritis, where only acute and time-limited dysenteric syndromes occurring within 1 month before the onset of the arthritis are considered as criteria for diagnosis (item 8; Table 1).

However, an association between gut disease and arthritis can be observed in other conditions:

1. arthritis associated with or complicating chronic inflammatory bowel diseases;
2. arthritis leading to the discovery of minor and previously not mentioned gut symptoms;
3. arthritis associated with purely histological gut lesions.

The whole of this puzzle is more or less included within the definition of spondylarthropathies, but a description of each entity, emphasizing the clinical differences, is nevertheless needed. Apparently unrelated to the spondylarthropathies are the arthritis of Whipple's disease, the bypass arthritis–dermatitis syndrome, collagenous colitis, and gluten-sensitive enteropathy.

Issues concerning the bowel flora, diet, and arthritis are mainly speculative and are discussed in Chapter 5.5.1.

Arthritis associated with inflammatory bowel disease

Ulcerative colitis

The association between colitis and arthritis was first described at the end of the nineteenth century (White 1895). It was dismissed as coincident rheumatoid arthritis until 1958, when closer studies showed it to be a seronegative arthritis with distinctive features (Bywaters and Ansell 1958). Different series have shown incidences of arthritis varying from 2.5 to 22 per cent. An incidence of 11.5 per cent in a large unselected series of 269 cases (Wright and Watkinson 1966) is typical. Ulcerative colitis itself occurs in 50/100 000 individuals in England and Scandinavia, but may be less prevalent in other geographical areas.

Crohn's disease

The prevalence of Crohn's disease now exceeds that of ulcerative colitis. The disease has become more frequent in the last 30 years, mainly in northern Europe (Moll 1985), and is making an appearance in southern Europe and in Africa. Arthritis occurs in 10 to 20 per cent in most series.

Arthritis in both ulcerative colitis and Crohn's disease shows many similarities, and both are considered as members of the spondylarthropathy group. Item 9 (present or past history of inflammatory bowel disease) of the criteria shown in Table 1 allows for their classification in this group.

A familial history of inflammatory bowel disease (item 11) associated with any of the other items (sum > 6), in the absence of such disease in the patient, allows only for the diagnosis of spondylarthropathy.

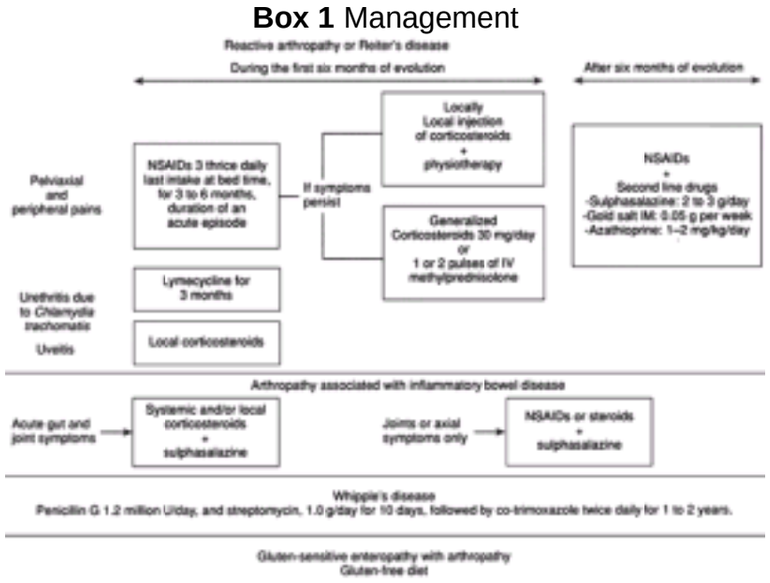
As in other spondylarthropathies, four clinical syndromes, more or less combined, may be described: a peripheral arthritis syndrome, an enthesopathic syndrome, a spinal or axial syndrome, and an extra-articular syndrome.

Peripheral arthritis

Mono- or asymmetrical oligoarthritis is coincident with the onset of inflammatory bowel disease or arises during the course of the disease. There is a close temporal association between exacerbations of the gut and joint disorders. Enteropathic arthritis remits after removal of the diseased bowel segment.

The knees are most commonly involved, closely followed by the ankles. Other joints may be affected, but less frequently. Individual attacks are self-limiting, 50 per cent lasting less than 6 months and only 20 per cent persisting longer than 1 year. Even when recurrent they do not lead to permanent joint damage or deformity (Haslock and Wright 1973). A few cases of erosive large-joint arthritis associated with granulomatous synovitis have been reported in Crohn's disease (Toubert *et al*. 1985). Otherwise the synovial tissue shows non-specific inflammatory changes as in synovial fluid (more than 4000 cells/mm ^{3}, the majority polymorphs).

Box 1 Management



Enthesopathic pain, particularly in the heel, may either be isolated or associated with peripheral arthritis or axial involvement.

Axial involvement

In contrast to peripheral arthritis, sacroiliitis and/or spondylitis are not clearly associated either with the onset of inflammatory bowel disease or with gut exacerbations. Spondylitis is often present years before the onset of colitis or ileitis.

The frequency of ankylosing spondylitis (5 per cent), and asymptomatic sacroiliitis (14 per cent) (Wright and Watkinson 1965) is similar in males and females. HLA-B27 antigen is present in 50 per cent of these patients. HLA-Bw62 is increased in Crohn's disease and in Crohn's arthritis (Mielants and Veys 1990).

Other osteoarticular manifestations

Clubbing of fingers has been reported (Kitis *et al*. 1979) but pulmonary osteoarthropathy may be only coincidental, as are many other more common musculoskeletal disorders.

In Crohn's disease, psoas or retroperitoneal abscess may be complicated by septic arthritis of the hip (Kyle 1971).

Extra-articular syndrome (Greenstein et al. 1976)

Uveitis is observed in 10 per cent of the cases and occurs more frequently in arthritic patients.

Erythema nodosum appears more frequently in Crohn's disease (6.6 per cent) than in ulcerative colitis (4 per cent), while pyoderma gangrenosum is more frequent in ulcerative colitis (5 per cent compared to 1.3 per cent) (Greenstein *et al*. 1976). In the same study, out of 160 patients with Crohn's disease, 17 had both cutaneous and joint lesions, 23 had only arthritis, and 13 only skin lesions. Psoriasis was observed in 1.5 per cent of the patients.

Aphthous stomatitis (4 per cent) is the oral lesion most frequently seen, often associated with erythema nodosum.

Gallstones, kidney stones, and liver disease are not associated with arthritis.

Arthritis with mild, previously unnoticed gut symptoms and asymptomatic gut lesions

When patients suffering from features of spondylarthropathy are asked about gastrointestinal symptoms, some previously unnoticed dysfunctions may be revealed, such as mild but recurrent episodes of diarrhoea, anal fistula, episodes of unexplained weight loss, or a low serum cholesterol. Ileocolonoscopy in those patients often shows Crohn's disease-like lesions (aphthoid ulceration, histiocytic microgranuloma, pyloric metaplasia, and sarcoid-like granuloma). These instances of subclinical inflammatory bowel disease, mainly Crohn's disease, are underestimated (Hogan *et al*. 1980).

Mielants and Veys (1990) have found similar lesions in patients with spondylarthropathies with a negative history for gut symptoms and in patients with ankylosing spondylitis whose next of kin suffered from proven Crohn's disease. In all patients with Crohn's disease-like lesions, the gut inflammation persisted together with the joint inflammation and was more severe in nature, whereas most acute lesions disappeared.

Systematic ileocolonoscopy has been performed in 30 patients suffering from spondylarthropathies without any gut symptoms and in 18 non-arthritic patients who required this examination. Macroscopic and microscopic lesions were classified according to the descriptions of Mielants and Veys (1990). Mild lesions suggestive of gut inflammation were observed in 10 of the 30 patients and in none of the reference group ($p < 0.02$). Microscopically, there were minimal signs of inflammation (lymphocyte and plasma-cell infiltration) at the ileum level in 18 out of 19 patients compared to 11 out of 18 in the reference group ($p < 0.05$), and at the rectum level in 15 out of 19 compared to 8 out of 17 ($p < 0.05$) (Dougados *et al*. 1987).

Whipple's disease

Whipple's disease is a rare, multisystem disorder first described in 1907. Most of 500 documented cases originate from Europe and the United States. Despite its rarity, some reports mention the occurrence of Whipple's disease among individuals living in close proximity to one another (Maizel *et al*. 1970). Three patients in the same village within a 10-year period and two patients who inhabited the same farmhouse, although several years apart, were reported by Capron *et al*. (1973) and Capron *et al*. (1975).

Clinical features

Clinical manifestations (Rubinow 1988) attributable to virtually every organ system may occur in Whipple's disease and frequently may appear years or sometimes decades before the onset of diarrhoea and malabsorption. Polyarthritis, prolonged fever, weight loss, lymphadenopathy, arterial hypotension, hyperpigmentation, polyserositis, and cardiac and pulmonary symptoms are seen with varying frequency. Central nervous abnormalities including personality changes, memory loss, ataxia, presenile dementia, myoclonus, spastic paresis, seizures, ophthalmoplegia, papilloedema, retrobulbar neuritis, and deafness may appear early in the course of the disease. Anaemia and leucopenia are common haematological manifestations. Only on rare occasions have these features been recognized as prodromes of Whipple's disease.

Polyarthritis is the most common prodromal feature: 60 to 90 per cent of patients have articular involvement, which is the sole initial manifestation of the disease in 50 per cent of them.

The attacks of arthritis are characteristically acute in onset, often transient, and intermittent. They last from a few hours only to a few days and usually remit

spontaneously. More rarely, the attacks may be of longer duration or even continue relentlessly for several years. Migratory oligoarthritis is more frequent than symmetrical polyarthritis or monoarthritis. Some patients complain only of recurrent arthralgia, whereas in others mild to florid synovitis may be observed. In order of frequency the joints affected are the knees, ankles, wrists, elbows, and small joints of the hands and shoulders. Residual joint deformity is usually absent or mild. The attacks of arthritis do not parallel the occurrence of diarrhoea, and the arthritis appears to subside several months to 2 years before the onset of diarrhoea and weight loss. The arthritis usually resolves within 2 months of instituting antibiotic therapy, and the reappearance of joint symptoms may herald a relapse due to premature discontinuation of that therapy. Subcutaneous nodules and clubbing are infrequent features.

Findings in synovial fluid vary with the clinical severity of the arthritis from more than 20 000 cells/mm ^{3}, mainly polymorphs, to lower white counts with a high percentage of mononuclear cells. The histopathological appearances of the synovial membrane parallel the inflammatory response found in the synovial fluid. Large, foamy, vacuolated cells containing periodic acid–Schiff (**PAS**)-positive granules and bacilli in various stages of degradation have been observed during acute episodes of arthritis in two patients, and non-caseating granulomatous lesions with histological features of sarcoidosis (Rouillon *et al*. 1993).

The involvement of axial joints is not as frequent as peripheral arthritis. When strict criteria were applied, only 4 of 95 patients could be classified as either definite ankylosing spondylitis or sacroiliitis. The spondylarthropathy criteria have not been tested on enough patients with Whipple's disease to know how many would be classified in this group of diseases. HLA-B27 antigen is present in over 30 per cent of patients according to Dobbins (1985) and in 1 out of 6 patients described by Khan (1982).

The ability to attribute early arthritic features to Whipple's disease requires a high index of suspicion and astute clinical acumen. Many patients are managed for many years as palindromic rheumatism, atypical rheumatoid arthritis, systemic lupus, or sarcoidosis. With the onset of abdominal pain, diarrhoea, steatorrhea, and weight loss the disease is more readily recognizable.

The diagnosis depends on the presence of PAS-positive macrophages in the jejunum or extraintestinal sites and eventually electron-microscopic demonstration of typical bacilli. Bacterial identification is made possible by applying the polymerase chain reaction to amplify the 16 *s* ribosomal RNA gene sequences in tissue samples (Relman *et al*. 1992).

Pathogenesis

Different organisms easily cultured by routine microbiological methods for the bowel flora have been proposed as Whipple's bacillus. None of them combines the size and the Gram-positive coloration of the intracellular bacilli. By applying the polymerase chain reaction to amplify 16- *s* ribosomal RNA gene sequences, a previously uncultured new bacterium, *Tropheryma whippelii*, was discovered. It has some homology with the actinomycetes. Other bacteria have also been described (Harmsen *et al*. 1994). A dysfunction of monocytes and macrophages was shown by Bjerknes *et al*. (1985): phagocytosis was normal but no intracellular degradation of *E. coli, Streptococcus pyogenes*, or zymogen particles was observed. Degradation was similarly impaired 3 and 9 months after therapy; intracellular binary fission could take place after phagocytosis. Conceivably, impaired degradation of killed or viable bacteria may lead to macrophage overload and impair their antigen-presenting functions.

Treatment

The efficacy of antibiotic therapy in Whipple's disease makes this condition, once universally fatal, totally distinct from arthritis associated with inflammatory bowel disease.

Tetracycline has been the mainstay of most treatment schedules but despite initial resolution in nearly all cases of gastrointestinal and extraintestinal features, within a week to a month after beginning tetracycline (1 g/day), some patients developed late, irreversible lesions of the central nervous system (Knox *et al*. 1976). The present recommendations for the treatment of newly diagnosed patients include parenteral penicillin G, 1.2 million units/day, and streptomycin, 1.0 g/day for 10 days, followed by co-trimoxazole twice daily for 1 to 2 years (Dobbins 1985).

Arthritis–dermatitis syndrome associated with bypass surgery (Utsinger et al. 1988)

Arthritis occurs in 8 to 36 per cent of patients (mostly females) after jejunocolonic bypass. The onset may be at any time in the first 3 years after bypass. The joints commonly affected include the metacarpophalangeal and proximal interphalangeal, wrists, knees, and ankles, sometimes with a symmetrical, rheumatoid-like presentation. The pattern of involvement is most often one of brief remissions. Chronic arthritis, juxta-articular erosions, and spondylitis are very rare. White-cell counts in synovial fluid have ranged from 500 to 39 000 cells/mm ^{3}, of which 10 to 98 per cent are polymorphonuclear leucocytes. Synovial biopsy demonstrates chronic synovitis with lymphocytic predominance but no lymphoid follicles.

Rheumatoid factor and antinuclear antibodies are absent, and the arthritis has no consistent histocompatibility association.

Dermatitis

Skin lesions accompany this arthritis in over 80 per cent of patients. The most common lesions are urticarial or papulovesicular. Lesions progress in 24 to 72 h from initially discrete macules of 2 to 12 mm diameter to pustulovesicles. Lesions are often at different stages of development. Other skin lesions include necrobiosis lipoidica, erythema nodosum, and various forms of cutaneous vasculitis. Jorizzo *et al*. (1983) suggest that the bowel bypass syndrome may occur without actual bypass.

This syndrome is associated with elevated levels of circulating immune complexes and cryoprotein precipitates (Utsinger *et al*. 1988). The antibodies in the complexes are more often to *E. coli* than to *Bacillus fragilis* or to group D streptococci.

Collagenous colitis

This condition was first described by Lindström (1976) and of some 40 patients to date, 10 per cent have had arthritis. Intestinal features consist of intermittent or persistent diarrhoea without bleeding. The majority of the patients are adult women, with a mean age of 56 years. The diagnostic finding is the linear deposition of hyaline material, 1 to 100 µm thick, consisting principally of collagen III, in the subepithelium of the colon. Lymphocytes, plasma cells, and a few eosinophils are also found, and sometimes an inflammatory exudate is present. The histological findings are distinct enough to differentiate the condition from other diseases of the colon (Rams *et al*. 1987). The aetiology is unknown. Treatment is symptomatic. Sulphasalazine and mepacrine have been recommended, with, in some cases, efficacy for both colitis and arthritis (Combe *et al*. 1988).

Gluten-sensitive enteropathy

There are some descriptions of arthropathy in cases of gluten-sensitive enteropathy (Adelizzi *et al*. 1982; Bourne *et al*. 1985; Pinals 1986). Occult coeliac disease in adults (Chakravarty and Scott 1992) and in children (Simoes and Amor 1992) is also associated with arthropathies. Effusive arthritis was present in knees and ankles, and responded promptly to a gluten-free diet. Arthralgia, sclerodactyly, and finger contracture have been described in a 60-year-old man, 10 years after the initial diagnosis of coeliac disease.

Chapter References

Adelizzi, R. A., Pecora, A. A., and Chiesa, J. C. (1982). Celiac disease, case report with an associated arthropathy. *American Journal of Gastroenterology*, **77**, 481–5.

Aho, K., Ahvonen, P., Lassus, A., Sievers, K. and Tiilikainen, A. (1973). HLA antigen 27 and reactive arthritis. *Lancet*, **ii**, 157.

Aho, K., Ahvonen, P., Lassus, A., Sievers, K., and Tiilikainen, A. (1974). HLA 27 in reactive arthritis. A study of Yersinia arthritis and Reiter's disease. *Arthritis and Rheumatism*, **17**, 521–6.

Amor, B. (1979). Reiter's syndrome: long term follow up data. *Annals of the Rheumatic Diseases*, **38** (Suppl. 1), S32–3.

Amor, B. (1985). Chlamydia and Reiter's syndrome. In *Advances in inflammation research. The spondyloarthropathies*, Vol. 9, (ed. M. Ziff and S. B. Cohen), pp. 203–10. Raven, New York.

Amor, B., Feldmann, J. L., Delbarre, F., Hors, J., Beaujan, M. M., and Dausset, J. (1974). HLA antigen W27. A genetic link between ankylosing spondylitis and Reiter's syndrome ? (Letter). *New England Journal of Medicine*, **290**, 572.

Amor, B., Kahan, A., Dougados, M., and Delrieu, F. (1984). Sulfasalazine and ankylosing spondylitis. *Annals of Internal Medicine*, **101**, 878.

Amor, B., Dougados, M., and Mijiyawa, M. (1990). Critères de classification des spondylarthropathies. *Revue du Rhumatisme et des Maladies Ostéoarticulaires*, **57**, 85–9.

Amor, B., Santos, R. S., Nahal, R., Listrat, V., and Dougados, M. (1994). Predictive factors for the long term outcome of spondyloarthropathies. *Journal of Rheumatology*, **21**, 1883–7.

Bardin, T., Enel, C., and Lathrop, M. G. (1990). Treatment by tetracycline or erythromycin of urethritides allows significant prevention of post-venereal arthritic flares in Reiter's syndrome patients. *Arthritis and Rheumatism*, **33**, S26 (Abstr. 97).

Bauer, W. and Engelman, E. P. (1942). A syndrome of unknown aetiology characterized by urethritis. conjunctivitis and arthritis (so-called Reiter's syndrome). *Transactions of the Association of American Physicians, Philadelphia*, **57**, 307–13

Benhamou, C. L. *et al*. (1987). Arthrites costo-vertébrales inférieures de la pelvispondylite rhumatismale. Révélation pseudo urologique. *Revue du Rhumatisme et des Maladies Ostéoarticulaires*, **54**, 203–7.

Beutler, A., Nanagara, R., Schumacher, H. R., and Hudson, A. (1994). Demonstration of *Chlamydia trachomatis* by in situ hybridization in synovial membranes of patients with chronic reactive arthritis. *Arthritis and Rheumatism*, **37**, S233 (Abstr. 438).

Bjerknes, R., Laerum, O., and Degaard, S. (1985). Impaired bacterial degradation by monocytes and macrophages from a patient with treated Whipple's disease. *Gastroenterology*, **89**, 1139–46.

Bourne, J. T. *et al*. (1985). Arthritis and coeliac disease. *Annals of the Rheumatic Diseases*, **44**, 592–8.

Braun, J., Laitko, S., Treharne, J., Eggens, U., Wu, P., Distler, A., and Sieper, J. (1994). *Chlamydia pneumoniae*, a new causative agent of reactive arthritis and undifferentiated oligoarthritis. *Annals of the Rheumatic Diseases*, **53**, 100–5.

Brewerton, D. A. *et al*. (1973*a*). Ankylosing spondylitis and HLA 27. *Lancet*, **i**, 904–8.

Brewerton, D. A. *et al*. (1973*b*). Reiter's disease and HLA-27. *Lancet*, **ii**, 996–8.

Bywaters, E. G. L. and Ansell, B. M. (1958). Arthritis associated with ulcerative colitis. A clinical and pathological study. *Annals of the Rheumatic Diseases*, **17**, 169–83.

Calin, A. (1986). A placebo controlled, cross-over study of azathioprine in Reiter's syndrome. *Annals of the Rheumatic Diseases*, **45**, 653–5.

Capron, J. P. *et al*. (1973). Maladie de Whipple: deux cas. *La Nouvelle Presse Médicale*, **2**, 2478.

Capron, J. P., Thevenin, A., Delamarre, J., *et al*. (1975). La maladie de Whipple. Etude de 3 cas et remarques épidémiologiques et radiologiques. *Lille Médica*l, **20**, 842–5.

Chakravarty, K. and Scott, D. G. I. (1992). Oligoarthritis. A presenting feature of occult coeliac disease. *British Journal of Rheumatology*, **31**, 349–50.

Chamot, A. M., Benhamou, C. L., Kahn, M. F., Beraneck, L., Kaplan, G., and Prost, A. (1987). Le syndrome acné pustulose hyperostose ostéite (SAPHO). Résultats d'une enquête nationale. 85 observations. *Revue du Rhumatisme et des Maladies Ostéoarticulaires*, **54**, 187–96.

Combe, B., Pierrugues, R., Bories, P., Barneon, G., Michel, H., and Sany, J. (1988). Arthrites associées à la colite collagène. A propos de 4 cas traités par la mépacrine. *Revue du Rhumatisme et des Maladies Ostéoarticulaires*, **55**, 929–32.

Csonka, G. W. (1958). The course of Reiter's syndrome. *British Medical Journa*l, **1**, 1088–90.

Dobbins, W. O., 3d (1985). Whipple's disease, an historical perspective. *Quarterly Journal of Medicine*, **56**, 523–31.

Dougados, M., Boumier, P., and Amor, B. (1986). Sulphasalazine in ankylosing spondylitis : a double blind controlled study in 60 patients. *British Medical Journal*, **293**, 911–14.

Dougados, M., *et al*. (1987). Iléocolonoscopie systématique au cours des spondyloarthropathies séronégatives. *Revue du Rhumatisme et des Maladies Ostéoarticulaires*, **54**, 279–83.

Dougados, M., *et al*. (1991). The European Spondylarthropathy Study group preliminary criteria for the classification of spondylarthropathy. *Arthritis and Rheumatism*, **34**, 1218–27.

Dunlop, E. M. C., Al-Hussaini, M. K., Garland, J. A., Treharne, J. D., Harper, I. A., and Jones, B. R. (1965). Infection of urethra by TRIC agent in men presenting because of 'non specific' urethritis. *Lancet*, **i**, 1125–8.

Fiessinger, N. and Leroy, E. (1916). Contribution à l'étude d'une épidémie de dysenterie dans la Somme (juillet–octobre 1916 *). Bulletins et Mémoires de la Société Médicale des Hôpitaux de Paris*, **40**, 2030–69.

Ford, D. K. and Rasmussen, G. (1964). Relationships between genito-urinary infection and complicating arthritis. *Arthritis and Rheumatism*, **7**, 220–7.

Ford, D. K., da Roza D., and Schulzer, M. (1985). Lymphocytes from the site of disease but not blood lymphocytes indicate the cause of arthritis. *Annals of the Rheumatic Diseases*, **44**, 701–10.

Freidank, H. M., Terreri, M. T., Peter, H. H., and Bredt, W. (1993). Comparison of serological tests for the detection of antibodies against *Chlamydia trachomatis* and *Chlamydia pneumoniae* in rheumatological patients. *International Journal of Medical Microbiology, Virology, Parasitology and Infectious Diseases*, **279**, 518–25.

Gounelle, H. and Marche, J. (1941). La maladie rhumatismale post-dysentérique. *Revue du Rhumatisme et des Maladies Ostéoarticulaires*, **8**, 355–401; 415–65.

Granfors, K. *et al*. (1989). Yersinia antigens in synovial-fluid cells from patients with reactive arthritis. *New England Journal of Medicine*, **320**, 216–21.

Greenstein, A. J., Janowitz, H. D., and Sachar, D. B. (1976). The extraintestinal complications of Crohn's disease and ulcerative colitis. A study of 700 patients. *Medicine (Baltimore)*, **55**, 401–12.

Hammer, R. E., Maika, S. D., Richardson, J. A., Tang, J. P., and Taurog, J. D. (1990). Spontaneous inflammatory disease in transgenic rats expressing HLA-B27 and human beta-2 m: an animal model of HLA-B27 associated human disorders. *Cell*, **63**, 1099–112.

Hannu, T. J. and Leirisalo-Repo, M. (1988). Clinical picture of reactive salmonella arthritis. *Journal of Rheumatology*, **15**, 1668–71.

Harkness, A. H. (1950). *Non-gonococcal urethritis*. Churchill Livingstone, Edinburgh.

Harmsen, D., Heeseman, J., Brabletz, T., Kirchner, T., and Muller-Hermelink, H. K. (1994). Heterogeneity among Whipple's-disease-associated bacteria. *Lancet*, **343**, 1288.

Haslock, I. and Wright, V. (1973). The musculo-skeletal complications of Crohn's disease. *Medicine (Baltimore)*, **52**, 217–25.

Hermann, E., Yu, D. T. Y., Meyer zum Büschenfelde, K. H., and Fleischer, B. (1993). HLA-B27-restricted CD8 T cells derived from synovial fluids of patients with reactive arthritis and ankylosing spondylitis. *Lancet*, **342**, 646–50.

Hogan, W. J., Hensley, G. T., and Geenen, J. E. (1980). Endoscopic evaluation of inflammation bowel disease. *Medical Clinics of North America*, **64**, 1083–102.

Jorizzo, J. L. *et al*. (1983). Bowel bypass syndrome without bypass syndrome. Bowel-associated dermatosis syndrome. *Archives of Internal Medicine*, **143**, 457–61.

Keat, A. C., Thomas, B. J., Taylor Robinson, D., Pegrum, G. D., Maini, R. N., and Scott, J. T. (1980). Evidence of *Chlamydia trachomatis* infection in sexually acquired reactive arthritis. *Annals of the Rheumatic Diseases*, **39**, 431–7.

Keat, A., Thomas, B., Dixey, J., Osborn, M., Sonnex, C., and Taylor-Robinson, D. (1987). *Chlamydia trachomatis* and reactive arthritis: the missing link. *Lancet*, **i**, 72–4.

Khan, M. A. (1982). Axial arthropathy in Whipple's disease. *Journal of Rheumatology*, **9**, 928–9.

Khan, M. Y. and Hall, W. H. (1965). Progression of Reiter's syndrome to psoriatic arthritis. *Archives Internal Medicine*, **116**, 911–18.

Kitis, G., Thompson, H., and Allan, R. N. (1979). Finger clubbing in inflammatory bowel disease. Its prevalence and pathogenesis. *British Medical Journal*, **2**, 825–8.

Knox, D. L., Bayless, T. M., and Pittman, F. E. (1976). Neurologic disease in patients with treated Whipple's disease. *Medicine (Baltimore)*, **55**, 467–76.

Kyle, J. (1971). Psoas abscess in Crohn's disease. *Gastroenterology*, **61**, 149–55.

Lauhio, A., Leirisalo-Repo, M., Lahdevirta, J., Saikku, P., and Repo, H. (1991). Double-blind, placebo-controlled study of three-month treatment with lymecycline in reactive arthritis, with special reference to Chlamydia arthritis. *Arthritis and Rheumatism*, **34**, 6–14.

Leirisalo, M. *et al*. (1982). Follow-up study on patients with Reiter's disease and reactive arthritis, with special reference to HLA B27. *Arthritis and Rheumatism*, **25**, 249–59.

Lindström, C. G. (1976). 'Collagenous colitis' with watery diarrhoea—a new entity? *Pathologia Europaea (Bruxelles)*, **11**, 87–9.

Locht, H., Kihlstrom, E., and Lindstrom, F. D. (1993). Reactive arthritis after Salmonella among medical doctors—study of an outbreak. *Journal of Rheumatology*, **20**, 845–8.

Maizel, H., Ruffin, J. M., and Dobbins, W. O., III (1970). Whipple's disease. A review of 19 patients from one hospital and a review of the literature since 1950. *Medicine (Baltimore)*, **49**, 175–205.

Marche, J. (1946). Contribution à l'étude de la maladie rhumatismale post-dysentérique et de certains types de maladies rhumatismales; les syndromes arthro-oculo-uréthro-parotidiens. *Revue du Rhumatisme et des Maladies Ostéoarticulaires*, **13**, 91–5.

Marche, J. (1954). Syndrome de Fiessinger Leroy Reiter et spondylarthrite ankylosante. Parentés et place nosologique. *Revue du Rhumatisme et des Maladies Ostéoarticulaires*, **21**, 320–8.

Masbernard, A. (1959). Le syndrome de Fiessinger–Leroy–Reiter. Enseignements fournis par l'étude de 80 cas observés en Tunisie. *Revue du Rhumatisme et des Maladies Ostéoarticulaires*, **26**, 21–45.

Mielants, H. and Veys, E. M. (1990). The gut in the spondyloarthropathies. *Journal of Rheumatology*, **17**, 7–10.

Moll, J. M. H. (1985). Inflammatory bowel disease. *Clinics in Rheumatic Disease*, **11**, 87–111.

Moll, J. M. H., Haslock, I., Macrae, I. F., and Wright, V. (1974). Associations between ankylosing spondylitis, psoriatic arthritis, Reiter's disease, the intestinal arthropathies and Behçet's syndrome. *Medicine, (Baltimore*), **53**, 343–64.

Montgomery, M. M., Poske, R. M., Barton, E. M., Foxworthy, D. T., and Baker, L. A. (1959). The mucocutaneous lesions of Reiter's syndrome. *Annals of Internal Medicine*, **51**, 99–109.

Noer, H. R. (1966). An 'experimental' epidemic of Reiter's syndrome. *Journal of the American Medical Association*, **198**, 693–8.

Oates, J. K. and Hancock, J. A. (1959). Neurological symptoms and lesions occurring in the course of Reiter's syndrome disease. *American Journal of the Medical Sciences*, **238**, 79–84.

Paronen, I. (1948). Reiter's disease: a study of 344 cases observed in Finland. *Acta Medica Scandinavica*, **131** (Suppl. 212), 1–114.

Pinals, R. S. (1986). Arthritis associated with gluten-sensitive enteropathy. *Journal of Rheumatology*, **13**, 201–4.

Putterman, C, and Rubinow, A. (1993). Reactive arthritis associated with *Clostridium difficile* colitis. *Seminars in Arthritis and Rheumatism*, **22**, 420–6.

Rahman, M. U., Cheema, M. A., Schumacher, H. R., and Hudson, A. P. (1992). Molecular evidence for the presence of chlamydia in the synovium of patients with Reiter's syndrome. *Arthritis and Rheumatism*, **35**, 521–9.

Rams, H. Roger, A. I., and Ghandur-Mnaymneh, L. (1987). Collagenous colitis. *Annals of Internal Medicine*, **106**, 108–13.

Reiter, H. (1916). Über eine bisher unerkannte Spirochäten Infektion (spirochaetosis arthritica). *Deutsche Medizinische Wochenschrift (Leipzig)*, **42**, 1535–6.

Relman, D. A., Schmidt, T. M., MacDermott, R. P., and Falkow, S. (1992). Identification of uncultured bacillus of Whipple's disease. *New England Journal of Medicine*, **327**, 293–301.

Rowe, I. F. and Keat, A. C. S. (1989). Human immunodeficiency virus infection and the rheumatologist. *Annals of the Rheumatic Diseases*, **48**, 89–91.

Rouillon, A., Menkes, C. J., Gester, J. C., Perez-Sawka, I., and Forest, M. (1993). Sarcoid-like forms of Whipple's disease. Report of 2 cases. *Journal of Rheumatology*, **20**, 1070–2.

Rubinow, A. (1988). Whipple's disease. In *Infections in the rheumatic diseases* (ed. L. Espinosa, D. L. Goldenberg, F. Arnett, and G. Alarcon), pp. 361–6. Grune and Stratton, Orlando, FA.

Sairanen, E., Paronen, I., and Mahonen, H. (1969). Reiter's syndrome. *Acta Medica Scandinavica*, **185**, 57–63.

Schachter, J., Barnes, M. G., Jones, J. P., Jnr, Engleman, E. P., and Meyer, K. F. (1966). Isolation of bedsoniae from the joints of patients with Reiter's syndrome. *Proceedings of the Society for Experimental Biology and Medicine*, **122**, 283–5.

Schlosstein, L., Terasaki, P. I., Bluestone, R., and Pearson, C. M. (1973). High association of an HLA antigen, W27, with ankylosing spondylitis. *New England Journal of Medicine*, **288**, 704–6.

Sieper, J., Braun, J., Reichardt, M., and Eggens, U. (1993 *a*). The value of specific antibody detection and culture in the diagnosis of reactive arthritis. *Clinical Rheumatology*, **12**, 245–52.

Sieper, J., Braun, J., Wu, P., Hauer, R., and Laitko, S. (1993 *b*). The possible role of shigella in sporadic enteric reactive arthritis. *British Journal of Rheumatology*, **32**, 582–5.

Simoes, M. and Amor, B. (1992). Juvenile arthritis and coeliac disease. *British Journal of Rheumatology*, **31**, 791.

Simon, A. K., Seipelt, E., Wu, P., Wenzel, B., Braun, J., and Sieper, J. (1993). Analysis of cytokine profiles in synovial T cell clones from chlamydial reactive arthritis patients: predominance of the Th1 subset. *Clinical and Experimental Immunology*, **94**, 122–6.

Steere, A. C. *et al*. (1983). The early clinical manifestations of Lyme disease. *Annals of Internal Medicine*, **99**, 76–82.

Szanto, E. and Hagenfeldt, K. (1979). Sacroiliitis and salpingitis. Quantitative 99m-Tc pertechnetate scanning in the study of sacroiliitis in women. *Scandinavian Journal of Rheumatology*, **8**, 129–35.

Tamura, N., Kobayashi, S., Hashimoto, H., and Hirose, S. (1993). Reactive arthritis induced by *Vibrio parahaemolyticus*. *Journal of Rheumatology*, **20**, 1062–3.

Taurog, J. D. *et al*. (1994). The germfree state prevents development of gut and joint inflammatory disease in HLA-B27 transgenic rats. *Journal of Experimental Medicin*e, **180**, 2359–64.

Tauxe, R. V. *et al*. (1987). *Yersinia enterocolitica* infections and pork: the missing link. *Lancet*, 1129–32.

Temine, P., Privat, Y., Marchand, J. P., Chandon, J. P., Follana, J., and Kopp, F. (1972). Les étapes évolutives d'un cas mortel de syndrome de Fiessinger–Leroy–Reiter. *Bulletin de la Société Française de Dermatologie et de Syphiligraphie (Paris)*, **79**, 334–5.

Toubert, A., Dougados, M., and Amor, B. (1985). Erosive granulomatous arthritis in Crohn's disease. (Letter). *Arthritis and Rheumatism*, **28**, 958–9.

Utsinger, P. D., Spalding, D. M., Weiner, S. R., and Clarke, J. (1988). Intestinal immunology and rheumatic disease: inflammatory bowel disease and intestinal bypass arthropathies. In *Infections in the rheumatic diseases* (ed. L. Espinosa, D. Goldenberg, F. Arnett, and G. Alarcon), pp. 317–41. Grune and Stratton, Orlando, FA.

Valtonen, J. M., Koskimies, S., Miettinen, A., and Valtonen, V. V. (1993). Various rheumatic syndromes in adult patients associated with high antistreptolysin O titers and their differential diagnosis with rheumatic fever. *Annals of the Rheumatic Diseases*, **52**, 527–30.

White, M. H. (1895). Colitis. *Lancet*, **i**, 538.

White, D. G., Woolf, A. D., Mortimer, P. P., Cohen, B. J., Blake, D. R., and Bacon, P. A. (1985). Human parvovirus arthropathy. *Lancet*, **i**, 419–21.

Will, R. K., Amor, B., and Calin, A. (1990). The changing epidemiology of rheumatic diseases: should ankylosing spondylitis now be included? *British Journal of Rheumatology*, **29**, 299–300.

Willkens, R. F. *et al*. (1981). Reiter's syndrome: evaluation of preliminary criteria for definitive disease. *Arthritis and Rheumatism*, **24**, 844–9.

Winchester, R. *et al*. (1987). The co-occurrence of Reiter's syndrome and acquired immunodeficiency. *Annals of Internal Medicine*, **106**, 19–26.

Wright. V. (1978). Reiter's disease. In *Copeman's textbook of the rheumatic diseases*, (5th edn.) (ed. J. T. Schott), pp. 549–66. Churchill Livingstone, Edinburgh.

Wright, V. and Watkinson, G. (1965). Sacro-iliitis and ulcerative colitis. *British Medical Journal*, **2**, 675–80.

Wright, V. and Watkinson, G. (1966). Articular complications of ulcerative colitis. *American Journal of Proctology*, **17**, 107–15.

5.6.1 Pauciarticular-onset juvenile chronic arthritis

David D. Sherry and Elizabeth D. Mellins and Barbara S. Nepom

Pauciarticular-onset juvenile chronic arthritis is the most commonly encountered subset of the chronic childhood arthritides, accounting for 40 to 50 per cent of all patients with juvenile chronic arthritis. By definition its onset occurs before the age of 16 years, active synovitis of at least one joint is present continuously for a minimum of 6 weeks by American Rheumatism Association criteria (Brewer *et al*. 1977) or for 3 months by European League Against Rheumatism criteria (Wood 1978), and a total of four or fewer joints are involved during the first 6 months of disease. Some of the characteristics which make this disease entity quite distinct from other forms of juvenile or adult arthritis include its striking predilection for preschool-age girls, the tendency for involvement of large joints excluding the hip and shoulder, the frequent occurrence of chronic anterior uveitis, the presence of antinuclear antibodies, and unique immunogenetic associations.

In our discussion we do not include a group of patients sometimes called pauciarticular juvenile chronic arthritis type 2. This designation refers to a condition predominantly affecting older boys (over 8 years old), with lower extremity arthritis, in whom HLA B27 is frequently found. These children generally have enthesitis (Rosenberg and Petty 1982) and may have more than four joints involved. The pattern of disease in these children is more consistent with spondylarthropathy than pauciarticular-onset juvenile chronic arthritis (see Chapter 5.5.2).

Epidemiology

Pauciarticular-onset juvenile chronic arthritis is a rheumatic disease distinctly of childhood, and the vast majority of patients are female toddlers. Overall, the disease affects an estimated 20 to 30 per 100 000 children (Gare *et al*. 1987). The female to male ratio is 4:1 (Petty 1979), and may be as high as 7.5:1 among children with iridocyclitis and chronic arthritis (Spalter 1975). The age at onset peaks sharply between 1 and 3 years, but ranges from as young as a few months of age to the teenage years (see Fig. 1 and Cassidy and Petty 1995). Only a limited number of cases in adults have been reported (Chaouat *et al*. 1990; Kenesi-Laurent *et al*. 1991), and the absence of HLA DR5 and of eye involvement in patients in the latter series raises the possibility that pauciarticular arthritis in adults represents a distinct disease entity. No unique racial or geographic associations have been described for this type of arthritis.

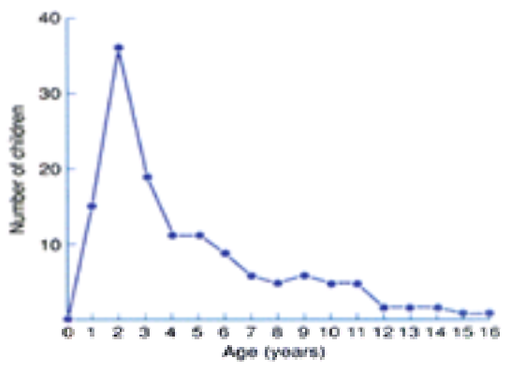


Fig. 1 Age at onset in 137 consecutive children with pauciarticular-onset juvenile chronic arthritis.

Clinical features

Juvenile chronic arthritis is always a diagnosis of exclusion, as there are no pathognomonic signs, symptoms, or laboratory investigations. Nevertheless, the classic clinical picture is quite recognizable: an otherwise healthy female toddler who has arthritis in only a few joints, such as one knee and one ankle (see Fig. 2(a)). Clinical features are usually mild compared with other forms of juvenile chronic arthritis, reactive arthritis, or joint infections (Morrissy 1990; Sherry 1990; Cassidy and Petty 1995). Non-articular inflammation is rare, except for chronic asymptomatic uveitis (discussed in detail below). Many children complain of little pain and are brought to the physician because joint swelling was noted by a parent. In one study, 26 per cent of patients presented without any pain whatsoever (Sherry *et al*. 1990). Those with symptoms usually complain of morning stiffness, gelling, or pain with use, but typically are not incapacitated and limit their activities only modestly. Constitutional signs and symptoms such as fever, malaise, anorexia, or organomegaly are not a part of pauciarticular-onset juvenile chronic arthritis and, if present, virtually exclude this diagnosis.

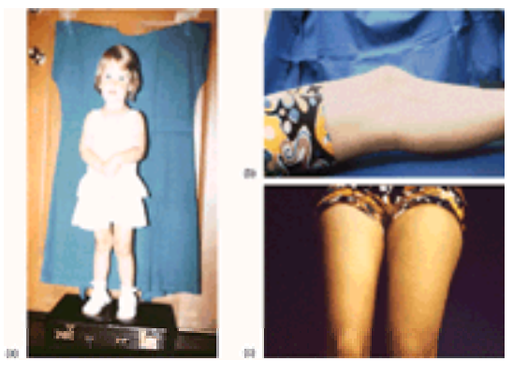


Fig. 2 Clinical features of pauciarticular-onset juvenile chronic arthritis. (a) Archetypical patient with pauciarticular-onset juvenile chronic arthritis. Note generally healthy young girl with involvement of a single knee. (b) Unilateral flexion contracture as seen frequently in pauciarticular-onset juvenile chronic arthritis. To evaluate

for subtle flexion contraction, especially in children with hypermobility, lift both heels equally high and observe for unequal knee height. (c) Thigh atrophy in unilateral disease. This may be a permanent sequela.

In 137 consecutive patients with this condition, we have found that almost half had involvement of the knee, with the ankle joint being the next most frequently affected (Table 1). In general, the joints most frequently involved are reported to be, in order, the knee, ankle, and elbow (Schaller and Wedgwood 1972; Ansell 1977). Involvement of other joints is less frequent, but not uncommon. Children with arthritis in just one or two small joints of the hand do not have the usual pattern of joint involvement, but in our experience, most do not progress to polyarticular changes. Involvement of the shoulder is exceedingly rare. Occasional patients with this pauciarticular-onset arthritis will eventually develop disease of the temporomandibular joint (Strabrun 1991) or cervical spine; the latter can produce torticollis. Arthritis of the hip is so rare in this condition that when that joint is involved at presentation, the diagnosis is suspect and should be made only after extensive evaluation for other causes. There will also be an occasional child who has five or six joints affected over the first 6 months, but whose disease never evolves into the typical polyarticular pattern with symmetrical involvement of many joints including the small joints of the hands. Although meeting criteria for polyarticular-onset juvenile arthritis, these children have a disease more like the pauciarticular-onset disease in terms of their risk for chronic uveitis and long-term prognosis.

	%
Single joint	62
Two joints	31
Three joints	4
Four joints	3
Knee	47
Ankle	21
Small hand joint	12
Wrist	5
Elbow	3
Hip	3

Table 1 Pattern of joint involvement at presentation in 137 consecutive cases of pauciarticular-onset juvenile chronic arthritis

Laboratory features

No particular laboratory abnormalities are diagnostic of pauciarticular-onset juvenile chronic arthritis. Acute-phase reactants are usually normal to slightly elevated. Occasionally one will see a child with synovitis of a few joints and a markedly elevated erythrocyte sedimentation rate. This should prompt an extensive search for infection or occult inflammation, such as inflammatory bowel disease or leukaemia. Reactive or viral arthritis may also produce high erythrocyte sedimentation rates. A rare condition, congenital hyperfibrinogenaemia, will cause a persistently elevated erythrocyte sedimentation rate without other laboratory features of inflammation.

Antinuclear antibodies (**ANA**) are present in 40 to 75 per cent of children with pauciarticular-onset chronic arthritis, depending on the analytical technique employed. These are directed to a heterogeneous mixture of nuclear antigens and usually give a homogeneous pattern, although a speckled pattern can also be observed. Antibodies to histones are relatively common; one report identified them in 42 per cent of patients with uveitis-negative, pauciarticular-onset disease (Malleson *et al*. 1992). ANA titres are usually low, less than or equal to three dilutions beyond the threshold of normal. Higher titres, especially in the older patient, should be investigated for specific autoantigens. For example, antibodies to DNA and extractable nuclear antigens would lead to the suspicion of systemic lupus erythematosus or mixed connective tissue disease. Antinuclear antibodies may develop over time; therefore, an initially negative ANA should be repeated within 1 year, owing to its critical importance in determining uveitis risk (see below). Most studies find no connection between the titre of ANA and disease activity. There is no evidence that ANA precede the development of pauciarticular-onset juvenile chronic arthritis; in fact, most healthy children with an isolated positive ANA will not develop any rheumatic disease when followed over several years (Cabral *et al*. 1992).

Rheumatoid factor is distinctly rare, occurring in less than 5 per cent of children with pauciarticular-onset disease. When it is present, it portends a polyarticular, prolonged course with a high risk of erosive arthritis. Non-classical rheumatoid factors (IgG and IgA) have been reported, but the significance of these is unknown.

Synovial fluid or a synovial biopsy are usually obtained only as an aid to excluding other diagnoses. Synovial fluid most often contains less than 25 000 white blood cells/mm^3, with most of these being polymorphonuclear. The glucose concentration in synovial fluid is within 10 per cent of the serum glucose concentration; protein concentration is elevated.

A synovial biopsy has features similar to those of adult rheumatoid arthritis, with hyperplasia and hypertrophy of the synovial lining, and vascular endothelial hyperplasia with lymphocytic and plasma cell infiltration. Progression of these inflammatory changes to pannus formation and cartilaginous and eventually bony erosion, although uncommon in pauciarticular-onset arthritis, can occur and is indistinguishable pathologically from other forms of juvenile or adult rheumatoid arthritis (Cassidy and Petty 1995).

Radiographs taken early in the disease process reveal joint effusion or soft tissue swelling. Over time, juxta-articular osteoporosis occurs, followed in more severe cases by joint space narrowing and ultimately erosions. However, many bones are cartilaginous in young children (such as the carpals) or have cartilaginous epiphyses. As these ossify, the articular surface may appear quite irregular or multiple ossification centres may occur, giving the appearance of an erosion. Bilateral views are necessary to assess joint space narrowing or growth abnormalities. The expertise of a paediatric radiologist familiar with the spectrum of normal variants in childhood musculoskeletal imaging should be sought.

Course

In most children with pauciarticular-onset arthritis, the disease remains pauciarticular. Although synovitis may develop in new joints over time, the total number of affected joints remains below five. A relatively short course of active arthritis, usually 2 to 5 years, is typical. Some of these children will have a subsequent episode of chronic arthritis, sometimes many years after the original episode. Each episode seems to mimic the first in terms of joints affected and duration of disease.

An important minority will eventually develop arthritis in many joints, yet only about 20 per cent of all children with disease of pauciarticular onset will later be classified as functional stage III or IV (Stoeber 1981; Cush and Fink 1987). There is a growing impression among paediatric rheumatologists that a larger number of children with pauciarticular-onset chronic arthritis may have long-term, active synovitis than was formerly appreciated (Cush and Fink 1987), and this number may not be reflected in the functional classification data. One North American study showed that patients with three or four arthritic joints at onset, or who had involvement of the ankle, wrist, or smaller joints, were more likely to develop a polyarticular course (Cush and Fink 1987; Wallace and Levinson 1991). However, in a European study, those with monoarticular disease developed polyarticular arthritis slightly more frequently than those with the oligoarticular type (Dequecker and Mardjuadi 1982). It is generally only those converting to a polyarticular course who develop marked joint destruction, similar to those with polyarticular-onset juvenile chronic arthritis. A patient whose disease remains pauciarticular for 5 years is unlikely to progress to polyarticular involvement.

Complications

Articular complications

Children with pauciarticular-onset chronic arthritis are at risk of developing juxta-articular muscle atrophy, bony enlargement of the joint, or leg length inequality (Fig. 2(b and c)). Muscle atrophy may be dramatic. When knee synovitis starts before the age of 3 years, quadriceps muscle atrophy may continue years after disease is past (Vostrejs and Hollister 1988). Bony overgrowth around the affected joint is primarily of cosmetic concern. In those children with unilateral knee involvement, bony

overgrowth may lead to a significant leg length discrepancy. When disease begins before 3 years of age, a longer leg develops on the affected side (Vostrejs and Hollister 1988), whereas when disease onset is after the age of 9 years, premature epiphyseal closure can occur, resulting in a shorter leg on the involved side (Simon *et al*. 1981). Mechanically, a longer leg on the involved side may be a problem, since the child will flex the arthritic knee to keep the pelvis level, thus exacerbating a flexion contracture. Rarely, a severe flexion contracture may progress to subluxation. Flexion contracture of the elbow can lead to an apparent foreshortening of the arm, causing a dwarf-like appearance. Effectively controlling joint inflammation can minimize, but not necessarily eliminate, these complications.

Synovial cysts, especially in the popliteal space, are not uncommon (Szer et al.1992). Intra-articular triamcinolone hexacetonide (1 mg/kg) is curative in most, because these cysts communicate with the adjacent joint. Acute onset of intense limb pain and swelling suggests a ruptured cyst.

Ocular complications

In addition to musculoskeletal complications, a chronic, insidious and potentially sight-threatening form of uveitis can occur in children with pauciarticular-onset chronic arthritis. Described first in a case report in 1910 (Ohm 1910), the association of chronic uveitis and juvenile arthritis is now firmly established and available data suggest it is observed worldwide (for historical review see Rosenberg 1987). The ocular inflammation of juvenile chronic arthritis characteristically affects the anterior uveal tract (the iris and ciliary body); involvement of the posterior uveal tract (the choroid) is infrequent (Key and Kimure 1975).

Prevalence studies of uveitis in patients with pauciarticular-onset juvenile chronic arthritis in the United States and Britain suggest that approximately 20 per cent develop chronic uveitis (reviewed in Sherry *et al*. 1991). Uveitis is also seen in a small fraction of patients with polyarticular onset (5 per cent) and more rarely in those with systemic-onset juvenile chronic arthritis (Key and Kimure 1975; Chylack 1977). The risk of uveitis is associated with the pauciarticular mode of onset of arthritis and not with the later extent of articular disease. In two independent studies, polyarticular disease developed in about 45 per cent of pauciarticular-onset patients with uveitis (Kanski and Shun-shin 1984; Wolf *et al*. 1987). Children with uveitis thus appear to contribute disproportionately to the group with pauciarticular-onset whose disease becomes polyarticular.

In addition to pauciarticular onset of arthritis, other risk factors for the uveitis associated with juvenile chronic arthritis have been identified (Table 2). Young girls who are seropositive for ANA and seronegative for rheumatoid factors are at highest risk. Indeed, the risk of developing uveitis for ANA-positive females whose pauciarticular-onset disease began before the age of 2 years is estimated at greater than 95 per cent (Chylack *et al*. 1979). It is not known why ANA are a serological marker of those at risk. Interestingly, these antibodies are not found in patients with chronic anterior uveitis without juvenile chronic arthritis, although this disease has otherwise similar characteristics, including a female preponderance (Rosenberg and Romanchuk 1990). In children with uveitis associated with chronic arthritis, the titre of the ANA is usually intermediate (less than 1:640) and does not correlate with the severity of ocular disease (Kanski 1977). No differences in reactivity to defined nuclear antigens are found in the ANA of children with and without uveitis (Malleson *et al*. 1989). One study of children with pauciarticular-onset chronic arthritis reported a correlation between antibodies to a novel 15-kDa nuclear antigen and uveitis, although some individuals with this antibody did not show ANA positivity in other tests (Neuteboom *et al*. 1992).

Female gender	
Female–male ratio	4.7–7.5 : 1[a]
Young onset age of arthritis	
Mean onset age (years)	4.6
Oligoarthritis	
Percentage with pauciarticular-onset arthritis	86.6
Percentage with polyarticular-onset arthritis	12.6
Percentage with systemic-onset arthritis	0.8
Antinuclear antibody positivity	
Percentage positive	80–90[b]

[a]Reviewed in Cassidy and Petty (1995).
[b]The percentage of uveitis patients with detectable antinuclear antibodies has risen to 90+ per cent since 1975 when sensitive assays became available (Wolf *et al*. 1987).
Table modified with permission from Rosenberg (1987).

Table 2 Risk factors associated with uveitis in juvenile chronic arthritis

The uveitis is commonly asymptomatic, even in the face of substantial damage to the eye. In reported series, pain and redness of the eye occur in up to 25 per cent of patients; visual disturbance, photophobia, and headache are less frequent (reviewed in Cassidy and Petty 1995). Detection of disease generally requires slit-lamp ophthalmological examination. Routine surveillance (slit lamp) of at-risk patients is thus essential, at a frequency determined by the presence of risk factors (Table 3). Biomicroscopic signs of active disease include the presence of inflammatory cells and protein flare in the anterior chamber of the eye and fresh keratitic precipitates. On the initial rheumatological evaluation, one should look carefully for evidence of earlier uveitis, such as pupillary irregularities and punctate corneal deposits (Fig. 3). If present, immediate ophthalmic evaluation is appropriate. If not, regular slit-lamp examinations should be arranged.

Systemic onset:	Yearly
Pauciarticular or polyarticular onset	
Under 7 years of age at onset:	
ANA positive:	Every 3–4 months for 4 years, then: Every 6 months for 3 years, then: Yearly
ANA negative:	Every 6 months for 7 years, then: Yearly
Seven years of age or older at onset:	Every 6 months for 4 years, then: Yearly

These guidelines are for patients who do not develop uveitis during this time. For those patients who develop uveitis, ophthalmological treatment and surveillance will be determined by the ophthalmologist; most patients will require more frequent long-term surveillance.
[a]Taken from the American Academy of Pediatrics Section on Rheumatology and Section of Ophthalmology (1993).

Table 3 Uveitis surveillance: recommended frequency of slit lamp examinations [a]

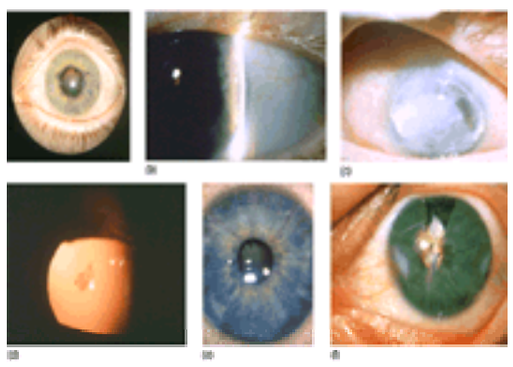


Fig. 3 Ocular involvement in juvenile chronic arthritis. (a) Signs of uveitis visible on clinical examination. This eye has developed pupillary irregularities from synechiae, which adhere the iris to the lens capsule. The dilation of the pupil reveals numerous areas of adhesion. The clouding of the pupillary reflex is caused by keratitic precipitates, which are clumps of inflammatory cells on the posterior surface of the cornea. (b) The white lacy area represents paralimbal band keratopathy, caused by deposition of calcium in Bowman's layer. (c) Extremely severe, untreated band keratopathy. (d) Small posterior subcapsular cataract, seen by retroillumination. Cataract formation may result from either steroid therapy or disease. (e) Chronic iritis with more advanced, complicated cataract. Posterior synechiae are also visible and can induce secondary glaucoma. (f) Ocular findings after surgical intervention for complicated cataract. The pupil has been secondarily scarred

down, after removal of cataract membrane. Ocular inflammation predisposes to this degree of scarring, which severely compromises vision. Current surgical techniques attempt to avoid these complications. Note the horseshoe iridectomy, performed for glaucoma, as well as the paralimbal band keratopathy at 3 and 9 o'clock.

In most children, uveitis and arthritis develop at different times. Uveitis is documented before onset of arthritis in 10 per cent of affected children (Rosenberg 1987), but asymptomatic (and undetected) uveitis may precede arthritis more often. Among routinely screened children with chronic arthritis, uveitis is usually detected within 7 years (median, 2 years) of the onset of arthritis (Kanski and Shun-shin 1984). However, uveitis has developed up to 34 years after joint symptoms (Cassidy *et al*. 1977).

This uveitis is a chronic condition, rarely lasting less than 2 years and often more than 15 years (Smiley 1976). In a study of 20 patients, the course was remitting and relapsing in 60 per cent, persistent in 20 per cent, and limited to a single episode in 20 per cent (Rosenberg 1987). Both eyes are involved in roughly two-thirds of patients (Rosenberg 1987), but both are not necessarily inflamed at the same time. Moreover, children with uveitis of one eye rarely develop involvement of the second eye after more than 1 year of unilateral disease (Kanski 1990). Eye and joint disease evolve independently, and the overall severity of each is likely to differ (Leak and Ansell 1987). In many cases, uveitis persists years after the arthritis has become inactive (Key and Kimure 1975).

A variety of ocular complications can occur in the uveitis associated with juvenile chronic arthritis (Fig. 3). Posterior synechiae, which are fibrous bands adhering the iris to the lens, give the pupil a star-burst or irregular appearance. Band keratopathy, a layer of calcium deposits in Bowman's membrane of the cornea, is another characteristic sequela. The pathogenesis of this condition is unknown and active iritis may persist for years without the development of keratopathy. Other complications include cataracts, glaucoma, and rarely, phthisis bulbi. Approximately 40 per cent of affected eyes progress to 20/200 visual acuity or below, and approximately 10 per cent of affected eyes will become blind (Rosenberg 1987).

Differential diagnosis

Conditions that may simulate pauciarticular-onset juvenile chronic arthritis can be classified into four major categories: monoarticular conditions, short-lived inflammatory conditions, spondylarthropathies, and complaints of pain without joint inflammation. A list of these conditions, along with some of the features distinguishing them from pauciarticular-onset arthritis, is given in Table 4 (see also related chapters in this text).

Table 4 Differential diagnosis of pauciarticular-onset juvenile chronic arthritis

Monoarticular conditions

Monoarticular conditions are the most difficult to sort out and include several requiring immediate intervention. However, it is not uncommon for careful examination of a child with monoarticular complaints to reveal involvement of other joints. When a single joint is involved, pauciarticular-onset juvenile chronic arthritis is a frequent cause, but the conditions described in Table 4 should also be considered.

Short-lived inflammatory conditions

These conditions cause inflammation in one or several joints which can simulate pauciarticular-onset juvenile chronic arthritis, and include inflammatory reactions to a number of viral and bacterial infections. However, the inflammation is relatively fleeting, usually lasting 1 to 4, but occasionally up to 8 weeks. When evaluating a child within the first weeks of disease, one need not withhold anti-inflammatory treatment in order to establish a diagnosis. Only rarely would juvenile chronic arthritis respond so well as to go into complete remission in such a short time.

Spondylarthropathies

These conditions are discussed at length in Chapter 5.5.2, and are contrasted with pauciarticular-onset juvenile chronic arthritis in Table 4. Briefly, unlike pauciarticular-onset disease, they occur predominantly in adolescents, boys are much more commonly affected than girls, the arthritis is usually limited to the lower extremity and involves both large and small joints, enthesitis is common, and HLA B27 is highly associated (Rosenberg and Petty 1982). Heel pain is often noted in spondylarthropathies but is uncommon in pauciarticular-onset arthritis. Furthermore, the iritis that occurs with the spondylarthropathies, in contrast to that of pauciarticular-onset disease, is typically acute and painful, leading to scleral injection and photophobia.

Pain complaints without joint inflammation

Many children present with musculoskeletal pain that does not emanate from the joint. Younger children in particular may not localize their pain accurately. These conditions can be persistent and underlying arthritis may be suspected, but true arthritis is never seen. The most common examples are listed in Table 4.

Treatment

As pauciarticular-onset juvenile chronic arthritis often affects the joints and eyes of very young children, a team approach which provides expertise in paediatric rheumatology, ophthalmology, physical therapy, and sometimes psychosocial issues is optimal. The overall guiding principle in treating these children is to keep the joint(s) as normal as possible while the disease is active so that once it becomes quiescent, the child is left with minimal complications. A corollary to this is to treat ocular inflammation early and thoroughly.

Control of intra-articular inflammation is paramount (Emery 1993). Initially non-steroidal anti-inflammatory drugs (**NSAIDs**) are given (see Table 5 and Ansell 1983; Silver 1988; Duffy *et al*. 1989; Hollingworth 1993). While a few studies have demonstrated the effectiveness of particular NSAIDs, such as aspirin, tolmetin, and naproxen (Levinson *et al*. 1977; Moran *et al*. 1979; Kvien *et al*. 1984), there are no consistent findings as to relative efficacy and tolerance. Therefore, the choice of non-steroidal drug is largely empirical, often dictated by the availability of liquid preparations for small children or by individual response. Twice-daily preparations enhance compliance.

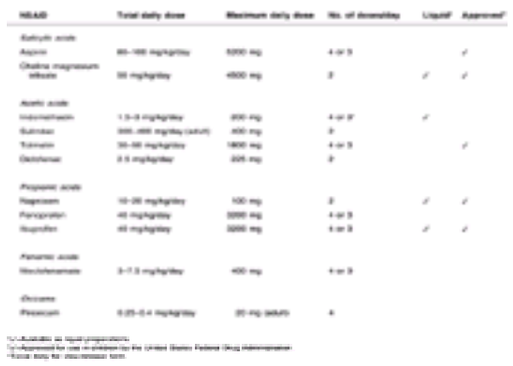


Table 5 Non-steroidal anti-inflammatory drug (NSAID) use in children with juvenile chronic arthritis

Most NSAIDs are tolerated well by children, with little clinical evidence of gastritis. Chemical hepatitis is their most common untoward effect, with aspirin the leading cause (Bernstein *et al*. 1977). We recommend close monitoring when liver function tests such as aspartate aminotransferase reach a level greater than three times normal, and we generally will stop the medication when the level rises to four times normal.

Pseudoporphyria, a skin disorder characterized by skin fragility, vesiculation, and scarring, has recently been reported as a side-effect of NSAIDs, particularly naproxen, among patients with juvenile chronic arthritis (Levy *et al*. 1990; Lang and Finlayson 1994). Both excessive sun exposure and fair complexion seem to increase risk for this complication. Scarring can also occur in the absence of blistering (Wallace *et al*. 1994).

The association of aspirin with Reye's syndrome, though controversial, is of concern, and it is prudent to interrupt aspirin therapy when influenza or varicella infection is suspected. It is not necessary or practical, however, to stop the drug during every viral syndrome in young children. It has been recommended that children treated with aspirin receive an annual influenza immunization (Committee on Infectious Diseases 1994).

Idiosyncratic reactions to all of the NSAIDs can occur. Neurological complications such as depression or personality changes are uncommon, and can be difficult to detect in very young children. Close attention to parental concerns is warranted.

While children often take NSAIDs for years without side-effects, potentially severe complications can occur, such as blood dyscrasias or renal papillary necrosis. Therefore we recommend careful monitoring during their use, including a complete blood count with platelet count, aspartate aminotransferase, alanine aminotransferase, blood urea nitrogen, creatinine, and urinalysis at least twice a year while on a stable dose.

Often an individual child will not respond, or not respond optimally, to a particular preparation. It can be difficult to determine when to move on to a different NSAID, and especially to a disease-modifying drug. Our guidelines are to try three NSAIDs, preferably from three different biochemical classes, at an optimal dose for 2 months each before moving on to other therapy (Lovell *et al*. 1984).

Intra-articular steroids are very useful and can be given in a number of different situations (Allen *et al*. 1986). These injections are warranted in some very young children who have great difficulty taking oral medications, especially those with only a few involved joints. If a single joint is affected, we are more likely to try intra-articular injection early. We will also consider steroid injection when the joint swelling is extreme and unlikely to respond rapidly to NSAID use. Injections can also be very useful if NSAID therapy is not possible owing to severe allergy or toxicity such as renal papillary necrosis.

In addition, we consider synovial cysts an indication for intra-articular injection. Synovial cysts usually completely resolve on injection of the adjacent joint with a steroid preparation (Allen *et al*. 1986). We will also consider intra-articular steroids if the arthritis is severe (intense symptoms or flexion contracture) or prolonged. Unacceptable duration of arthritis must be determined individually, but in our clinic we will inject a joint if it is markedly swollen over 3 months on optimum NSAID therapy. Finally, medical therapy will sometimes control inflammation successfully except for one recalcitrant joint; in that case, intra-articular injection may be a useful adjunct.

Some joints may not respond to an initial injection but respond well to a second attempt. If successful, joint injection can be repeated to a maximum of three injections per year per joint, or a total maximum of six injections per joint. If the arthritis continues beyond this time, more aggressive treatment is required. Intra-articular steroids may lead to localized cutaneous atrophy, hypopigmentation, or intra-articular calcifications, but these are rarely of clinical significance. We use triamcinalone hexacetonide up to 1 mg/kg for the larger joints. Triamcinalone acetonide may also be used.

Some children with prolonged or intense arthritis, or with evolution to a more severe polyarticular course, will require more aggressive medical treatment. In these cases, we frequently use sulphasalazine because of its ease of administration, relative safety, and effectiveness in chronic arthritis (Gedalia *et al*. 1993). Its use is especially attractive in those children whose disease is not erosive, but who have persistent and recalcitrant inflammation. Long-term sequelae such as degenerative joint disease are of concern in patients with persistent inflammation, even in the absence of radiographic changes.

For those children who develop destructive joint changes, more aggressive therapy may be indicated. As in children with polyarticular or systemic-onset chronic arthritis, early recognition and aggressive treatment of destructive disease is critical for optimal outcome. While pauciarticular-onset juvenile chronic arthritis most often carries a favourable prognosis, it is important to monitor patients closely to identify those who will progress to more severe disease. For example, radiographs of affected joints should be obtained yearly in patients with persistent arthritis. Unlike typical adult rheumatoid arthritis, radiographic signs of destructive joint disease may not become evident in juvenile chronic arthritis until many years into the disease course.

For those patients requiring further disease-modifying drugs, there are several options (Rosenberg 1989; Gabriel and Levinson 1990 ; Giannini *et al*. 1993), although few studies have been done in pauciarticular-onset juvenile chronic arthritis. There is some controversy as to the effectiveness of hydroxychloroquine, and one study found no difference between it and placebo (Van Kerckhove *et al*. 1988). Similarly, good data on the use of gold and penicillamine are scant. Methotrexate has been used in some pauciarticular-onset patients with good results (Truckenbrodt and Häfner 1986; Wallace *et al*. 1989; Giannini *et al*. 1992); its safety and efficacy make it an increasingly attractive option in this setting. We currently will often use methotrexate if NSAIDs and sulphasalazine have failed. Methotrexate is usually given at a dose between 0.3 and 0.6 mg/kg per week (usually leaning toward the higher dose) as a single oral, subcutaneous, or intramuscular dose. An occasional patient is given as much as 1 mg/kg per week. We monitor blood counts, liver enzymes, and renal function monthly. The role of liver biopsy in ascertaining toxicity is controversial and presently we do not recommend it (Walker *et al*. 1993).

In those patients whose arthritis completely resolves, the question of when to stop medication arises. The length of treatment depends on the severity and duration of the arthritis; in general, the more difficult it is to achieve remission, the longer the treatment will continue after remission. For example, if the disease lasts 6 to 12 months, we will discontinue medications 3 to 6 months after remission. If the disease lasts longer, medications should be continued for 6 to 12 months past remission. In pauciarticular-onset juvenile arthritis, remission must be a clinical diagnosis, as many children will never have morning stiffness, increased erythrocyte sedimentation rate, or even joint pain. We define remission as the complete clinical absence of active synovitis, in addition to the lack of the conventional signs and symptoms mentioned above. Thus we continue to treat children if they have joint swelling, even if they have no complaints and are fully functional.

Motor activity is so critical to all aspects of childhood, including growth, development, and social interactions, that we are very aggressive in the use of physical therapy to keep the range of motion and muscular strength as normal as possible. However, age-appropriate exercise programmes require experience and creativity. Young children may be particularly difficult to treat because they unconsciously substitute the use of unaffected muscle groups during exercise; for the same reason, normal play activities are not an acceptable alternative to directed physical therapy. Many of these children benefit from seeing a physical therapist experienced in the treatment of these diseases. This can also help decrease parent–child conflicts over exercising.

Although not usually needed, orthotic devices can be very useful in appropriate patients. A few children have tenacious flexion contractures that require serial night splinting or even serial casting to correct. Serial casting is usually carried out three times a week for up to a month. Casting under anaesthesia is done infrequently but can be of great help in difficult situations. Where there is length inequality, shoe lifts on the short side improve gait and encourage full knee extension on the long

side; this also helps keep the quadriceps muscle strong, especially the vastus medialis, which contracts only during the last 10° of knee extension.

Surgical intervention is rarely necessary. Synovectomy for chronic arthritis is controversial and long-term benefit is limited. In one well-designed study, synovectomy did not prevent progressive joint destruction in adult rheumatoid arthritis (Arthritis Foundation Committee on Evaluation of Synovectomy 1988). Although the complications of arthroscopic synovectomy are much lower than with the open procedure, the effort required in rehabilitation makes this procedure inappropriate for younger children. Chemical and irradiation synovectomies in children have not been adequately studied and we have no experience with these modalities. The rare child with subluxation of the knee may require surgical correction; long-term outcome is uncertain regardless of surgery. Leg length inequality can, rarely, persist into adolescence and is amenable to growth-stopping procedures on the long leg (Simon *et al*. 1981).

In some patients, eye involvement proves to be more of a therapeutic challenge than the joint disease. Corticosteroid eye drops and mydriatics to prevent synechiae are the typical initial regimen and are generally thought to be effective in preserving vision in eyes with minimal inflammation (Wolf *et al*. 1987). None the less, in one study, 42 per cent of children had not responded to topical steroids after 6 months, despite early detection and treatment of disease (Chylack 1977). Unresponsive disease is usually treated with subtenon injections of steroid or with oral prednisone; however, these increase the risk of cataract formation and glaucoma. Adjunctive use of NSAIDs may permit a reduction in steroid dose (Olson *et al*. 1988). Experience with immunosuppressive therapy such as azathioprine or chlorambucil is limited (reviewed in Kanski 1990); cyclosporin A has been used in severe, refractory uveitis. Band keratopathy may require chelation with EDTA or corneal scraping. Surgical intervention may also be necessary for cataracts and glaucoma. In the past, surgical treatment has been only marginally successful in these children, but results with microsurgical techniques and laser therapy are improving (Flynn *et al*. 1988; Kanski 1990).

The psychosocial aspects of pauciarticular-onset juvenile chronic arthritis bear some mention, especially as they may affect the ability to deliver therapy. The vast majority of patients are young and resilient and do quite well psychologically. They are not particularly limited and are able to carry out developmental tasks without difficulty. However, control issues may become a source of persistent strife between child and parent. This is especially true of the young child who may dislike taking medicine or exercising and the adolescent who is beginning to individuate from the family and deny imperfections. Moreover, even though pauciarticular-onset arthritis may be a mild condition in the spectrum of chronic childhood diseases, it can be a source of substantial distress in individual families. The degree of other stresses in the family, particularly parental dysfunction, may contribute more to the state of psychological health of children with juvenile chronic arthritis than the disease itself (Daltroy *et al*. 1992). Therefore, appropriate attention to these issues can potentially have a dramatic impact on the well-being of the child.

The child with pauciarticular-onset chronic arthritis is best cared for by an interdisciplinary team consisting of members who are experienced with the complications of this disease and intimately familiar with each other's roles (Brewer *et al*. 1989). This will enhance both family education and team communication. This team should include physicians (a paediatric rheumatologist, ophthalmologist, and orthopaedist), nurses, physical and occupational therapists, a social worker, and, as needed, other health professionals such as a nutritionist. It is with such a team that a uniform, consistent plan of therapy can be initiated and appropriately altered if complications arise. Preventive measures, especially those dealing with psychological or behaviour factors, can be instituted.

Prognosis

Over 80 per cent of children with pauciarticular-onset chronic arthritis suffer little or no musculoskeletal disability at 15-year follow-up (Ansell and Wood 1976; Stoeber 1981; Dequecker and Mardjuadi 1982). The small percentage of those who do poorly is exclusively made up of those children whose disease follows a polyarticular course (Cush and Fink 1987). The majority of the children who become polyarticular do so within the first 5 years of onset. It has been hoped that certain HLA genes might be associated with the subset of children who progress to a polyarticular course and thus predict those at high risk, but at this time there is no consensus that particular genes are helpful in predicting disease course in an individual child (see below).

The outcome in the associated uveitis varies from remission without residua to significant visual loss. A critical determinant affecting outcome is the extent of disease at initial examination (Wolf *et al*. 1987). Eyes which are normal or have mild uveitis at first evaluation do significantly better than those with posterior synechiae (Table 6). In addition, early onset of uveitis in relationship to arthritis correlates with poor outcome (Wolf *et al*. 1987). Ocular prognosis is of greatest concern if uveitis is documented before the onset of arthritis, as progression to symptomatic disease implies significant injury to the eye (Wolf *et al*. 1987). These results strongly suggest that early intervention beneficially influences outcome; indeed, if untreated, uveitis may cause irreversible injury and visual loss in as little as 2 years (Wolf *et al*. 1987). However, there may also be a subgroup of patients with a more benign form of the disease (Smiley 1976). Kanski (1977) reported that 8 out of 26 eyes with continuing active uveitis for more than 10 years remained free of complications or visual loss. Unfortunately, no particular features distinguish such a subset of children at disease onset. The overall prognosis for vision among patients with the uveitis associated with juvenile arthritis has apparently improved in recent years (Sherry *et al*. 1991; Cassidy and Petty 1995). This observed decrease in disease severity is probably due to more comprehensive detection of uveitis, including benign disease, and to more timely treatment.

Cumulative complications	Initial examination	
	Mild uveitis[a] (% of eyes)	Advanced uveitis[b] (% of eyes)
None	64	
Posterior synechiae	12	100
Cataract	28	81[c]
Band keratopathy	5	77[c]
Glaucoma	17	45[c]
Phthisis bulbi	0	13
Final visual acuity < 20/200	3	58[c]

[a]Normal, cells, flare, or keratitic precipitates (*n* = 58 eyes).
[b]Posterior synechiae (*n* = 31 eyes).
[c]$p < 0.001$.
Modified from (Wolff et al. 1987).

Table 6 Frequency of complications of chronic uveitis

Aetiology

The aetiology of pauciarticular-onset juvenile chronic arthritis is unknown. Any satisfactory model of pathogenesis must account for the two most striking features of this disease: the preponderance of young female patients and the frequent association of uveitis. The former suggests a possible contribution of an X-linked gene; the latter may reflect tropism of an infectious agent or involvement of an autoantigen common to the eye and the joint.

One attractive hypothesis is that dysregulation of the immune response is important in disease aetiology. This notion is supported by the observation that disease susceptibility is conferred by particular HLA haplotypes (see below). There is also some evidence of aberrant immune reactivity in these patients (reviewed in Miller 1990). Examples include the presence of ANA, elevated levels of circulating immune complexes, and altered *in vitro* immunoglobulin synthesis. Children with pauciarticular-onset chronic arthritis have been found to have an increased C3d/C3 ratio, suggesting activation of the alternate complement pathway. A correlation between C3d/C3 ratios and the titre of IgG antibodies to lipid A was also observed in pauciarticular-onset disease, suggesting these antibodies may play a role in disease pathology (Olds and Miller 1990). In addition, IgA deficiency and pauciarticular-onset disease have been associated. These findings may represent primary immune abnormalities that contribute to pathogenesis or may be secondary to the disease process. Other indices of immune function in these patients, such as lymphocyte subpopulation ratios and responses to mitogens, are usually normal.

The possibility that juvenile chronic arthritis represents a chronic infection or that it is initiated by an environmental trigger has prompted the search for candidate micro-organisms. No single pathogen has been consistently identified with the development of this disease (Phillips 1988). Rubella virus has been isolated from lymphoreticular cells of 7 of 19 children with chronic rheumatic disease, including 2 of 6 with pauciarticular-onset disease (Chantler *et al*. 1985). Persistence of rubella virus in these patients may reflect an aetiological role for the virus or a state of immunodeficiency in the patients. In two small series, antibody to peptidoglycan, a constituent of bacterial cell walls, was elevated in 25 to 50 per cent of pauciarticular patients with chronic uveitis (Burgos-Vargas *et al*. 1986; Moore *et al*. 1989). This finding may be relevant to disease aetiology because humoral immunity to streptococcal cell wall preparations has been implicated in the pathogenesis of chronic synovitis in animal models (Greenblat *et al*. 1980). Alternatively, these antibodies may reflect a state of altered immune reactivity in children with juvenile chronic arthritis. In this regard, it is of interest that defective antibody responses to immunization with bacteriophage were observed in a study of children with each type of

juvenile chronic arthritis (Ilowite *et al*. 1987).

The aetiology of the associated uveitis is likewise unknown, but appears to be immune mediated. Histopathological studies show non-granulomatous, inflammatory infiltration, including plasma cells and lymphocytes (Sabetes *et al*. 1979). Ocular fluids from affected eyes contain elevated immunoglobulin levels (Sabetes *et al*. 1979; Rahi *et al*. 1977) and ANA (Rahi *et al*. 1977). The antigenic stimulus that initiates or maintains this process has not been identified. Inflammation of both the joint and eye occur in several diseases (e.g. Kawasaki's disease, seronegative spondylarthropathies, inflammatory bowel disease). The possibility that collagen acts as an autoantigen at both sites has been investigated in patients with juvenile chronic arthritis, but studies have failed to demonstrate a heightened immune response to collagen in children with uveitis (Rosenberg 1987). In one study, 29 per cent of patients with pauciarticular-onset juvenile chronic arthritis and uveitis were found to have serum antibodies to a low-molecular-weight fraction of bovine iris proteins. However, the presence of these antibodies did not correlate with severity of uveal inflammation (Hunt *et al*. 1993). Paradoxically, 30 per cent of children with this chronic arthritis and uveitis manifest a humoral response to a retinal antigen, S (Petty *et al*. 1987). Immunization with this protein induces an acute uveitis in animal models, but its role in chronic uveitis is unclear, as there is at present no animal model of chronic uveitis.

Immunogenetics

Several lines of evidence indicate that genetic factors are involved in disease susceptibility (summarized in Maksymowych and Glass 1988; Cassidy and Petty 1995). Multiple cases of pauciarticular-onset juvenile chronic arthritis within one family are unusual, but some have been reported. In addition, concordance for disease is increased in twins; in many cases, clinical disease manifestations between twins are similar (Clemens *et al*. 1985). Most compelling, a large number of studies indicate that certain genes within the *HLA* complex contribute to disease susceptibility (the *HLA* association studies discussed below are reviewed in Nepom 1991; Nepom and Glass 1992; De Inocencio *et al*. 1993; Fernandez-Vina *et al*. 1994).

These reported *HLA* associations are quite complex, in contrast to those of many other rheumatic diseases where a single allele or family of alleles confers disease risk. In pauciarticular-onset juvenile chronic arthritis, not only are a number of different alleles associated with disease, but products of different loci are involved. The sometimes confusing results of these different studies are further confounded by variations in *HLA* typing methods, inclusion criteria, and ethnic and geographic populations. None the less, the reproducibility of many *HLA* associations by many different investigators substantiates their importance in susceptibility to pauciarticular-onset juvenile chronic arthritis.

Initial studies described *HLA* class I associations with pauciarticular disease, with *HLA A2*, the most consistently identified allele. The consensus of more recent studies using precise DNA-based HLA typing techniques is that alleles of the class II loci are most strongly associated. However, there is some evidence that the *HLA-A2 *0201* allele remains significantly associated with disease even after class II associations are taken into account (Fernandez-Vina *et al*. 1994).

Among *HLA* class II genes, several alleles of the *DRB1* locus are the most reproducibly associated with disease. *DR8* (primarily the *DRB1*0801* allele in many studies) shows the strongest correlation, with relative risks of approximately 4 to 12. In a number of different North American and northern and southern European populations, *DR8* accounts for approximately 25 to 50 per cent of patients. The *DR8* association is intriguing, since this allele is infrequent (3 to 10 per cent) in control Caucasian populations, and has not been associated with other autoimmune conditions. Among Caucasians, *DRB1*0801* is strongly linked to the *DQ* genes *DQA1*0401* and *DQB1*0402*, but the association in some studies is stronger with *DR* than *DQ*.

DR5 is also consistently associated with pauciarticular-onset juvenile chronic arthritis, with a relative risk of about 2 to 7 in most studies, slightly lower than that for *DR8*. Several *DR5*+ alleles are present among patients, but *DRB1*1104* alleles are most significantly increased (Melin-Aldana *et al*. 1992; Haas *et al*. 1994).

DR6 has been variably reported to be increased in some populations. The *DRB1*1301* allele appears primarily responsible for this association.

In contrast, *DR4*, which is the *HLA* type most highly associated with both adult and juvenile rheumatoid factor-positive rheumatoid arthritis, is almost never seen in pauciarticular-onset juvenile chronic arthritis. *DR7* also seems to provide protection for this disease.

Interestingly, pauciarticular-onset juvenile chronic arthritis was one of the first diseases shown to be associated with an allele of the *DPB1* locus. Several groups have shown that *DPB1*0201* contributes to disease susceptibility, an effect independent of the *DRB1* contribution and not due to linkage with a particular haplotype. Possessing *DPB1*0201* may not by itself provide enough susceptibility to lead to disease expression; instead, it adds to the risk conferred by the *DRB1* susceptibility alleles.

Gene dosage may play a role in determining the magnitude of risk for this disorder. In addition to the additive contribution of *DPB1* to *DRB1* alleles, some studies have reported increased frequencies of patients heterozygous for two *DRB1* susceptibility alleles. Possessing *A*0201* may also increase risk when added to *DR* and *DQ* susceptibility alleles. Thus, having more than one susceptibility allele may confer additional susceptibility or modulate disease.

Despite quite reproducible *HLA* associations with disease, one cannot rule out the possibility that they are due to involvement of genes linked to these alleles, rather than to the *HLA* genes themselves. Recent studies have examined possible associations with pauciarticular-juvenile chronic arthritis of candidate non- *HLA* genes known to reside near or within the major histocompatibility complex that may contribute to immune processes. So far neither *TAP* nor *LMP* alleles have been reproducibly associated with this disorder (reviewed in Ploski and Førre 1994), although some studies report positive correlations (Donn *et al*. 1994). Other candidate polymorphic genes in the MHC include the genes for tumour necrosis factor, heat-shock protein 70, certain complement components, and other antigen processing genes.

Genes linked to the MHC clearly do not account entirely for disease expression, so, in addition to a hypothesized role for environmental agents, contributing non-*HLA*-linked genes have been sought. The T-cell receptor (**TCR**) for antigen is a leading candidate, but so far no associations with germline TCR polymorphisms have been consistently recognized (Nepom *et al*. 1991). Skewed usage of TCR genes within synovial compartment T cells, however, has been reported (Sioud *et al*. 1992; Bernstein *et al*. 1993). Some groups have found an association of a particular subset of patients with a null allele of the TCR *V*b 6.1 gene (Maksymowych *et al*. 1992; Charmley *et al*. 1994), while others have not (Ploski *et al*. 1993). Another group has noted an increase of null alleles of several different TCR genes among patients (Barron *et al*. 1993). The relevance of these findings to disease susceptibility awaits further confirmation.

Finally, McDowell *et al*. reported an association of certain disease subsets with an allele of the interleukin-1a gene, a lymphokine encoded on chromosome 2, and not linked to the *HLA* region (McDowell *et al*. 1995). Interestingly, this polymorphism occurs in the promoter region of the gene, suggesting intriguing mechanistic possibilities.

Many investigators have hoped that disease-associated alleles will allow prediction of clinical manifestations of pauciarticular-onset juvenile chronic arthritis, such as which patients are at increased risk for uveitis or progression to polyarticular disease. While an association of iritis, ANA, and *DR5* has been observed by some groups, other studies have not consistently borne this out. *DR8* has also been variably linked to the occurrence of uveitis. Similarly, attempts to link *HLA* alleles to severity of joint disease have been confusing. *DR8, DR5*, and a *DR6* haplotype have all been reported to be increased among patients with mild or persistently pauciarticular arthritis. Some reports correlate genes on the *DQA1*0101*/*DRB1*0101* haplotype with severe arthritis. Another links an interleukin-1a allele with chronic iridocyclitis (McDowell *et al*. 1995). Although at present none of these associations is strong enough for use in predicting specific manifestations in an individual patient, it is hoped that further elucidation of immunogenetic associations with specific aspects of this disorder will ultimately lead to improved understanding of disease pathogenesis as well as clinical utility in the management of patients.

Summary

The typical child with pauciarticular-onset juvenile chronic arthritis is a female toddler with chronic arthritis in a couple of large joints. Up to 70 per cent have antinuclear antibodies present, which select out the 20 per cent who will develop chronic asymptomatic uveitis. Education, physical therapy, and non-steroidal agents are usually required, but complications or severe disease may necessitate further therapy. Ocular disease requires prompt, intensive therapy. Many children have joint disease that is active only a few years and 80 per cent suffer no major functional disability at 15 years after onset. The remaining 20 per cent can have widespread destructive arthritis.

Illustrative cases

Case 1

A 5-year-old girl presents with a 1-month history of a swollen, but not very painful, left knee. She is not stiff in the mornings or after prolonged sitting. There is no history of trauma, piercing injury, tick bite, rash, or travel to an area where Lyme disease is endemic. Family history is negative for arthritis, low back pain, enthesitis, acute iritis, inflammatory bowel disease, and psoriasis.

Physical examination is normal except for increased intra-articular fluid in the left knee. Her leg lengths are closely checked for inequality and they are equal. She does not have a knee flexion contracture. When lying supine with both legs elevated to 45°, the left leg is not as hyperextensible as the right. Her left thigh circumference is 1 cm less than the right when measured at a level 7 cm above the patella. All other joints are normal. Her pupils and corneas are closely observed in a darkened room and reveal no evidence of scarring from chronic uveitis.

> At this point we have a young girl who generally feels well, with an isolated swollen left knee. The lack of bony overgrowth indicates that this process has not been prolonged. Because she has no systemic signs, the diagnosis of pauciarticular-onset juvenile chronic arthritis is high on the list. However, a reactive arthritis cannot be ruled out because of the brevity of the episode to date. Lyme disease is unlikely in the absence of travel to an endemic area or a history of exposure.
>
> Laboratory evaluations that would be helpful at this point includes a complete blood count, platelet count, erythrocyte sedimentation rate, urinalysis, ANA, and rheumatoid factor. Had any abnormalities in her eyes been noted, she would have been sent that day to an ophthalmologist; since there were no gross changes, she is instructed to see an ophthalmologist within 2 weeks.

She is begun on an oral anti-inflammatory medication.

Laboratory test results from her first visit show normal blood counts, erythrocyte sedimentation rate, and urinalysis. Rheumatoid factor is absent; ANA are present at 1:80. Report of the slit-lamp examination from the ophthalmologist is normal. Over the next 2 months there is only slight improvement.

At this point the diagnosis of pauciarticular-onset juvenile chronic arthritis is more assured, owing to continuation of symptoms for more than 6 weeks and normal laboratory values except for the positive ANA, which is typical for this condition. In spite of some improvement and minimal functional impairment, however, the arthritis is still active, so other NSAIDs will be tried. Quarterly slit-lamp examinations are scheduled with her ophthalmologist.

After another 2 months on optimal NSAID doses, the patient returns with continued disease.

Persistent swelling in a single joint over this period of time encourages one to think of other diagnoses such as pigmented villonodular synovitis. In addition, if the diagnosis is pauciarticular-onset juvenile chronic arthritis, it would now be reasonable to treat the joint more specifically. Therefore, synovial fluid is removed from her knee and, since infection is not a consideration, intra-articular triamcinolone hexacetonide is given. Her NSAID is continued. Analysis of the fluid shows 12 000 white blood cells (90 per cent polymorphonuclear cells) and 2000 red blood cells.

Her knee becomes much less swollen over the next 2 months, but then fluid reaccumulates.

The relatively large number of red blood cells in the synovial fluid and the recurrent nature of the arthritis makes pigmented villonodular synovitis a concern. Therefore a synovial biopsy is taken. The biopsy shows acute and chronic inflammation, which rules out the diagnosis of pigmented villonodular synovitis. A second injection of intra-articular corticosteroid is given and she subsequently does well. She has no pain or limitation of range of motion. She has a slit-lamp examination of her eyes every 3 months to look for asymptomatic uveitis.

Case 2

At 2 years of age this girl develops arthritis in her knees and one thumb. Laboratory test results show a sedimentation rate of 25 mm/h and normal blood counts. She has a negative rheumatoid factor but does have ANA at a titre of 1:160. She is initially treated with an oral anti-inflammatory medication. An ophthalmologist evaluates her eyes every 3 months for signs of uveitis.

> This little girl shows a classic presentation for pauciarticular-onset juvenile chronic arthritis: involvement of fewer than five joints, primarily large joints, with an essentially normal laboratory evaluation except for the typical positive ANA.
>
> After initial improvement arthritis develops in a fourth joint, an elbow. In addition, about 1 year after presentation, evidence of uveitis is noted during routine ophthalmological surveillance; she is treated with local corticosteroids and mydriatics.
>
> Because of ongoing synovitis, she is placed on sulphasalazine. This is a useful choice when arthritis is clearly not under control and a disease-modifying medication is desired, but the situation does not seem to warrant a cytotoxic drug.
>
> This example also points out the critical importance of routine ophthalmological surveillance. Her uveitis is noted early, when simple local measures will usually bring it under control and prevent potentially devastating visual impairment. In this child's case the uveitis and progression of arthritis occurred concurrently, but more often the two do not occur simultaneously.
>
> Her arthritis is well controlled on sulphasalazine and a NSAID, and her uveitis is brought under control. However, at 5 years of age she develops recurrent synovitis. Over a 2-month period, she begins to complain of morning stiffness for the first time, and she develops arthritis in multiple small joints of her hands, feet, wrists, neck, hips, and ankles. Repeat blood samples and radiographs are obtained.
>
> Her laboratory tests remain unchanged, with continued positive ANA and negative rheumatoid factor. Radiographs of multiple joints show juxta-articular osteoporosis but no erosions. Now this child has moved into the category of pauciarticular-onset juvenile chronic arthritis with evolution to a polyarticular course. She is at great risk for progressive, erosive disease because of the widespread nature of her joint involvement, in spite of the lack of radiographically visible erosions. Children, in contrast to most adults, may not show erosions until many years into their disease. Typically, the ANA will remain positive and the rheumatoid factor negative throughout the course. She now requires the addition of methotrexate to her regimen to help control the arthritis, as well as formal physiotherapy two or three times a week to maintain strength, joint range of motion, and functional activities. As her prognosis is now significantly altered, it must be discussed again at length with her parents.
>
> Over the ensuing years, this girl continues to have occasional flares of synovitis, although her arthritis is generally kept under moderate control with methotrexate. She has developed erosions in a number of joints and will probably have permanent sequelae from her arthritis, such as the eventual need for joint replacement. She had two further episodes of uveitis, but both were noted early and treated aggressively and successfully. She continues to have ophthalmological screening every 3 months.

Chapter References

Allen, R.C., Gross, K.R., Laxer, R.M., Malleson, P.N., Beauchamp, R.D., and Petty, R.E. (1986). Intraarticular triamcinolone hexacetonide in the management of chronic arthritis in children. *Arthritis and Rheumatism*, **29**, 997–1001.

American Academy of Pediatrics Section on Rheumatology and Section of Ophthalmology (1993). Guidelines for ophthalmologic examinations in children with juvenile rheumatoid arthritis. *Pediatrics*, **92**, 295–6.

Ansell, B.M. (1977). Joint manifestations in children with juvenile chronic polyarthritis. *Arthritis and Rheumatism*, **20**, 204–6.

Ansell, B.M. (1983). The medical management of chronic arthritis in childhood. *Annals of the Academy of Medicine* (*Singapore*), **12**, 168–73.

Ansell, B.M. and Wood, P.H.N. (1976). Prognosis in juvenile chronic polyarthritis. *Clinics in the Rheumatic Diseases*, **2**, 397–412.

Arthritis Foundation Committee on Evaluation of Synovectomy (1988). Multicenter evaluation of synovectomy in the treatment of rheumatoid arthritis. Report of the results at the end of five years. *Journal of Rheumatology*, **15**, 764.

Barron, K.S., Deulofeut, J.D., Reveille, J.D., and Robinson, M.A. (1993). T cell receptor (TCR) *V*b null alleles in juvenile rheumatoid arthritis (JRA). *Arthritis and Rheumatism*, **36** (Suppl.), S263.

Bernstein, B.H., Singsen, B.H., King, K.K., and Hanson, V. (1977). Aspirin-induced hepatotoxicity and its effect on juvenile rheumatoid arthritis. *American Journal of Diseases of Children*, **131**, 659–63.

Bernstein, B.H., Miltenberg, A.M.M., van Laar, J.M., Hertzberger, R., and Breedveld, F.C. (1993). T cell receptor rearrangements in juvenile rheumatoid arthritis: a search for oligoclonality. *Clinical and Experimental Rheumatology*, **11**, 209–13.

Brewer, E.J., Jr. *et al*. (1977). Current proposed revision of JRA criteria. *Arthritis and Rheumatism*, **20**, 195–9.

Brewer, E.J., Jr., McPherson, M., Magrab, P.R., and Hutchins, V.L. (1989). Family-centered, community-based, coordinated care for children with special health care needs. *Pediatrics*, **83**, 1055–60.

Burogs-Vargas, R., Howard, A., and Ansell, B.M. (1986). Antibodies to peptidoglycan in juvenile onset ankylosing spondylitis and pauciarticular onset juvenile arthritis associated with chronic iridocyclitis. *Journal of Rheumatology*, **13**, 760–2.

Cabral, D.A., Petty, R.E., Fung, M., and Malleson, P.N. (1992). Persistent antinuclear antibodies in children without identifiable inflammatory rheumatic or autoimmune disease. *Pediatrics*, **98**, 441–4.

Cassidy, J.T. and Petty, R.E. (1995). *Textbook of pediatric rheumatology*, (3rd edn). W.B. Saunders Company, Philadelphia.

Cassidy, J.T., Sullivan, D.B., and Petty, R.E. (1977). Clinical patterns of chronic iridocyclitis in children with JRA. *Arthritis and Rheumatism*, **20**, 224–7.

Chantler, J.K., Tingle, A.J., and Petty, R.E. (1985). Persistent rubella virus infection associated with chronic arthritis in children. *New England Journal of Medicine*, **313**, 1117–23.

Chaouat, D., Chaouat, Y., and Aron-Rosa, D. (1990). Pauciarticular juvenile chronic arthritis with ocular involvement and antinuclear antibody presenting in an adult woman. *British Journal of Rheumatology*, **29**, 236–237.

Charmley, P., Nepom, B.S., and Concannon, P. (1994). HLA and T cell receptor b-chain DNA polymorphisms identify a distinct subset of pauciarticular juvenile rheumatoid arthritis patients. *Arthritis and Rheumatism*, **37**, 695–701.

Chylack, L.T., Jr. (1977). The ocular manifestations of juvenile rheumatoid arthritis. *Arthritis and Rheumatism*, **20**, 217–23.

Chylack, L.T., Dueker, D.K., and Philaja, D.J. (1979). Ocular manifestations of juvenile rheumatoid arthritis: pathology, fluorescein iris angiography and patient care patterns. In *Juvenile rheumatoid arthritis* (ed. J.J. Miller), pp. 149–63. Publishing Sciences Group, Littleton.

Clemens, L.E., Albert, E., and Ansell, B.M. (1985). Sibling pairs affected by chronic arthritis of childhood: evidence for a genetic predisposition. *Journal of Rheumatology*, **12**, 108.

Committee on Infectious Diseases (1994). *1994 Red Book: Report of the Committee on Infectious Diseases* (23rd edn.), p. 280. American Academy of Pediatrics, Elk Grove Village, IL.

Cush, J.J. and Fink, C.W. (1987). Clinical outcome of pauciarticular onset juvenile arthritis. *Arthritis and Rheumatism*, **30** (Suppl.), S34.

Daltroy, L.H. *et al*. (1992). Psychosocial adjustment in juvenile arthritis. *Journal of Pediatic Psychology*, **17**, 277–89.

De Inocencio, J., Giannini, E.H., and Glass, D.N. (1993). Can genetic markers contribute to the classification of juvenile rheumatoid arthritis? *Journal of Rheumatology*, **20** (Suppl. 40), 12–18.

Dequecker, J. and Mardjuadi, A. (1982). Prognostic factors in juvenile chronic arthritis. *Journal of Rheumatology*, **9**, 909–15.

Donn, R.P., Davies, E.J., Holt, P.L., Thomson, W., and Ollier, W. (1994). Increased frequency of TAP2B in early onset pauciarticular juvenile chronic arthritis. *Annals of the Rheumatic Diseases*, **53**, 261–4.

Duffy, C.M., Laxer, R.M., and Silverman, E.D. (1989). Drug therapy for juvenile arthritis. *Comprehensive Therapy*, **15**, 48–59.

Emery, H.M. (1993). Treatment of juvenile rheumatoid arthritis. *Current Opinion in Rheumatology*, **5**, 629–733.

Fernandez-Vina, M., Fink, C.W., and Stastny, P. (1994). HLA associations in juvenile arthritis. *Clinical and Experimental Rheumatology*, **12**, 205–14.

Flynn, H.W., Davis, J.L., and Culbertson, W.W. (1988). Pars plena lensectomy and vitrectomy for complicated cataracts in juvenile rheumatoid arthritis. *Ophthalmology*, **95**, 1114–19.

Gabriel, C.A. and Levinson, J.E. (1990). Advanced drug therapy in juvenile rheumatoid arthritis. *Arthritis and Rheumatism*, **33**, 587–90.

Gare, B.A. *et al*. (1987). Incidence and prevalence of juvenile chronic arthritis: a population survey. *Annals of the Rheumatic Diseases*, **46**, 277–81.

Gedalia, A., Barash, J., Press, J., and Buskila, D. (1993). Sulphasalazine in the treatment of pauciarticular-onset juvenile chronic arthritis. *Clinical Rheumatology*, **12**, 511–14.

Giannini, E.H. *et al*. (1992). Methotrexate in resistant juvenile rheumatoid arthritis. *New England Journal of Medicine*, **326**, 1043–9.

Giannini, E.H. *et al*. (1993). Comparative efficacy and safety of advanced drug therapy in children with juvenile rheumatoid arthritis. *Seminars in Arthritis and Rheumatism*, **23**, 34–46.

Greenblat, J.J., Hunter, N., and Schwab, J.H. (1980). Antibody response to streptococcal cell wall antigens associated with experimental arthritis in rats. *Clinical and Experimental Immunology*, **42**, 450–7.

Haas, J P., Truckenbrodt, H., Paul, C., Hoza, J., Scholz, S., and Albert, E.D. (1994). Subtypes of HLA-DRB1*03, *08, *011, *12, *13 and *14 in early onset pauciarticular juvenile chronic arthritis (EOPA) with and without iridocyclitis. *Clinical and Experimental Rheumatology*, **12** (Suppl. 10), S7–S14.

Hollingworth, P. (1993). The use of non-steroidal anti-inflammatory drugs in paediatric rheumatic diseases. *British Journal of Rheumatology*, **32**, 73–7.

Hunt, D.W., Petty, R.E., and Millar, F. (1993). Iris protein antibodies in serum of patients with juvenile rheumatoid arthritis and uveitis. *International Archives of Allergy and Immunology*, **100**, 314–18.

Ilowite, N.T., Wedgewood, R.J., Rose, L.M., Clark, E.A., Lindgren, C.G., and Owen, M.J. (1987). Impaired in vivo and in vitro antibody responses to bacteriophage OX174 in juvenile rheumatoid arthritis. *Journal of Rheumatology*, **14**, 957.

Kanski, J.J. (1977). Anterior uveitis in juvenile rheumatoid arthritis. *Archives of Ophthalmology*, **95**, 1794–7.

Kanski, J.J. (1990). Uveitis in juvenile chronic arthritis. *Clinical and Experimental Rheumatology*, **8**, 499–503.

Kanski, J.J. and Shun-shin, G.A. (1984). Systemic uveitis syndromes in childhood: analysis of 340 cases. *Ophthalmology*, **91**, 1247–51.

Kenesi-Laurent, M.A., Kaplan, G., and Kahn, M.F. (1991). Oligoarthrites de l'adulte avec facteurs anti-nuclearies. Originalite du syndrome, rapports avec l'oligoarthrite juvenile. *Revue des Rhumatisme et Maladies Osteo-Articulaire*, **58**, 1–6.

Key, S.N. and Kimure, S.J. (1975). Iridocyclitis associated with juvenile rheumatoid arthritis. *American Journal of Ophthalmology*, **80**, 425–9.

Kvien, T.K., Høyeraal, H.M., and Sandstad, B. (1984). Naproxen and acetylsalicylic acid in the treatment of pauciarticular and polyarticular juvenile rheumatoid arthritis. *Scandinavian Journal of Immunology*, **13**, 342–50.

Lang, B.A. and Finlayson, L.A. (1994). Napoxen-induced pseudoporphyria in patients with juvenile rheumatoid arthritis. *Journal of Pediatrics*, **124**, 639–42.

Leak, A.M. and Ansell, B.M. (1987). The relationship between ocular and articular disease activity in juvenile rheumatoid arthritis complicated by chronic anterior uveitis. *Arthritis and Rheumatism*, **30**, 1196–7.

Levinson, J.E. *et al*., (1977). Comparison of tolmetin sodium and aspirin in the treatment of juvenile rheumatoid arthritis. *Journal of Pediatrics*, **91**, 799–804.

Levy, M.L., Barron, K.S., Eitchenfield, A., and Honig, P.J. (1990). Naproxen-induced pseudoporphyria in juvenile chronic arthritis. *Journal of Pediatrics*, **117**, 660–4.

Lovell, D.J., Giannini, E.H., and Brewer, E.J., Jr. (1984). Time course of response to nonsteroidal anti-inflammatory drugs in juvenile rheumatoid arthritis. *Arthritis and Rheumatism*, **27**, 1433–7.

Maksymowych, W.P. and Glass, D.N. (1988). Population genetics and molecular biology of the childhood chronic arthropathies. *Baillière's Clinical Rheumatology*, **2**, 649–71.

Maksymowych, W.P. *et al*. (1992). Polymorphism in a T-cell receptor viable gene is associated with susceptibility to a juvenile rheumatoid arthritis subset. *Immunogenetic*, **35**, 257–62.

Malleson, P., Petty, R.E., Fung, M., and Candido, E.P.M. (1989). Reactivity of antinuclear antibodies with histones and other antigens in juvenile rheumatoid arthritis. *Arthritis and Rheumatism*, **32**, 919–23.

Malleson, P.N., Fung, M.Y., Petty, R.E., Mackinnon, M.J., and Schroeder, M.-L. (1992). Autoantibodies in chronic arthritis of childhood: relations with each other and with histocompatibility antigens. *Annals of the Rheumatic Diseases*, **51**, 1301–6.

McDowell, T.L., Symons, J.A., Ploski, R., Forre, Æ., and Duff, G.W. (1995). A genetic association between juvenile rheumatoid arthritis and a novel interleukin-1a polymorphism. *Arthritis and Rheumatism*, **38**, 221–8.

Melin-Aldana, H. *et al*. (1992). Human leukocyte antigen-DRB1*1104 in the chronic iridocyclitis of pauciarticular juvenile rheumatoid arthritis. *Journal of Pediatrics*, **121**, 56–60.

Miller, J.J. (1990). Immunologic abnormalities of juvenile arthritis. *Clinical Orthopaedics and Related Research*, **259**, 23–30.

Moore, T.L., El-Najdawi, E., and Dorner, R.W. (1989). Antibody to streptococcal cell wall peptidoglycan-polysaccharide polymers in sera of patients with juvenile rheumatoid arthritis but absent in isolated immune complexes. *Journal of Rheumatology*, **16**, 1069–73.

Moran, H. *et al*., (1979). Naproxen in juvenile chronic polyarthritis. *Annals of the Rheumatic Diseases*, **38**, 152.

Morrissy, R.T. (ed.) (1990). *Lovell and Winter's pediatric orthopaedics*, (3rd edn). Lippincott, Philadelphia.

Nepom, B. (1991). The immunogenetics of juvenile rheumatoid arthritis. In *Rheumatic disease clinics of North America* (ed. B.H. Athreya). W.B. Saunders, Philadelphia.

Nepom, B.S. and Glass, D.N. (1992). Juvenile rheumatoid arthritis and HLA: Report of the Park City III Workshop. *Journal of Rheumatology*, **19**, 70–4.

Nepom, B.S., Malhotra, U., Schwarz, D.A., Nettles, J.W., Schaller, J.G., and Concannon, P. (1991). HLA and T cell receptor polymorphisms in pauciarticular juvenile rheumatoid arthritis. *Arthritis and Rheumatism*, **34**, 1260–7.

Neuteboom, G.H., Hertzberger-ten Cate, R., de Jong, J., and van den Brink, H.G. (1992). Antibodies to a 15 kDa nuclear antigen in patients with juvenile chronic arthritis and uveitis. *Investigative Ophthalmology and Visual Science*, **33**, 1657–60.

Ohm, J. (1910). Bandformige Hornhauttrubung bei einem neunjahrigen Mädchen und ihre Behandlung mit subkonjunktivalen Jodkaliumeinspritzungen. *Klinische Monatsblatter für Augenheilkunde*, **48**, 243.

Olds, L.C. and Miller, J.J. (1990). C3 activation products correlate with antibodies to lipid A in pauciarticular juvenile arthritis. *Arthritis and Rheumatism*, **33**, 520.

Olson, N.Y., Lindsley, C.B., and Godfrey, W.A. (1988). Nonsteroidal anti-inflammatory drug therapy in chronic childhood iridocyclitis. *American Journal of Diseases of Children*, **142**, 1289–92.

Petty, R.E. (1979). Epidemiology of juvenile rheumatoid arthritis. In *Juvenile rheumatoid arthritis* (ed. J.J. Miller), pp. 11–31. Publishing Sciences Group, Littleton.

Petty, R.E., Hunt, D.W.C., Rollins, D.F., Schroeder, M.L., and Puterman, M.L. (1987). Immunity to soluble retinal antigen in patients with uveitis accompanying juvenile rheumatoid arthritis. *Arthritis and Rheumatism*, **30**, 287–93.

Phillips, P.E. (1988). Evidence implicating infectious agents in rheumatoid arthritis and juvenile rheumatoid arthritis. *Clinical and Experimental Rheumatology*, **6**, 87–94.

Ploski, R. and Førre, Æ. (1994). Non-HLA genes and susceptibility to juvenile chronic arthritis. *Clinical and Experimental Rheumatology*, **12** (Suppl. 10), S15–S17.

Ploski, R., Hansen, T., and Førre, Æ. (1993). Lack of association with T-cell receptor TCRBV6S1*2 allele in HLA-DQA1*0101-positive Norwegian juvenile chronic arthritis patients. *Immunogenetics*, **38**, 444–5.

Rahi, A.S.H., Kanski, J.J., and Fielder, A. (1977). Immunoglobulins and antinuclear antibodies in aqueous humor from patients with juvenile rheumatoid arthritis (Still's disease). *Transactions of the Ophthalmological Society of the United Kingdom*, **97**, 217–22.

Rosenberg, A.M. (1987). Uveitis associated with juvenile rheumatoid arthritis. *Seminars in Arthritis and Rheumatism*, **16**, 158–73.

Rosenberg, A.M. (1989). Advanced drug therapy for juvenile rheumatoid arthritis. *Pediatrics*, **114**, 171–8.

Rosenberg, A.M. and Petty, R.E. (1982). A syndrome of seronegative enthesopathy and arthropathy in children. *Arthritis and Rheumatism*, **25**, 1041–7.

Rosenberg, A.M. and Romanchuk, K.G. (1990). Antinuclear antibodies in arthritis and nonarthritic children with uveitis. *Journal of Rheumatology*, **17**, 60–1.

Sabetes, R., Smith, T., and Apple, D. (1979). Ocular histopathology in juvenile rheumatoid arthritis. *Annals of Ophthalmology*, **11**, 733–7.

Schaller, J. and Wedgwood, R.J. (1972). Juvenile rheumatoid arthritis: a review. *Pediatrics*, **50**, 940–53.

Sherry, D.D. (1990). Limb pain in childhood. *Pediatrics in Review*, **12**, 39–46.

Sherry, D.D., Bohnsack, J., Salmonson, K., Wallace, C.A., and Mellins, E. (1990). Painless juvenile rheumatoid arthritis. *Journal of Pediatrics*, **116**, 921–3.

Sherry, D.D., Mellins, E.D., and Wedgwood, R.J. (1991). Decreasing severity of chronic uveitis in children with pauciarticular arthritis. *American Journal of Diseases of Children*, **145**, 1026–8.

Silver, R.M. (1988). Nonsteroidal anti-inflammatory drugs in the management of juvenile arthritis. *Journal of Clinical Pharmacology*, **28**, 566–70.

Simon, S., Whiffen, J., and Shapiro, F. (1981). Leg-length discrepancies in monarticular and pauciarticular juvenile rheumatoid arthritis. *Journal of Bone and Joint Surgery*, **63A**, 209–15.

Sioud, M., Kjeldsen-Kragh, J., Suleyman, S., Vinje, O., Natvig, J.B., and Førre, Ø. (1992). Limited heterogeneity of T cell receptor variable region gene usage in juvenile rheumatoid arthritis synovial T cells. *European Journal of Immunology*, **22**, 2413–18.

Smiley, W.K. (1976). The eye in juvenile chronic polyarthritis. *Clinical Rheumatology*, **2**, 413.

Spalter, H.F. (1975). The visual prognosis in juvenile rheumatoid arthritis. *Transactions of the American Ophthalmological Society*, **73**, 554.

Stoeber, E. (1981). Prognosis in juvenile chronic arthritis. *European Journal of Pediatrics*, **135**, 225–8.

Strabrun, A.E. (1991). Impaired mandibular growth and micrognathic development in children with juvenile rheumatoid arthritis. A longitudinal study of lateral cephalographs. *European Journal of Orthodontics*, **13**, 423–34.

Szer, I.S., Klein-Gitelman, M., DeNardo, B.A., and McCauley, R.G.K. (1992). Ultrasonography in the study of prevalence and clinical evolution of popliteal cysts in children with knee effusions. *Journal of Rheumatology*, **19**, 458–62.

Truckenbrodt, H. and Häfner, R. (1986). Methotrexate therapy in juvenile rheumatoid arthritis: a retrospective study. *Arthritis and Rheumatism*, **29**, 800–7.

Van Kerckhove, C., Giannini, E.H., and Lovell, D.J. (1988). Temporal patterns of response to D-penicillamine, hydroxychloroquine, and placebo in juvenile rheumatoid arthritis patients. *Arthritis and Rheumatism*, **31**, 1252–8.

Vostrejs, M. and Hollister, J.R. (1988). Muscle atrophy and leg length discrepancies in pauciarticular juvenile rheumatoid arthritis. *American Journal of Diseases of Children*, **142**, 343–5.

Walker, A.M. *et al*. (1993). Determinants of serious liver disease among patients receiving low-dose methotrexate for rheumatoid arthritis. *Arthritis and Rheumatism*, **36**, 329–35.

Wallace, C.A. and Levinson, J.E. (1991). Juvenile rheumatoid arthritis: outcome and treatment for the 1990s. In *Rheumatic disease clinics of North America* (ed. B.H. Athreya), pp. 891–906. W.B. Saunders, Philadelphia.

Wallace, C.A., Bleyer, W.A., Sherry, D.D., Salmonson, K.L., and Wedgwood, R.J. (1989). Toxicity and serum levels of methotrexate in children with juvenile rheumatoid arthritis. *Arthritis and Rheumatism*, **32**, 677–81.

Wallace, C.A., Farrow, D., and Sherry, D.D. (1994). Increased risk of facial scars in children taking non-steroidal antiinflammatory drugs. *Journal of Pediatrics*, **125**, 819–22.

Wolf, M.D., Lichter, P.R., and Ragsdale, C.G. (1987). Prognostic factors in the uveitis of juvenile rheumatoid arthritis. *Ophthalmology*, **94**, 1242–8.

Wood, P.H.N. (1978) Special meeting on: nomenclature and classification of arthritis in children. In *The care of rheumatic children* (ed. E. Munther), p. 47. EULAR, Basel.

5.6.2 Systemic-onset juvenile chronic arthritis

Ronald M. Laxer and Rayfel Schneider

Introduction

While the earliest English language description of systemic-onset juvenile chronic arthritis dates back to 1897 (Still 1897), Espinel has argued eloquently that Caravaggio's 'Il amore dormiente', painted in the early 1600s, is really the first description of this fascinating entity (Espinel 1994). The painting depicts a young boy with jaundice, recessed jaw, distended abdomen, muscle atrophy, and multiple joint deformities. It was, however, Still's classic paper that highlighted the unique features of systemic-onset juvenile chronic arthritis and differentiated it from both rheumatoid arthritis in adults and other forms of chronic arthritis of childhood. Although a great deal has been written about the clinical aspects of systemic-onset juvenile chronic arthritis, there has been little investigation into the aetiology, pathogenesis, and treatment of this severe, and potentially fatal, form of juvenile chronic arthritis.

While arthritis is required to confirm the diagnosis of systemic-onset juvenile chronic arthritis, true joint inflammation may not be present at the onset of the disease. In fact, patients with otherwise classic features of systemic-onset juvenile chronic arthritis have developed arthritis as late as 9 years after the onset (Calabro *et al*. 1976). In addition to chronic arthritis, the American College of Rheumatology classification criteria require 2 weeks of intermittent fever spikes to at least 103°F (Brewer *et al*. 1977), while the European League Against Rheumatology criteria (Ansell 1978) require at least one other feature (rash, adenopathy, hepatosplenomegaly, pericarditis) in addition to fever and arthritis.

Epidemiology

Variations in classification criteria for juvenile chronic arthritis and the reports of both population-based studies and specialty clinic or hospital-based studies account for the wide prevalence range (8–220 cases/100 000) and wide incidence range (2.6–13.9 cases/100 000) reported for juvenile chronic arthritis (Hochberg 1981; Gewanter *et al*. 1983; Towner *et al*. 1983; Andersson Gare *et al*. 1987). Although difficult to determine, the prevalence probably approximates to 50 cases/100 000 and systemic-onset disease accounts for approximately 10 to 20 per cent of these patients with a reported range of 7 to 43 per cent (Hanson *et al*. 1977; Andersson Gare and Fasth 1992).

In contrast to the two to three fold female predominance for all juvenile chronic arthritis, there is an almost equal sex incidence in systemic-onset disease. However, females may be more commonly affected when the disease begins after age 10 years (Ansell and Wood 1976). Although systemic-onset disease may occur at any age from the neonatal period to adolescence, in two-thirds of patients the onset is under 5 years of age (Ansell 1977) and the mean age at onset is 4 to 6 years (Schaller 1977).

The non-articular features of systemic-onset juvenile chronic arthritis make a viral aetiology an attractive hypothesis, but there is little direct evidence to support this. Seasonal variation in disease onset, with a higher incidence in the late spring, summer, and autumn, has been reported (Lindsley 1987), but is not corroborated by others except for a similar variation in a specific geographic region in a Canadian study (Feldman *et al*. 1996).

HLA associations

HLA studies have demonstrated genetic heterogeneity in patients with systemic-onset disease (Maksymowych and Glass 1988; Nepom 1991). Inconsistent and weak associations have been reported for a variety of class I antigens. Stronger associations have been reported for class II antigens including DR5 (relative risk 4–7) and DR8 (relative risk 4–5), which have both been associated with pauciarticular juvenile chronic arthritis, and DR4 (relative risk 2–5), which has been associated with rheumatoid factor positive polyarticular disease (De Inocencio *et al*. 1993). Unlike rheumatoid factor positive arthritis, there has not been a striking association with DR4 homozygosity and there are conflicting reports regarding the prognostic significance of DR4 in systemic disease (Singh *et al*. 1989; Bedford *et al*. 1992). As with other HLA associations, the DR4 association with systemic disease does not appear to hold true for all systemic disease populations (Nepom and Glass 1992). An association with Dw7 has also been reported (Stasny and Fink 1979).

Clinical features (Table 1)

Very common
Spiking fever (with chills and sweats)
Evanescent rash
Myalgias
Arthralgias
Arthritis
Common
Generalized lymphadenopathy
Hepatosplenomegaly
Polyserositis
Anorexia
Weight loss
Rare
Myocarditis
Coagulopathy
Ocular involvement
Central nervous system involvement
Haemophagocytic syndrome
Primary pulmonary disease
Renal involvement
Amyloidosis

Table 1 Clinical features during the course of systemic-onset juvenile chronic (rheumatoid) arthritis

Extra-articular manifestations

The hallmark of systemic-onset juvenile chronic arthritis is the systemic toxicity and extra-articular features that occur during 'attacks' or flares of disease.

Fever

Fever is the one extra-articular manifestation that is absolutely essential to make a diagnosis. A typical fever pattern occurs in the majority of cases, and is described as a quotidian, or occasionally double quotidian, fever pattern (Fig. 1). During typical flares, the temperature will spike rapidly in the late afternoon or early evening, to at least 39°C, only to return fairly quickly to normal, or often below normal, even without antipyretics. This fever pattern is very different from the fever of other connective tissue diseases, acute rheumatic fever, malignancy, or infectious fevers, which tend to be more persistent or while spiking do not return to below the baseline. Occasionally, the classic fever pattern is only established when anti-inflammatory treatment is begun. Characteristically, the patient will appear toxic during the spike of fever, may have chills and rigors, and will complain of severe arthralgias and myalgias. The bed sheets are often soaked as the fever resolves. Frequently, rash (see below) is present only during the fever spikes. When the fever subsides, the patients are usually much more comfortable and appear less toxic. The fever must be present for at least 2 weeks to satisfy classification criteria, and may persist for months even with treatment. Rarely, fever may follow the development of arthritis (Prieur *et al*. 1984).

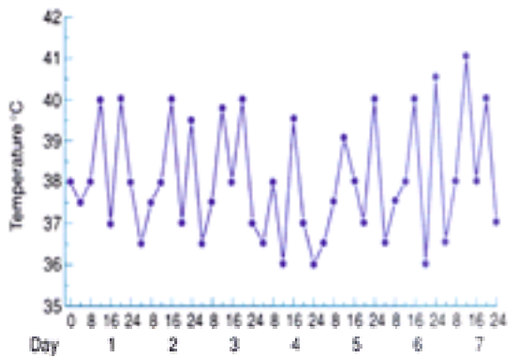


Fig. 1 Temperature chart of a patient demonstrating daily (quotidian) or double-daily fever spikes with rapid return to below the baseline of 37°C.

Rash

The rash of systemic-onset juvenile chronic arthritis is a salmon-pink colour, and is most prominent over the chest, abdomen, back, and intertriginous areas. Distal involvement of the extremities is uncommon but can occur, as can a facial rash. Individual lesions are 3 to 5 mm in diameter and often demonstrate central clearing within the surrounding pink border (Fig. 2(a)). The rash may coalesce into large lesions. The rash is usually macular but urticarial lesions may occur. While the rash typically comes and goes with the spikes of fever (evanescent), it may sometimes be persistent and even occur without fever or other systemic manifestations. The rash may also demonstrate the Koebner phenomenon—the exaggeration of rash in areas of minor trauma. In systemic-onset juvenile chronic arthritis this may be induced by the bed sheets 'squeezing' the skin with a resultant linear distribution of rash (Fig. 2(b)). If the rash is pruritic (10 per cent of patients), then scratching may also induce the Koebner phenomenon. As the rash does tend to come and go, it is unusual to obtain skin biopsies. Histologically, the lesion typically shows a sparse, perivascular infiltrate with a predominance of polymorphonuclear leucocytes (Isdale and Bywaters 1956).

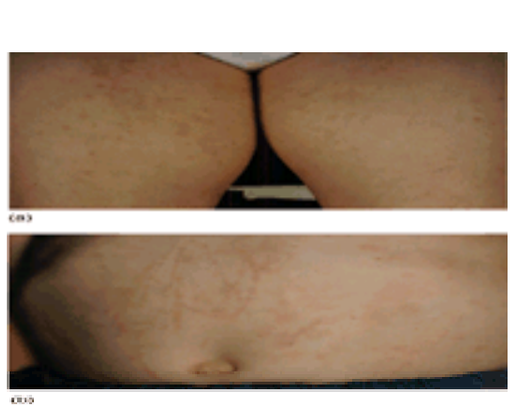


Fig. 2 (a) Rash of systemic-onset juvenile arthritis showing characteristic salmon-pink macular eruptions with central clearing. (b) Koebner phenomenon (appearance of exaggeration of rash in areas of minor trauma).

Reticuloendothelial involvement

Reticuloendothelial hyperplasia with hepatomegaly, splenomegaly, and generalized lymphadenopathy is a common feature of systemic-onset juvenile chronic arthritis. Although the organomegaly is usually asymptomatic, rarely patients with active systemic disease may develop massive hepatomegaly accompanied by a severe abdominal pain (Schaller *et al*. 1970). Mild elevations of serum transaminases occur frequently and are usually not clinically significant. The degree of transaminase elevation does not correlate with the extent of liver enlargement. Aspartate aminotransferase is more frequently elevated than alanine aminotransferase and significant hyperbilirubinaemia is rare. Liver histology in systemic-onset juvenile chronic arthritis is characterized by non-specific periportal inflammatory cell infiltration and Kupfer cell hyperplasia. Fatty change, intrahepatic cholestasis, and fibrosis have been reported less frequently (Tesser *et al*. 1982; Hadchouel *et al*. 1985; Agarwal *et al*. 1994). While transaminase elevations may be seen during active systemic disease, such elevations tend to occur rather sporadically and unpredictably (Rachelefsky *et al*. 1976). This may make it difficult to determine whether hypertransaminasaemia reflects disease activity or the effect of treatment with potentially hepatotoxic medications including salicylates, non-steroidal anti-inflammatory drugs (**NSAIDs**), gold, or methotrexate. Chronic liver disease does not occur in systemic-onset juvenile chronic arthritis but has been described in adult-onset Still's disease (Tesser *et al*. 1982).

Patients with systemic-onset juvenile chronic arthritis may rarely develop hepatomegaly with acute, severe hepatic dysfunction and even fulminant hepatic failure (Hadchouel *et al*. 1985). This potentially life-threatening syndrome has been associated with neurological involvement (drowsiness and coma) and disseminated intravascular coagulation and bleeding. (For a further description see below under ' Less common features'.)

Serositis

Involvement of serosal surfaces is one of the hallmarks of systemic-onset juvenile chronic arthritis (Yousefzadeh and Fishman 1979). Pericardial involvement is most common and pleuritis is more common than peritonitis. Pericarditis is frequently asymptomatic and is best detected by 2D echocardiography. In one series of children with all forms of juvenile chronic arthritis, pericardial involvement was seen in 45 per cent of autopsy cases, although it was only clinically recognized in 7 per cent, most of whom had systemic-onset juvenile chronic arthritis (Lietman and Bywaters 1963). In another series, 9 out of 57 patients had symptomatic pericardial involvement, which was isolated in five cases but associated with myocarditis in the remaining four (Goldenberg *et al*. 1992). In the majority, cardiac involvement occurred within the first year but it can occur at any time during the disease course, particularly in association with other systemic manifestations. Most children will have echocardiographic evidence of pericarditis during systemic flares but may not have any clinical signs of pericarditis (Brewer *et al*. 1977). Occasionally, patients present only with pericarditis and fever before a diagnosis of systemic juvenile chronic arthritis is considered. Therefore, pericardiocentesis may be performed, both for relief of symptoms and diagnosis. Manifestations include tachycardia (out of keeping with the fever), anterior chest pain, and a pericardial friction rub. More severe

involvement may lead to dyspnoea, tachypnoea, and even right-sided congestive heart failure. Rarely, cardiac tamponade and constrictive pericarditis may result (Yancey *et al*. 1981; Pearl 1982; Goldenberg *et al*. 1990). The electrocardiogram may be normal, show ST wave elevation, or non-specific ST segment changes. The chest radiograph may show enlargement of the cardiothoracic silhouette.

Pleuritis is much less common than pericarditis. Occasionally, large pleural effusions, associated almost invariably with pericarditis, may dominate the clinical picture. Sterile peritonitis may result in severe abdominal pain (Bhettay and Thomson 1985).

Arthritis

Most patients have arthritis at disease onset and those who do not commonly have arthralgias or myalgias with the fever spikes. Patients with only arthralgias or no joint symptoms pose substantial diagnostic challenges. To complicate this, the duration from the onset of fever to the development of arthritis may range from just a few weeks to several years. Most patients, however, develop arthritis within 3 months of disease onset (Ansell 1977). Chronic, persistent arthritis evolves in half to two-thirds of patients (Schaller and Wedgwood 1972; Calabro *et al*. 1976).

In our series of 38 patients followed for at least 2 years, pauciarthritis was present in 55 per cent at onset and polyarthritis in 35 per cent, while 10 per cent had no objective signs of arthritis. Although more patients had a pauciarticular onset, the majority of those with persistent arthritis evolved to a polyarticular disease. In more than 75 per cent of patients, the wrists, knees, and ankles are involved. Although somewhat less commonly affected, involvement of the cervical spine, hips, and temporomandibular joints is quite characteristic. In fact, hip involvement occurs in about 50 per cent of patients, is almost always bilateral, and is usually associated with polyarthritis. The hip and wrist joints are the most frequent sites of progressive and advanced destructive changes (Fig. 3(a, b), and Fig. 4), which may occur as early as the first year after onset (Svantesson *et al*. 1983). Almost one-third of patients who have hip involvement may require total hip arthroplasty (Hayem *et al*. 1994). Of the small joints, the hands are more commonly affected than the feet. Cricoarytenoid arthritis with resultant laryngeal stridor has been described (Jacobs and Hui 1977). Tenosynovitis of the carpus and tarsus is common.

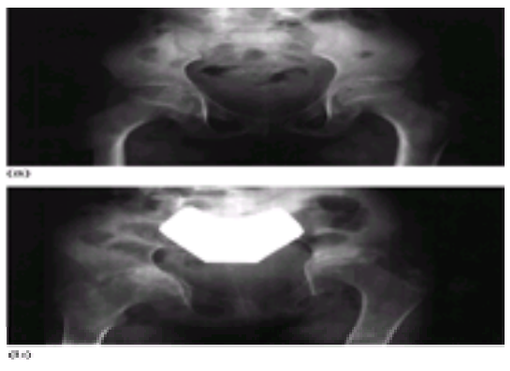


Fig. 3 Girl with systemic-onset juvenile chronic arthritis since 7 years. Hip radiographs show: (a) At 33 months after disease onset, osteopoenia, joint space narrowing, and erosions. (b) At 4.5 years after disease onset, increased loss of joint space, protrusio acetabulae, subchondral irregularity and erosions, and sclerosis of both sides of the joints. There is also flattening of the left femoral head.

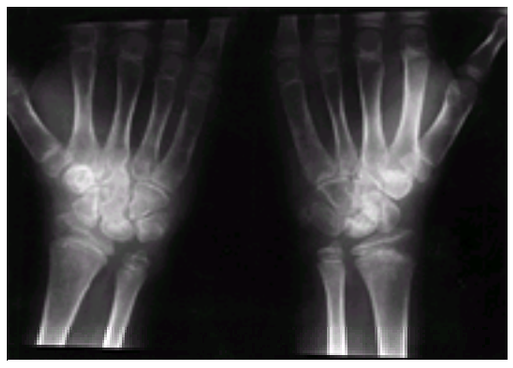


Fig. 4 Same patient as in Fig. 3. Wrist radiographs show advanced changes 2 years after disease onset with moderate narrowing of the carpus, sclerosis, carpal irregularity and erosions, and deformity of the distal radial epiphysis.

Less common features

Many organ systems can be involved in addition to those mentioned above. Pulmonary disease (other than pleuritis) is rare, but can involve the pulmonary parenchyma (Calabro *et al*. 1976; Athreya *et al*. 1980; Wagener *et al*. 1981; Zaglul *et al*. 1982). Primary pulmonary hypertension has been reported (Padeh *et al*. 1991).

Myocarditis, although rare, can be a serious event with considerable mortality. It has been reported to occur in up to 12 per cent of a series of patients with systemic-onset juvenile chronic arthritis from Brazil, and symptoms include tachycardia, dyspnoea, and congestive heart failure (Goldenberg *et al*. 1992). Typically, this occurs during systemic flares of disease and usually together with pericarditis. Myocarditis should be suspected in patients who have persistent tachycardia (out of keeping with fever or anaemia), cardiomegaly, and congestive heart failure. The diagnosis can be made by electrocardiography (showing increased PR interval and low voltages) and echocardiography (showing reduced ventricular function). Because of the high mortality associated with myocarditis, there must always be a high index of suspicion for its development. Although it has been suggested that digitalis may promote arrhythmias (Miller 1977), a recent report did describe using digitalis with good effect (Goldenberg *et al*. 1992). Although cardiac murmurs are common (resulting from anaemia and fever) valvular disease itself is almost never seen, and this helps differentiate systemic-onset juvenile arthritis from acute rheumatic fever with carditis, where valvular inflammation is prominent. However, there are rare reports of valvular abnormalities (Kramer *et al*. 1983; Heyd and Glaser 1990).

Over the last dozen years, several reports have described a syndrome in patients with systemic-onset juvenile chronic arthritis marked by fever, hematocytopenias, hepatic dysfunction, encephalopathy, and disseminated intravascular coagulation with bleeding. This syndrome has been reported under a variety of names. It has occurred following viral infections as well as following changes in medical therapy, but can also occur spontaneously. Most recently, Stephan *et al*. reported three patients with systemic-onset juvenile chronic arthritis (and one with polyarticular juvenile arthritis) who had 'macrophage activation syndrome' (Stephan *et al*. 1993). This seems to be an appropriate term to include the 'consumptive coagulopathy' described by Silverman *et al*. (Silverman *et al*. 1983), a syndrome of 'acute haemorrhagic, hepatic, and neurological manifestations' described by Hadchouel *et al*. (Hadchouel *et al*. 1985), and the several reports of haemophagocytic syndromes that have followed viral infections (Heaton and Moller 1985; Morris *et al*. 1985) in children with systemic-onset juvenile chronic arthritis. The features of this syndrome include persistent high fever (different from the quotidian fever of systemic-onset juvenile chronic arthritis), lymphadenopathy and hepatosplenomegaly, bruising and mucosal bleeding, hepatic dysfunction, drowsiness, and even coma. The laboratory features include anaemia, neutropenia, thrombocytopenia, and evidence of disseminated intravascular coagulation (low fibrinogen levels, raised levels of fibrin degradation products). Deficiency of clotting factors may result in a raised PT and PTT. In fact, children with systemic-onset juvenile chronic arthritis may have increased levels of factor VIII related antigen, fibrinopeptide A, fibrinogen, and fibrin split products, with normal levels of platelet factor IV, indicative of a vasculopathy (Scott *et al*. 1984), even in the absence of this syndrome. Bone marrow and lymph nodes usually show histiocytic consumption of red cells and platelets (Fig. 5). This syndrome is associated with considerable morbidity and mortality. Early recognition and supportive management is vital in reducing mortality. While some patients do recover with expectant management alone, treatment recommendations have included both corticosteroids and cyclosporin (Hadchouel *et al*. 1985; Stephan *et al*. 1993).

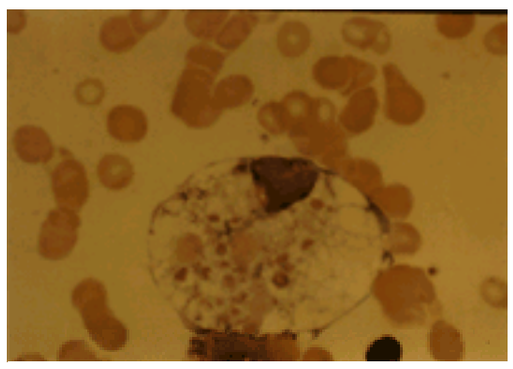


Fig. 5 Bone marrow examination of a patient with macrophage activation syndrome showing histiocytic phagocytosis of red blood cells and platelets (by courtesy of Dr A. Poon).

Central nervous system manifestations are dominated by irritability and lethargy during the fever spikes. True organic brain syndrome is rare and no case series has actually examined patients for central nervous system involvement. However, occasional cases of central nervous system vasculitis have been documented and electrocardiographic changes may be seen (Jan *et al*. 1972). Autopsy series have shown perivascular infiltrates of chronic inflammatory cells in the brain. Two patients with systemic-onset juvenile chronic arthritis treated with long-standing corticosteroids developed epidural lipomatosis and presented with signs of spinal cord compression requiring emergency laminectomy (Arroyo *et al*. 1988).

Renal involvement may occur as a complication of treatment or may indicate the onset of amyloidosis. Although mild abnormalities, including proteinuria and mild haematuria (Antilla 1972), may be seen, particularly with fever, significant renal disease is not a component of systemic-onset juvenile chronic arthritis and its presence should raise suspicion about the diagnosis. Significant proteinuria in the presence of long-standing systemic-onset juvenile chronic arthritis is an indication to exclude amyloidosis with appropriate tissue biopsies.

Ocular involvement is distinctly unusual relative to other forms of juvenile chronic arthritis, but asymptomatic uveitis does occur. It is recommended that these patients be screened annually (Anonymous 1993). Several cases of tenosynovitis of the superior oblique muscle (Brown's syndrome) have been reported (Wang *et al*. 1984; Moore and Morin 1985).

Amyloidosis

Amyloidosis is a serious complication of all subtypes of juvenile chronic arthritis and is associated with significant morbidity and mortality. The clinician should be particularly suspicious of the diagnosis in the systemic subtype in which it occurs most frequently. Although rarely reported in North America, 9 to 10 per cent of patients with systemic-onset disease in European series have developed this complication (Ansell and Wood 1976; Stoeber 1981; Svantesson *et al*. 1983). The reason for this discrepancy in incidence is unclear. No HLA allele has been associated with amyloidosis; however, a restriction fragment length polymorphism related to the amyloid-P component gene has been associated with the development of amyloidosis in patients with systemic-onset disease (Woo *et al*. 1987). Serum amyloid-A protein is usually elevated in amyloidosis but is not predictive of its development (Scheinberg and Benson 1980); however, persistent elevation of the C-reactive protein level may predict the development of amyloidosis (Gwyther *et al*. 1982).

David (David *et al*. 1993), in the largest series of patients with juvenile chronic arthritis and amyloidosis, reported that 57 per cent of these patients had systemic-onset disease. The interval between the onset of disease and the diagnosis of amyloidosis varied widely from 1.5 to 25 years. Ninety per cent of patients had active synovitis at the time of diagnosis. The presenting clinical features and causes of death in this series are summarized in Table 2. The age at disease onset and the duration of disease activity were not related to the development of amyloidosis (Schnitzer and Ansell 1977). The diagnosis should be suspected in patients with persistent proteinuria. The clinical features are generally accompanied by laboratory evidence of an acute-phase reaction with anaemia, thrombocytosis, elevation of the erythrocyte sedimentation rate and C-reactive protein, hypergammaglobulinaemia, and hypoalbuminaemia. Confirmation of amyloidosis is most reliably achieved by renal, rectal, or even subcutaneous fat biopsy. Scintigraphy using a radio-iodinated serum amyloid-P component has been shown to be a useful, non-invasive technique for detecting amyloid deposits in both suspected and occult sites and may be useful in monitoring the response to therapy (Hawkins *et al*. 1993).

	Percentage
Clinical features at diagnosis	
Proteinuria	100
Active juvenile chronic arthritis	90
Oedema	53
Hypertension	25
Abdominal pain	22
Hepatomegaly	22
Splenomegaly	19
Diarrhoea	13
Renal failure	3
Ascites	3
Causes of death	
Renal failure	82
Infection	12
Malignancy	3
Other	3

Modified from Ansell *et al*. (1980)

Table 2 Amyloidosis in juvenile chronic arthritis

Treatment is aimed at controlling the underlying inflammatory process since no therapy has been effective in removing amyloid deposits. In the British series (David *et al*. 1993), chlorambucil was found to have a significant impact on survival. Ten years after amyloidosis was detected, 80 per cent of patients treated with chlorambucil survived compared with less than 25 per cent of patients who did not receive cytotoxic therapy. After 15 years of follow-up, two-thirds of chlorambucil-treated patients were still alive. It should be noted that not all patients with systemic disease are chlorambucil-responsive (Deschenes *et al*. 1990) and its potential toxicities include malignancy and infertility. There is no conclusive role for other immunosuppressive agents.

Growth and nutrition

Children with systemic-onset juvenile chronic arthritis frequently have abnormalities of growth, as documented by Still himself (Fig. 6). The systemic features of the disease are associated with hypercatabolism, resulting in breakdown of tissue stores, and also lead to anorexia. Poor energy and nutrient intake has been noted in several studies (Bacon *et al*. 1990; Mortensen *et al*. 1990). In addition, the frequent requirement for corticosteroids will have an effect on growth. Daily intake of corticosteroids equal to or greater than 5 mg/m^2 will result in growth delay (Blodgett *et al*. 1956). While alternate-day steroid therapy will suppress growth to a lesser degree than daily treatment (Byron *et al*. 1983), this is often difficult to achieve in severely ill systemic-onset juvenile chronic arthritis patients. In addition, Bernstein *et al*. have documented that patients with systemic-onset juvenile chronic arthritis treated with corticosteroids had lower growth velocities than a similar group of systemic lupus erythematosus patients treated with steroids, suggesting that the disease itself has a growth suppressing effect (Bernstein *et al*. 1977).

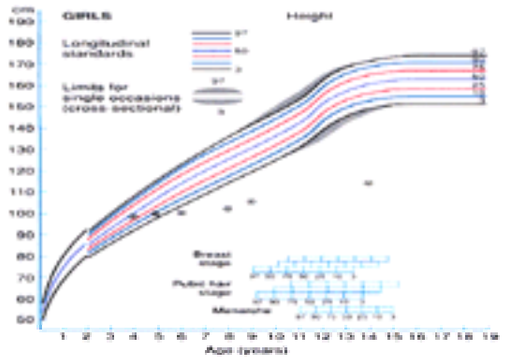


Fig. 6 Growth curve of a 15-year-old girl with severe systemic-onset juvenile arthritis, requiring long-term, high-dose prednisone treatment, showing severe growth delay.

The role of growth hormone in children with systemic-onset juvenile chronic arthritis is unclear. Levels of growth hormone have been reported as both normal and reduced (Butenandt *et al*. 1974; Allen *et al*. 1991). Low levels of insulin-like growth factors, which mediate the effects of growth hormone, have been reported (Bennett *et al*. 1988; Aitman *et al*. 1989). Several reports have studied the effects of treatment with growth hormone on children with different forms of juvenile chronic arthritis (Butenandt 1979; Svantesson 1991). Most recently, using doses of 12 and 24 IU/m^2, given three times per week, increased height velocities were documented. These were more significant in children receiving 24 IU/m^2 per week, and less marked in systemic-onset versus patients with either pauciarticular or polyarticular-onset disease. It is unclear from this study whether the ultimate height reached will be altered with growth hormone treatment (Davies *et al*. 1994). Most importantly, suppression of disease activity with medications other than corticosteroids and adequate nutrition must be achieved. Currently, treatment with growth hormone should be reserved for patients in prospective studies and for children whose growth is significantly below the third percentile.

Laboratory features

There are no specific laboratory features that are diagnostic of systemic-onset juvenile chronic arthritis. Rather, the common laboratory abnormalities reflect an activation of the acute-phase response, and taken together, are supportive of a diagnosis of systemic-onset juvenile chronic arthritis when other disorders are excluded by appropriate history, physical, and laboratory investigations.

The characteristic haematological abnormalities include anaemia, thrombocytosis, and leucocytosis (Table 3). The anaemia typically is an anaemia of chronic disease. This results in a normochromic normocytic smear, with haemoglobin levels ranging from 90 to 105 g/l. Occasionally, in the face of very active systemic toxicity, haemoglobin values will drop quickly, to values as low as 50 g/l. Frequently, the anaemia of chronic disease is compounded by the effects of medications and nutritional deficiency. For example, occult blood loss secondary to non-steroidal anti-inflammatory drugs and poor iron intake may lead to iron deficiency, thus accentuating the anaemia and resulting in hypochromia and microcytosis (Harvey *et al*. 1987). Despite blood loss and poor iron intake, however, serum ferritin is usually increased and is therefore not helpful in detecting iron deficiency as ferritin is an acute-phase reactant (Craft *et al*. 1977; Pelkonen *et al*. 1986). Bone marrow examination in patients with systemic-onset juvenile chronic arthritis (Fig. 7) usually shows a reactive marrow, with an increased number of plasma cells and with stainable iron. The mechanisms underlying the anaemia of systemic-onset juvenile chronic arthritis are unclear but may reflect an abnormal response to cellular mediators of haematopoiesis (Prouse *et al*. 1987; Silverman *et al*. 1988). Erythroid aplasia, similar to the syndrome of transient erythroblastopenia of childhood, has been reported (Rubin *et al*. 1978). Rarely, acute severe pancytopenia has developed in association with presumed viral haemophagocytic syndromes (Heaton and Moller 1985; Morris *et al*. 1985). Severe reaction to medications have been implicated in cases of consumptive coagulopathy (Silverman *et al*. 1983; Hadchouel *et al*. 1985).

Common
Anaemia of chronic disease
Iron deficiency anaemia
Neutrophilic leucocytosis
Thrombocytosis
Rare
Acute haemolysis
Haemophagocytic syndrome (disease or drug-induced)
Disseminated intravascular coagulation (disease or drug-induced)
Erythroid aplasia
Other nutritional deficiency anaemias
Leucopenia (disease or drug-induced)
Thrombocytopenia (disease or drug-induced)
Coagulopathy

Table 3 Haematological abnormalities in systemic-onset juvenile chronic arthritis

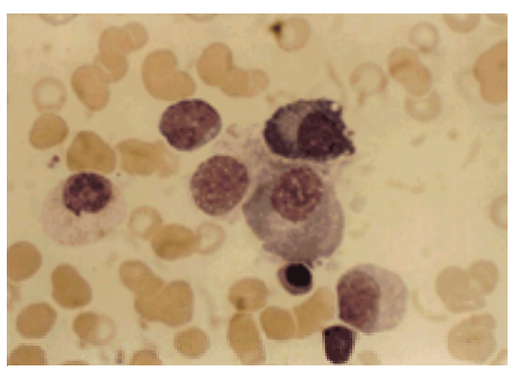


Fig. 7 Bone marrow examination of a patient with newly-diagnosed systemic-onset juvenile arthritis showing reactive plasmacytosis (magnification×1250, illustration by courtesy of Dr A. Poon).

Leucocytosis and thrombocytosis are also hallmarks of systemic-onset juvenile chronic arthritis, so much so that normal counts should always raise suspicion about the diagnosis. Typically, a peripheral blood smear will show a 'left shift' with an increase in the number of immature neutrophils. These neutrophils appear activated, with vacuoles and toxic granules, suggesting infection. White blood cell counts as high as 50×10^9/l may be seen. Similarly, thrombocytosis is characteristic of active disease. Rarely, both leucopenia and/or thrombocytopenia can occur, either as isolated events (Sherry and Kredich 1985) or as part of a disseminated intravascular coagulation-like syndrome.

The erythrocyte sedimentation rate is raised, often to greater than 100 mm/h (Westergren). Polyclonal hypergammaglobulinaemia is often observed, although not necessarily at onset (Petty *et al*. 1977); however, both transient and persistent IgA deficiency have been reported (Pelkonen *et al*. 1983). Other indicators of an acute-phase reaction include an elevated C-reactive protein (Gwyther *et al*. 1982) and significant hypoalbuminaemia, which may be multifactorial in aetiology (poor dietary intake, reduced hepatic synthesis, and intestinal leak). Serum complement levels are often raised, again indicating an acute-phase response (Hoyeraal and Mellbye 1974), and may help differentiate systemic-onset juvenile chronic arthritis from systemic lupus erythematosus. Renal function is normal, although mild proteinuria (which may be fever related) and red and white blood cells in the urine are occasionally observed (Antilla 1972). Elevation of serum transaminases are

frequent (see 'Reticuloendothelial involvement' above). Recently, dyslipoproteinaemia has been observed in all juvenile chronic arthritis subtypes, particularly those with systemic-onset juvenile chronic arthritis. This abnormality was felt to reflect active disease and may be mediated by cytokines (Ilowite *et al*. 1989).

While most children with systemic-onset juvenile chronic arthritis are seronegative for antinuclear antibody and rheumatoid factor, up to 37 per cent may be antinuclear antibody positive (Pauls *et al*. 1989; Siamopoulou-Mavridou *et al*. 1991) and 5 per cent rheumatoid factor positive (Cassidy *et al*. 1986; Lang and Shore 1990). No particular antinuclear antibody specificities have been consistently identified. Hidden rheumatoid factor has been reported in upwards of 50 per cent of patients (Moore *et al*. 1984). Immune complexes are found in up to 80 per cent of patients when assessed by a variety of methods (Moore *et al*. 1982). Evidence for complement activation may be found (Miller *et al*. 1986) and levels of complement receptor 1 (CR1) (complement C3b receptors) on erythrocytes were reduced in patients with systemic-onset juvenile chronic arthritis (Thomsen *et al*. 1987).

Abnormalities of immunoregulation, cell number and function, and cytokines have been reported. Unfortunately, most studies include patients with all types of juvenile chronic arthritis and do not specifically address systemic-onset disease alone. Studies of cellular immunity have shown B lymphocyte dysfunction (Tsokos *et al*. 1987; Barron *et al*. 1989), which may result from abnormal T-suppressor-cell function (Alarcon-Riquelme *et al*. 1988; Silverman *et al*. 1990a). In addition, anti-T-cell antibodies have been found in many children with systemic-onset juvenile chronic arthritis (Borel *et al*. 1984).

The clinical and laboratory features of systemic-onset juvenile chronic arthritis are very suggestive of a cytokine mediated process. The fever, skin rash, hypergammaglobulinaemia, hypoalbuminaemia, raised erythrocyte sedimentation rate, and fibrinogen that are characteristic of systemic-onset juvenile chronic arthritis may all be explained by an immune response involving the cytokines interleukin-1 and -6 (**IL-1** and **IL-6**) and tumour necrosis factor-a. In fact, several studies do suggest abnormalities in cytokine production and regulation. Levels of IL-1 are raised and in one study correlated with disease activity (Martini *et al*. 1986). Levels of IL-1b were uniquely raised in systemic-onset juvenile chronic arthritis as opposed to other types of juvenile chronic arthritis (Mangge *et al*. 1995). Prieur *et al*. found a naturally occurring inhibitor to IL-1 during the febrile phase in the urine of patients with systemic-onset juvenile chronic arthritis (Prieur *et al*. 1987). Increased levels of sIL-2R, indicative of immune activation, have been reported by various authors (Silverman *et al*. 1991; Fassbender *et al*. 1992; Lipnick *et al*. 1993). While raised sIL-2R levels are found in all types of juvenile chronic arthritis, levels do appear to be higher in patients with systemic onset, and correlate with active disease. Similarly, raised levels of serum IL-6 have correlated both with disease activity and thrombocytosis in patients with systemic-onset juvenile chronic arthritis (de Bennedetti *et al*. 1991). In addition, reduced levels of sIL-6 receptor but increased levels of IL-6/sIL-6R complexes have been reported (de Bennedetti *et al*. 1994). In preliminary studies, we have found increased levels of soluble tumour necrosis factor receptor P55 and P75, as have others (Mangge *et al*. 1995). Other abnormalities include elevated levels of tumour necrosis factor-a which correlated with disease activity (Mangge *et al*. 1995), elevated soluble CD8 levels (Lipnick *et al*. 1993), and, more recently, increased level of soluble phospholipase A_2 which correlated with active disease (Pruzanski *et al*. 1994). As mentioned, very few of these abnormalities are unique to systemic-onset juvenile chronic arthritis but they are more prominent relative to other subtypes, probably indicating a greater degree of immune activation. Results from studies of the interferon system are conflicting and have not shed light on the pathogenesis of systemic-onset juvenile chronic arthritis (Bacon *et al*. 1983; Arvin and Miller 1984).

Radiological features

The radiological features of juvenile chronic arthritis have been comprehensively reviewed (Reed and Wilmot 1991) but only two studies have detailed the radiological features specific for the systemic-onset subtype (Cassidy and Martel 1977; Lang *et al*. 1995). Table 4 shows the frequency of radiological abnormalities described by Lang *et al*. (Lang *et al*. 1995). The wrists were the most common sites of early and advanced radiological changes followed by the ankles, knees, tarsal joints, hips, and metacarpophalangeal joints. Similar findings were reported by Cassidy *et al*. (Cassidy and Martel 1977) with a significantly greater frequency of early periosteal new bone formation (50 per cent) and epiphyseal and vertebral compression fractures (attributed to the use of corticosteroids). Cervical spondylitis with narrowing, irregularity, and fusion of the apophyseal joints is also common. Ankylosis most commonly affects the C2–C3 level but may include the entire cervical spine (Fig. 8). Cervical spine ankylosis has been reported to be more frequent in systemic-onset disease than other subtypes (Espada *et al*. 1988). Other typical sites of ankylosis are the wrist and tarsus. It is noteworthy that metaphyseal rarefaction, a typical radiological finding of acute leukaemia, has been described in systemic-onset arthritis (Martel *et al*. 1962).

Radiological feature	Percentage
Soft tissue swelling/osteopenia	81
Joint space narrowing	50
Growth abnormalities	48
Erosions	43
Subluxation	21
Ankylosis	19
Joint destruction	14
Protrusio acetabulae	10
Periosteal new bone formation	10

Modified from Lang *et al*. (1995).

Table 4 Frequency of radiological abnormalities in 42 patients with systemic-onset juvenile chronic arthritis

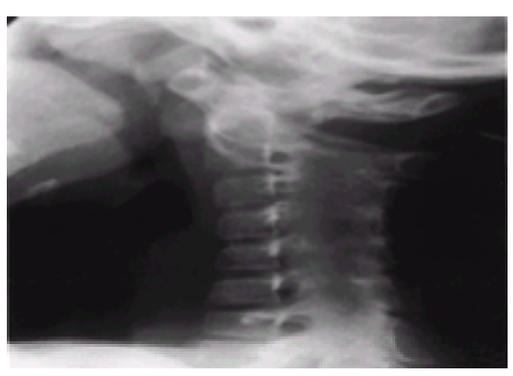


Fig. 8 Boy with systemic-onset juvenile chronic arthritis since 3.5 years. Cervical spine radiograph taken 5 years after disease onset shows ankylosis of the apophyseal joints (C2-C5).

A striking finding in the study by Lang *et al*. (Lang *et al*. 1995) was the early appearance of destructive changes. One-third of patients had erosions and joint space narrowing, 8 per cent had hip subluxation, and one patient developed ankylosis within 2 years of disease onset. These changes were sometimes seen within the first year of disease. Subtle changes of subchondral irregularity and sclerosis seemed to portend the development of erosions. These early destructive changes may be followed by progressive polyarticular disease.

Unusual radiological findings include large humeral cysts, soft tissue calcification unrelated to intra-articular corticosteroid injections (Lang *et al*. 1995), and shoulder synovial cysts (Barbaric and Young 1972). A ring of proliferative osteophytes at the junction of the femoral head and neck, similar to that described in patients with ankylosing spondylitis, has also been described (Mitnick and Mitnick 1980).

Differential diagnosis

When children present with a clinical picture that includes evening fever spikes, evanescent rash in association with the fever, polyarthritis, lymphadenopathy, hepatosplenomegaly, and polyserositis the diagnosis of systemic-onset juvenile arthritis is generally quite straightforward. However, many of these signs and

symptoms are often lacking and the clinician must be astute enough to recognize clues that lead to a diagnosis of systemic-onset juvenile arthritis. One of the most important clues is the variation in the clinical signs and symptoms that can occur over a 24-h period. Children may appear very well throughout most of the day, only to look toxic at the time of a fever spike. It is at this time that the rash usually appears. In addition, arthralgias and myalgias can be extremely severe with the fever spikes.

Of utmost importance is that in order to make a diagnosis of juvenile arthritis, a number of exclusions must be made. In the differential diagnosis of systemic-onset juvenile arthritis these include primarily infectious and postinfectious disorders, other inflammatory diseases, and malignancies in addition to some other more rare febrile disorders of childhood.

Infections

Bacterial and viral infections must be searched for diligently. Bacterial infections that may be difficult to diagnose and are associated with prolonged fever include 'hidden' processes such as abscesses and osteomyelitis or diseases associated with intermittent bacteraemia, such as subacute bacterial endocarditis. In the child with localized pain, technecium bone scanning may be indicated to search for osteomyelitis. Abscesses must be searched for by nuclear scanning, ultrasound, or computed tomography (**CT**) scans. Careful examination, assessment of the cardiovascular system for changing murmurs, search for mucocutaneous findings, and frequent blood cultures will lead to a diagnosis of subacute bacterial endocarditis. Mycobacterial and other granulomatous infectious diseases should be excluded with a thorough history of potential exposure to these organisms and appropriate radiographical and serological investigations.

Viral infections may present with fevers, rash, lymphadenopathy, and hepatosplenomegaly, and can therefore closely resemble systemic-onset juvenile arthritis. Important candidates in the differential diagnosis include Epstein–Barr virus, rubella, adenovirus, and hepatitis B in the prehepatitic phase. Serological assays will be helpful in confirming the diagnosis of a viral infection. As with most other diseases, the fever associated with viral illnesses tends to be persistent, rather than quotidian, and patients are not nearly as toxic as patients with systemic-onset juvenile arthritis during the fever spikes. Recently, parvovirus has been implicated as perhaps being of aetiological importance in adult rheumatoid arthritis and juvenile arthritis (Nocton *et al*. 1993). However, it has not been associated with a picture of systemic-onset juvenile arthritis.

Connective tissue diseases

'Connective tissue diseases' must be strongly considered in the differential diagnosis. The most important to consider are systemic lupus erythematosus, juvenile dermatomyositis, and systemic vasculitis. Features common to both systemic lupus erythematosus and systemic-onset juvenile arthritis include constitutional symptoms (fever, malaise, anorexia), arthritis/arthralgia, serositis, rashes, lymphadenopathy, and hepatosplenomegaly. However, vasculitic rashes seen in systemic lupus erythematosus are not typically seen in systemic-onset juvenile arthritis, and the rashes of systemic lupus erythematosus are not evanescent. Mucous membrane involvement is common in systemic lupus erythematosus but not in juvenile chronic arthritis. Fever tends to be persistent and not intermittent in systemic lupus erythematosus. Central nervous system and renal disease are almost never seen in juvenile arthritis but are common in systemic lupus erythematosus. Anaemia may be common to both, but leucopenia and thrombocytopenia are seen in systemic lupus erythematosus and almost never in juvenile chronic arthritis, where elevation of the white blood cell and platelet count are the rule. Specific autoantibodies are the hallmark of systemic lupus erythematosus but essentially absent in juvenile chronic arthritis. Reduced C3 and C4 complement levels in systemic lupus erythematosus result from immune complex deposition and consumption. In juvenile chronic arthritis, C3 and C4 tend to be elevated as a reflection of the acute-phase response. For further differences see Chapter 5.7.2.

Juvenile dermatomyositis may present as a systemic illness with arthritis but the heliotrope rash, Gottron's papules, and proximal muscle weakness should differentiate this presentation from systemic-onset juvenile arthritis. While myalgias are very common in patients with systemic-onset juvenile arthritis during fever spikes, true myositis does not occur. Serum levels of muscle enzymes will be helpful in differentiating the two.

Systemic vasculitis may present with fevers, malaise, anorexia, weight loss, arthralgias, and myalgias. The laboratory changes may be similar to those of systemic-onset juvenile arthritis in showing a marked elevation of the acute-phase response. The presence of nodules, vascular bruits, hypertension, mononeuritis multiplex, and cerebral disease help differentiate this from systemic-onset juvenile arthritis. Signs of internal organ involvement (e.g. lung and kidney) will also help in differentiating the two.

Malignancy

Malignancy forms one of the most important differential diagnostic categories—in particular, acute lymphoblastic leukaemia, lymphoma, and neuroblastoma (Schaller 1972). Important differentiating features are that children with malignancy have much more bone pain than patients with juvenile arthritis, and the pain tends to be persistent. Night-time pain is an important symptom in malignant disease (Ostrov *et al*. 1993) but both skin rash and serositis are uncommon. A low to normal white blood cell in the face of what appears to be systemic-onset juvenile arthritis is suggestive of a bone marrow infiltrative process. Lymphoma and neuroblastoma may present with fever and arthritis. Neuroblastoma is more common in young children and may not always be associated with a palpable abdominal mass. Screening of the urine for catecholamines and examining the chest and abdomen by CT scanning is helpful in excluding these. All malignancies may be especially difficult to diagnose in children who may have received even a very short course of corticosteroids for a presumptive diagnosis of juvenile chronic arthritis.

Postinfectious syndromes

Postinfectious syndromes must be considered in the differential diagnosis. The classic postinfectious syndrome is acute rheumatic fever following a group A b-haemolytic streptococcal throat infection. As with systemic-onset juvenile chronic arthritis, acute rheumatic fever may be associated with fever, rash, arthritis, and pericarditis, in association with laboratory markers of acute inflammation. However, the arthritis tends to involve only one or two joints at a time and then 'migrate' to other joints. Painful joints are prominent in acute rheumatic fever but much less so in systemic-onset juvenile arthritis, especially during the afebrile periods. Furthermore, the joints in acute rheumatic fever are often red, but are not in patients with systemic-onset juvenile arthritis. Pericarditis in acute rheumatic fever occurs only in the setting of endocarditis, which is extremely rare in systemic-onset juvenile chronic arthritis. Hepatosplenomegaly and lymphadenopathy are very rare in acute rheumatic fever. The rash, erythema marginatum, is uncommon and not as evanescent as the rash of systemic-onset juvenile arthritis. In acute rheumatic fever, a preceding history of streptococcal infection is necessary. One may have to rely on serology (antistreptolysin-O, antihyaluronidase) as evidence of infection, and a four-fold change in titre is necessary. At least 25 per cent of children with systemic-onset juvenile arthritis will have an increased antistreptolysin-O titre during the febrile phase, which is a reflection of the hypergammaglobulinaemia seen in these patients.

Reactive arthritis following bacterial infections involving the gastrointestinal or genitourinary tracts must be included in the differential diagnosis, and a history of preceding infections must be carefully addressed. Fever may occur in some cases of childhood reactive arthritis. However, the extra-articular features are very different from those of systemic-onset juvenile arthritis and involve the eyes, oral mucosa, and entheses. Inflammatory bowel disease may present with fever and arthritis before the onset of gastrointestinal symptoms. In addition, growth delay may be an early clue to this disorder.

Others

In the child under 1 year of age, chronic infantile neurological, cutaneous, and articular syndrome (**CINCA**) must be entertained in the differential diagnosis. It can be differentiated by its very young age of onset, associated chronic meningitis, uveitis, mental retardation, and epiphyseal changes (see Chapter 5.13.6). Juvenile sarcoidosis may also present with fever and rash; however, the rash is not transient and marked tenosynovial inflammation occurs. A biopsy of the skin or synovium will lead to the diagnosis (Hafner and Vogel 1993). Familial Mediterranean fever, an autosomal recessive disorder, presents with recurrent febrile episodes that generally last 48 to 72 h. These attacks may be associated with polyserositis and arthritis and a marked elevation of the acute-phase response accompanies the attacks. The short-lived, relapsing and remitting nature of the attacks, absence of chronic arthritis, and positive family history (when present) in a person of Mediterranean descent are suggestive of familial Mediterranean fever (Gedalia *et al*. 1992). The hyperimmunoglobulinaemia D syndrome begins in children below the age of 1 year and is associated with arthritis, lymphadenopathy, and splenomegaly. In contrast to systemic-onset juvenile chronic arthritis, there do not appear to be any residua of the arthritis, and, in contrast to familial Mediterranean fever, amyloidosis has not been reported to occur (Drenth *et al*. 1994).

The diagnostic workup must be focused to exclude all the entities considered in the American Rheumatism Association classification criteria (Brewer *et al*. 1977), particularly those mentioned above. It is vitally important to ensure that the child's temperature is recorded and plotted every 4 h to document the intermittent fever pattern of systemic-onset juvenile chronic arthritis. The recommended diagnostic investigations and clues to alternative diagnoses are summarized in Table 5 and

Table 6.

Table 5 Tests in the initial diagnostic investigation of suspected systemic-onset juvenile chronic (rheumatoid) arthritis

Leucopenia, thrombocytopenia
Child looks ill even during afebrile episodes
Pain out of keeping with degree of synovitis
Bony tenderness
'Hard' hepatosplenomegaly/lymphadenopathy
Recent antibiotic usage
Monoarthritis
Persistent diarrhoea
Significant weight loss

Table 6 Manifestations that raise suspicion of another diagnosis

Disease course and prognosis

Patients with systemic-onset juvenile chronic arthritis may follow a monocyclic course with complete remission within 2 years of disease onset, a polycyclic course characterized by exacerbations of systemic disease activity, or a course of persistent polyarthritis (Calabro *et al*. 1976). The mean duration of active disease is 5 to 6 years but some patients have persistent disease activity well into their adult years (Schaller and Wedgwood 1972; Svantesson *et al*. 1983). Fifty per cent of patients will have recurrent episodes of fever of variable duration (Schaller 1977) and one-third of patients may have fever for more than 1 year after disease onset (Svantesson *et al*. 1983). Although some North American series report that few patients will have active, systemic disease for more than 5 years (Calabro *et al*. 1976; Baum *et al*. 1980), Prieur *et al*. (Prieur *et al*. 1984) and Häfner and Truckenbrodt (Häfner and Truckenbrodt 1986) report that 25 to 30 per cent of their patients in Europe had persistent systemic symptoms 10 to 15 years after disease onset. Occasionally, patients may have recurrences of active disease following prolonged remissions.

Although the outlook for the extra-articular manifestations is ultimately good, most series report that progressive destructive arthritis occurs in at least one-third of patients (Schaller 1977; Mozziconacci *et al*. 1983; Cabane *et al*. 1990) and accounts for the principal complications of this disease, with most of these patients in functional class III or IV after long-term follow-up. This subgroup of patients would seem to have the worst outcome of all patients with juvenile chronic arthritis. A recent report of 14 patients with systemic-onset disease who were followed for more than 10 years is more optimistic (David *et al*. 1994). These patients had relatively good functional and psychological outcomes, with none in modified Steinbrocker classes III or IV, but total hip replacement surgery was required in 57 per cent.

It has proved difficult to predict reliably the disease course and outcome for the individual patient. Several long-term studies have suggested the following to be possible predictors of poor outcome: disease onset under 5 to 7 years of age, female sex, persistent disease activity for 1 to 5 years (which may be associated with amyloidosis), cardiac disease, thrombocytosis, accelerated radiological changes, and raised IgA levels (Mozziconacci *et al*. 1983; Svantesson *et al*. 1983; Hull 1988). In our own series, prognostic indicators of destructive arthritis were identified within 6 months of disease onset. The two most highly predictive factors found were persistence of systemic symptoms and a platelet count of more than 600×10^9/l. Three-quarters of our patients with both of these predictors 6 months after disease onset developed joint destruction. In addition, these patients more commonly had pericarditis, hepatosplenomegaly, and hypoalbuminaemia at disease onset than those with a more benign course. They also more commonly developed polyarthritis, leucocytosis (>12×10^9/l) and anaemia (< 100 g/l) at 6 months after the onset of disease (Schneider *et al*. 1992).

Mortality

There are more deaths amongst patients with systemic-onset juvenile chronic arthritis than any other subtype of juvenile chronic arthritis. The reported mortality of 14 per cent in several large European studies (Ansell and Wood 1976; Stoeber 1981; Hafner and Truckenbrodt 1986) is significantly higher than the 2.9 per cent mortality reported for all types of juvenile chronic arthritis (Baum and Gutkowska 1977). The major causes of death associated with systemic-onset arthritis are shown in Table 7. Amyloidosis has been associated with almost half the deaths in European patients with juvenile chronic arthritis and with only 13 per cent of deaths amongst North American patients; this may account for the higher mortality in European series. Infection is the most common cause of death in North American patients and is often associated with corticosteroid therapy. Acute hepatic failure has emerged as a significant cause of death, and in a review of several series accounted for 17 per cent of juvenile chronic arthritis deaths, most in those with systemic-onset disease (Boone 1977). Most deaths occur within the first 10 years of disease onset.

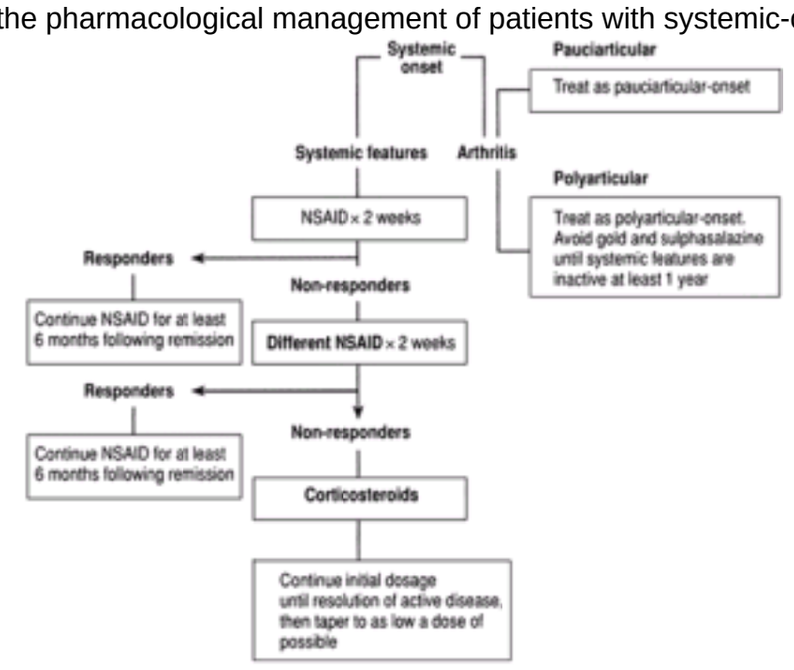
Amyloidosis with renal failure
Infection
Hepatic failure
Myopericarditis
Haematological disorders

Table 7 Major causes of death in systemic-onset juvenile chronic arthritis

Management (Box 1)

The lack of randomized placebo-controlled trials in patients with systemic-onset juvenile arthritis makes it difficult to evaluate the efficacy of any therapy in this disease. Because of increased medication-related toxicity in these patients, all therapeutic interventions must be carefully monitored. The approach to the management of systemic-onset juvenile chronic arthritis must be a co-ordinated one involving all members of the health-care team. In addition to the articular and extra-articular features of the disease, growth and psychosocial development must be continuously addressed and monitored. The significant financial burdens placed upon families with a chronically ill child must not be overlooked.

Box 1 Guidelines to the pharmacological management of patients with systemic-onset juvenile arthritis



Extra-articular features

Initial attempts to control fever should be with NSAIDs. Many paediatric rheumatologists in North America have moved away from salicylates in view of their increased hepatotoxicity relative to other NSAIDs and the potential for Reye's syndrome (Rennebohm *et al*. 1985). Success in controlling fevers has been achieved with both ibuprofen and indomethacin (Brewer 1977). Doses as high as 60 mg/kg in four divided doses of ibuprofen (B.M. Ansell, personal communication) or indomethacin 2 to 3 mg/kg per day seem to be more effective, with less toxicity than acetylsalicylic acid, and should be used prior to treating with corticosteroids for fever alone. Tolmetin sodium in doses up to 40 mg/kg per day may also be effective (Gewanter and Baum 1981). Because NSAIDs are protein bound, it is essential that the serum albumin be measured and the dose adjusted downward if hypoalbuminaemia is present. Failing to do so will result in excessive free drug which can result in toxicity. A minimum of a 1-week trial of an NSAID should be given before it is deemed to have failed; if the patient is not too ill, a second NSAID trial with a different preparation should be attempted.

In our experience, at least 50 per cent of patients will have an inadequate response to NSAIDs. When NSAIDs fail to control systemic toxicity, corticosteroid treatment is indicated. If used in high enough doses, steroids will virtually always result in resolution of fever and systemic toxicity. However, the significant side-effects of daily corticosteroid therapy are well known and as they do not limit the duration of active disease or alter the long-term prognosis, they must be used judiciously. At times, patients, parents, and physicians may have to be willing to settle for some fever provided that the peaks are not too high and not associated with severe systemic toxicity or anaemia. While alternate day dosing is preferable, in our experience, patients who require steroids require at least 1 mg/kg per day, and often in divided doses. Occasionally, symptoms may be so severe that the daily dose may be better administered in three or even four doses over a 24-hour period for a short period of time. High-dose, intravenous pulse methylprednisolone may also be used with severe flares but we have found only very short-term benefits from this treatment.

While the presence of lymphadenopathy, hepatosplenomegaly, and rash usually correlate with more active systemic symptoms and are an indication of active disease, these alone do not justify an increase in treatment.

Anaemia is common in patients with systemic-onset juvenile chronic arthritis and is usually a result of chronic disease. Maintenance doses of iron supplementation may raise the haemoglobin concentration slightly (Koerper *et al*. 1978) but are rarely of great benefit. When the disease is extremely active patients will not respond to oral iron therapy, even if they are iron deficient. Intravenous iron oxide saccharate (Martini *et al*. 1994) at a cumulative dose calculated to reach an ideal haemoglobin of 125 g/l has been recommended as a safe and effective treatment for severe and persistent anaemia in some patients with systemic-onset juvenile arthritis who are unresponsive to oral iron treatment. In one small study, erythropoietin was also effective in raising the haemoglobin value in systemic-onset juvenile arthritis patients with the anaemia of chronic disease (Fantini *et al*. 1992).

Nutritional deficiencies of folic acid and vitamin B12 may need to be addressed. Rarely, during acute systemic flares, the haemoglobin concentration may fall rapidly to as low as 40 to 50 g/l, necessitating blood transfusion (particularly in the face of cardiac compromise) in addition to corticosteroid therapy.

Attempts to reduce steroid toxicity have led to the use of intravenous immunoglobulin. Three separate studies have now reported on the use of intravenous immunoglobulin in systemic-onset juvenile arthritis. In the first, eight patients with severe juvenile chronic arthritis were treated for at least 6 months, in an open-label study. An impressive improvement in both articular and extra-articular disease and in laboratory abnormalities, as well as in a reduction in the dose of steroids, were noted (Silverman *et al*. 1990b). A second study did not seem to show any long-term benefit of intravenous immunoglobulin, although there was an improvement in the laboratory markers of active disease (Prieur *et al*. 1990). The recently completed Pediatric Rheumatology Collaborative Study Group trial did not show a statistically significant improvement in patients treated with intravenous immunoglobulin compared to placebo, although there was a trend towards overall improvement in the intravenous immunoglobulin-treated group (Silverman *et al*. 1994). However, many patients entered the trial so early in the disease course that spontaneous remissions may have occurred. Future studies should be directed at patients who early in their course fall into poor prognostic groups. At the time of writing, intravenous immunoglobulin should still be considered experimental in the treatment of systemic-onset juvenile chronic arthritis and should probably be reserved for patients whose systemic symptoms are not controlled by steroids, or who have significant steroid toxicity.

Serositis, as with the fever of systemic-onset juvenile chronic arthritis, will often respond to NSAIDs. Indomethacin seems to be especially effective for the treatment of pericarditis (Sherry *et al*. 1982). Corticosteroids in low to moderate doses (0.5–1 mg/kg per day) are usually sufficient to control serositis if NSAIDs are not effective. We have found intravenous pulse methylprednisolone 30 mg/kg per day (maximum 1 g) daily for 3 days to be rapidly effective and without toxicity, although of only short-term benefit. If there is significant compromise of cardiac function, pericardiocentesis may be required. Some authors recommend a pleural or pericardial drain for several days. Ventricular tachycardia complicating pericardiocentesis has resulted in a few deaths (Goldenberg *et al*. 1990). There is a report of one patient who, despite these measures, required the emergency placement of a pericardial window (Alukal *et al*. 1984). The efficacy of intrapericardial corticosteroids has not been substantiated and we have not found this necessary.

Myocarditis may occasionally be of such severity that congestive heart failure ensues. Treatment with high-dose oral or intravenous pulse corticosteroids, together with other supportive measures, is indicated. The use of digoxin may result in arrhythmias and sudden death if inflammation is not adequately controlled (Miller 1977), but a recent report did document the efficacy of digoxin (Goldenberg *et al*. 1992).

Articular disease

Many children with systemic-onset juvenile chronic arthritis have arthritis that is only problematic during flares of systemic disease. These children generally have a good outcome and respond well to management of the systemic components of the illness. However, the subset of children with persistent polyarticular disease, even in the absence of systemic manifestations, has progressive erosive disease which is difficult to treat with the standard first and second line agents. Unfortunately, no adequately controlled studies have been conducted in this group of children and our current approach is based largely on anecdotal experience. Currently, it is very difficult to predict at onset the course that patients will follow. It would seem that the earlier definitive treatment is instituted, the more effective treatment will ultimately be. Therefore, prognostic indicators are very important in determining the pharmacotherapeutic approach (Svantesson *et al*. 1983; Schneider *et al*. 1992).

The general approach to the management of the chronic arthritis of systemic-onset juvenile arthritis assumes the same principles as for the management of arthritis in other forms of juvenile arthritis. However, the increased drug-related toxicity is perhaps unique to patients with systemic-onset juvenile arthritis. This may be seen with salicylates, NSAIDs, and disease modifying antirheumatic drugs, including gold (Silverman *et al*. 1983; Hadchouel *et al*. 1985) and sulfasalazine (Cassidy 1990; Hertzberger Ten Cate and Cats 1991; Caspi *et al*. 1992). In fact, many authors consider that both gold and sulfasalazine are contraindicated in patients with systemic-onset juvenile arthritis, particularly if the disease is systemically active. Currently, methotrexate is the drug of choice for most patients with juvenile arthritis, including systemic-onset juvenile arthritis, needing second-line agents (Giannini *et al*. 1992). To date, it has not been associated with increased toxicity in patients with systemic-onset juvenile arthritis. However, it is also unclear how effective methotrexate actually is for the arthritis of systemic-onset juvenile arthritis, and there is no evidence that it has any effect on systemic symptoms. In fact, there has been a suggestion that systemic-onset patients may be less responsive to methotrexate than patients with other types of juvenile arthritis (Halle and Prieur 1991). In a retrospective review, 63 per cent of patients were deemed to have responded to weekly, low-dose oral methotrexate by 6 months. Analysis showed that early treatment (within 2 years of onset), before the development of radiographical lesions at the time of starting methotrexate, may improve the response (Ravelli *et al*. 1994).

For patients who do not respond to methotrexate, other alternatives must be sought. Chlorambucil has been used in patients with systemic-onset juvenile chronic arthritis who also develop amyloidosis (Deschenes *et al*. 1990), with significantly improved survival (David *et al*. 1993), but systemic features may still not be well controlled (Manners and Ansell 1986). In addition, the risk of leukaemia seems to be particularly increased with this alkylating agent (Palmer and Ansell 1984). Azathioprine may be somewhat effective (Kvien *et al*. 1986). The use of cyclophosphamide has been limited in the literature to case reports (Skoglund *et al*. 1971; Walters *et al*. 1972).

Early studies of cyclosporin in systemic-onset juvenile chronic arthritis (Bjerkhoel and Forre 1988; Ostensen *et al*. 1988) showed only minimal effect on synovitis and systemic symptoms persisted in several patients. The toxicity of cyclosporin seemed to outweigh the benefits. However, a more recent study using somewhat lower doses (unfortunately, uncontrolled) did document efficacy at recommended doses of 5 mg/kg per day in two divided doses (Pistoia *et al*. 1993). Improvement in terms of reduction of arthritis, fever, and prednisone dose was noted as early as 1 month after starting treatment. It is unclear whether cyclosporin is truly remitting, if patients will be able to discontinue cyclosporin, and what the long-term efficacy and toxicity is.

A preliminary study of the use of recombinant g-interferon in nine patients showed clinical improvement in seven, and an overall marked improvement in laboratory abnormalities (Pernice *et al*. 1989). The immunomodulatory agent thymopentin was effective for the systemic features, but not well-established arthritis, in a small series of children with systemic-onset juvenile chronic arthritis (Bardare *et al*. 1990).

A preliminary, open trial reported the results of repetitive use of pulse treatment with intravenous methylprednisolone (30 mg/kg per day for three consecutive days) with cyclophosphamide (0.4 g/m^2) on the third day, together with methotrexate at a dose of 10 mg/m^2 per week. This 3-day regimen was repeated every 3 months if the disease activity persisted and the patients required oral corticosteroids. Cyclophosphamide was included in subsequent pulses only if the patient had extra-articular disease. While there was an improvement noted in both articular and extra-articular disease, the laboratory features did not show statistically significant improvement. These preliminary results are impressive, and ideally would need to be confirmed in a randomized, placebo-controlled trial. Given the difficulty of conducting these types of trials in juvenile arthritis, the long-term follow-up of these patients must be reported before this approach can be recommended (Shaikov *et al*. 1992).

Adult-onset Still's disease

An entity quite similar in clinical and laboratory manifestations to systemic-onset juvenile chronic arthritis, but occurring in adults, was reported by Bywaters (Bywaters 1971). The so-called 'adult-onset Still's disease' has subsequently been described at all ages and with a worldwide distribution. Females outnumber males slightly. Seventy-five per cent of reported cases range between the ages of 16 and 35 years at the onset, with the overall incidence decreasing with age. Cases have been reported up to age 70. The fever pattern is identical to that of systemic-onset juvenile chronic arthritis (Fig. 1). Approximately 90 per cent have a rash that follows the fever course and may demonstrate a Koebner phenomenon; the rash may be pruritic. One important feature not well appreciated in systemic-onset juvenile chronic arthritis is a complaint of a severe sore throat, during flares of disease (Bujak *et al*. 1973; Esdaile *et al*. 1980; Ohta *et al*. 1987). While arthritis is not necessarily present at the onset, arthralgias are present in virtually all patients with the fever spikes. Although initial series described the arthritis as being quite mild, chronic arthritis with disability may be a sequel in up to 20 per cent of cases (Elkon *et al*. 1982; Cush *et al*. 1987; Cabane *et al*. 1990; Pouchot *et al*. 1991). As in the childhood form, hepatosplenomegaly and lymphadenopathy are reported in 40 to 75 per cent of cases. Weight loss of at least 10 per cent was recorded in approximately one-third of cases (Ohta *et al*. 1987). Pericarditis is the most common cardiac manifestation and tamponade may rarely occur. Both myocardial (Sachs *et al*. 1990) and endocardial (Taillan *et al*. 1989) involvement are rarely seen. Pulmonary disease appears to be more common than in systemic-onset juvenile chronic arthritis and is usually transient and mild, but severe restrictive lung disease has been observed (Corbett *et al*. 1983; Cantor *et al*. 1987). Abdominal pain may relate to hepatitis, adenitis, and sterile peritonitis, although this complaint is usually overshadowed by the other manifestations of adult-onset Still's disease (Bujak *et al*. 1973; Pollet *et al*. 1990). Abnormal liver function tests have been reported in up to 76 per cent of patients (Pouchot *et al*. 1991). Hepatic dysfunction may occur as part of a disseminated intravascular coagulation syndrome, may be related to medications, or as part of the underlying disease (Esdaile *et al*. 1979). Neurological involvement may rarely occur during systemic flares (Denault *et al*. 1990) or as a result of infection or complications of therapy (Wouters and van de Putte 1986). A number of ophthalmological manifestations, including inflammatory orbital pseudotumour (Cush *et al*. 1985), panophthalmitis (Bujak *et al*. 1973), and Brown's syndrome (Kaufman *et al*. 1987), have been reported. One case of sensorineural hearing loss, responsive to prednisone, was observed (Markusse *et al*. 1988). Rarely, renal disease may occur (Wendling *et al*. 1990).

The disease may follow several courses (Cush *et al*. 1987; Pouchot *et al*. 1991). In one series, those with either a mono- or polycyclic systemic course had articular manifestations, primarily during systemic exacerbations, and a good functional outcome. On the other hand, patients with a chronic articular course (either monocyclic systemic or polycyclic systemic) do not fare as well. Those patients with a polyarticular onset, axial arthritis, need for steroids within 2 years of onset, a history suggestive of childhood attacks, and the presence of a juvenile chronic arthritis rash seem to be at a greater overall risk for progressive joint damage and an unfavourable outcome (Cush *et al*. 1987; Pouchot *et al*. 1991). Involvement of the carpus with ankylosis (Medsger and Christy 1975; Pouchot *et al*. 1991) is particularly common, as is tarsal ankylosis, and involvement of the cervical spine and hips with rapid progression of destructive disease. Distal interphalangeal involvement is also not unusual.

In comparison to same-sex siblings, patients with adult-onset Still's disease had significantly higher levels of pain, psychological disability, and physical disability. Despite these problems, educational achievement, occupational prestige, social functioning, social support, annual family income, and days lost from work did not differ between patients and same-sex siblings (Sampalis *et al*. 1995).

The treatment of adult-onset Still's disease should follow along the same lines as those of systemic-onset juvenile chronic arthritis. Initial attempts at fever control should be made with either high dose salicylates (100 mg/kg per day) or other NSAIDs, particularly indomethacin. These two agents may be effective together when neither alone gives sufficient therapeutic benefit (Esdaile *et al*. 1980; Wouters and van de Putte 1986; Cush *et al*. 1987). The majority of patients will ultimately require moderate to high-dose glucocorticoid therapy at some time during their course. At a median follow-up of 10 years, 50 per cent of patients still required treatment with second-line agents (gold, hydroxychloroquine, or methotrexate) and one-third of these continued to require low-dose prednisone (Sampalis *et al*. 1995). The effect of slow-acting antirheumatic drugs on the course of the articular disease has not been established and these agents do not have any effect on the systemic features.

The laboratory abnormalities, like those of systemic-onset juvenile chronic arthritis, are non-specific and include leucocytosis with neutrophilia, normochromic, normocytic anaemia, and thrombocytosis. Eosinophilia appears to be common in the Japanese cases (Ohta *et al*. 1987). Hypoalbuminaemia, hypergammaglobulinaemia, and increased serum complement levels are commonly observed. Raised serum levels of hepatic transaminases, while common, are usually transient and reflect active disease. No consistent HLA associations have been identified (Wouters *et al*. 1986; Pouchot *et al*. 1991).

Death may result in a very small number of patients from a wide variety of causes, including liver involvement and systemic amyloidosis (Ohta *et al*. 1987; Reginato *et al*. 1987).

Chapter References

Agarwal, R.K., O'Neil, K.M., and Bedi, D. (1994). Focal radiolucent hepatic lesions in a patient with juvenile rheumatoid arthritis. *Journal of Rheumatology*, **21**, 580–1.

Aitman, T.J. *et al*. (1989). Serum IGF-I levels and growth failure in juvenile chronic arthritis. *Clinical and Experimental Rheumatology*, **7**, 557–61.

Alarcon-Riquelme, M.E. *et al*. (1988). Immunoregulatory defects in juvenile rheumatoid arthritis. Comparison between patients with the systemic or polyarticular forms. *Journal of Rheumatology*, **15**, 1547–50.

Allen, R.C., Jimenez, M., and Cowell, C.T. (1991). Insulin-like growth factor and growth hormone secretion in juvenile chronic arthritis. *Annals of the Rheumatic Diseases*, **50**, 602–6.

Alukal, M.K., Costello, P.B., and Green, F.A. (1984). Cardiac tamponade in systemic juvenile rheumatoid arthritis requiring emergency pericardiectomy. *Journal of Rheumatology*, **11**, 222–5.

Andersson Gare, B. and Fasth, A. (1992). Epidemiology of chronic arthritis in Southwestern Sweden: a 5 year, prospective population study. *Pediatrics*, **90**, 950–8.

Andersson Gare, B. *et al*. (1987). Incidence and prevalence of juvenile chronic arthritis: a population survey. *Annals of the Rheumatic Diseases*, **46**, 277–81.

Anonymous (1993). American Academy of Pediatrics Section on Rheumatology and Section on Ophthalmology: Guidelines for ophthalmologic examinations in children with juvenile rheumatoid arthritis. *Pediatrics*, **92**, 295–6.

Ansell, B.M. (1977). Juvenile chronic polyarthritis. Series 3. *Arthritis and Rheumatism*, **20** (Suppl), 176–8.

Ansell, B.M. (1978). Heberden oration, 1977. Chronic arthritis in childhood. *Annals of the Rheumatic Diseases*, **37**, 107–20.

Ansell, B., M. and Wood, P.H.N. (1976). Prognosis in juvenile chronic polyarthritis. *Clinics in Rheumatic Diseases*, **2**, 397–412.

Antilla, R. (1972). Renal involvement in juvenile rheumatoid arthritis. *Acta Paediatrica Scandinavica*, **227** (suppl), 1–73.

Arroyo, I.L., Barron, K.S., and Brewer, E.J., Jr (1988). Spinal cord compression by epidural lipomatosis in juvenile rheumatoid arthritis. *Arthritis and Rheumatism*, **31**, 447–51.

Arvin, A.M. and Miller, J.J. (1984). Acid labile alpha-interferon in sera and synovial fluids from patients with juvenile arthritis. *Arthritis and Rheumatism*, **27**, 582–5.

Athreya, B.H. *et al*. (1980). Pulmonary manifestations of juvenile rheumatoid arthritis. A report of eight cases and review. *Clinics in Chest Medicine*, **1**, 361–74.

Bacon, T.H., de Vere-Tyndall, A., Tyrrell, D.A., Denman, A.M., and Ansell, B.M. (1983). Interferon system in patients with systemic juvenile chronic arthritis: *in vivo* and *in vitro* studies. *Clinical and Experimental Immunology*, **54**, 23–30.

Bacon, M.C. *et al*. (1990). Nutritional status and growth in juvenile rheumatoid arthritis. *Seminars in Arthritis and Rheumatism*, **20**, 97–106.

Barbaric, Z.L. and Young, L.W. (1972). Synovial cysts in juvenile rheumatoid arthritis. *American Journal of Roentgenology*, **116**, 655–60.

Bardare, M., Corona, F., Ogliari, M.T., and Cohen, E. (1990). Thymopentin in the treatment of juvenile chronic arthritis. *Clinical and Experimental Rheumatology*, **8**, 89–93.

Barron, K.S., DeCunto, C.L., Montalvo, J.F., Orson, F.M., and Lewis, D.E. (1989). Abnormalities of immunoregulation in juvenile rheumatoid arthritis. *Journal of Rheumatology*, **16**, 940–8.

Baum, J. and Gutkowska, G. (1977). Death in juvenile rheumatoid arthritis. *Arthritis and Rheumatism*, **20**, 253–5.

Baum, J. *et al*. (1980). Juvenile rheumatoid arthritis. A comparison of patients from the USSR and USA. *Arthritis and Rheumatism*, **23**, 977–84.

Bedford, P.A., Ansell, B.M., Hall, P.J., and Woo, P. (1992). Increased frequency of DR4 in systemic onset juvenile chronic arthritis. *Clinical and Experimental Rheumatology*, **10**, 189–93.

Bennett, A.E., Silverman, E.D., Miller, J.J., and Hintz, R.L. (1988). Insulin-like growth factors I and II in children with systemic onset juvenile arthritis. *Journal of Rheumatology*, **15**, 655–8.

Bernstein, B.H. *et al*. (1977). Growth retardation in juvenile rheumatoid arthritis (JRA). *Arthritis and Rheumatism*, **20**, 212–6.

Bhettay, E. and Thomson, A.J.G. (1985). Peritonitis in juvenile chronic arthritis. *South African Medical Journal*, **68**, 605–6.

Bjerkhoel, F. and Forre, O. (1988). Cyclosporin treatment of a child with severe systemic juvenile rheumatoid arthritis. *Scandinavian Journal of Rheumatology*, **17**, 483–6.

Blodgett, F.M., Burgin, L., Iezzoni, D., Sribetz, D., and Talbot, N.B. (1956). Effects of prolonged cortisone therapy on the statural growth, skeletal maturation and metabolic status of children. *New England Journal of Medicine*, **254**, 636–41.

Boone, J.E. (1977). Hepatic disease and mortality in juvenile rheumatoid arthritis. *Arthritis and Rheumatism*, **20**, 257–8.

Borel, Y., Morimoto, C., Cairns, L., Mantzouranis, E.S., Strelkaukas, A.J. and Schlossman, S.F. (1984). Anti-T cell antibody in juvenile rheumatoid arthritis. *Journal of Rheumatology*, **11**, 56–61.

Brewer, E.J. (1977). Nonsteroidal antiinflammatory agents. *Arthritis and Rheumatism*, **20**, 513–25.

Brewer, E.J., Jr *et al*. (1977). Current proposed revision of JRA Criteria. JRA Criteria Subcommittee of the Diagnostic and Therapeutic Criteria Committee of the American Rheumatism Section of The Arthritis Foundation. *Arthritis and Rheumatism*, **20**, 195–9.

Bujak, J.S., Aptekar, R.G., Decker, J.L., and Wolff, S.M. (1973). Juvenile rheumatoid arthritis presenting in the adult as fever of unknown origin. *Medicine*, **52**, 431–44.

Butenandt, O. (1979). Rheumatoid arthritis and growth retardation in children: treatment with human growth hormone. *European Journal of Pediatrics*, **130**, 15–28.

Butenandt, O., Kelch, A., and Rajmann, E. (1974). Growth hormone studies in patients with rheumatoid arthritis with or without glucocorticoid therapy. *Zeitschrift für Kinderheilkunde*, **118**, 53–62.

Byron, M.A., Jackson, J., and Ansell, B.M. (1983). Effect of different corticosteroid regimens on hypothalamic-pituitary-adrenal axis and growth in juvenile chronic arthritis. *Journal of the Royal Society of Medicine*, **76**, 452–7.

Bywaters, E.G. (1971). Still's disease in the adult. *Annals of the Rheumatic Diseases*, **30**, 121–33.

Cabane, J. *et al*. (1990). Comparison of long term evolution of adult onset and juvenile onset Still's disease, both followed up for more than 10 years. *Annals of the Rheumatic Diseases*, **49**, 283–5.

Calabro, J.J., Holgerson, W.B., Sonpal, G.M., and Khoury, M.I. (1976). Juvenile rheumatoid arthritis: A general survey of 100 patients observed for 15 years. *Seminars in Arthritis and Rheumatism*, **5**, 257–98.

Cantor, J.P., Pitcher, W.D., and Hurd, E. (1987). Severe restrictive pulmonary defect in a patient with adult-onset Still's disease. *Chest*, **92**, 939–40.

Caspi, D., Fuchs, D., and Yaron, M. (1992). Sulphasalazine induced hepatitis in juvenile rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **51**, 275–6.

Cassidy, J.T. (1990). Management of JCA: slow-acting anti-rheumatic drugs. In *Pediatric rheumatology update* (eds P. Woo, P.H. White, and B.M. Ansell), pp. 66–80. Oxford University Press, New York.

Cassidy, J.T. and Martel, W. (1977). Juvenile rheumatoid arthritis: clinicoradiologic correlations. *Arthritis and Rheumatism*, **20**, 207–11.

Cassidy, J.T. *et al*. (1986). A study of classification criteria for a diagnosis of juvenile rheumatoid arthritis. *Arthritis and Rheumatism*, **29**, 274–81.

Corbett, A.J., Zizic, T.M., and Stevens, M.B. (1983). Adult-onset Still's disease with an associated severe restrictive pulmonary defect: a case report. *Annals of the Rheumatic Diseases*, **42**, 452–4.

Craft, A.W., Eastham, E.J., Bell, J.I., and Brigham, K. (1977). Serum ferritin in juvenile chronic polyarthritis. *Annals of the Rheumatic Diseases*, **36**, 271–3.

Cush, J.J., Leibowitz, I.H., and Friedman, S.A. (1985). Adult-onset Still's disease and inflammatory orbital pseudotumor. *New York State Journal of Medicine*, **85**, 110–1.

Cush, J.J., Medsger, T.A., Jr, Christy, W.C., Herbert, D.C., and Cooperstein, L.A. (1987). Adult-onset Still's disease. Clinical course and outcome. *Arthritis and Rheumatism*, **30**, 186–94.

David, J., Vouyiouka, O., Ansell, B.M., Hall, A., and Woo, P. (1993). Amyloidosis in juvenile chronic arthritis: a morbidity and mortality study. *Clinical and Experimental Rheumatology*, **11**, 85–90.

David, J. *et al*. (1994). The functional and psychological outcomes of juvenile chronic arthritis in young adulthood. *British Journal of Rheumatology*, **33**, 876–81.

Davies, U.M., Rooney, M., Preece, M.A., Ansell, B.M., and Woo, P. (1994). Treatment of growth retardation in juvenile chronic arthritis with recombinant human growth hormone. *Journal of Rheumatology*, **21**, 153–8.

de Bennedetti, B.F. *et al*. (1991). Correlation of serum interleukin-6 levels with joint involvement and thrombocytosis in systemic juvenile rheumatoid arthritis. *Arthritis and Rheumatism*, **34**, 1158–63.

de Bennedetti, B.F. *et al*. (1994). Serum soluble interleukin 6 (IL-6) receptor and IL-6/soluble IL-6 receptor complex in systemic juvenile rheumatoid arthritis. *Journal of Clinical Investigation*, **93**, 2114–9.

De Inocencio, J., Giannini, E.H., and Glass, D.N. (1993). Can genetic markers contribute to the classification of juvenile rheumatoid arthritis? *Journal of Rheumatology*, **40** (suppl.), 12–18.

Denault, A., Dimopoulos, M.A., and Fitzcharles, M.A. (1990). Meningoencephalitis and peripheral neuropathy complicating adult Still's disease. *Journal of Rheumatology*, **17**, 698–700.

Deschenes, G., Prieur, A.M., Hayem, F., Broyer, M., and Gubler, M.C. (1990). Renal amyloidosis in juvenile chronic arthritis: evolution after chlorambucil treatment. *Pediatric Nephrology*, **4**, 463–9.

Drenth, J.P.H., Haagsma, C.J., and Van der Meer, J.W.M. (1994). Hyperimmunoglobulinemia D and periodic fever syndrome. The clinical spectrum in a series of 50 patients. *Medicine*, **73**, 133–44.

Elkon, K.B. *et al*. (1982). Adult-onset Still's disease. Twenty-year followup and further studies of patients with active disease. *Arthritis and Rheumatism*, **25**, 647–54.

Esdaile, J.M., Tannenbaum, H., Lough, J., and Hawkins, D. (1979). Hepatic abnormalities in adult onset Still's disease. *Journal of Rheumatology*, **6**, 673–9.

Esdaile, J.M., Tannenbaum, H., and Hawkins, D. (1980). Adult Still's disease. *American Journal of Medicine*, **68**, 825–30.

Espada, G., Babini, J.C., Maldonado-Cocco, J.A., and Garcia-Morteo, O. (1988). Radiologic review: the cervical spine in juvenile rheumatoid arthritis. *Seminars in Arthritis and Rheumatism*, **17**, 185–95.

Espinel, C.H. (1994). Caravaggio's 'Il Amore Dormiente': a sleeping cupid with juvenile rheumatoid arthritis. *Lancet*, **344**, 1750–2.

Fantini, F., Gattinara, M., Gerloni, V., Bergomi, P., and Cirla, E. (1992). Severe anemia associated with active systemic-onset juvenile rheumatoid arthritis successfully treated with recombinant human erythropoietin: a pilot study. *Arthritis and Rheumatism*, **35**, 724–6.

Fassbender, K., Michels, H., Vogt, P., Aeschlimann, A., and Muller, W. (1992). Soluble interleukin-2 receptors in children with juvenile chronic arthritis. *Scandinavian Journal of Rheumatology*, **21**, 120–3.

Feldman, B.M. *et al*. (1996). Seasonal onset of systemic-onset juvenile rheumatoid arthritis. *Journal of Pediatrics*, **129**, 513–18.

Gedalia, A., Adara, A., and Gorodischer, R. (1992). Familial mediteranean fever in children. *Journal of Rheumatology*, **19** (Suppl 35), 1–9.

Gewanter, H.L. and Baum, J. (1981). The use of tolmetin sodium in systemic onset juvenile rheumatoid arthritis. *Arthritis and Rheumatism*, **24**, 1316–19.

Gewanter, H.L., Roghmann, K.J., and Baum, J. (1983). The prevalence of juvenile arthritis. *Arthritis and Rheumatism*, **26**, 599–603.

Giannini, E.H. *et al*. (1992). Methotrexate in resistant juvenile rheumatoid arthritis. Results of the U.S.A.-U.S.S.R. double-blind, placebo-controlled trial. *New England Journal of Medicine*, **326**, 1043–9.

Goldenberg, J. *et al*. (1990). Cardiac tamponade in juvenile chronic arthritis: report of two cases and review of publications. *Annals of the Rheumatic Diseases*, **49**, 549–53.

Goldenberg, J. *et al*. (1992). Symptomatic cardiac involvement in juvenile rheumatoid arthritis. *International Journal of Cardiology*, **34**, 57–62.

Gwyther, M., Schwarz, H., Howard, A., and Ansell, B.M. (1982). C-reactive protein in juvenile chronic arthritis: an indicator of disease activity and possibly amyloidosis. *Annals of the Rheumatic Diseases*, **41**, 259–62.

Hadchouel, M., Prieur, A.M., and Griscelli, C. (1985). Acute hemorrhagic, hepatic and neurologic manifestations in juvenile rehumatoid arthritis: possible relationship to drugs or infection. *Journal of Pediatrics*, **106**, 561–6.

Hafner, R. and Truckenbrodt, H. (1986). Course and prognosis of systemic juvenile chronic arthritis—retrospective study of 187 patients. *Klinische Padiatrie*, **198**, 401–7.

Hafner, R. and Vogel, P. (1993). Sarcoidosis of early onset. A challenge for the pediatric rheumatologist. *Clinical and Experimental Rheumatology*, **11**, 685–91.

Halle, F. and Prieur, A.M. (1991). Evaluation of methotrexate in the treatment of juvenile chronic arthritis according to the subtype. *Clinical and Experimental Rheumatology*, **9**, 297–302.

Hanson, V., Kornreich, H., Bernstein, B., King, K.K., and Singsen, B. (1977). Prognosis of juvenile rheumatoid arthritis. *Arthritis and Rheumatism*, **20**, 279–84.

Harvey, A.R., Pippard, M.J., and Ansell, B.M. (1987). Microcytic anaemia in juvenile chronic arthritis. *Scandinavian Journal of Rheumatology*, **16**, 53–9.

Hawkins, P.N. *et al*. (1993). Serum amyloid P component scintigraphy and turnover studies for diagnosis and quantitative monitoring of AA amyloidosis in juvenile rheumatoid arthritis. *Arthritis and Rheumatism*, **36**, 842–51.

Hayem, F., Calede, C., Hayem, G., and Kahn, M.F. (1994). (Involvement of the hip in systemic-onset forms of juvenile chronic arthritis. Retrospective study of 28 cases). *Revue du Rhumatisme, edition Francaise*, **61**, 583–9.

Heaton, D.C. and Moller, P.W. (1985). Still's disease associated with Coxsackie infection and haemophagocytic syndrome. *Annals of the Rheumatic Diseases*, **44**, 341–4.

Hertzberger, Ten Cate R. and Cats, A. (1991). Toxicity of sulfasalazine in systemic juvenile chronic arthritis. *Clinical and Experimental Rheumatology*, **9**, 85–8.

Heyd, J. and Glaser, J. (1990). Early occurrence of aortic valve regurgitation in a youth with systemic-onset juvenile rheumatoid arthritis. *American Journal of Medicine*, **89**, 123–4.

Hochberg, M.C. (1981). Adult and juvenile rheumatoid arthritis: current epidemiologic concepts. *Epidemiologic Reviews*, **3**, 27–44.

Hoyeraal, H.M. and Mellbye, O.J. (1974). High levels of serum complement factors in juvenille rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **33**, 243–7.

Hull, R.G. (1988). Outcome in juvenile arthritis. *British Journal of Rheumatology*, **27**, 66–71.

Ilowite, N.T., Samuel, P., Beseler, L., and Jacobson, M.S. (1989). Dyslipoproteinemia in juvenile rheumatoid arthritis. *Journal of Pediatrics*, **114**, 823–6.

Isdale, I.C. and Bywaters, E.G.L. (1956). The rash of rheumatoid arthritis and Still's disease. *Quarterly Journal of Medicine*, **25**, 377–87.

Jacobs, J.C. and Hui, R.M. (1977). Cricoaretynoid arthritis and airway obstruction in juvenile rheumatoid artyhritis. *Pediatrics*, **59**, 292–4.

Jan, J.E., Hill, R.H., and Low, M.D. (1972). Cerebral complications in juvenile rheumatoid arthritis. *Canadian Medical Association Journal*, **107**, 623–5.

Kaufman, L.D., Sibony, P.A., Anand, A.K., and Gruber, B.L. (1987). Superior oblique tenosynovitis (Brown's syndrome) as a manifestation of adult Still's disease. *Journal of Rheumatology*, **14**, 625–7.

Koerper, M.A., Stempel, D.A., and Dallman, P.R. (1978). Anemia in patients with juvenile rheumatoid arthritis. *Journal of Pediatrics*, **92**, 930–3.

Kramer, P.H., Imboden, J.B., Jr, Waldman, F.M., Turley, K., and Ports, T.A. (1983). Severe aortic insufficiency in juvenile chronic arthritis. *American Journal of Medicine*, **74**, 1088–91.

Kvien, T.K., Hoyeraal, H.M., and Sandstad, B. (1986). Azathioprine versus placebo in patients with juvenile rheumatoid arthritis: a single center double blind comparative study. *Journal of Rheumatology*, **13**, 118–23.

Lang, B.A. and Shore, A. (1990). A review on the current concepts on the pathogenesis of juvenile rheumatoid arthritis. *Journal of heumatology*, **17** (Suppl 21), 1–15.

Lang, B.A., Schneider, R., Reilly, B.J., Silverman, E.D., and Laxer, R.M. (1995). Radiologic features of systemic onset juvenile rheumatoid arthritis. *Journal of Rheumatology*, **22**, 168–73.

Lietman, P.S. and Bywaters, E.G.L. (1963). Pericarditis in juvenile rhaumatoid arthritis. *Pediatrics*, **32**, 855–60.

Lindsley, C.B. (1987). Seasonal variation in systemic onset juvenile rheumatoid arthritis. *Arthritis and Rheumatism*, **30**, 838–9.

Lipnick, R.N., Sfikakis, P.P., Klipple, G.L., and Tsokos, G.C. (1993). Elevated soluble CD8 antigen and soluble interleukin-2 receptors in the sera of patients with juvenile rheumatoid arthritis. *Clinical Immunology and Immunopathology*, **68**, 64–7.

Maksymowych, W.P. and Glass, D.N. (1988). Population genetics and molecular biology of the childhood chronic arthropathies. *Baillière's Clinical Rheumatology*, **2**, 649–71.

Mangge, H. *et al*. (1995). Serum cytokines in juvenile rheumatoid arthritis. Correlation with conventional inflammation parameters and clinical subtypes. *Arthritis and Rheumatism*, **38**, 211–20.

Manners, P.J. and Ansell, B.M. (1986). Slow-acting antirheumatic drug use in systemic onset juvenile chronic arthritis. *Pediatrics*, **77**, 99–103.

Markusse, H.M., Stolk, B., van der Mey, A.G., de Jonge-Bok, J.M., and Heering, K.J. (1988). Sensorineural hearing loss in adult onset Still's disease. *Annals of the Rheumatic Diseases*, **47**, 600–2.

Martel, W., Holt, J.F., and Cassidy, J.T. (1962). Roentgenologic manifestations of juvenile rheumatoid arthritis. *American Journal of Roentgenology*, **88**, 400–23.

Martini, A. *et al*. (1986). Enhanced interleukin 1 and depressed interleukin 2 production in juvenile arthritis. *Journal of Rheumatology*, **13**, 598–603.

Martini, A. *et al*. (1994). Intravenous iron therapy for severe anaemia in systemic-onset juvenile chronic arthritis. *Lancet*, **344**, 1052–4.

Medsger, T.A., Jr and Christy, W.C. (1975). Selective carpo-metacarpal arthritis with early ankylosis in adult onset Still's disease. *Arthritis andRheumatism*, **18**, 526–7.

Miller, J.J. (1977). Myocarditis in juvenile rheumatoid artyhritis. *American Journal of Diseases of Childhood*, **131**, 205–9.

Miller, J.J., Olds, L.C., Silverman, E.D., Milgrom, H., and Curd, J.G. (1986). Different patterns of C3 and C4 activation in the varied types of juvenile arthritis. *Pediatric Research*, **20**, 1332–7.

Mitnick, J. and Mitnick, H. (1980). Kidney enlargement and cervical spine disease in a child. *Journal of the American Medical Association* **243**, 465–6.

Moore, A.T. and Morin, J.D. (1985). Bilateral acquired inflammatory Brown's syndrome. *Journal of Pediatric Ophthalmology and Strabismus*, **22**, 26–30.

Moore, T.L., Sheridan, P.W., Traycoff, R.B., Zuckner, J., and Dorner, R.W. (1982). Immune complexes in juvenile rheumatoid arthritis: a comparison of four methods. *Journal of Rheumatology*, **9**, 395–401.

Moore, T.L. *et al*. (1984). Autoantibodies in juvenile arthritis. *Seminars in Arthritis and Rheumatism*, **13**, 329–36.

Morris, J.A., Adamson, A.R., Holt, P.J.L., and Davson, J. (1985). Still's disease and the virus-associated haemophagocutic syndrome. *Annals of the Rheumatic Diseases*, **44**, 349–53.

Mortensen, A.L., Allen, J.R., and Allen, R.C. (1990). Nutritional assessment of children with juvenile chronic arthritis. *Journal of Paediatrics and Child Health*, **26**, 335–8.

Mozziconacci, P., Prieur, A.M., Hayem, F., and Oury, C. (1983). (Articular prognosis of the systemic form of chronic juvenile arthritis (100 cases)). *Annales de Pediatrie*, **30**, 553–6.

Nepom, B. (1991). The immunogenetics of juvenile rheumatoid arthritis. *Rheumatic Disease Clinics of North America*, **17**, 825–42.

Nepom, B.S. and Glass, D.N. (1992). Juvenile rheumatoid arthritis and HLA: report of the Park City III workshop. *Journal of Rheumatology Supplement*, **33**, 70–4.

Nocton, J.J., Miller, L.C., Tucker, L.B., and Schaller, J.G. (1993). Human parvovirus B19-associated arthritis in children. *Journal of Pediatrics*, **122**, 186–90.

Ohta, A., Yamaguchi, M., Kaneoka, H., Nagayoshi, T., and Hiida, M. (1987). Adult Still's disease: review of 228 cases from the literature. *Journal of Rheumatology*, **14**, 1139–46.

Ostensen, M., Hoyeraal, H.M., and Kass, E. (1988). Tolerance of cyclosporine A in children with refractory juvenile rheumatoid arthritis. *Journal of Rheumatology*, **15**, 1536–8.

Ostrov, B.E., Goldsmith, D.P., and Athreya, B.H. (1993). Differentiation of systemic juvenile rheumatoid arthritis from acute leukemia near the onset of disease. *Journal of Pediatrics*, **122**, 595–8.

Padeh, S., Laxer, R.M., Silver, M.M., and Silverman, E.D. (1991). Primary pulmonary hypertension in a patient with systemic-onset juvenile arthritis. *Arthritis and Rheumatism*, **34**, 1575–9.

Palmer, R.G. and Ansell, B.M. (1984). Acute leukaemia related to chlorambucil therapy for juvenile chronic arthritis. *Clinical and Experimental Rheumatology*, **2**, 81–3.

Pauls, J.D., Silverman, E., Laxer, R.M., and Fritzler, M.J. (1989). Antibodies to histones H1 and H5 in sera of patients with juvenile rheumatoid arthritis. *Arthritis and Rheumatism*, **32**, 877–83.

Pearl, W. (1982). Pericarditis in juvenile rheumatoid arthritis. *Pediatrics*, **70**, 154–5.

Pelkonen, P., Savilahti, E., and Makela, A.L. (1983). Persistent and transient IgA deficiency in juvenile rheumatoid arthritis. *Scandinavian Journal of Rheumatology*, **12**, 273–9.

Pelkonen, P., Swanljung, K., and Siimes, M.A. (1986). Ferritinaemia as an indicator of systemic disease activity in children with systemic juvenile rheumatoid arthritis. *Acta Paediatrica Scandinavica*, **75**, 64–8.

Pernice, W., *et al*. (1989). Therapy for systemic juvenile rheumatoid arthritis with g-interferon. A pilot study of nine patients. *Arthritis and Rheumatism*, **32**, 643–6.

Petty, R.E., Cassidy, J.T., and Sullivan, D.B. (1977). Serologic studies in juvenile rheumatoid arthritis: a review. *Arthritis andRheumatism*, **20**, 260–7.

Pistoia, V. *et al*. (1993). Cyclosporin A in the treatment of juvenile chronic arthritis and childhood polymyositis-dermatomyositis. Results of a preliminary study. *Clinical and Experimental Rheumatology*, **11**, 203–8.

Pollet, S.M., Vogt, P.J., and Leek, J.C. (1990). Serous peritonitis in adult Still's syndrome. *Journal of Rheumatology*, **17**, 98–101.

Pouchot, J. *et al*. (1991). Adult Still's disease: manifestations, disease course, and outcome in 62 patients. *Medicine*, **70**, 118–36.

Prieur, A.M., Bremard-Oury, C., Griscelli, C., and Mozziconacci, P. (1984). Pronostic des formes systemiques d'arthrite chronique juvenile. *Archives Francaise de Pediatrie*, **41**, 91–7.

Prieur, A.M., Griscelli, C., Kauffman, M.T., and Dayer, J.M. (1987). Specific interleukin-1 inhibitor in serum and urine of children with systemic juvenile chronic arthritis. *Lancet*, **2**, 1240–2.

Prieur, A.M., Adleff, A., Debre, M., Boulate, P., and Griscelli, C. (1990). High dose immunoglobulin therapy in severe juvenile chronic arthritis: long-term follow-up in 16 patients. *Clinical and Experimental Rheumatology*, **8**, 603–8.

Prouse, P.J. *et al*. (1987). Anaemia in juvenile chronic arthritis: serum inhibition of normal erythropoeisis *in vitro*. *Annals of the Rheumatic Diseases*, **46**, 127–34.

Pruzanski, W. *et al*. (1994). Phospholipase A2 in juvenile rheumatoid arthritis: correlation to disease type and activity. *Journal of Rheumatology*, **21**, 1951–4.

Rachelefsky, G.S. *et al*. (1976). Serum enzyme abnormalities in juvenile rheumatoid arthritis. *Pediatrics*, **58**, 730–6.

Ravelli, A. *et al*. (1994). Factors associated with response to methotrexate in systemic-onset juvenile chronic arthritis. *Acta Paediatrica*, **83**, 428–32.

Reed, M.H. and Wilmot, D.M. (1991). The radiology of juvenile rheumatoid arthritis. A review of the English language literature. *Journal of Rheumatology*, **31** (suppl.), 2–22.

Reginato, A.J., Schumacher, H.R., Jr, Baker, D.G., O'Connor, C.R., and Ferreiros, J. (1987). Adult onset Still's disease: experience in 23 patients and literature review with emphasis on organ failure. *Seminars in Arthritis and Rheumatism*, **17**, 39–57.

Rennebohm, R.M., Heubi, J.E., Daugherty, C.C., and Daniels, S.R. (1985). Reye syndrome in children receiving salicylate therapy for onnective tissue disease. *Journal of Pediatrics*, **107**, 877–88.

Rubin, R.N., Walker, B.K., Ballas, S.K., and Travis, S.F. (1978). Erythroid aplasia in juvenile rheumatoid arthritis. *American Journal of Diseases of Children*, **132**, 760–2.

Sachs, R.N., Talvard, O., and Lanfranchi, J. (1990). Myocarditis in adult Still's disease. *International Journal of Cardiology*, **27**, 377–80.

Sampalis, J.S. *et al*. (1995). A controlled study of the long-term prognosis of adult Still's disease. *American Journal of Medicine*, **98**, 384–8.

Schaller, J. (1972). Arthritis as a presenting manifestation of malignancy in children. *Journal of Pediatrics*, **81**, 793–7.

Schaller, J.G. (1977). Juvenile rheumatoid arthritis: Series 1. *Arthritis and Rheumatism*, **20**, 165–70.

Schaller, J. and Wedgwood, R.J. (1972). Juvenile rheumatoid arthritis: a review. *Pediatrics*, **50**, 940–53.

Schaller, J., Beckwith, B., and Wedgwood, R.J. (1970). Hepatic involvement in juvenile rheumatoid arthritis. *Journal of Pediatrics*, **77**, 203–10.

Scheinberg, M.A. and Benson, M.D. (1980). SAA amyloid protein levels in amyloid-prone chronic inflammatory disorders. Lack of association with amyloid disease. *Journal of Rheumatology*, **7**, 724–6.

Schneider, R. *et al*. (1992). Prognostic indicators of joint destruction in systemic-onset juvenile rheumatoid arthritis. *Journal of Pediatrics*, **120**, 200–5.

Schnitzer, T.J. and Ansell, B.M. (1977). Amyloidosis in juvenile chronic polyarthritis. *Arthritis and Rheumatism*, **20**, 245–52.

Scott, J.P., Gerber, P., Matyjowski, M.C., and Pachman, L.C. (1984). Evidence for intravasculas coagulation in systemic onset, but not polyarticular, juvenile rheumatoid arthritis. *Arthritis and Rheumatism*, **28**, 256–61.

Shaikov, A.V. *et al*. (1992). Repetitive use of pulse therapy with methylprednisolone and cyclophosphamide in addition to oral methotrexate in children with systemic juvenile rheumatoid arthritis–preliminary results of a longterm study. *Journal of Rheumatology*, **19**, 612–16.

Sherry, D.D. and Kredich, D.W. (1985). Transient thrombocytopenia in systemic onset juvenile rheumatoid arthritis. *Pediatrics*, **76**, 600–3.

Sherry, D.D., Patterson, M.W.H., and Petty, R.E. (1982). The use of indomethacin in the treatment of pericarditis in childhood. *Journal of Pediatrics*, **100**, 995–8.

Siamopoulou-Mavridou, A. *et al*. (1991). Autoantibodies in Greek juvenile chronic arthritis patients. *Clinical and Experimental Rheumatology*, **9**, 647–52.

Silverman, E.D., Miller, J.J., Bernstein, B., and Shafai, T. (1983). Consumption coagulopathy associated with systemic juvenile rheumatoid arthritis. *Journal of Pediatrics*, **103**, 872–6.

Silverman, E.D., Laxer, R.M., Estrov, Z., and Freedman, M.H. (1988). Anemia of systemic onset juvenile arthritis: inhibition of erythropoeisis by interleukin-2 responsive T lymphocytes. *Arthritis and Rheumatism*, **31**, S12.

Silverman, E.D., Somma, C., Khan, M.M., Melmon, K.L., and Engleman, E.G. (1990 *a*). Abnormal T suppressor cell function in juvenile rheumatoid arthritis. *Arthritis and Rheumatism*, **33**, 205–11.

Silverman, E.D. *et al*. (1990*b*). Intravenous gammaglobulin therapy in systemic juvenile rheumatoid arthritis. *Arthritis and Rheumatism*, **33**, 1015–22.

Silverman, E.D., Laxer, R.M., Nelson, D.L., and Rubin, L.A. (1991). Soluble interleukin-2-receptor in juvenile rheumatoid arthritis. *Journal of Rheumatology*, **18**, 1398–402.

Silverman, E.D. *et al*. (1994). Intravenous immunoglobulin in the treatment of systemic juvenile rheumatoid arthritis: a randomized placebo controlled trial. *Journal of Rheumatology*, **21**, 2353–8.

Singh, G. *et al*. (1989). Histocompatibility antigens in systemic-onset juvenile rheumatoid arthritis. *Arthritis and Rheumatism*, **32**, 1492–3.

Skoglund, R.R., Schanberger, J.E., and Kaplan, J.M. (1971). Cyclophosphamide therapy for severe juvenile rheumatoid arthritis. *American Journal of Diseases of Children*, **121**, 531–3.

Stasny, P. and Fink, C.W. (1979). Different HLA-D associations in adult and juvenile rheumatoid arthritis. *Journal of Clinical Investigation*, **63**, 124–30.

Stephan, J.L. *et al*. (1993). Macrophage activation syndrome and rheumatic disease in childhood: a report of four new cases. *Clinical and Experimental Rheumatology*, **11**, 451–6.

Still, G.F. (1897). On a form of chronic joint disease in children. *Medical-chirurgical Transactions*, **80**, 47–59.

Stoeber, E. (1981). Prognosis in juvenile chronic arthritis. Follow-up of 433 chronic rheumatic children. *European Journal of Pediatrics*, **135**, 225–8.

Svantesson, H. (1991). Treatment of growth failure with human growth hormone in patients with juvenile chronic arthritis. A pilot study. *Clinical and Experimental Rheumatology*, **6**, 47–50.

Svantesson, H., Akesson, A., Eberhardt, K., and Elborgh, R. (1983). Prognosis in juvenile rheumatoid arthritis with systemic onset. A follow-up study. *Scandinavian Journal of Rheumatology*, **12**, 139–44.

Taillan, B. *et al*. (1989). Adult onset still's disease complicated by endocarditis with fatal evolution. *Clinical Rheumatology*, **8**, 541.

Tesser, J.R., Pisko, E.J., Hartz, J.W., and Weinblatt, M.E. (1982). Chronic liver disease and Still's disease. *Arthritis and Rheumatism*, **25**, 579–82.

Thomsen, B.S. *et al*. (1987). Complement C3b receptors on erythrocytes in patients with juvenile rheumatoid arthritis. *Arthritis and Rheumatism*, **30**, 967–71.

Towner, S.R., Michet, C.J., Jr, O'Fallon, W.M., and Nelson, A.M. (1983). The epidemiology of juvenile arthritis in Rochester, Minnesota 1960–1979. *Arthritis and Rheumatism*, **26**, 1208–13.

Tsokos, G.C. *et al*. (1987). Cellular immunity in patients with systemic juvenile rheumatoid arthritis. *Clinical Immunology and Immunopathology*, **42**, 86–92.

Wagener, J.S., Taussig, L.M., DeBenedetti, C., Lemen, R.J., and Loughlin, G.M. (1981). Pulmonary function in juvenile rheumatoid arthritis. *Journal of Pediatrics*, **99**, 108–10.

Walters, D., Robinson, R.G., Dick-Smith, J.B., Corrigan, A.B., and Webb, J. (1972). Poor response in two cases of juvenile rheumatoid arthritis to treatment with cyclophosphamide. *Medical Journal of Australia*, **2**, 1070.

Wang, F.M., Wertenbaker, C., Behrens, M.M., and Jacobs, J.C. (1984). Acquired Brown's syndrome in children with juvenile rheumatoid arthritis. *Ophthalmology*, **91**, 23–6.

Wendling, D., Hory, B., and Blanc, D. (1990). Adult Still's disease and mesangial glomerulonephritis. Report of two cases. *Clinical Rheumatology*, **9**, 95–9.

Woo, P., O'Brien, J., Robson, M., and Ansell, B.M. (1987). A genetic marker for systemic amyloidosis in juvenile arthritis. *Lancet*, **2**, 767–9.

Wouters, J.M. and van de Putte, L.B. (1986). Adult-onset Still's disease; clinical and laboratory features, treatment and progress of 45 cases. *Quarterly Journal of Medicine*, **61**, 1055–65.

Wouters, J.M., Reekers, P., and van de Putte, L.B. (1986). Adult-onset Still's disease. Disease course and HLA associations. *Arthritis andRheumatism*, **29**, 415–18.

Yancey, C.L., Doughty, R.A., Cohlan, B.A., and Athreya, B.H. (1981). Pericarditis and cardiac tamponade in juvenile rheumatoid arthritis. *Pediatrics*, **68**, 369–73.

Yousefzadeh, D.K. and Fishman, P.A. (1979). The triad of pneumonitis, pleuritis, and pericarditis in juvenile rheumatoid arthritis. *Pediatric Radiology*, **8**, 147–50.

Zaglul, H., Carswell, F., and Simpson, R.M. (1982). A case of persistent pulmonary functional abnormalities in systemic-onset juvenile chronic polyarthritis. *European Journal of Pediatrics*, **138**, 315–16.

5.6.3 Rheumatoid factor-negative polyarthritis in children ('seronegative' polyarthritis)

Anne-Marie Prieur

Polyarticular onset occurs in about 30 per cent of all patients with juvenile chronic arthritis. By definition, it occurs before the age of 16 years with the involvement of a minimum of five joints; the duration varies according to the adopted set of diagnostic criteria. The most accepted diagnostic criteria are either those of the European League Against Rheumatism (**EULAR**) (Wood 1978) or those of the American College of Rheumatology (**ACR**), previously the American Rheumatism Association (**ARA**) (Brewer *et al*. 1977). In the EULAR criteria, polyarthritis with positive rheumatoid factor is excluded, while in the ACR proposal, patients with positive rheumatoid factor are included in the group of polyarticular onset. This helps to explain why this disease is named juvenile chronic arthritis in Europe, but juvenile rheumatoid arthritis in North America. Arthritis must be present for at least 3 months in the EULAR criteria, or for 6 weeks in the ACR criteria. These proposed diagnostic criteria do not take into account the type of course, which obviously is extremely important for the prognosis and outcome. Pauciarticular onset with a polyarticular course will not be considered in this chapter.

The following description will not include the group of so-called 'seropositive' patients. We prefer, for obvious reasons, not to speak of seropositivity or seronegativity, but of rheumatoid factor (**RF**) -positive or -negative. The RF-positive group represents the early onset of adult rheumatoid arthritis and all data available in adults are also true in children (see Chapter 5.4.3). It is uncommon and occurs in less than 10 per cent of the whole polyarticular group in children. Polyarticular onset that is RF negative is the most common and is extremely heterogeneous. Some subgroups are easy to identifiy, but others need further evaluation to establish convenient and precise criteria for classification. Four subgroups can be identified in the RF-negative group (Table 1). The following discussion is based on the author's experience of more than 15 years in a paediatric rheumatology clinic with an annual referral of more than 1200 patients, one-third of them being first-time referrals.

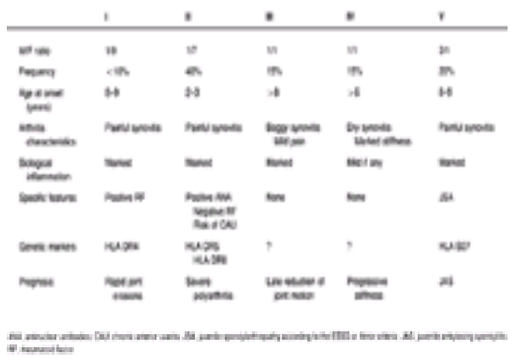


Table 1 A proposed classification of subgroups of juvenile chronic arthritis with polyarticular onset

Frequent manifestations of RF-negative polyarthritis

Extra-articular manifestations

Extra-articular manifestations may be present. Fever can be observed in one-third of patients. It is generally low grade, or a high fever often of short duration. By definition, long-lasting swinging fever as described in systemic onset is never observed in the polyarticular type. In very young children, a transient rash lasting 1 or 2 days can occur in the early stages. There is generally no lymphadenopathy, hepatosplenomegaly, or visceral involvement.

Joint manifestations

Joint manifestations dominate the clinical presentation. Generally, the mode of onset is rapidly progressive. The child is referred to the specialist often after several weeks of worsening joint stiffness and/or swelling.

A joint is a complex organ (Chapter 2.5) made up of synovial membrane, articular cartilage, fibrous capsule, ligaments, tendons with their site of attachment on bone (the enthesis), bursas, and tendon sheaths. In children, the joints also have growth cartilage. All these structures may be inflamed during the rheumatic process and chronic inflammation results in joint alterations and deformities. Chronic swelling and limited motion results in local demineralization. Muscular wasting is a consequence of the limitation of motion and it reduces joint motility. Chronic hyperaemia may induce local accelerated growth and growth cartilage fusion. Cartilage and bone erosions are late manifestations. There is potential for cartilage generation in children, which may delay severe functional impairment.

Upper limb involvement

Deformities in the hands and wrists are the consequence of an imbalance between the modified cartilage surfaces, local accelerated growth of bones, and modification of the forces between tendons and ligaments. Reduced carpal length, joint space reduction, and erosions can lead to carpal fusion. A carpal dislocation due to bone lysis is rare but can be observed in severe cases. Radial or ulnar growth reduction induces radial or ulnar deviation. Small joint and associated tendon involvement results in boutonnière or swan-neck deformities of the fingers (Fig. 1).

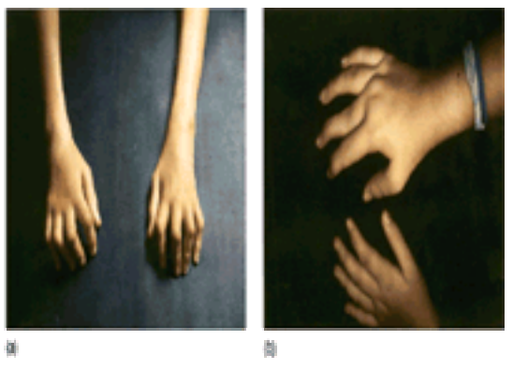


Fig. 1 Boutonnière deformity (a) and swan-neck deformity (b) of fingers.

Flexion contracture is the first manifestation of elbow involvement. It may be mild and of little functional significance. Limitation of supination is very common. The shoulders are frequently involved in severe cases, leading to limitation of movment, particularly abduction and external rotation of the glenohumeral joint. In severe cases, growth disturbances and humeral head modifications are observed (Fig. 2).

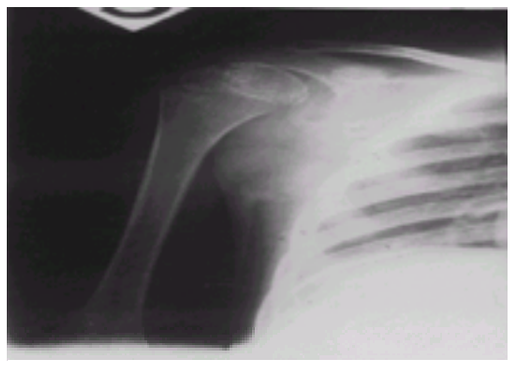


Fig. 2 Shoulder radiograph showing osteoporosis, joint space narrowing, and bony irregularity.

Lower limb involvement

Lower limb involvement may have significant functional consequences. Individual deformities can affect the performance of other lower limb joints. The foot is a complex joint with many articular surfaces in several planes. Foot deformities may also be increased by hip and knee involvement. Tenosynovitis and bursitis are nearly as common as ankle synovitis (Fig. 3). Subtalar joint involvement most often results in valgus deformity or, less commonly, in a varus deformity. As in the carpal area, bony fusion of the tarsus may occur. Mid-tarsal involvement affects the equilibrium of the foot. Metatarsal joint involvement leads to valgus toe deformities.

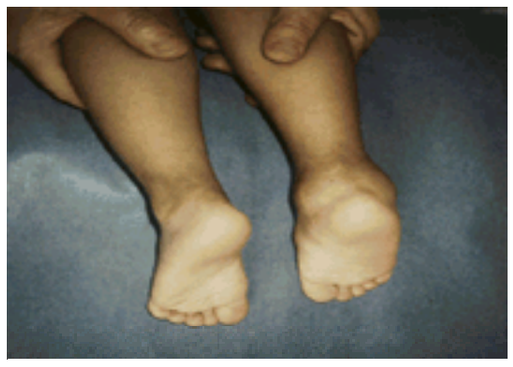


Fig. 3 Retromalleolar tenosynovitis, predominantly on the right side.

Flexion contracture of the knee is common, being particularly rapid and severe in the very young child. In the absence of correction, it may rapidly induce a posterior subluxation of the tibia due to capsular retraction (Fig. 4a). The overgrowth of the epiphysis induces a flexion contracture (Fig. 4b). In severe cases, the patella may fuse to the anterior femoral surface.

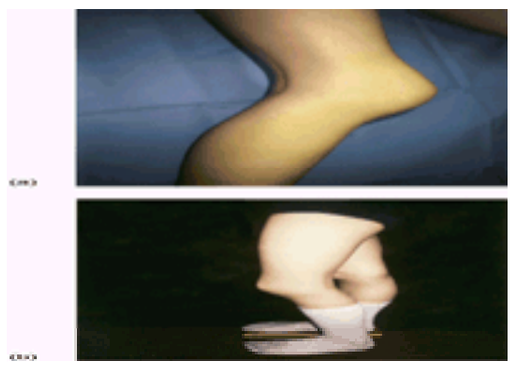


Fig. 4 Posterior tibial subluxation (a). Increased growth of the left lower limb due to chronic inflammation of the knee (b).

Hip involvement is characterized by flexion contracture and limitation of motion, particularly of abduction and rotation. This is secondary to the muscle spasm, mainly of the adductors, induced by local synovitis. Chronic inflammation induces anatomic modifications of the bone, with femoral head overgrowth, principally in the external part, and decreased development of the neck, which appears shorter and wider (Fig. 5(a)). Later, aseptic necrosis of the femoral head may occur. Acetabular modifications with erosions or protrusions are possible (Fig. 5(b)). Joint abnormalities at one site in the lower limb may have a reciprocal effect in aggravating deformities at other sites. For example, hip contracture induces a compensatory lumbar lordosis, knee overgrowth increases flexion contractures, and valgus of the knee increases talus deformity.

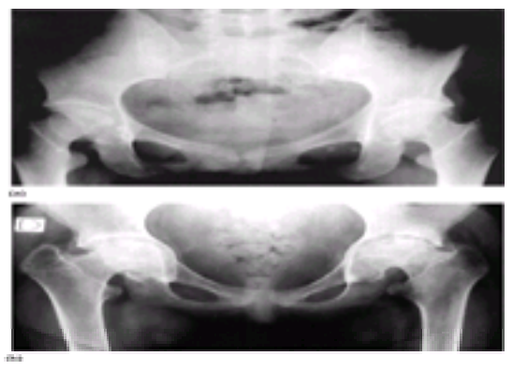


Fig. 5 Chronic involvement of the hip with shortened femoral neck, irregular head, and joint space narrowing (a); with acetabular protrusion (b).

Spine

Spine involvement is generally clinically expressed at the cervical level. Torticolis and limitation of motion are common (Fig. 6). Radiological changes develop progressively with apophyseal joint fusion most often of the C2–C3 vertebrae, but also of the other cervical spaces. Intervertebral instability below (Fig. 7) and above the fused cervical segment may occur. Subluxation of the atlas on the axis can be observed. A surgical arthrodesis is indicated in case of spinal cord compression. Anaesthetists should be aware of possible difficulties during intubation. Although rarely mentioned, synovitis of the thoracic and lumbar apophyseal joints is possible. A high frequency of spinal scoliosis is described in these children.

Fig. 6 Torticolis due to cervical spine involvement.

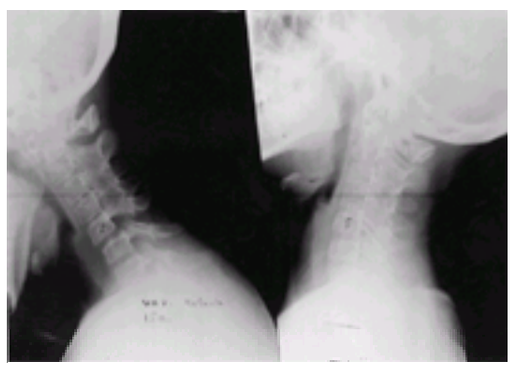


Fig. 7 Cervical spine fusion with underlying hypermobility.

Temporomandibular joint

Temporomandibular joint involvement is common. It is often discovered at routine examination. It reduces the normal growth of the mandible and results in micrognathia (Fig. 8). Dental malocclusion is common and may warrant surgical correction when growth is completed.

Fig. 8 Microretrognathia.

Laboratory abnormalities

Laboratory features are non-specific in RF-negative polyarthritis. Acute-phase reactants and erythrocyte sedimentation rate can be elevated or normal. The leucocyte count is normal or increased, as is the platelet count. Low blood-cell counts are unusual and suggest an alternative diagnosis.

Specific features of subgroups of RF-negative polyarthritis

Polyarthritis positive for antinuclear antibodies

Although less frequent than the oligoarticular type with positive antinuclear antibodies, this group of patients represents about 40 per cent of the polyarticular onset in our clinic. In a series of 136 children with arthritis positive for antinuclear antibodies, 21 were polyarticular at onset (Peralta and Prieur 1990). As in the

oligoarticular-onset group with antinuclear antibodies, there is a female preponderance, although the proportion of boys is higher than in the oligoarticular type. The disease is observed in very young children, two-thirds of the children reviewed in my series (Peralta and Prieur 1990) being less than 3 years old at onset. There is a risk of eye involvement as in the oligoarticular type (21 compared with 42 per cent). To my knowledge, no prognostic study of this particular group of polyarticular onset has been reported. In my experience, very young children with symmetrical involvement of both large and small joints within the first 3 months have a poorer outcome than those with symmetrical polyarticular involvement predominantly of large joints but not necessarily at the same time.

The presence of antinuclear antibodies in the serum characterizes this subgroup of patients. The most frequently used substrate for diagnosis is the HEp-2 cell. The titres of antinuclear antibodies are usually low and there are no obvious correlations between the titre of the autoantibody and the severity of the disease. The specificity of these antinuclear antibodies is antihistone (Malleson *et al*. 1989; Østensen *et al*. 1989; Pauls *et al*. 1989; Monestier *et al*. 1990), but there are discrepancies in the frequencies of histone subtypes recognized by the autoantibodies. Anti-H1 and anti-H3 are the most commonly described (Pauls *et al*. 1989). No correlation has been observed with the presence of uveitis except in one study in which a higher frequency of H3-reacting autoantibodies was found in children with uveitis (Østensen *et al*. 1989). The autoreactivity to histone fragments is demonstrated either to the C-terminal of H1 (Monestier *et al*. 1990) or to several fragments within H3 and H4 (Tuaillon *et al*. 1990; Leak *et al*. 1993). In a recent study, autoantibodies from sera of 138 children with juvenile chronic arthritis were assessed with 34 histone peptides covering the full length of the four core histones, and two peptides from H1. No correlation was found either with disease subtype or activity, or with the presence of chronic anterior uveitis (Stemmer *et al*. 1995).

Polyarthritis with boggy synovitis

Some children may present with very thick pannus involving joints in a symmetrical manner. Tenosynovitis is common. Pain remains mild and functional impairment is late. Laboratory tests show a high erythrocyte sedimentation rate (greater than 40 mm/h), marked leucocytosis, and increased levels of acute-phase reactants. There are no autoantibodies. Boys and girls are equally affected.

'Dry' polyarthritis

This group of patients is roughly the counterpart of the previous group. The diagnosis is often delayed as there is no joint swelling but very progressive stiffness. These joint manifestations lead to gait abnormalities including limping. It is not unusual for these children to be referred to a paediatric neurologist, or even to a psychiatrist! However, osteoarticular examination reveals a limited range of motion because of articular stiffness but without obvious synovitis. Joint stiffness seems to be due to capsular and tendon contraction. Muscle wasting occurs. The laboratory profile is normal or mildly inflammatory. There are no autoantibodies.

Polyarthritis with spondylarthropathy features (see Chapter 5.5.2)

Some older patients, particularly boys, can present with polyarthritis (more than four joints), most often in the lower limbs. A careful clinical examination and the presence of relatives with spondylarthropathy suggests the diagnosis of undifferentiated spondylarthropathy. They meet the criteria proposed by the European Spondylarthropathy Study Group or those of Bernard Amor (see Chapter 5.5.5) (Amor *et al*. 1990; Dougados *et al*. 1991). Generally in children, the only manifestation is peripheral arthritis with no spinal involvement. Joint pain is often marked. Non-steroidal anti-inflammatory drugs (**NSAIDs**) induce rapid pain relief. Biological evidence of inflammation is more pronounced than in the oligoarticular type of spondylarthropathy. There are no autoantibodies in this group of patients, except for the occasional presence of antinuclear antibodies in the group with psoriatic arthritis. In rare cases, there may be inflammatory bowel disease or reactive arthritis (Reiter's syndrome) with eye and mucosal manifestations. The risk of developing a spondylitis within 5 to 10 years is high. Psoriatic arthritis is considered by many authors to be a spondylarthropathy. However, a higher frequency of girls and a younger age at onset means that its place in the spondylarthropathy group is questionable (Table 2).

	Sex ratio (% male)	Age at onset (years)
Juvenile ankylosing spondylitis	84	10.7
Reiter's syndrome	75	10.1
Undifferentiated spondylarthropathies	66	11
Inflammatory bowel disease	62	11.4
Juvenile psoriatic arthritis	45	8.8

[a]According to the criteria proposed by the European Spondylarthropathy Study Group and Bernard Amor (Chapter 5.5.5).

Table 2 Sex ratio and age at onset in the different groups of spondylarthropathies in children [a]

Differential diagnosis (Table 3)

Table 3 Differential diagnosis of RF-negative polyarthritis

Polyarthritis related to an infectious agent

Viruses may induce a joint reaction, particularly in very young children. It may be preceded by an upper respiratory tract infection, with or without fever. Synovial fluid analysis shows a majority of lymphocytes. Joint involvement generally lasts less than 2 weeks. Parvovirus B19, hepatitis, and rubella are among the most common causes of arthritis. Rubella vaccination may be followed by a polyarthritis which can last several weeks. Streptococcal infection or acute rheumatic fever may cause a migratory arthritis (see Chapter 5.3.12). Other signs of rheumatic fever, such as carditis, chorea, and eruption, may be present. Treatment consists of anti-inflammatory drugs, aspirin and/or corticosteroids, and penicillin for at least 5 years. Although this condition now occurs very rarely, we should be aware of its possible resurgence (Kaplan 1990). Lyme borreliosis can be considered when the child has had a tick bite in an endemic area. Typical clinical symptoms are a flu-like illness with erythema chronicum migrans. Some weeks later other symptoms can occur, including neurological, cardiac, ocular, and articular manifestations. Joint involvement is less frequent in Europe than in North America. The diagnosis and treatment is described in Chapter 5.3.4.

Acute infections of the joint or bone are usually easily diagnosed (see Chapter 5.3.2). However, a polyarthritis-like disease can occur in immunodeficiencies (see Chapter 5.3.11). The most frequent immunodeficiency in which this is observed is the X-linked humoral deficiency or Bruton's disease, which occurs in young boys with a history of upper respiratory tract, pulmonary, or gut infection. The diagnosis is made by the absence of serum immunoglobulins. A multifocal, bacterial joint

infection should be ruled out first and treated. However, non-bacterial joint swelling is possible, probably resulting from chronic virus infection. Immunoglobulin infusion generally improves chronic joint involvement. Other immunodeficiencies such as acquired humoral deficiencies, or Wiskott–Aldrich and ataxia telangiectasia, can also be complicated by non-bacterial arthritis. In the latter, joint manifestations are rarely the initial symptom.

Autoimmune rheumatic disorders

Autoimmune rheumatic disorders such as systemic lupus erythematosus must also be considered. Polyarthritis is nearly always present in systemic lupus erythematosus and is one of the eleven diagnostic criteria of the ACR (see Chapter 5.7.2). The presence of skin manifestations, serositis, and renal involvement should prompt laboratory investigations to confirm the diagnosis. Cytopenia, anti-DNA antibodies, and decreased complement are the usual findings. Polyarthritis is also prominent in overlap syndromes and mixed connective tissue disease (see Chapter 5.12.1). Polymyositis or dermatomyositis are usually easily diagnosed, but a joint component is possible.

Systemic vasculitic syndromes

Systemic vasculitic syndromes (see Chapter 5.11.8) are not exceptional. Kawasaki disease is generally observed in very young children, but the joint symptoms are of secondary importance to the extra-articular manifestations. Polyarteritis nodosa may cause a very severe and painful polyarthritis associated with myalgia. As well as the painful joints, skin involvement includes nodules, with typical changes of medium-sized arteries on biopsy. Cutaneous and articular features without visceral involvement are possible. The differential diagnosis of RF-negative polyarthritis from other other forms of systemic vasculitis such as Wegener's granulomatosis, lymphomatoid granulomatosis, Henoch–Schönlein purpura, and hypocomplementaemic vasculitis is generally easy. Arthritis can occur in relapsing polychondritis. The diagnosis is generally obvious with inflammation of auricular cartilage and nasal chondritis with a saddle-nose deformity.

Behçet's syndrome (see Chapter 5.11.7) can occur in children. These patients may have polyarthritis. The diagnosis is made on the basis of possible familial clustering and ethnic origin, when genital and oral ulcerations are present. Inheritance is autosomal recessive. These children can also develop ocular, intestinal, and neurological manifestations (Koné-Paut *et al*. 1995). Familial Mediterranean fever (see Chapter 5.13.2) occurs in Sephardic Jews, Armenians, Greeks, Turks, and Arabs. Inheritance is also autosomal recessive and the gene has been mapped on the short arm of chromosome 16. Clinical manifestations include fever, skin rash, serositis, abdominal pain, and possibly arthritis. Colchicine is an effective therapy in preventing attacks and the occurrence of renal amyloidosis.

Sarcoid arthritis produces a distinctive syndrome in children under the age of 4 years, sometimes in the first year of life. The triad of skin rash, uveitis, and huge proliferation of boggy synovium with tenosynovitis is highly suggestive of sarcoid arthritis. Skin or synovial biopsy shows the typical sarcoid granulomata. Joint involvement is relatively painless. Visceral involvement is rare. A familial association has been described (Blau 1985). Vasculopathy has been reported and some authors suggest that this entity should be considered as a 'familial granulomatous arteritis'. Systemic corticosteroids are often necessary to control uveitis and arthritis. Polyarthritis can be observed in the adult type of sarcoidosis in older children.

Haematological disorders

In early childhood, joint involvement in sickle-cell disease may result in hand–foot syndrome. Swelling of the hands and feet is extremely painful. Later, migratory arthritis can be observed. Bone pain can be due either to bone infarction or to osteomyelitis. The diagnosis is based on ethnic origin and haemoglobin electrophoresis. Constitutional bleeding disorders such as haemophilias can be manifested by joint haemorrhage. Generally, the deficiency of clotting factor is identified. Modern therapeutic management aims to prevent haemarthrosis and joint destruction.

Acute leukaemia may induce bone pain and joint swelling due to infiltrates of leukaemic cells. The diagnosis must be confirmed rapidly by examination of a bone marrow smear and specific therapy started without delay. Neuroblastoma with bone metastasis must also be considered in a young child with osteoarticular pain.

Other causes of polyarticular manifestations

Patients with chronic, recurrent, multifocal osteomyelitis can present with bone pain. There is no joint swelling and careful examination localizes the pain to the metaphyseal area. Radiography confirms the diagnosis by the presence of osteolytic lesions in the metaphysis. The lesion is sterile and antibiotics have no effect. The course may be long and relapsing.

Patients with hip pain are sometimes referred as having possible polyarthritis. In an adolescent, the possibility of lamellar coxitis should be ruled out. It is generally unilateral and the laboratory screen is normal, while radiography shows narrowing of the hip space. The cause is unknown and the prognosis is unpredictable. Similarly, Legg–Calvé–Perthes' disease induces osteonecrosis of the femoral head in young boys around the age of 7 years. It is generally unilateral, rarely bilateral. Radiographs are normal at onset, but scintigraphy or MRI may show necrosis at this stage. Orthopaedic management is required. In younger children, transient synovitis of the hip is frequent. It occurs around 6 years of age and may be preceded by upper respiratory tract infection. An increased erythrocyte sedimentation rate is frequent. Bed rest with traction and NSAIDs are necessary. Recovery occurs within a few days. Occasionally hip dysplasia is evident. Some patients develop recurrent episodes of transient synovitis of the hip.

Diabetic arthropathy is observed in uncontrolled diabetes. Exceptionally it may be the first manifestation of diabetes. Diabetic cheiroarthropathy with progressive flexion contractures of the fingers, as a result of increased deposition of collagen in the tissues, is well recognized. Idiopathic juvenile osteoporosis is a painful disease. There are no joint manifestations. Usually, there are no abnormalities of phosphocalcium metabolism. Recovery occurs in adolescence, sometimes with sequellae resulting from spontaneous fractures or vertebral collapse.

Immunogenetics

Most of the immunogenetic studies in polyarticular-onset juvenile arthritis separate the group of 'seropositive' and 'seronegative' forms. RF-positive polyarticular arthritis in children and adults are both associated with HLA DRB1 (Ploski *et al*. 1993). Several authors have observed an association between RF-negative polyarticular arthritis in children and HLA DR8, mainly involving the HLA DRB1*0801 subtype (Morling *et al*. 1985; Hall *et al*. 1989; Fernandez-Vina *et al*. 1990; Barron *et al*. 1992; Ploski *et al*. 1993). Ethnic differences may be observed in some populations, such as the studies on Italian (Fantini *et al*. 1987) and Czech children (Cerna *et al*. 1994) in which no correlation with HLA DR8 could be found.

HLA DR8, in fact, is associated with several forms of juvenile arthritis including early-onset oligoarticular forms (Malagon *et al*. 1991) and juvenile spondylarthropathies (Ploski *et al*. 1995). In contrast, HLA DPB1*0201 is associated with the oligoarticular, but not with the polyarticular forms. The frequency of HLA DPB1*0301 is increased significantly in patients with polyarticular arthritis and juvenile spondylarthropathies. This association between HLA DRB1*0801 and HLA DPB1*0301 in polyarticular-onset disease has been confirmed by several authors. This combination, rarely observed in the normal population or in the oligoarticular types of disease, suggests an interaction between these two alleles, conferring an increased susceptibility to the polyarticular expression of the disease.

However, as seen in Table 4, there are discrepancies between the studies, probably due to various factors including ethnic background, the absence of clinical homogeneity of the patients, the definition of each clinical subgroup, and the number of patients studied. It is evident that uniformly agreed criteria for the improved subgrouping of patients is still required.

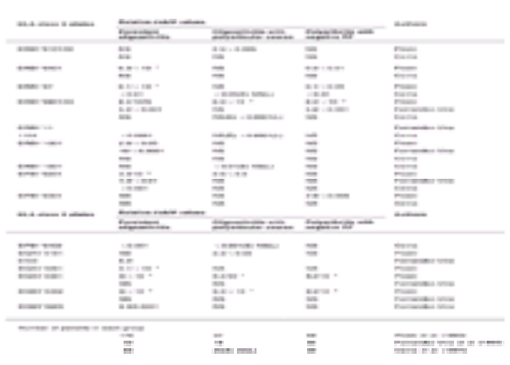


Table 4 HLA associations in RF-negative polyarthritis compared with oligoarthritis, with or without a polyarticular course, in three different geographical areas

Management

The principles of management are those of any childhood chronic disease of the joints, namely to offer pain relief, maintain satisfactory joint function, minimize drug toxicity, and allow the child to grow and be educated as normally as possible. It is sometimes difficult to meet all these requirements and complete trust between the family and the paediatric team must be established.

Non-steroidal anti-inflammatory drugs (NSAIDs)

The use of NSAIDs remains the basis of treatment. In children, only a certain number of drugs have been studied and their licence for use varies in different countries. Acetylsalicylic acid and other NSAIDs such as diclofenac, ibuprofen, and naproxen are available for use in most countries. Recommended doses are indicated in Table 5. In general, all NSAIDs have similar efficacy but tolerance may vary. Aspirin induces the most frequent side-effects, as shown in a comprehensive study (Barron *et al*. 1982). The mode of administration may be important for compliance. Ideally, aspirin must be administered every 4 h, while other NSAIDs can be given 1 to 3 times over 24 h. The few available pharmacokinetic studies in older children have shown only minor differences to adults. The half-life in synovial fluid is longer than in plasma (Hallé *et al*. 1991). NSAIDs are usually rapidly effective, but some authors have found that they take several weeks to have full effect (Lovell *et al*. 1984). The most frequent side-effects are gastrointestinal (mainly abdominal pain), cutaneous (urticaria, rash, and hypersensitivity), and haematological (anaemia). Rare cases of renal damage and disorders of the central nervous system (headache and dizziness) have been reported. Allergic reactions can also occur. Macrophage activation syndrome is exceptional in polyarticular disease, but it must be borne in mind, particularly when aspirin is used (Prieur and Stephan 1994).

Name	Dose per kg body weight per day (maximum daily dose)
Acetylsalicylic acid	50–100 mg (4 g)
Naproxen	15–20 mg (1.5 g)
Ibuprofen	40–60 mg (2.4 g)
Ketoprofen	3–6 mg (300 mg)
Fenoprofen	40–50 mg (3.2 g)
Flurbiprofen	3–4 mg (300 mg)
Diclofenac	2–3 mg (200 mg)
Sulindac	4–6 mg (400 mg)
Meclofenamate	4–7 g (300 g)
Piroxicam	0.3 mg
Indomethacin	1–3 mg (200 mg)
Tolmetin	20–30 g (1.6 g)

Table 5 Non-steroidal anti-inflammatory drugs in children

Slow-acting antirheumatic drugs

Slow-acting antirheumatic drugs are indicated in the polyarticular forms of juvenile chronic arthritis. The most commonly used agents are gold salts and D-penicillamine. Other thiol derivatives can be used, such as tiopronine and sulphasalazine. Synthetic antimalarials are used, particularly in Scandinavian countries. These treatments all take several weeks to work and carry a risk of side-effects that are now well documented in adults. Their use in children must take account of age and the clinical presentation. An improvement is usually observed during the first 6 months, and there is no point in continuing treatment if no result is obtained after this period. Table 6 summarizes the main preparations, their doses, and side-effects.

Name	Dose	Side-effects
Auranofin (oral)	0.3 mg/kg body weight per day (max. 9 mg)	Diarrhoea Same as above but less severe
D-Penicillamine (oral)	5, for 1 month then 10 mg/kg body weight per day (if necessary 15 mg/kg per day)	Rash Cytopenia Proteinuria Pneumopathy Autoimmunity: myasthenia, lupus
Tiopronine (oral)	10 for 1 month, then 20 mg/kg body weight per day	Same as D-Penicillamine
Sulphasalazine (oral)	20–30 mg/kg body weight per day	Rash Cytopenia Increased transaminases
Hydroxychloroquine (oral)	5–7 mg/kg body weight per day (max. 200 mg)	Retinopathy Cutaneous symptoms Neuropathies Cytopenia

Table 6 Slow-acting antirheumatic drugs in children

Several prospective studies have been carried out in the paediatric setting. D-Penicillamine has been the subject of three comparative trials, which yielded somewhat contradictory results. A French study comparing D-penicillamine with placebo showed that the active drug had a degree of efficacy (Prieur *et al*. 1985), while an American–Russian study comparing D-penicillamine, hydroxychloroquine, and placebo showed no difference in efficacy or tolerability among the three groups (Brewer *et al*. 1986). The frequency of side-effects in these two studies was acceptable, whereas a Norwegian team reported side-effects in 25 per cent of patients on D-penicillamine in a comparative study with gold salts (Kvien *et al*. 1985). Certain studies indicate that gold salts given by the oral route were slightly more effective than placebo (Giannini *et al*. 1990), an effect that persisted beyond 5 years (Giannini *et al*. 1991a) but did not reach statistical significance. Among the thiol derivatives, tiopronine had similar efficacy and tolerability to D-penicillamine when studied by the author. Several paediatric trials of sulphasalazine have been performed. Sulphasalazine appears to be beneficial in certain conditions, especially spondylarthropathies (Job-Deslandre and Menkes 1991). Published results on slow-acting antirheumatic drugs are only mediocre. A meta-analysis of 6-month efficacy on disease activity confirmed the lack of spectacular improvement (Giannini *et al*. 1991b). In addition, most of these studies did not distinguish between the different forms of juvenile chronic arthritis. There is no evidence that combinations of slow-acting antirheumatic drugs are effective in children, although this has not been studied in children in a systematic fashion.

Steroids

Steroids are a very powerful tool but often have unacceptable side-effects in children. They can be given systemically or locally.

Systemic steroids generally are not indicated in the polyarticular form of juvenile chronic arthritis. In severe crippling cases, they can be used to improve the functional status. The side-effects of systemic steroids are the main problem in paediatric use. Rapid weight gain can only be avoided by a strict diet, which must be not only sodium-free but also restricted in slow-resorption carbohydrates. The most troublesome cutaneous side-effects are permanent striae, but these are rare if weight gain is controlled by dieting. Arterial hypertension and diabetes can develop. Osteoporosis can lead to very painful vertebral collapse which necessitates immobilization in a corset. The use of deflazacort, available in some countries, apparently lessens the impact of steroids on bone metabolism (Loftus *et al*. 1993). Aseptic joint necrosis, especially of the hips, is far from rare and complicates underlying diseases that can also involve the joints. The onset of steroid-induced cataracts must be monitored closely. Finally, above all in children, daily steroid therapy arrests growth. This is overcome by alternate-day dosing when possible (Prieur 1993).

Local steroid therapy can be applied to the joints and eyes.

Intra-articular steroid injections have modified totally the prognosis of joint manifestations. They can be used in polyarticular forms when stiffness and flexion contractures cannot be controlled by general therapy. Among the many available products, only triamcinolone hexacetonide gives satisfactory results. It is fairly potent and must therefore be used with care to avoid local complications (mainly cutaneous atrophy at the injection site). Some attempts at multiple joint injection have given encouraging results (Pugh *et al*. 1995)

This procedure is straightforward for large joints such as the knee, but general anaesthesia may be necessary for small and/or tight joints, especially in the very young child. There is no lower age limit for this type of treatment as long as the precautions for use are respected; this also means that an experienced practitioner must treat the joints of small children. In general, the volume injected must be adapted to the volume of the joint, and the product must not be injected 'under pressure'. The joint must be rested with a splint for 3 days. Complications are rare and mainly consist of cutaneous atrophy at the injection site or asymptomatic intra-articular calcifications. Needless to say, the potency of this preparation contraindicates its use for injecting tendon sheaths, where only water-soluble steroids can be used.

Uveitis, a frequent complication of polyarticular forms of juvenile chronic arthritis with antinuclear antibodies, is treated with steroid-based eye drops and mydriatic agents.

Cytotoxic drugs

Cytotoxic treatments may be of value in severe forms. However, their efficacy is often difficult to determine objectively because of the small number of sound, prospective, multicentre trials. Too many specialists continue to use a given treatment in 'selected cases', meaning that the results are uninterpretable. Ideally, as in oncology, multicentre protocols should be established to accelerate the assessment of the efficacy and tolerability of experimental treatments.

Methotrexate is one of the most widely used immunomodulatory agents. The use of methotrexate in children with rheumatic diseases was proposed when its efficacy and acceptable tolerability were established in adults. The recommended dose, based on a double-blind, placebo-controlled trial, is 10 mg/m ^{2} per week (Giannini *et al*. 1992), but certain authors propose higher doses for particularly resistant forms (Wallace *et al*. 1991)

The tolerability of methotrexate in children is acceptable (Graham *et al*. 1992). In combination with NSAIDs, the level of methotrexate in the blood can increase (Dupuis *et al*. 1990). A few cases of severe liver damage have been reported (Keim *et al*. 1990). Side-effects, which are generally mild, must be monitored every 1 or 2 months. Many patients (almost 30 per cent) show increased transaminase levels. Values usually return to normal when the dose is reduced or treatment is suspended. After a few months of treatment, nausea is fairly frequent immediately after methotrexate intake, and these patients may benefit from a switch to the intramuscular route. Oral aphthae, cytopenia, and pulmonary manifestations are observed rarely.

Methotrexate seems to have clear efficacy in some cases, especially in pauciarticular juvenile chronic arthritis with the presence of antinuclear antibodies (Hallé and Prieur 1991). A European, multicentre, double-blind trial is underway to confirm this observation.

Azathioprine at a dose of 2 to 2.5 mg/kg body weight per day, is used by certain specialists. It is generally well tolerated. However, the possible long-term risk of cancer must be borne in mind, and this agent should probably only be used in severe cases (Silman *et al*. 1988). In juvenile chronic arthritis, azathioprine appears to be slightly more effective than placebo (Kvien *et al*. 1986).

Alkylating agents can be considered in exceptional cases when classical treatments fail to control disease progression. However, chlorambucil (0.2 mg/kg) has a very high mutagenic risk (Prieur *et al*. 1979). It is thus only recommended for life-threatening complications such as secondary amyloidosis, which hardly ever complicates polyarticular juvenile chronic arthritis. Similarly, the indication for cyclophosphamide in the polyarticular forms is only for exceptional cases.

Cyclosporin is effective in adult rheumatoid arthritis at doses below 5 mg/kg body weight per day (Dougados *et al*. 1988). Studies involving patients with juvenile chronic arthritis have failed to show marked improvement (Østensen *et al*. 1988), but further trials are required. The paediatric side-effects of cyclosporin are identical to those observed in adults.

Surgery (see Chapter 6.2)

Surgery plays an increasing role in the management of chronic rheumatic diseases when the medical means are insufficient. The orthopaedic surgeon must participate in the therapeutic discussion. The technical approaches are numerous. Arthroscopy is now efficient for small joints, with adapted equipment. It allows intra-articular examination, biopsies, and synovectomy. Surgery in the form of soft tissue release, osteotomies, surgical realignment, and arthrodesis may be necessary to treat fixed joint deformities. Surgical treatment of growth deformities is mandatory when functional impairment and secondary mechanical problems in adjacent joints develop. Finally, joint arthroplasty must be considered when joint destruction and joint failure lead to major handicap. Anaesthetists must be aware of cervical spine involvement and temporomandibular arthritis, which may make intubation difficult.

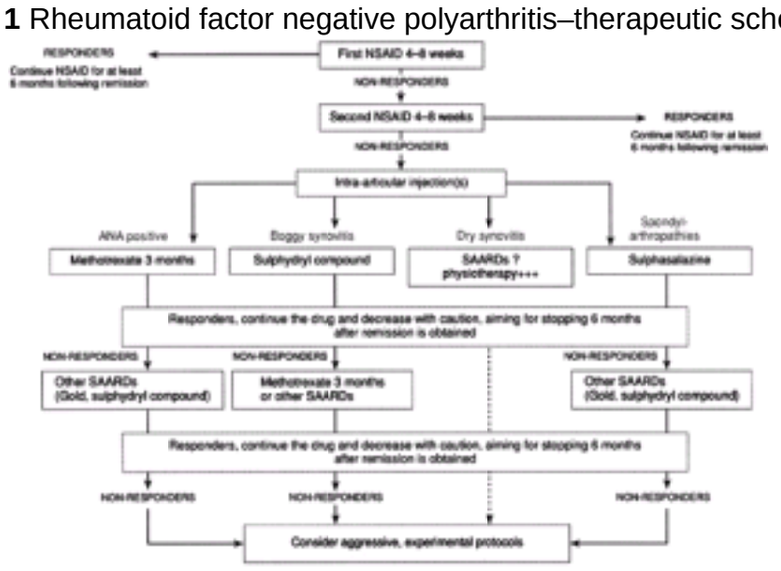
Rehabilitation (see Chapter 6.4)

Physical treatments are necessary for any child with chronic arthritis. The techniques involve applied heat (water, paraffin, or hot packs) to induce muscle relaxation. Flexion deformities must be stretched out either actively by the child or passively by the child, parent, or therapist. Correct positioning of the joint must be obtained at rest with appropriate splinting. Prone positions should be recommended for reading or watching television. Each joint requires a specific technique of rehabilitation to prevent deformities. The treatment must be adapted to the severity of the disease, the impact of each joint on total function, and radiological changes. In cases of surgical intervention, both preoperative and postoperative rehabilitation are necessary.

Conclusion

The management of polyarticular juvenile arthritis involves a common approach which is adapted according to the different subtypes of arthritis as shown in the practical guidelines presented in Box 1.

Box 1 Rheumatoid factor negative polyarthritis–therapeutic scheme



Chapter References

Amor, B., Dougados, M., and Mijiyawa, M. (1990). Critères de classification des spondylarthropathies. *Revue du Rhumatisme*, **57**, 85–9.

Barron, K.S., Person, D.A., and Brewer, E.J. (1982). The toxicity of non steroidal antiinflammatory drugs in juvenile rheumatoid arthritis. *Journal of Rheumatology*, **9**, 149–55.

Barron, K. *et al*. (1992). DNA analysis of HLA-DR, DQ, and DP alleles in children with polyarticular juvenile rheumatoid arthritis. *Journal of Rheumatology*, **19**, 1611–16.

Blau, E.B. (1985). Familial granulomatous arthritis, iritis and rash. *Journal of Pediatrics*, **107**, 689–93.

Brewer, E.J., Jr *et al*. (1977). Current proposed revision of JRA criteria. *Arthritis and Rheumatism*, **20**, 195–9.

Brewer, E.J, Giannini, E.H., Kuzmina, N., and Alekseev, L. (1986). Penicillamine and hydroxychloroquine in the treatment of juvenile rheumatoid arthritis. Results of the USA–USSR double blind placebo controlled trial. *New England Journal of Medicine*, **314**, 1269–70.

Cerna, M. *et al*. (1994). Class II alleles in juvenile arthritis in Czech children. *Journal of Rheumatology*, **24**, 159–64.

Dougados, M., Awada, H., and Amor, B. (1988). Cyclosporin A in rheumatoid arthritis, a double blind, placebo controlled study in 52 patients. *Annals of the Rheumatic Diseases*, **47**, 463–9.

Dougados, M. *et al*. (1991). The European spondylarthropathy study group preliminary criteria for the classification of spondylarthropathy. *Arthritis and Rheumatism*, **34**, 1218–27.

Dupuis, L.C. *et al*. (1990). Methotrexate–non steroidal antiinflammatory drug interaction in children with arthritis. *Journal of Rheumatology*, **17**, 1469–73.

Fantini, F. *et al*. (1987). HLA phenotypes in Italian children affected with juvenile chronic arthritis. *Clinical and Experimental Rheumatology*, **5**, 17.

Fernandez-Vina, M.A., Finck, C.W., and Stastny, P. (1990). HLA antigens in juvenile arthritis. Pauciarticular and polyarticular juvenile arthritis are immunogenetically distinct. *Arthritis and Rheumatism*, **33**, 1787–94.

Giannini, E.H. *et al*. (1990). Auranofin in the treatment of juvenile rheumatoid arthritis—results of the USA–USSR double-blind, placebo controlled cooperative trial.

Giannini, E.H. *et al*. (1991*a*). Auranofin therapy in juvenile rheumatoid arthritis: results of the five-year open label extension trial. *Journal of Rheumatology*, **18**, 1240–2.

Giannini, E.H. *et al*. (1991*b*). Meta-analysis of antirheumatic drug trials in juvenile rheumatoid arthritis. *Arthritis and Rheumatism*, **34**, S152.

Giannini, E.H. *et al*. (1992). Methotrexate in resistant juvenile rheumatoid arthritis. *New England Journal of Medicine*, **326**, 1043–9.

Graham, L.D., Myones, B.L., Rivas-Chalon, R.F., and Pachman, L.M. (1992). Morbidity associated with long-term methotrexate therapy in juvenile rheumatoid arthritis. *Journal of Pediatrics*, **120**, 468–73.

Hall, P.J. *et al*. (1989). HLA and complement C4 antigens in polyarticular onset seronegative juvenile chronic arthritis: association of early onset with HLA-DRw8. *Journal of Rheumatology*, **16**, 55–9.

Hallé, F. and Prieur, A-M. (1991). Evaluation of methotrexate in the treatment of juvenile chronic arthritis according to the subtype. *Clinical and Experimental Rheumatology*, **9**, 297–302.

Hallé, F. *et al*. (1991). Pharmacokinetics of pirprofen in children with juvenile chronic arthritis. *European Journal of Drug Metabolism and Pharmacokinetics*, **16**, 29–34.

Job-Deslandre, C. and Menkes, C.J. (1991). Sulphasalazine in the treatment of juvenile spondylarthropathy. *Arthritis and Rheumatism*, **34**, S153.

Kaplan, E.K. (1990). Rheumatic fever. *Current Opinion in Rheumatology*, **2**, 836.

Keim, D., Ragadale, C., Heidelberger, K., and Sullivan, D. (1990). Hepatic fibrosis in the use of methotrexate for juvenile arthritis. *Journal of Rheumatology*, **17**, 846–8.

Koné-Paut, I., Yurdakul, S., Bahabri, S., Shaefae, N., Ozen, S., and Bernard, J-L. (1995). Epidemiological features of Behçet's disease in children: an international collaborative study of 86 cases. *Arthritis and Rheumatism*, **38**, S336.

Kvien, T.K., Hoyeraal, H.M., and Sandstad, B. (1985). Azathioprine versus placebo in patients with juvenile rheumatoid arthritis: a single center double-blind comparative study. *Journal of Rheumatology*, **13**, 118–23.

Kvien, T.K., Hoyeraal, H.M., and Sandstad, B. (1986). Gold sodium thiomalate and D-penicillamine. A controlled, comparative study in patients with pauciarticular and polyarticular juvenile rheumatoid arthritis. *Scandinavian Journal of Rheumatology*, **14**, 346–54.

Leak, A.M., Tuaillon, N., Muller, S., and Woo, P. (1993). Study of antibodies to histones and histone synthetic peptide in pauciarticular juvenile chronic arthritis. *British Journal of Rheumatism*, **32**, 426–31.

Loftus, J.K., Reeve, J., and Hesp, R. (1993). Deflazacort in juvenile chronic arthritis. *Journal of Rheumatology*, **20** (Suppl. 37), 40–2.

Lovell, D.J., Giannini, E.H., and Brewer, E.J., Jr (1984). Time course of response to non steroidal antiinflammatory drugs in juvenile rheumatoid arthritis. *Arthritis and Rheumatism*, **27**, 1433–7.

Malagon, C. *et al*. (1991). The iridocyclitis of early onset pauciarticular juvenile rheumatoid arthritis: outcome in immunogenetically characterized patients. *Journal of Rheumatology*, **19**, 160–3.

Malleson, P., Petty, R.E., Fun, M., and Candido, P.M. (1989). Reactivity of antinuclear antibodies with histones and other antigens in juvenile rheumatoid arthritis. *Arthritis and Rheumatism*, **32**, 919–23.

Monestier, M. *et al*. (1990). Antihistone antibodies in antinuclear antibody-positive juvenile arthritis. *Arthritis and Rheumatism*, **33**, 1836–41.

Morling, N. *et al*. (1985). HLA antigen frequencies in juvenile chronic arthritis. *Scandinavian Journal of Rheumatology*, **14**, 209–16.

Østensen, M., Høyeraal, H.M., and Kåss, E. (1988). Tolerance of cyclosporin A in children with refractory rheumatoid arthritis. *Journal of Rheumatology*, **15**, 1536–8.

Østensen, M., Fredriksen, K., Kåss, E., and Rekvig, O.P. (1989). Identification of antihistone antibodies in subsets of juvenile chronic arthritis. *Annals of the Rheumatic Diseases*, **48**, 114–17.

Pauls, J.D., Silverman, E., Laxer, R.M., and Fritzler, M.J. (1989). Antibodies to histones H1 and H5 in sera of patients with juvenile rheumatoid arthritis. *Arthritis and Rheumatism*, **32**, 877–83.

Peralta, J.L. and Prieur, A.-M. (1990). Arthrite chronique juvénile avec présence d'anyicorps antinucléaires sériques. *Archives de Pediatrie*, **47**, 497–502.

Ploski, R. *et al*. (1993). HLA class II alleles and heterogeneity of juvenile rheumatoid arthritis: DRB1*0101 may define a novel subset of the disease. *Arthritis and Rheumatism*, **36**, 465–72.

Ploski, R. *et al*. (1995). Association to HLA-DRB1*08, HLA-DPB1*0301 and homozygosity for HLA linked proteasome gene in juvenile ankylosing spondylitis. *Human Immunology*, **44**, 88–96.

Prieur, A-M. (1993). The place of corticosteroid therapy in juvenile chronic arthritis in 1992. *Journal of Rheumatology*, **30**, (Suppl. 37), 32–4.

Prieur, A-M. and Stephan, J-L. (1994). Macrophage activation syndrome in pediatric rheumatic diseases. *Revue du Rhumatism* (*English Edn*), **6**, 385–8.

Prieur, A-M., Balafrej, M., Griscelli, C., and Mozziconacci, P. (1979). Résultats et risques long terme des traitements immunosuppresseurs dans l'arthrite chronique juvénile. *Revue du Rhumatism*, **46**, 85–90.

Prieur, A-M. *et al*. (1985). Evaluation of D-penicillamine in juvenile chronic arthritis. A double-blind multicentre study. *Arthritis and Rheumatism*, **28**, 376–82.

Pugh, M.T., Grosse, R., and Southwood, T.R. (1995). Multiple joint corticosteroid injections in polyarticular juvenile chronic arthritis: a single-blind one year folloiw-up study. *Arthritis and Rheumatism*, **38**, S283.

Silman, A.J. *et al*. (1988). Lymphoproliferative cancer and other malignancy in patients with azathioprine. A 20 year follow-up study. *Annals of the Rheumatic Diseases*, **47**, 988–92.

Stemmer, C., Tuaillon, N., Prieur, A-M., and Muller, S. (1995). Mapping of B-cell epitopes by antibodies to histones in subsets of juvenile chronic arthritis. *Clinical and Experimental Immunopathology*, **76**, 82–9.

Tuaillon, N. *et al*. (1990). Antibodies from patients with rheumatoid arthritis and juvenile chronic arthritis analysed with core histone synthetic peptides. *International Archives of Allergy and*

Immunology, **91**, 297–305.

Wallace, C., Sherry, D., and Salmonson, K. (1991). Treatment of juvenile rheumatoid arthritis with higher dose methotrexate. *Arthritis and Rheumatism*, **34**, S152.

Wood, P.H.N. (1978). Special meeting on nomenclature and classification of arthritis in children. In *The care of rheumatic children* (ed. E. Munthe), p 47. European League Against Rheumatism, Basel.

5.7.1 Systemic lupus erythematosus in adults

David A. Isenberg and Angela C. Horsfall

Introduction

Systemic lupus erythematosus has taken the mantle of syphilis as the great mimic of other conditions. It is probably better to think of it as a group of related disorders rather than a single disease entity. By analogy it might also be compared to the Hydra monster of ancient Greek mythology (Fig. 1). This beast, one of the offspring of Echidne and Typhon was notoriously unpleasant and possessed numerous heads. It was said that cutting off one head led to the growth of two or three others. Lupus presents in many unpleasant guises and the successful treatment of, say, joint pain in lupus may be followed by the emergence of skin rash or pleuropericardial involvement. There is another analogy with lupus as no one could agree precisely how many heads the Hydra had or about their appearance—similarly there is disagreement as to how disease activity in lupus should be assessed.

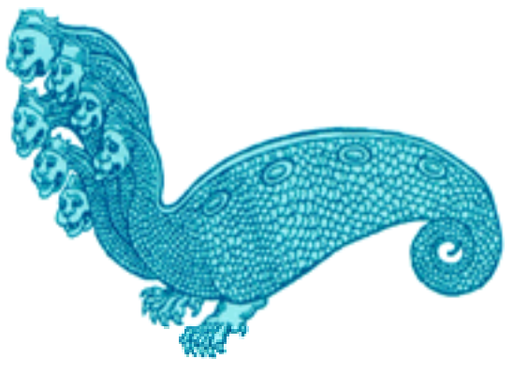


Fig. 1 The Hydra—a useful analogy of lupus (see text).

In this chapter we will detail the clinical features of lupus, analyse its serology, appraise the experimental models of the disease, and review studies of its immunopathology and treatment. Drug-induced forms of lupus are considered in Chapter 5.19.1 but the disease in the male and in elderly people is highlighted here, and several controversial areas are discussed. The major historical aspects of lupus, reviewed in detail elsewhere, are indicated in Table 1.

Table 1 Historical aspects of lupus

Definition and classification of lupus

Systemic lupus erythematosus is perhaps best defined as a clinical syndrome with a complex, multifactorial aetiology, characterized by inflammation and the involvement of most of the body's organs or systems. It is subject to many remissions and exacerbations and, although the musculoskeletal system and skin are invariably affected, frequently gives rise to manifestations in the kidney, heart, lungs, and central nervous system. The diversity among its clinical features is matched by an apparent diversity among the autoantibodies detectable in the serum.

The American Rheumatism Association (now the American College of Rheumatology, **ACR**) has published two sets of criteria in 1971 and 1982, which have been widely adopted. Strictly speaking the criteria are for the classification of the disease rather than use as a diagnostic tool, although in practice there is blurring of this distinction. The 1982 revised criteria are set out in Table 2.

Table 2 Revised criteria of the American Rheumatism Association for the classification of systemic lupus erythematosus (SLE)

The variable clinical and serological expression of lupus makes it easy to obtain a distorted view of the disease. Patients present to a variety of specialists, each of whom will see a different spectrum of the disease. Thus bias amongst the reporting physicians must be borne in mind when assessing reports about lupus. A broad overview of the cumulative percentage incidence of the features of systemic lupus in five large series is shown in Table 3. It must be remembered that the incidence of some clinical features varies between ethnic groups. A recent study of 137 Chinese patients with systemic lupus in Hong Kong reported a relatively low incidence of arthritis (71 per cent) compared with other groups but a high incidence of renal involvement (70 per cent) (Lee *et al*. 1993).

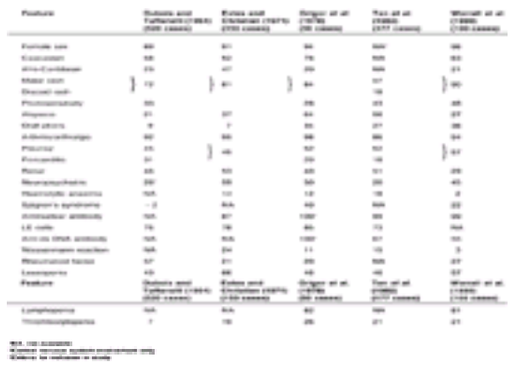


Table 3 Cumulative percentage incidence of systemic lupus features in 100 patients in comparison with other large studies

Epidemiology and natural history

Lupus is a worldwide disease. Although it has been estimated that approximately 1 in 250 black women in the United States and the West Indies, about 1 in 1000 Chinese, and 1 in 4300 Caucasians in New Zealand have systemic lupus, there are some curiously conflicting data. In particular, it seems that lupus is rare in most parts of Africa (Fessel 1988). Two recent studies from urban centres in the United Kingdom have highlighted the significant variation in the prevalence of lupus among different ethnic groups sharing much the same environment. The study from Nottingham (Hopkinson *et al*. 1993), while noting a prevalence of 45.4/100 000 per year in women (3.7/100 000 per year in men), implied that the numbers of lupus patients of Afro-Caribbean and Asian origin were overrepresented, but did not quote any figures. In contrast, the study from Birmingham (Johnson *et al*. 1995) reported prevalence rates of 36.2, 90.6 and 206/100 000 among women of Caucasian, Asian, and Afro-Caribbean origin respectively.

The important genetic contribution to the aetiology of lupus is emphasized by the study of twin concordance reported by Deapen *et al*. (1992). Of 107 twin pairs studied, concordance among monozygotic pairs was found to be 24 per cent compared with 2 per cent among dizygotic pairs. While the figure for monozygotic twins is lower than previously reported, it is still 12 times that for the dizygotic pairs and may in fact be an underestimate, as not all of the twins were examined personally by the authors and long-term follow-up of the pairs was restricted.

It is widely agreed that lupus is approximately 10 to 20 times more common in women than men. There is also little dispute that the overwhelming majority of patients with lupus will develop their disease between the ages of 15 and 40 years.

Although, in the early part of the century, lupus was considered as a serious and frequently fatal disease, perceptions of it have changed considerably. This seems to reflect the easier identification of milder cases with the introduction of widely available tests for measuring antinuclear antibodies. The introduction of corticosteroids and immunosuppressive drugs, dialysis, and renal transplantation has improved the chances of survival in the more serious cases. However, as will be discussed, lupus continues to cause considerable morbidity and 10 to 20 per cent of patients succumb from either the disease, a side-effect of its treatment, or both within 10 years of follow-up.

Clinical features

Non-specific features

Lupus, in common with many other chronic diseases, is accompanied by a variety of non-specific or general features. Of these, lethargy is frequently the most disabling and the least likely to attract the sympathy of the physician! It is, however, invariably present in active lupus (Wysenbeek *et al*. 1993) and often requires counselling. Careful examination frequently reveals the presence of lymphadenopathy, especially in the axillae, which persists long after patients have gone into remission. On occasion it may be so prominent in the neck or under the arms that a biopsy has to be taken to exclude any more sinister pathology. Patients with active lupus may experience weight loss and, on occasion, nausea. Most of these general features will improve when treatment is commenced.

Musculoskeletal involvement

Arthralgia occurs in about 90 per cent of patients with systemic lupus. The joint pain is polyarticular and frequently symmetrical, episodic, and flitting in nature. It is accompanied in about half the patients by early morning stiffness. Very often the patient's symptoms outweigh the objective signs, and major synovial effusions are rare. Severe clinically overt arthritis with joint deformity is probably confined to 5 to 10 per cent of these patients (Spronk *et al*. 1992). Furthermore, unlike patients with rheumatoid arthritis, the deformities (Fig. 2) are usually related to an intense tenosynovitis and less frequently to synovial hypertrophy, with or without bone erosion. These deformities in the hands are known as Jaccoud's arthropathy, which is generally a reversible subluxation.

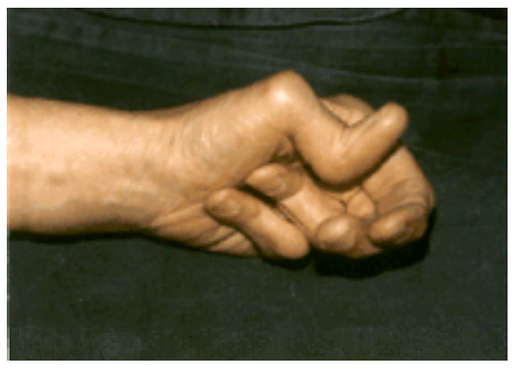


Fig. 2 Severe hand deformity in a lupus patient due to chronic tenosynovitis, not an erosive arthritis.

Joint involvement is relatively mild and, therefore, it is uncommon for 'lupus joints' to be examined histologically but a 'lupus synovium' has been described with a characteristic minor cellular inflammation, occasional haemotoxylin bodies, and non-specific vasculitis and perivasculitis. Electron microscope studies have revealed cytoplasmic inclusions in vascular endothelial cells on occasion. Synovial fluid examination usually reveals a low white-cell count (less than 3000 cells/mm ^{3}), in which mononuclear cells predominate. The fluid is occasionally positive for rheumatoid factor or antinuclear antibodies. Immune complex deposition has been described in synovial tissues and is thought to be responsible for the observed inflammatory lesions. Approximately 2 to 3 per cent of patients classified as having lupus also meet the ACR's criteria for rheumatoid arthritis, with clear erosive disease.

An occasional complication of joint involvement in lupus is spontaneous tendon rupture, which is generally confined to the patella or Achilles tendon (Furie and Chartash 1988). Its aetiology is not well understood, but in some cases at least it seems to be related to inflammatory changes in and around the tendon as a result of the underlying disease process. It is also possible that corticosteroid therapy is in part responsible.

Other, less common, musculoskeletal features include subcutaneous nodules (present in about 5 per cent of patients with systemic lupus and indistinguishable from those found in patients with rheumatoid arthritis), calcinosis (less common in lupus than in scleroderma or dermatomyositis), chondritis, and avascular necrosis. The last of these features occurs in 5 to10 per cent of lupus patients, most cases being associated with prior corticosteroid therapy, and in many instances the avascular necrosis occurs at multiple sites. The remainder are caused usually by small vessel vasculitis or fat emboli.

Myalgia, muscle weakness, and tenderness have been reported in up to 60 per cent of patients with lupus, although a true myositis is confined to about 5 per cent of these patients (Isenberg and Snaith 1981). Treatment with corticosteroids and chloroquine may cause a myopathy, but in the main the myalgia experienced by patients seems to be a complication of adjacent joint involvement.

Histologically a vacuolar myopathy has been described in lupus. This is identified by the presence of plump, swollen sarcolemmal nuclei with other prominent vacuolated nuclei, centrally located within the muscle fibre. Immunoglobulin deposition is often seen in the muscles of patients with lupus (Isenberg 1983) but this is irrespective of whether they have clinically overt muscle disease and seems to relate better to minor fibre damage as a secondary event, rather than to a primary inflammatory myopathy.

Dermatological involvement

Although lupus takes its name from the classic butterfly rash found over the bridge of the nose and malar bones (see Fig. 3(a)), this is actually found in only about one-third of patients with systemic lupus. There are, however, numerous other forms of dermatological involvement (Pistiner *et al*. 1991, and see Table 4). These include maculopapular rashes and discoid lesions (Fig. 3(b)), splinter haemorrhages, dilated capillaries at the nail base, bullous lesions, angioneurotic oedema, and buccal and nasal ulceration (Fig. 3(c)). Vasculitis affecting the fingers and toes is common and frequently causes pain and tenderness (Fig. 3(d)). Many lupus rashes are clearly photosensitive and thus confined to the light-exposed areas. Not surprisingly, photosensitivity is commonest in white females. It is thus most important to advise patients with systemic lupus to avoid sun exposure and use sunscreen even when driving with an arm exposed through an open window.

Frequency	Manifestation
Common (20–50% approx.)	Butterfly rash Photosensitivity Non-specific maculopapular lesions Chronic discoid lesion Non-scarring alopecia Purpura/petechiae
Less common (5–20% approx.)	Mucous membrane lesions Urticaria Diffuse hyperpigmentation Leg ulcer Subcutaneous nodules
Occasional (<5%)	Periorbital oedema Jaundice Severe scarring alopecia Pruritus Bullae Panniculitis Psoriaform lesions

Table 4 Cutaneous manifestations in systemic lupus

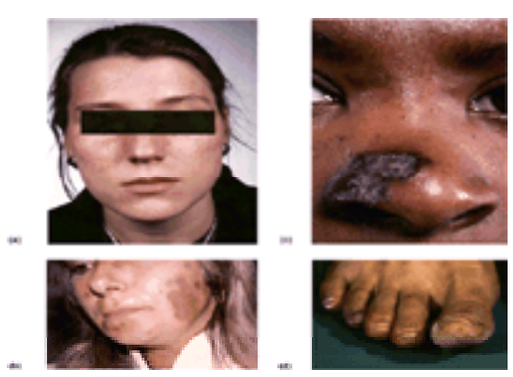


Fig. 3 (a) Butterfly rash; (b) a discoid lupus rash; (c) nasal ulceration; (d) vasculitis affecting the toes, with associated hyperkeratosis.

Alopecia is a common feature of lupus. Although the lack of precision in interpreting and reporting it led to this feature being deleted from the American Rheumatism Associations's 1982 revised criteria for lupus, it remains an important component of the disease. It is usually diffuse but non-scarring, and more often a cause of limited upset. In contrast severe, scarring alopecia is uncommon but can be quite devastating for those who develop it.

Among the less common but important dermatological features of lupus are diffuse hyperpigmentation, usually most prominent on the light-exposed and extensor surfaces of the body, and lupus panniculitis. The latter is a form of lipoatrophy which usually develops as a relapsing, nodular, non-suppurative lesion. The nodules resemble those seen in Weber–Christian disease. The skin overlying the nodules may ulcerate, and scars invariably remain after healing takes place.

A variant of systemic lupus known as subacute cutaneous lupus erythematosus has been described. Early reports suggested that there were two types, annular and papulosquamous, and that the condition was associated with anti-Ro antibodies and HLA DR3. More recent evidence (Callen and Klein 1988) suggests that this group of patients is less distinctive but generally easily controlled.

Immunoglobulin deposition at the dermal/epidermal junction has been recognized for some 40 years. These immunoglobulins are usually of the IgG or IgM isotype. Intriguingly these depositions may be identified in areas of skin which are not light exposed (such as the buttocks) and which have no rash. This forms the basis of the so-called lupus band test (see Fig. 4). Complement components may also be found at the dermal/epidermal junction.

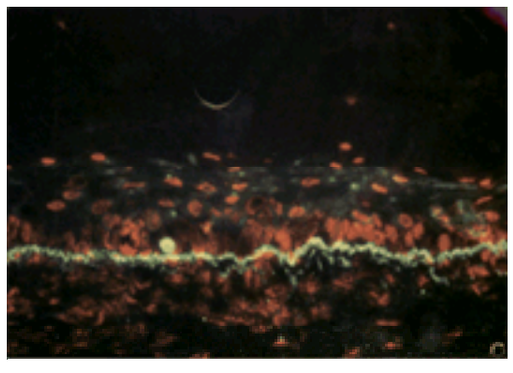


Fig. 4 Lupus band tests—the linear green staining is due to the deposition of IgM identified by a fluoresceinated antihuman antibody; the red counterstain is propidium iodide.

Cardiovascular and pulmonary involvement

Involvement of the heart in lupus has been recognized for nearly 100 years since Osler first described pancarditis. The names of Libman and Sachs are often thought of as synonymous with cardiac involvement, following their classic descriptions of non-bacterial verrucous endocarditis. There have now been multiple studies confirming the effects of lupus on the heart. These may be subdivided into pericardial, myocardial, or endocardial/valvular involvement.

Pericardial disease

Pericardial disease is the most common component of heart involvement in lupus. However, the frequency with which pathology is implied by various tests far outweighs the number of patients with clinically evident involvement. A pericardial rub is also more common than significant accumulations of pericardial fluid. Mandell (1987) reviewed 22 studies of pericardial involvement in lupus and concluded that whereas 29 per cent of the patients had clinical evidence of the disease, echocardiography revealed abnormalities in 37 per cent and necropsy studies showed that 66 per cent of patients had pericardial involvement. Abnormalities of the electrocardiogram, notably of the T waves, are detectable in up to 75 per cent of patients. Echocardiographic findings more commonly show pericardial thickening than large effusion. The results of pericardial fluid analyses may mimic those seen in bacterial pericarditis, and on occasions this latter diagnosis is difficult to distinguish. Studies of pericardial fluid have shown the presence of an antinuclear antibody, LE cells, and even hypocomplementaemia.

When large pericardial effusions have been discovered there may be further underlying complicating factors, such as uraemia and viral or bacterial infection. Constrictive pericarditis can develop within a few weeks of the first appearance of a pericardial effusion. There is some evidence that corticosteroid therapy may contribute to the development of constriction.

Histological abnormalities vary from occasional foci of fibrinoid degeneration and inflammatory infiltrates, to far more extensive lesions. Immune complex components have also been found throughout the pericardial tissue, even in areas where histologically the pericardium looks normal. There is no diagnostic pathological finding for lupus pericarditis, with the possible exception of haematoxylin bodies.

Myocardial disease

True myocardial involvement is less frequent than pericardial disease. Again, however, the results of investigations and necropsy studies suggest that involvement is much more common than is suspected clinically. Clinical myocarditis, usually defined by combinations of unexplained tachycardia, congestive heart failure, arrythmias, prolongation of the PR interval on electrocardiography, or cardiomegaly without pericardial effusion or valvular disease, occurs in up to 15 per cent of patients with lupus.

Echocardiographic studies have suggested that myocardial function can reversibly deteriorate in parallel with flares of generalized lupus activity. Perhaps not surprisingly the ability to correct abnormalities is more obvious in normotensive than in hypertensive patients. This probably reflects transient myocarditis or increased local ischaemia secondary to small vessel obstruction due to vasculitis, for example. It is thought that the risk of myocardial ischaemia is increased by concomitant steroid therapy.

Histological studies of myocardium have indicated that a mild non-specific perivascular infiltration with lymphocytes and neutrophils is relatively common. Intimal proliferation of the smaller intramyocardial arteries is also commonly reported, together with hyalinized vessels that may reflect either previous arteritis or primary thrombosis. The latter is of particular interest in view of the recognized links with antiphospholipid antibodies. The propensity for corticosteroid therapy to increase the risk factors, such as hypertension, hypercholesterolaemia, and obesity, for coronary artery disease has been emphasized recently (Petri *et al*. 1994). In contrast hydroxychloroquine was associated with a lowered serum cholesterol.

Valves

Conduction defects and rhythm disturbances are recognized as occasional features of lupus, but have rarely been found in more than 10 per cent of patients.

Systolic murmurs have been recorded in up to a third of patients with lupus, but in the majority of cases this probably represents the hypodynamic circulation secondary to the chronic anaemia often found in these individuals. In contrast, diastolic murmurs are rather rare.

The classic endocarditis described by Libman and Sachs (1924), although identified in up to 50 per cent of cases at autopsy, rarely causes clinically significant lesions. Histologically small (1 to 4 cm) vegetations (verrucae) comprising proliferating and degenerating valve tissue with fibrin and thrombi are seen. A recent prospective echocardiographic study of 132 consecutive patients with lupus reported a prevalence of valvular lesions of 22.7 per cent (Khamashta *et al*. 1990). These lesions were most commonly found adjacent to the edges of the mitral and aortic valves and have been shown to contain immunoglobulin and complement components, notably within the walls of the small junctional vessels in the active portions of the verrucous endocardial lesions. These deposits might therefore represent immune complexes deposited via the circulation. In the report referred to above, the valve vegetations were associated with the presence of antiphospholipid antibodies. Leung *et al*. (1990) also found a correlation between antiphospholipid antibodies and both valvular abnormalities and isolated left ventricular dysfunction in a study using M-mode, 2-D, and Doppler echocardiography.

In Mandell's review (Mandell 1987) of the reports of haemodynamically significant valvular disease in lupus, aortic incompetence and mitral regurgitation were the most frequently found among the paucity of published case reports.

Bacterial endocarditis has been reported on a number of occasions in patients with lupus. It has not been determined whether the most likely cause is a consequence of the underlying immunopathology of the disease or the predisposition to infection. However, reports of bacterial endocarditis in these patients do antedate

corticosteroid therapy, suggesting that in some patients at least it is the primary immunopathology which predisposes to secondary bacterial infection. As most of the lesions of Libman–Sachs endocarditis are too small to be assessed accurately by echocardiography, any vegetations which can be identified in a patient with lupus who is febrile should certainly raise the possibility of bacterial endocarditis.

Pulmonary disease

The nature and features of pulmonary disease are indicated in Table 5. By far the commonest feature is pain due to pleurisy, which affects approximately half of lupus patients at some time. This pain may be uni- or bilateral and is usually present at the costophrenic angles anteriorly or posteriorly. Pleural effusions, usually small, also uni- or bilateral, straw-coloured, with a protein level generally greater than 3 g/dl, high mononuclear cell count but normal glucose level, are found in a quarter of the patients. Low levels of pleural fluid complement are quite common compared with effusions in other conditions, such as heart failure and cancer, but a positive antinuclear antibody is recorded very infrequently.

1. Pleuritic pain/pleuritis is present in 20–60%
2. Pleural effusions are found in 20–30%. These are usually small volume, straw coloured, white cell count 3–5000/mm^3 mostly mononuclear cells and lymphocytes, glucose levels approximate to those in the blood (unlike rheumatoid arthritis in which the levels are lower), antinuclear antibodies may be detected, protein content varies from 2.75 g–6.4 g%.
3. Chest radiographs and lung function tests (and autopsy findings) invariably indicate a greater degree of pulmonary involvement than is evident clinically.
4. Interstitial fibrosis, pulmonary vasculitis, and interstitial pneumonitis are found in up to one-fifth of the patients.
5. Pulmonary hypertension is unusual and has been linked to the presence of antiphospholipid antibodies.

Table 5 Major pulmonary manifestations in lupus

Parenchymal involvement attributable to lupus has been reported in 18 per cent of patients (Haupt *et al*. 1981). These patients had interstitial fibrosis, pulmonary vasculitis, and interstitial pneumonitis. However, these authors argued that many pulmonary lesions such as alveolar haemorrhage, alveolar wall necrosis, and oedema, previously attributed to direct lupus involvement, are probably secondary to factors such as concurrent infection, congestive heart failure, renal failure, and oxygen toxicity. It must be remembered that the immunosuppressive therapy required by many lupus patients does predispose them to concurrent infection. A true lupus pneumonitis is recognized but is rare (less than 2 per cent).

Almost as uncommon is clinically symptomatic diffuse interstitial lung disease. Its manifestations resemble those found in patients with scleroderma and rheumatoid arthritis who may also develop this complication. Thus the slow onset of a chronic non-productive cough with shortness of breath is usually present. Occasionally a more acute presentation occurs after an episode of acute lupus pneumonitis (Weinrib *et al*. 1990).

There is great interest in antiphospholipid antibodies and thrombotic events in systemic lupus. It is accepted that around 10 per cent of patients develop thrombophlebitis and/or a pulmonary embolus. Antiphospholipid antibodies should be sought in patients with these presenting symptoms and also the small number who present with pulmonary hypertension.

Well recognized, but also uncommon, are patients with the so-called 'shrinking lung syndrome' (Hoffbrand and Beck 1965). It is evident that diaphragmatic dysfunction makes a significant contribution to this syndrome.

Abnormal pulmonary function tests (described in detail in Chapter 1.3.5) and notably diminished total lung capacity and flow rates, often show more serious involvement than expected. Haemoptysis is unusual, although commoner than major pulmonary haemorrhage which, while rare, may be life-threatening.

In the relatively few cases studied, immune complex deposition has been correlated with histological evidence of inflammatory lesions in the pleural and pericardial membrane.

Renal involvement

In many published series, renal disease has been the most common cause of death in patients with lupus. However, the clinical symptoms suggesting renal involvement, notably ankle swelling, shortness of breath related to secondary heart failure, and 'frothy' urine, rarely become evident until substantial damage has been done. Thus careful monitoring of the blood pressure for hypertension, the urine for protein, red cells, or casts, and the plasma for raised creatinine and urea levels is most important. Sequira and Balean at The London Hospital appear to have been the first to recognize that nephritis was a component of systemic lupus in 1902, although serious renal disease was thought rare until the 1940s. Histologically, it has been shown that lupus nephritis appears in many guises. The World Health Organization (**WHO**) have subdivided renal lupus into five major categories according to biopsy-derived information. In addition endstage renal disease with sclerosed glomeruli is recognized (see Fig. 5). Kidneys with this appearance are non-functional. These categories are shown in Table 6 and Fig. 5. In brief, the major manifestations are defined as minimal or mesangial change; mild or focal proliferative; severe or diffuse proliferative, and membranous. The glomerulus appears to bear the brunt of the attack in lupus. The range of glomerular changes include swelling, proliferation of mesangial, endothelial, and parietal epithelial cells, with infiltration by monocytes and polymorphonuclear leucocytes. In addition immune complexes, foci of necrosis, and haemotoxylin bodies can be identified in the glomeruli of lupus patients. While the WHO score has been widely adopted, it has drawbacks. It does not, for example, consider tubulointerstitial disease and makes no allowances for varying degrees of severity within individual categories; neither does it recognize the recently described overlap of lupus nephritis and the multiple small thrombi asssociated with antiphospholipid antibodies (see Fig. 5).

World Health Organization classification	Deposits			Proliferation			
	Mesangial	Subendothelial	Subepithelial	Mesangial	Endocapillary	Extracapillary	Necrosis
Class I	–	–	–	–	–	–	–
Class IIa	+	–	–	+	–	–	–
Class IIb	+	–	–	+	–	–	–
Class III	++	+	±	+	+	+	+
Active					Focal	Focal	Focal
Inactive							
Class IV	++	++	+	+++	+++	+	+
Diffuse proliferative							
Membranous proliferative							
Class V							
Pure	+	+	+++	+	+	+	+
With I, II, IV	±	±	+++	+	±	±	±

Table 6 Classification of renal lupus

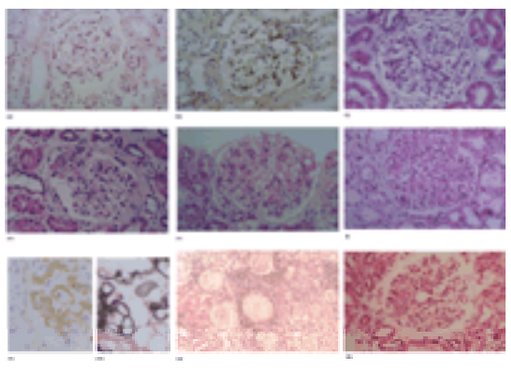


Fig. 5 Renal lupus (by courtesy of Dr M.H. Griffiths). (a) Systemic lupus erythematosus, no renal lesions, WHO class I. In addition to the mesangial immune deposit there is mesangial cell proliferation and the glomerulus appears hypercellular (haematoxylin and eosin). (b) Mesangial lupus nephritis, WHO class IIA. Immune complexes are demonstrable in the mesangium, seen here as granular brown deposits (immunoperoxidase technique for IgG). (c) Mesangial lupus nephritis, WHO class IIB. In addition to mesangial immune deposit there is mesangial cell proliferation and the glomerulus appears hypercellular (haematoxylin and eosin). (d) Focal proliferative lupus nephritis, WHO class III. Segmental proliferation (arrow) is seen in this glomerulus (haematoxylin and eosin). (e) Diffuse proliferative lupus nephritis, WHO class IV. The glomerular capillary walls are irregularly thickened forming, in places, classical 'wire loops' (arrow). There is variable cellular proliferation (haematoxylin and eosin). (f) Membranous lupus nephritis, WHO class V. There is diffuse capillary-wall thickening produced by abundant immune deposits in and on the basement membrane, seen in (i) as granular brown staining (immunoperoxidase technique for IgG). The basement membrane extends out between and around the deposits to produce a series of spikes and circles shown with silver staining in (ii) (hexamine silver technique). (g) Endstage renal disease showing completely sclerosed, non-functional glomeruli. (h) Hyaline thrombus in the glomerulus of a patient with systemic lupus and a high titre of antiphospholipid antibodies.

The role of renal biopsy in the management of lupus nephritis

There is a difference of opinion as to precisely when renal biopsy should be undertaken in patients with lupus, and about its value. The ability of lupus nephritis to transform from one variety to another and for the same biopsy to have more than one histological appearance is partly responsible for the conflict. In addition, few studies of the relationships between renal histology and clinical outcome have actually directly addressed the question as to what information the renal histology adds to the clinical data. Goulet *et al*. (1993) used regression tree techniques to show that combinations of serum creatinine, 24-h urine protein levels, nephrotic syndrome, and duration of prior renal disease provide accurate prognostic information about lupus nephritis without recourse to biopsy.

The biopsy itself is not without its problems, including quite heavy haematuria on occasion. It is obviously most important to assess the clotting capability of the patient before undertaking the biopsy. Fries *et al*. (1978) found that renal biopsy contains important prognostic information but that it was less than that of even the simplest clinical classification. Whiting-O'Keefe *et al*. (1982) applied a stepwise regression analysis to data collected over a 12-month period after renal biopsy in 130 patients with lupus to see if biopsy added any useful information to the clinical data. They found the histological classification did not add significantly to the predictive power of the 'before biopsy' model, but that certain features, notably the percentage of glomeruli which had undergone sclerosis in the presence of subendothelial deposits on electron microscopy, did increase the ability to predict the effect of 12 months of treatment of lupus nephritis. These authors felt that renal biopsy did not add important prognostic information over and above the clinical history examination and laboratory tests. A more recent study (Blanco *et al*. 1994) emphasized that chronicity markers, notably hyalinosis, tubular atrophy, and glomerular sclerosis on light microscopy and subepithelial, mesangial, and intramembranous deposits on electron microscopy, are the best indicators of a poor prognosis. More worryingly Schwarz *et al*. (1993) cast doubt upon the ability of pathologists to reproduce activity and chronicity scores accurately.

In contrast, Stamenkovic *et al*. (1986) described treatment based on renal histology of 56 patients with lupus. The mean follow-up period from first biopsy was 8.2 years, by which time 5.3 per cent were dialysis dependent, but nearly 95 per cent had resumed normal renal function. They found that the biopsy provided valuable information about the state of the kidney, independent of the clinical stage of the disease, and have argued that biopsy alone can improve predictions about 'renal survival' in lupus. McLaughlin *et al*. (1994) in a long-term follow-up study of 123 patients with systemic lupus who had a renal biopsy between 1970 and 1984, showed that the biopsy was helpful in assessing prognosis in patients with normal serum creatinine. In those with an elevated serum creatinine, the biopsy did not contribute additional information about the risk of dying.

Although it is not uniformly agreed, it can be recommended that patients with lupus who have haematuria and/or proteinuria and those with diminished glomerular filtration rate should be seriously considered for renal biopsy. However, as the above review of the controversy about the significance (and reproducibility in reporting) of renal biopsies indicates, the information they provide about prognosis should not be overestimated. The opinion of a pathologist with experience in assessing these biopsies is strongly advised.

Nervous system involvement

The first suggestion of involvement of the central nervous system in lupus was altered mental function, described at the end of the last century by Kaposi, and confirmed by Sir William Osler early this century.

Features of neurological disease range from the common, relatively harmless migraine headaches, to major psychotic episodes, and grand mal seizures, recognized in some lupus patients.

Manifestations of lupus affecting the nervous system can be subdivided into central or cerebral effects, peripheral lesions, and psychological aspects. Additional discussion of this topic is found in Chapter 1.2.1.2 and Chapter 1.3.4.

Central/cerebral involvement

Up to 40 per cent of lupus patients suffer from migraine, although this may be manifested by teichopsia alone. Of much greater concern are the grand mal seizures which may be an initial manifestation of lupus in perhaps 5 per cent of cases, but are present in up to 20 per cent of patients eventually. As with a number of other features it may be difficult to be certain whether the seizures represent true cerebral disease, or a manifestation of more general problems. They may, for example, be secondary to uraemia and other biochemical disturbances associated with renal involvement. Similarly, hemiplegia (and transverse myelitis) may be consequent upon primary neurological disease or could be secondary to hypertension, or associated with the more recently recognized antiphospholipid antibodies. Cerebellar disease in lupus appears to be uncommon as is aseptic meningitis. A variety of organic brain syndromes with impaired temporal–spatial orientation, poor memory, and intellectual deficit are all well recognized and remain difficult management problems.

There has been a resurgence of interest in a small group of patients with lupus who suffer from the movement disorders, chorea or ballismus. On occasion, chorea due to lupus may be difficult to differentiate from that due to rheumatic fever. However, in the more recently described patients, links with the presence of antiphospholipid antibodies have been stressed. This is discussed in more detail later in the chapter.

Ocular lesions in lupus are well recognized. These include conjunctivitis, episcleritis, and cytoid bodies (white patches seen on retinal examination). In addition, retinal haemorrhage, and occasional papilloedema (usually found in association with malignant hypertension) and macular degeneration have all been described. The potential retinal toxicity of antimalarial drugs is discussed later in the chapter.

Peripheral neuropathy

Approximately 10 per cent of patients with lupus develop a peripheral neuropathy in the course of their disease. These are usually sensory, occasionally sensorimotor. Cranial nerve involvement is rather less common, usually associated with active systemic disease and manifested by visual defects, tinnitus, vertigo, nystagmus, ptosis, and facial palsies. Feinglass *et al*. (1976) reported that the most commonly affected cranial nerves in their study were VII, III, VI, V, and IX in order

of decreasing frequency. Optic neuritis was also uncommon, although it may, on rare occasions, be a presenting feature.

Psychological aspects

It has been claimed that up to 70 per cent of patients with lupus suffer a variety of psychiatric abnormalities. However, this label includes depression and anxiety, and most studies have failed to separate the non-specific psychological stresses associated with a debilitating and sometimes painful disease like lupus, from those specifically caused by the disease itself. Whatever the precise cause, a recent report of seven suicides in patients with lupus serves to emphasize that depression must be taken very seriously in systemic lupus (Matsukawa *et al*. 1994).

A lack of significant correlation between indices of general disease activity and psychiatric morbidity has been acknowledged by several authors. However, as with other aspects of lupus, more detailed testing, in this case using psychometric tests and nuclear magnetic resonance imaging, has been claimed to identify subtle degrees of impairment which may not be immediately evident clinically. Thus using a variety of standardized neuropsychological tests Hanley *et al*. (1993) identified cognitive impairment in 21 per cent of their patients with lupus compared with 4 per cent of their rheumatoid and healthy controls. Shortall *et al*. (1995) using a wide ranging set of neurophysiological tests showed that mood and mood disorders in patients with lupus were unrelated to measures of disease activity but were associated with psychological and social factors.

Emotional lability, personality change, impairment of judgement, and difficulty in performing simple tests of cognitive function, such as recall of serial numbers, all suggest organic involvement in lupus. The major psychoses, notably paranoia, schizophrenia, and hypomania, are also well documented. During the 1950s and 1960s, after the introduction of corticosteroids, concern was expressed that large doses of these drugs given for therapeutic purposes might actually be responsible for some of the psychiatric manifestations. Later studies have tended to discount this possibility. For example, Feinglass *et al*. (1976) considered that only 2 of 140 patients had a steroid-induced psychosis.

Investigations of neurological disease

It seems generally agreed that examination of the cerebrospinal fluid in neuropsychiatric disease is not very useful. Patients with lupus may show moderately raised cell counts, some increase in protein and IgG levels, and occasionally low glucose in their cerebrospinal fluid. It has been suggested that C4 levels in the cerebrospinal fluid are low during active disease and return to normal as the patient improves. Its potential use, however, is limited by the need for serial determinations. A range of autoantibodies including anti-Sm, antineuronal, antilymphocytotoxic and, most recently, antiribosomal P has been linked to nervous system involvement. Invariably hopes raised in the initial reports have been dashed by later studies (reviewed in Hay and Isenberg 1993). For example, in 1987 in a retrospective study, antiribosomal P antibodies were reported to be highly specific for lupus psychosis (Bonfa *et al*. 1987). Since then there have been at least five other studies (reviewed by Teh and Isenberg 1994) with conflicting findings, and on balance there seems little value in a single measurement of antiribosomal P antibodies to identify patients with lupus psychosis, although serial estimations may be of value in a few individuals.

However, when brain infarction is suspected as a cause for a sudden cerebral event, it is clearly worthwhile testing for antiphospholipid antibodies (see later). The greatest value of cerebrospinal fluid examination is in ruling out concomitant infection, especially as many patients with lupus are treated with major immunosuppressive drugs.

Electroencephalographic studies have been shown by some to be helpful during flares of cerebral disease. The most commonly observed abnormality was diffuse slow-wave activity but focal changes have been found in some 30 per cent of the patients studied. Electroencephalographic changes tend to be associated with seizure activity or focal neurological signs, although on occasion these findings have been reported in patients with purely psychiatric symptoms. Other, more recent, studies have looked at visual, auditory, and somatosensory-evoked potentials. These tests have still to acquire widespread acceptance.

Computed tomography (**CT**) has been used to analyse patients with lupus for approximately 10 years. Unfortunately the results have been conflicting. Thus diffuse cerebral abnormalities have been found in some patients with active disease, but not in others. CT is, however, very helpful in distinguishing between cerebral infarction and cerebral haemorrhage. It has also been claimed that many patients with neuropsychiatric manifestations of lupus have increased cerebral atrophy as evidenced by enlarged sulci, either with or without ventricular enlargement. In fact corticosteroids may promote this atrophy.

The more recent introduction of nuclear magnetic resonance imaging (**MRI**) now offers a further means of investigating cerebral lupus. Figure 6 shows a patient who presented with severe depression, yet who had multiple small infarcts visible on MRI despite normal electroencephalographic and CT scans. High intensity spots are the most common MRI brain abnormality report in systemic lupus, present in at least one-third of patients (Ishikawa *et al*. 1994). However, these spots are neither specific for neuropsychiatric lupus nor do they show good correlation with central nervous involvement.

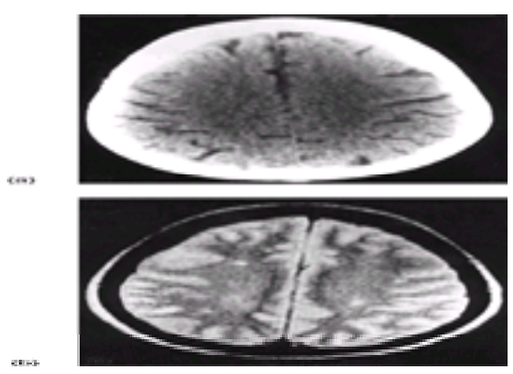


Fig. 6 Equivalent brain sections seen on CT scaning (a) and MRI scanning (b) in a patient with neuropsychiatric disease. Although some widening of the sulci is seen in (a), several discrete ischaemic areas are present in (b) which are not demonstrated on CT scanning.

Haemopoietic involvement

A normochromic, normocytic anaemia, the 'anaemia of chronic disease', is present in up to 70 per cent of patients with lupus. Levels of ferritin in these patients are usually normal. In some individuals other factors contribute to anaemia. Thus some patients have endstage renal disease and many patients are treated with non-steroidal anti-inflammatory drugs, which may cause gastric bleeding. Coombs' positive haemolytic anaemia occurs in approximately 10 per cent of all patients. Much less frequently, a microangiopathic haemolytic anaemia with disseminated intravascular coagulation has been described. It should be noted that a positive Coombs' test is not always associated with haemolysis. An association between pure red cell aplasia and systemic lupus has also now been established.

Leucopenia (< 4×10^9/l) and lymphopenia (< 1.5×10^9/l) are the most frequent abnormalities of the white blood-cell count in patients with lupus. Estimates for the former have ranged from approximately 45 to 65 per cent and for the latter up to 80 per cent. Both T and B lymphocytes are reduced while null cells are increased. Lymphocytotoxic antibodies have been found in over one-third of patients with lupus. In contrast, leucocytosis is rare in lupus in the absence of infection or major corticosteroid therapy.

There are at least three types of clinical presentation of thrombocytopenia associated with lupus. Of these, chronic thrombocytopenia (< 100×10^9/l) has been detected in approximately 20 per cent of most series. This is associated rarely with bleeding episodes in patients with lupus, unlike those unusual cases of acute thrombocytopenia where the fall in the platelet count may be both dramatic and life-threatening. Finally, some patients may present with what initially appears to be an idiopathic thrombocytopenia, usually treated successfully with corticosteroids, which only several years later is followed by other manifestations of the disease. A detailed overview of the haematological manifestations of lupus has been published recently (Keeling and Isenberg 1993).

Other clinical manifestations of lupus

Vascular lesions, notably Raynaud's phenomenon, cutaneous vasculitis, and ulcers and gangrene of the fingers and toes are all well recognized in lupus patients. Approximately one-third of patients have Raynaud's phenomenon, which may antedate the onset of the disease by several years. It is generally relatively easy to control with vasodilating drugs, but on rare occasions it may be associated with gangrene of the extremities (Fig. 7). Unlike the older population which suffers gangrene due to atherosclerosis, the potential for recovery in the patient with lupus is better as the patients are much younger. Active vasculitis in lupus may manifest as necrotic ulcers, small cutaneous infarction, or lupus profundus. Leg ulcers have been recorded in up to 5 per cent of patients with lupus, most commonly around or just above the malleoli.

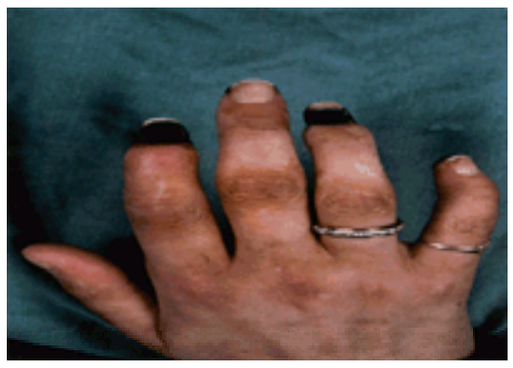


Fig. 7 Gangrene affecting two terminal digits in a young female patient with lupus.

Although many patients with lupus develop some form of gastrointestinal complaint during the course of their disease, it is often difficult to be certain whether the lupus itself, the drugs used in its treatment, or other unrelated causes are responsible. Certainly anorexia, nausea, vomiting, or diarrhoea will occur at some point during the history of disease in over half of the patients, but frequently the cause will turn out to be iatrogenic. Abdominal pain is found in 10 to 20 per cent of the patients. The causes range from mild non-specific gastroenteritis to life-threatening mesenteric vasculitis. On occasions, an aseptic peritonitis may occur which often requires laparotomy to exclude a perforated ulcer or gangrenous piece of valve.

Pharyngitis and dysphagia are occasional features of lupus as is pancreatitis. A problem with treating pancreatitis is that several of the drugs used to manage lupus (corticosteroids, azathioprine, and vasodilators) can precipitate an attack. The subject has been fully reviewed elsewhere (Watts and Isenberg 1989).

Hepatomegaly, which can be detected in approximately a quarter of patients with lupus, is usually a minor enlargement and rarely accompanied by major abnormalities in liver function tests. It is generally agreed that jaundice occurs in less than 5 per cent of patients with lupus. Equally rare is hepatic vasculitis and the Budd–Chiari syndrome, although this latter feature has also been linked to the presence of antiphospholipid antibodies. Splenomegaly is present in approximately 10 per cent of the patients, although the spleen is rarely greatly enlarged.

Apart from the autoantibodies detected in lupus (described later in this chapter) a number of other common blood tests are frequently abnormal in the patients. For example, hypoalbuminaemia has been described in up to 50 per cent of patients, and hypergammaglobulinaemia in up to 60 per cent. IgG and IgM levels, in particular, have been reported to be elevated in patients with lupus; IgA and IgE levels much less frequently so.

The erythrocyte sedimentation rate is increased in the vast majority of these patients, over 90 per cent in some series. Occasionally, however, normal levels are found even in the presence of active disease. In contrast, levels of C-reactive protein are generally, although not always, normal in patients with lupus except for those with a concurrent infection, an erosive arthritis or possibly serotosis. A normal level is thus of some value in helping to rule out infection in these patients although, as levels of C-reactive protein may reflect disease activity in certain individuals, its value is far from absolute.

Lupus and the risk of infection

As indicated in the section on the relationship between lupus and malignancy, the combined effect of the disease itself and its treatment is to render the immune system more prone to infection. It is often impossible to apportion responsibility for such infections to one or the other, but there are major consequences for outcome, discussed later in this chapter. In addition a variety of (generally) non-fatal infectious diseases, such as herpes zoster (Kahl 1994), salmonella (Abramson *et al*. 1985), and candida (Sieving *et al*. 1975), are quite common.

Diseases complicating lupus

Sjögren's syndrome is present in approximately 20 per cent of patients. The dryness of the eyes and mouth differs little from those cases of primary disease, although antibodies to Ro and La are present less frequently. Sjögren's syndrome is reviewed fully in Chapter 5.10. Autoimmune thyroid disease (generally hypo-rather than hyperthyroid) has been identified in 5 to 10 per cent of patients with lupus. However, antibodies to thyroglobulin and thyroid microsomes have been found in up to one-third of the patients. The treatment of under- or over-active autoimmune thyroid disease in patients who also have lupus is similar to those without it. Likewise myositis, also detectable in up to 5 per cent of patients with lupus, is treated no differently from patients with the idiopathic disease. A smaller percentage of patients with lupus (perhaps no more than 1 per cent) has myasthenia gravis. This may antedate or postdate the onset of the lupus itself.

Much has been written as to whether systemic lupus and rheumatoid arthritis can coexist. It is certainly very rare to have lupus glomerulonephritis occurring in patients with seropositive erosive rheumatoid arthritis. However, it is generally accepted that up to 5 per cent of patients do have an erosive arthropathy, suggesting that an overlap between these two conditions may exist in some individuals.

Lupus in special situations

Pregnancy and lupus

Even in healthy individuals, pregnancy results in drastic immunological changes. These have been the subject of much interest in the past few years. Oestrogens, for example, are thought to decrease T-suppressor cell function, while androgens have the opposite effect. Oestrogens also tend to increase immune complex clearance thus decreasing their renal deposition. The CD5 + B lymphocytes are also thought to be under oestrogen control. Progesterone is thought to have immunosuppressive properties and its production increases throughout pregnancy. Prolactin, another hormone associated with pregnancy, is a known modulator of lymphocyte responses to antigen in rodents, and human B and T cells are known to have receptors for it. It is thus evident that the effects of pregnancy on the immune system in general are complex and therefore not surprising that patients with lupus with their significantly disordered immune system may suffer deleterious effects during pregnancy.

There are many conflicting data about the effect of pregnancy on the patient with lupus. Despite earlier reports suggesting that sterility might be common among patients, more recent studies have noted that sterility and fertility were little changed by lupus. An exception to this are those patients in renal failure who do have reduced fertility. Similarly, whereas some reports of an increase in maternal mortality in the patients during pregnancy were described in the 1950s and 1960s, the current view is that the majority of pregnancies do not adversely affect the mother with lupus. Thus in a case–control study of 46 patients with 79 pregnancies, lupus flares occurred no more frequently than in non-pregnant controls (Urowitz *et al*. 1993). The frequency of non-renal complications during pregnancy is relatively low but is a little higher in the period immediately after parturition. Flares of disease, especially renal involvement, may require the introduction of or increase in corticosteroids, and on occasion the baby may have to be induced as early as 30 weeks of gestation.

In contrast to the relatively encouraging outcome for the mother with lupus, fetal outcome is much less certain. A combination of spontaneous abortion and still birth causes a fetal mortality of around 20 per cent. Up to 25 per cent of babies born to mothers with lupus may have to be delivered prematurely, for a combination of

reasons relating to fetal distress as well as maternal ill health.

In the past decade the link between recurrent spontaneous abortion and the presence of antiphospholipid antibodies has been established. This is discussed in detail elsewhere in the chapter. The precise mechanism of these fetal deaths remains uncertain.

Mothers with lupus who have antibodies to Ro and/or La also appear to be prone to develop the so-called neonatal lupus syndrome. It appears that approximately 1 in 20 women who have either of these antibodies will have a child with this syndrome, which is notable for its congenital conduction defects or skin rashes. This subject is discussed in detail in the next chapter on lupus in children.

Lupus in males

Although lupus in males, especially Caucasian males, is uncommon, many different groups have attempted to identify characteristics of male patients with lupus that distinguish them from women with the disease. A comparison of several reported series (Table 7) shows, however, that no clearly defined criteria have been identified. As discussed by Isenberg and Malik (1994) virtually every claim of a distinctive feature in one series is rebutted in others. In the United States, however, it has been reported that the prognosis is worse for male patients with lupus. Among males with lupus there appears to be no evidence of androgen deficiency, although in one large series 50 per cent had elevated plasma oestrogen levels (Miller *et al*. 1983), but corticosteroid administration might have been expected to decrease these values.

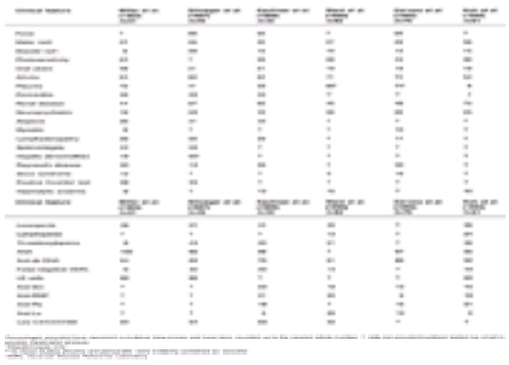


Table 7 Comparison of clinical features and serological abnormalities in six published series of male patients with systemic lupus (%)

Individuals with Klinefelter's syndrome, who have an unusual XXY karyotype, are more susceptible to systemic lupus. Abnormalities in oestradiol metabolism in these patients may be linked to the persistent oestrogenic stimulation, which might explain the predisposition.

Several families in which systemic lupus predominates in males have been described (Lahita *et al*. 1983) where sons may have inherited the disease from their fathers, analogous with disease in the BXSB mouse, an experimental model of lupus which is described later in this chapter.

Lupus in the elderly

It is clearly a matter of opinion at which point lupus in the young or middle aged becomes lupus in the elderly! Most reports have taken 50 or 55 years as a cut off, although there has been very little attempt to relate chronological age to the menopause, a fact which may well be important in the aetiology of the disease in these patients.

In most large series, lupus commencing in the sixth decade of life represents about 10 per cent of the study population. However, there are conflicting reports on the patterns of presentation, organ involvement, serological findings, and prognosis in this group. It appears that the clinical onset of lupus in the elderly is more insidious, milder, has a lower incidence of severe renal and neurological complications, a lower frequency of antibodies to double-stranded DNA, and hypocomplementaemia, but an increased frequency of serositis, interstitial lung disease, and antibodies to Ro and La. The last of these features suggests that the overlap between lupus and Sjögren's syndrome is frequent in an elderly population.

Among less frequent modes of presentation in the elderly, a polymyalgia rheumatica-like picture has been described, and neuropsychiatric manifestations which might easily be confused with other types of organic disease are important in this group. The time between onset of disease and presentation, and between presentation and diagnosis, are increased compared with younger patients, although this should change with increasing awareness that lupus can occur for the first time well into old age—the oldest case reported so far was diagnosed aged 87.

The antiphospholipid antibody syndrome and lupus

Associations between anticardiolipin antibodies and the lupus anticoagulant with systemic lupus erythematosus have attracted considerable interest in the past 15 years.

Anticardiolipin antibodies are part of an overlapping spectrum of antiphospholipid antibodies of which the lupus anticoagulant is part. Detailed analysis of anticardiolipin antibodies has distinguished two major varieties. In patients with infectious diseases the antibodies recognize epitopes on cardiolipin itself. However, in many patients with lupus, the antibodies are probably binding to a complex or neo-epitope formed by phospholipid and a plasma cofactor b $_2$-glycoprotein 1 (Galli *et al*. 1990). Thus b$_2$-glycoprotein 1 dependency was noted for anticardiolipin (40 per cent), antiphosphatidyl serine (20 per cent), and antiphosphatidylinositol (18 per cent) antibodies but not for syphilis or normal sera (Matsuda *et al*. 1994).

A list of the clinical features widely believed to be associated with patients with lupus who have these antiphospholipid antibodies is shown in Table 8. A metanalysis undertaken by Love and Santoro (1990) suggests caution in the interpretation of the published results. In their analysis of 29 published series, they estimated an average frequency of 34 per cent for the lupus anticoagulant and 44 per cent for anticardiolipin antibodies in studies representing over 1000 patients with lupus. However, anticardiolipin antibodies are also prevalent in patients with a wide variety of diseases other than idiopathic lupus, including drug-induced lupus, rheumatoid arthritis, and acute infection. In patients with lupus a statistically significant association has been shown between the presence of either antibody and a history of thrombosis, neurological disorders, and thrombocytopenia (see also Chapter 5.7.3). In a large prospective cohort study of 389 primiparous women assessed at study entry and delivery, 24 per cent were positive for antiphospholipid antibody, 15.8 per cent of whom had fetal loss compared with 6.5 per cent of antibody-negative patients (Lynch *et al*. 1994). Elevated IgG antiphospholipid antibody levels were statistically associated with recurrent fetal loss but not with low birth weight, neonatal distress, or maternal complications.

Venous and arterial thrombosis

Thrombocytopenia

Cerebral disease (including cerebrovascular accident, transient ischaemic attacks, chorea, amaurosis fugax)

Recurrent fetal loss

Pulmonary hypertension

Livedo reticularis

Table 8 Clinical features which have been linked to antiphospholipid antibodies in patients with lupus

A small cohort of patients with lupus with anticardiolipin antibodies has been shown to develop impaired renal function due to multiple small thrombi (Leaker *et al*. 1991). These patients have minimal proteinuria and only gradually increasing renal damage. This type of pathology may coincide with the more typical glomerulonephritis.

A number of contentious issues about antiphospholipid antibodies remain. For example, the precise links with other autoantibodies have been the subject of debate. It is widely accepted that anticardiolipin antibodies are associated with the biologically false-positive Venereal Disease Research Laboratory (VDRL) test. However, early studies undertaken with monoclonal antibodies which suggested significant overlap between those binding cardiolipin and DNA were not supported by studies in patients with lupus. Although low affinity (generally IgM) antibodies to DNA may bind cardiolipin, higher affinity (generally IgG) antibodies to DNA do not. This would imply separate subpopulations of anti-DNA and antiphospholipid antibodies. This is probably an oversimplification, as a single amino-acid substitution can convert an antiphospholipid antibody into an anti-DNA antibody, supporting the view that these antibodies are very closely related (Diamond and Scharff 1984).

Neither the lupus anticoagulant nor anticardiolipin antibodies appear to correlate with age, duration of disease, or a variety of well-known lupus clinical features, including polyarthritis, vasculitis, or serositis.

Lupus and malignancy

Given that the immune system is so disordered in patients with lupus and that many patients are treated with major immunosuppressive drugs, there has been much recent interest in whether there is an increased risk of malignancy in systemic lupus. Table 9 reviews several published series. A rather low frequency of malignancy change (with no obvious predilection for any particular site) is evident. However, in the most recent series quoted (Menon *et al*. 1993), 7 out of 150 patients developed a malignancy, 5 of whom have died (4 as a direct result of the tumour). An occasional association between systemic lupus and Hodgkin's lymphoma has been described (Bhalla *et al*. 1993).

Number of patients with lupus studied	Total number of deaths	Number of deaths from malignancy	Reference
365	68	0	Urman and Rothfield (1977)
428	94	0	Karsh *et al.* (1979)
609	128	4	Wallace *et al.* (1979)
1103	272	0	Rosner *et al.* (1982)
150	16	4	Menon *et al.* (1993)

Table 9 Deaths from malignancy in patients with systemic lupus erythematosus

Assessing lupus disease activity

The assessment of disease activity in lupus is clearly central to patient management, but until recently there has been no consensus on measurement. Liang *et al*. (1988) reviewed, more than 60 different systems (attempting to establish a disease activity index) that have been described in the literature. This, in itself, is good evidence that no one system has won general acceptance. It also reflects the continuing difficulty in determining whether lupus should be thought of as an individual disease or a group of closely related conditions in the absence of a gold standard, by which to judge disease activity.

In the past 12 years more determined attempts to compare and contrast some of the different activity indices have been undertaken. Thus the **SLAM** (systemic lupus activity measures), **SLEDAI** (systemic lupus erythematosus disease activity index), and **ECLAM** (European Community lupus activity measure) are three global score systems which have been shown to correlate well with each other (Vitali *et al*. 1992). They also correlate well with the **BILAG** system (British Isles Lupus Assessment Group) which was established to provide more detailed information about disease activity in each of eight organs or systems (Hay *et al*. 1993). The BILAG index is based upon the principle of the 'physician's intention to treat.' There has been an encouraging international effort to compare the SLAM, SLEDAI, and BILAG systems in combined studies of both 'paper' and real patients. These systems have repeatedly been shown to correlate with one another and to be reliable in evaluating disease activity in systemic lupus. Most recently they have been shown to be sensitive to change in disease activity over time (Gladman *et al*. 1994). Information about these three indices and a comparison between them, based on an assessment of real, as opposed to 'paper' patients, by seven different physicians is shown in Table 10.

	VAS[a]	SLEDAI[a]	SLAM[a]
SLEDAI	0.261	0.732	
SLAM	0.209	0.732	
BILAG[b]	0.152	0.763	0.797

[a]VAS, visual analogue scale completed by the observer. Figures are based on real patient assessments by seven different physicians.
[b]BILAG, British Isles Lupus Assessment Group. This activity index is based on the physician's 'intention to treat' the patient. Lupus activity is divided into eight areas: general features, locomotor system, nervous system, renal involvement, dermatological involvement, pleuropericardial disease, vasculitis, and haematological involvement. Within each system the patients are designated A (active), implying that major immunosuppressive therapy needs to be initiated or increased; B (benefit), the patient is known to have active disease but the therapy does not require alteration; C (contentment), remission or symptoms in that organ/system; D (discount), there is no current involvement in this organ/system; or E, there is no (evidence) of activity in the organ/system now or previously (see Hay *et al*. (1993) for further details).
[c]SLAM (systemic lupus activity measures) devised by Dr Liang (Boston) and SLEDAI (systemic lupus erythematosus disease activity index) described by Drs Linnaeus, Gladman, and Bombardier are two good global score indices (see Liang *et al*. (1988) for further discussion).

Table 10 Correlations (*r* values) between three indices of activity for systemic lupus erythematosus

Equally constructive have been attempts by 'lupologists' to agree an index that distinguishes damage (due to lupus or its treatment) from disease activity. The distinction in not simply an academic one. For example, a patient with shortness of breath may have active but reversible vasculitis which could improve with major immunosuppressive therapy. Alternatively the symptom may be due to fibrosis causing irreversible damage for which there would be no requirement for such treatment. Thus a **SLICC** (systemic lupus international collaborating clinics) damage index has been developed (see Table 11) and is currently being assessed by many groups in North America and Europe.

Table 11 Systemic lupus international collaborating clinics (SLICC) damage index [a]

Immunopathology of lupus

In this section we review the immunopathology of the disease, both in strains of lupus-prone mice and in humans, by describing the specificity of autoantibodies, dysregulation, and abnormalities at the cellular level, and the genetic background that predisposes to the autoimmune response. Since the first edition of this textbook much interest has focused on apoptosis and the genes or 'autogenes' encoding the proteins important in regulation of apoptosis. Programmed cell death, or apoptosis, is quite distinct from the death of cells by necrosis and is a normal feature of cellular regulation. The current state of knowledge in this area will be addressed in both experimental models and human disease.

Experimental models of lupus

Several strains of mice spontaneously develop clinical and serological symptoms that resemble lupus. The most frequently studied strains are summarized in Table 12 and reviewed by Yoshida *et al*. (1990). Comparison of inbred mouse strains with outbred humans may not be relevant genetically but experimental manipulations in mice have made an important contribution to understanding the immunopathology of lupus.

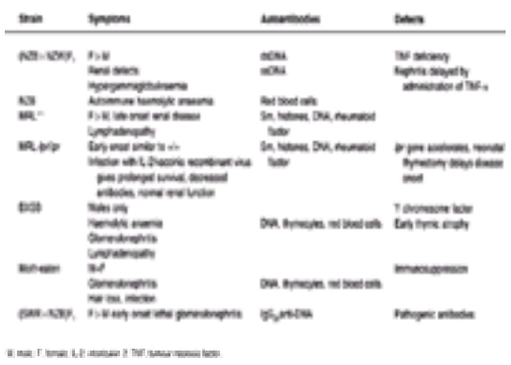


Table 12 Lupus-prone mouse strains

At the genetic level, three genes have been shown to influence the pattern of disease in susceptible strains. The MRL $^{+/+}$ mice spontaneously develop late-onset lupus with a 50 per cent survival time of 18 months. The MRL-*lpr/lpr* differs by only one gene (*lpr* gene) but this single gene accelerates the disease process to give a 50 per cent survival time of between 2 and 4 months. The introduction of the *gld* gene into non-autoimmune C3H/HeJ mice results in lymphoproliferation and autoimmune disease. In contrast the *xid* gene confers protection against autoimmunity in the NZB/W F_1 model by preventing terminal B-cell proliferation and the emergence of autoreactive B-cell clones.

F_1 hybrids between an autoimmune strain (NZB) and a non-autoimmune strain (SWR) have an accelerated autoimmune disease with a high incidence of nephritis. All females are dead by 1 year of age. These mice have IgG2b anti-dsDNA antibodies which are cationic and deposit in the glomerular basement membrane. The pathogenic antibodies are derived from the normal SWR parent and carry a nephritogenic idiotypic marker which is not found in the circulation of either parent. The normal parents have deleted 50 per cent of their T-cell receptor Vb chains and are I-E negative (equivalent to HLA DR in humans), thus they have peripheral T cells with I-E reactive T-cell receptors. The autoimmune parent and the F_1 offspring are I-E positive. Autoimmunity arises from the expression of 'forbidden' T-cell receptors by double-negative T-helper cells and suggests an abnormality in thymic selection/deletion.

Additional evidence from lupus mouse models shows that both cellular oncogenes (C-*myc*, N-*ras*, and C-*myb*) and retroviral genes may contribute towards the autoimmune process. Retroviruses have been implicated in murine lupus. A major envelope glycoprotein antigen, gp70, of type C murine retroviruses is present in the sera of all mice. This antigen can thus be regarded as autologous. The antigen may exist free in the circulation, similar to gp120 in HIV-positive patients, or it may exist as part of a xenotropic endogenous retrovirus. It can behave as an acute-phase reactant (stimulated by lipopolysaccharide, etc.). Circulating complexes of gp70 and anti-gp70 antibodies have been identified in the sera of lupus-prone mice, although free anti-gp70 antibodies cannot be detected (Izui *et al*. 1981). Levels of $gp70^{+}$ immune complexes correlated with nephritis, and gp70 was identified within immune deposits in the kidney (Maruyama *et al*. 1983). A recent report described the preliminary identification of the gp70 receptor as a 100 kDa glycosylated heterodimer on thymic leukaemia cells in mouse and humans.

The *Yaa* gene (Y-chromosome-linked autoimmune acceleration) accelerates disease in MRL $^{+/+}$ mice. These mice have increased levels of gp70–anti-gp70 immune complexes but no increase in the levels of circulating antibodies to DNA. Disease expression has been postulated to be controlled by at least four genes, three of which have been mapped to chromosomes 7 and 17.

Mutations in three autosomal recessive autoimmune genes have been identified in autoimmune strains of mice, reviewed by Mountz *et al*. (1994). All of these genes lead to defects in programmed cell death or apoptosis. The Fas protein has been intimately linked with apoptosis and is normally expressed on the cell surface (CD95). MRL-*lpr/lpr* mice have an endogenous retroviral DNA sequence integrated into the *Fas* gene which results in incorrect membrane expression of the Fas protein and loss of apoptosis. This abnormality would result in failure of self-reactive T cells to undergo apoptosis in the thymus and may account for the accumulation of double-negative T cells (CD4-CD8-) which infiltrate many tissues in these mice. The coining of the expression 'autogene' by Talal encompasses a group of non-MHC, non-immunoglobulin or T-cell receptor genes whose abnormal function contributes to the development of autoimmune disease (Talal 1994). *Fas* may be considered as the first identified autogene. *bcl-2* is an oncogene which inhibits apoptosis, particularly of T cells, and may therefore be considered as a second autogene. Transgenic mice expressing *bcl-2* within B cells show prolonged B-cell survival, production of autoantibodies, prolonged B-cell memory, and inhibition of clonal deletion of self-reactive B lymphocytes, while *bcl-2* transgenic thymocytes display enhanced survival rates in the absence of growth factors and in the presence of lymphotoxic factors (summarized in Aringer *et al*. 1994). The autoimmune consequence of the defect in the *Fas* molecule was highlighted by amelioration of disease in CD2-*Fas* transgenic MRL-*lpr/lpr* mice (Wu *et al*. 1994). The CD2 promoter and enhancer was used to restore *Fas* expression in these mice and resulted in greatly reduced features of autoimmune disease, autoantibody production, and the complete elimination of the development of lymphoproliferative disease. These findings force us to consider that the hyperproliferation and production of abnormal cells in autoimmune disease may be due to defects in the appropriate removal of cells rather than the traditional definition of defects of excessive proliferation and stimulation of certain cells.

Recently a non-autoimmune strain of mice has been used to study human systemic lupus. These mice carry the SCID gene (severe combined immunodeficiency disease) and consequently lack T and B lymphocytes. This deficiency allows human peripheral blood lymphocytes to be grafted into the peritoneum of the mice and survive for up to 7 months. When peripheral blood lymphocytes from patients with lupus were injected into SCID mice over 3 mg/ml of human IgG was found in the

circulation. Furthermore, deposits of mouse C3 were found in the renal glomeruli, indicating complement activation by immune complexes. However, no abnormal histology nor deterioration in renal function was associated with these deposits (Duchosal *et al*. 1990). Subsequently, lack of clinical or histological signs of vasculitis associated with lupus in either the skin or kidneys of SCID mice receiving xenografts of lymphocytes from patients with lupus was confirmed by another group (Ashany *et al*. 1992a). Production of autoantibodies by SCID–systemic lupus chimeras requires both T and B cells but is still insufficient to initiate pathological lesions (Geppert and Jasin 1990).

SCID mice have since been used to study aspects of murine lupus. SCID mice were reconstituted with spleen cells from young or old MRL- *lpr/lpr* mice or with *lpr* bone marrow. Seventy per cent of SCID mice receiving mature *lpr* splenocytes produced antibodies to Sm and dsDNA and one-third of these mice had extensive mononuclear and granulocytic infiltrates in lymphoid tissue, liver, and kidneys. However, histology revealed no glomerulonephritis, a marked reduction in the CD4-CD8- T-cells subset compared with the donor spleens, and the absence of massive lymphadenopathy which is associated with infiltration by double-negative T cells in the donor MRL-*lpr/lpr* mice. Nearly all recipients of young donor splenocytes developed acute graft-versus-host disease, dying within 4 weeks of cell transfer, while the majority of the recipients of bone marrow cells developed a wasting disease characterized by lymphoid atrophy and fibrosis in the absence of autoantibody production (Ashany *et al*. 1992b).

The role of self-reactive T cells in renal pathology has been demonstrated by cloning T cells from the interstitium of MRL- *lpr/lpr* mice with lupus nephritis. These T cells are unique and are regulated by the *lpr* gene. They have the phenotype CD4-CD8-abTCR +, proliferate in response to renal tubular cells, and secrete interferon-g which induces major histocompatibility (**MHC**) class II antigens and intercellular adhesion molecule 1 (**ICAM-1**) on renal tubular epithelial cells, making these cells capable of triggering T-cell hybridomas to proliferate and secrete interleukin 2 (**IL-2**) (Kelley *et al*. 1993). These double-negative T cells proliferate mostly in the liver and migrate to the periphery. Hepatic mononuclear cells vigorously secrete IL-6 which can be neutralized by antimouse IL-6 antibody and prevent the induction of ICAM-1 on hepatic sinusoidal endothelial cells seen in MRL- *lpr/lpr* mice (Ohteki *et al*. 1993). Upregulation of cell adhesion molecules (ICAM-1, VCAM-1) has also been demonstrated in kidneys of autoimmune mice and may mediate the adherence of pathogenic inflammatory cells in murine lupus nephritis (Wuthrich *et al*. 1990; Wuthrich 1992).

A form of 'experimental lupus' has been induced in normal strains of mice by two 1 µg injections of a human monoclonal anti-DNA antibody bearing the pathogenic 16/6 idiotype. Subsequently mouse IgG carrying the same idiotype was found in mesangial deposits in the kidneys. The mice developed proteinuria, an increased erythrocyte sedimentation rate, and leucopenia and had circulating autoantibodies reactive with a variety of autoantigens (Mendlovic *et al*. 1988). Independent confirmation of this data is awaited as it has not been confirmed by several other laboratories, implicating an environmental factor in the development of this lupus model (Williams and Isenberg 1994). However, transfer of splenocytes from Balb/c mice, in which systemic lupus had been induced by injection of 16/6Id, into SCID mice resulted in transfer of lupus disease, the characteristics of which were more severe if 16/6Id was given at the same time as cell transfer. The SCID recipients had high titres of autoantibodies, multiple immune complex deposits in the kidneys, and glomerular histological changes, including thickening of the glomerular basement membrane, mesangial proliferation, and some fibrotic changes (Segal *et al*. 1992).

Autoantibodies in systemic lupus

The serological hallmark of systemic lupus is the presence of circulating autoantibodies directed against a wide variety of nuclear, cytoplasmic, and plasma membrane antigens which are outlined in Table 13. In a recent study 98 per cent of patients had antinuclear antibodies and 56 per cent had elevated levels of antibodies to dsDNA (Worrall *et al*. 1990).

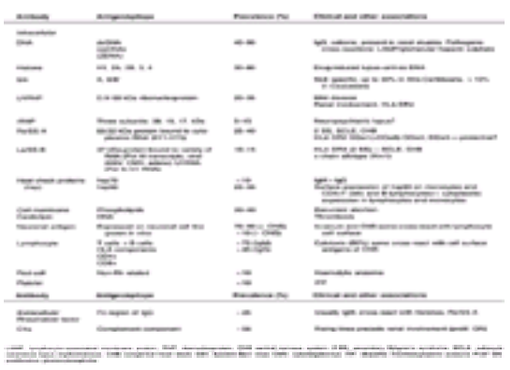


Table 13 Autoantibodies in systemic lupus

Antibodies to dsDNA are usually of the IgG isotype. Studies of kidney biopsies from patients with glomerulonephritis show deposition of IgG and complement components indicating a localized immune response and subsequent inflammation. Eluates from affected kidneys show that the IgG has specificity for dsDNA. In addition these antibodies are cationic in charge, clonally restricted, and high affinity (Kalunian *et al*. 1989). Not all circulating anti-DNA antibodies possess these characteristics, suggesting a subpopulation of 'pathogenic' antibodies preferentially localized in the kidney. Two mechanisms may account for this deposition. First, circulating immune complexes of DNA–anti-DNA antibodies may become trapped in the glomerulus, activate complement, and lead to local inflammation and tissue damage Second, non-complexed anti-DNA antibodies may cross-react with heparan sulphate, the glycosaminoglycan side-chain of heparan sulphate proteoglycan, which is a major constituent of glomerular basement membrane; this anionic side-chain is predominantly responsible for the negative charge of the membrane and may merely reflect antigen non-specific binding of cationic anti-DNA antibodies. Recently this binding was reported to be mediated via bound complexes containing both DNA and histones. Histones have very high affinity for heparan sulphate (Termaat *et al*. 1990) and antihistone antibodies are also a feature of systemic lupus. Evidence of direct pathological consequences of human anti-dsDNA antibodies has been described recently by Ehrenstein *et al*. (1995) using the SCID mouse model.

Heparin has been shown to inhibit the DNA binding of antibodies eluted from kidneys of humans and MRL- *lpr/lpr* mice but did not inhibit DNA binding of erum antibodies. Low, non-anticoagulant doses of heparin administered to MRL- *lpr/lpr* mice prevented the development of severe lupus nephritis and may be useful in treating humans with lupus (Naparstek *et al*. 1990).

Molecular biology techniques have revealed some interesting new data on the nature of the DNA antigen involved in lupus. DNA purified from immune complexes isolated from the sera of patients with lupus was found to be enriched in the deoxynucleotide bases guanosine and cytidine compared with normal genomic DNA. It was highly immunogenic in rabbits, unlike commercial DNA, and bound to anti-DNA antibodies with high avidity. Cloning and sequencing of this DNA revealed 82 per cent homology with the *gag*/*pol* fragment encoded by the ENV gene of the human immunodeficiency virus, HIV. This finding lends support to a viral aetiology of systemic lupus but a survey of sera from 100 patients with lupus revealed no antibodies to retroviral proteins (p24) (Krapf *et al*. 1989).

Recently a human monoclonal IgG antibody (B3) with specificity for dsDNA has been modelled on the computer, enabling a molecular map of the antigen–antibody interaction to be visualized (Fig. 8). This type of approach will facilitate site-directed mutagenesis experiments to determine the key amino acids for the recognition of DNA (Kalsi *et al*. 1996).

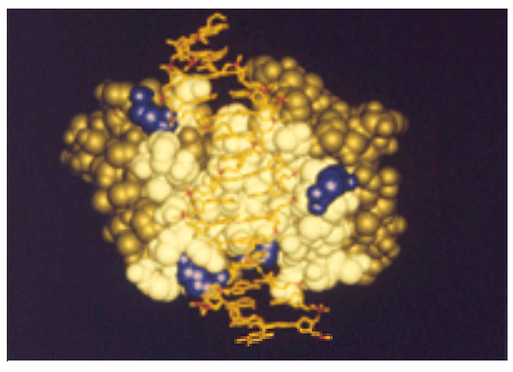


Fig. 8 Computer modelled possible docked complex between an IgG monoclonal anti-dsDNA antibody designated B3 and dsDNA. The light chain is shown on the left-hand side and the heavy chain on the right-hand side. Olive colour=framework residues; white=complementarity determining regions. The arginines (blue) at positions 227A, L54, and H53 all make interactions with the DNA phosphate backbone. (Reproduced from Kalsi *et al*. 1996, with permission from Pergamon Press.)

While antibodies to DNA occur in many patients, antibodies to some of the other nuclear antigens show an association with ethnic origin. Antibodies to Sm are more frequently associated with Afro-Caribbeans than Caucasians. This also relates to the HLA status of the individual (see later) in that many of the Afro-Caribbean patients carry the DR2 haplotype (Olsen *et al*. 1993). Antibodies to U1RNP are usually associated with mild disease, a lower incidence of renal involvement, and the MHC haplotype, DR4.

Antibodies to Ro and La are more frequently associated with Sjögren's syndrome secondary to a diagnosis of lupus. Occasionally these antibodies can be detected in saliva, but not in serum prior to the symptoms associated with salivary gland infiltration (Horsfall *et al*. 1989).

Claims have been made suggesting an association of autoantibodies to ribosomal P protein in patients with lupus with cerebral involvement and psychosis but these findings remain controversial.

Autoantibodies to C1q may be elevated in systemic lupus and are indicative of proliferative glomerulonephritis (Siegert *et al*. 1993).

Abnormalities and dysregulation at the cellular level

Systemic lupus is characterized by multiple functional defects among cells of the immune system—T and B lymphocytes, natural killer cells, and accessory cells (antigen presenting cells) (Table 14). Numbers of circulating lymphocytes may be altered profoundly. Hyperactive B cells may be increased in number with coexistent T lymphocytopenia. Numbers of both lymphocyte populations are extremely variable and fluctuate between normal and abnormal levels with respect to disease activity and duration.

Table 14 Cellular abnormalities and cytokine dysregulation

The increase in activated B cells contributes to the hypergammaglobulinaemia associated with reactivity to self antigens outlined in Table 13. On circulating B cells, receptors for the cytokine, IL-2, are increased while CR1 (the receptor for C3b) expression is decreased. There is increased cytoplasmic expression of the heat shock protein, **hsp**90, in B cells and CD4 + T cells (but not CD8 + T cells) compared with normal cells. Some of the excess hsp90 in patients with elevated levels is localized on the lymphocyte cell surface and therefore accessible to the immune system (Erkeller-Yuksel *et al*. 1992).

Among the two major T-cell populations, CD4 + (helper/inducer) and CD8 + (suppressor/cytotoxic), there is a marked reduction of a subset of T cells bearing the CD4 + and CD45R + phenotype. This population of cells helps to induce suppression by providing a signal to the CD8 + population. The reduction in this subset may explain the failure of T cells to suppress the hyperactive B cells. Anti-T-cell autoantibodies may be responsible for the depletion of this particular subset. A study has reported that the titre of anti-T-cell antibodies is directly proportional to the ratio of CD4/CD8 killing and that flares of disease are associated with an increase in CD4/CD8 killing and disease remission is accompanied by a corresponding decrease in the ratio. Parallel changes in anti-T-cell titres reflect the disease activity (Yamada *et al*. 1993). Many of the T cells with the classical ab T-cell receptor (**TCR**) chains lack the CD4 or CD8 phenotypes normally associated with these cells. As 'double negatives' they have probably escaped thymic deletion. A recent report described that of all activated T cells cloned *in vitro* from patients with lupus only 15 per cent actually provided 'help' for B cells to make pathogenic anti-DNA antibodies, i.e. antibodies of IgG isotype, cationic charge, specific for native DNA, and clonally restricted in spectrotype. The majority (83 per cent) of these T cells were CD4 + and expressed the classical ab TCR, responded to endogenous antigen presented by autologous B cells, and were class II restricted. The remaining 17 per cent were CD4-CD8- (double negative) and were not class II restricted. Of these, 70 per cent expressed the alternative g TCR and proliferated in response to endogenous heat shock or stress proteins of the hsp60 family expressed by lupus B cells (Rajagopalan *et al*. 1990). Sequencing of the T-cell receptors showed Vd and Vg gene usage not found in normal healthy adults and resembled that of fetal thymocytes early in ontogeny (Rajagopalan *et al*. 1992).

Increased DNA mutations have been observed in T cells in patients with lupus, which may result in T-cell death and increased release of non-degraded DNA by necrosis rather than apoptosis, in turn contributing to the production of anti-DNA antibodies (Gemelig-Meyling *et al*. 1992).

Several authors have suggested that a biochemical defect of T cells underlies the impairment of T-cell responses in systemic lupus. There are two possible defects in the T-cell cAMP pathway; one at the level of adenylate cyclase and another at the level of cAMP-dependent protein kinase. Cross-linking of cell surface receptors and lymphocyte movement to sites of inflammation (homing) have effects on the cAMP pathway. These events enhance the intracellular turnover of cAMP, promote occupancy of cAMP receptors, activate cAMP-dependent phosphorylation, and induce directed mobility of surface molecules to a pole of the cell (capping). The cAMP pathway thus mediates the mobility of certain transmembrane and glycolipid-anchored cell surface molecules resulting in ligands bound to T-cell membrane molecules to be selectively internalized or cleared from the cell surface by capping and endocytosis or by shedding.

In comparison with normal T cells, lupus T cells showed markedly abnormal capping of cell surface proteins (CD4 and 8) during active and inactive disease and showed decreased cAMP production in response to adenosine, associated with an inability to switch phenotype and express suppressor activity. Cell permeable cAMP did not bypass potential adenylate cyclase defects nor restore suppressor activity (reviewed by Kammer and Stein 1990).

The appearance of class II molecules, usually HLA DR, on T cells is taken as a marker of activation. Peripheral T cells with increased expression of HLA DP at the cell surface and as mRNA transcripts have been found in patients with lupus. The frequency of HLA DP expression exceeded that of HLA DR and correlated well with disease activity. The ratio of HLA DP + T cells is inversely proportional to the extent of IL-2 production during *in vitro* response to mitogen (Hishikawa *et al*. 1990; Kanai *et al*. 1993).

Some patients with lupus show decreased responses to immunization *in vivo* (primary immune response) and decreased responses to B-cell challenge *in vitro*. Other patients show normal or even increased responses to immunization. Disease activity and immunosuppressive therapy obviously influence the response in individual patients (reviewed by Turner-Stokes and Isenberg 1988).

Antigen-specific T cells have now been cloned from patients with lupus. Patients who have circulating antibodies to the ribosomal P2 protein have T cells which can proliferate *in vitro* to recombinant P2 and are inhibited in the presence of antibodies to MHC class II antigens. These T cells are CD4 + and thus may help B cells to produce antigen-specific antibodies (Crow *et al*. 1994). HLA DR restricted T cells have also been cloned from a patient with lupus and shown to induce IgG anti-DNA antibodies *in vitro* from high density (activated) B cells from DR-matched patients with lupus, and IgM anti-DNA antibodies from B cells from DR-matched normal

individuals (Murakami *et al*. 1992). An HLA DR4-restricted T-cell clone has been derived from a healthy individual and recognized the protein but not the RNA moiety of UsnRNP. The T-cell clone had a cytokine profile characteristic of Th1 cells (see the following section) and did not provide helper activity for autoantibody synthesis *in vitro* (Wolff-Vorbeck *et al*. 1994).

Cytokines

In the first edition of this textbook we wrote that 'the role of cytokines in the immunopathology of lupus is uncertain although likely to be of importance'. This prudent statement is underlined by the explosion of literature in this area and the demonstration of effective anticytokine antibody therapy of autoimmune disease. The characterization of different subsets of T-helper (**Th**) cells is directly related to their cytokine profiles and several autoimmune diseases may now be classified according to the predominance of a particular T-helper cell subset. We present a simplified summary of the major findings in this area and refer the reader to a recent excellent review by Horwitz and Jacob (1994).

Table 15 summarizes the major differences between Th1 and Th2 cells in terms of cytokine profiles and function, and outlines the cytokines found in serum or produced spontaneously after stimulation of peripheral blood mononuclear cells. Th1 cells support cell-mediated immunity while Th2 cells provide B-cell help and also suppress cell-mediated immunity. The balance between cytokines from Th1 and those from Th2 cells is thus critical in determining the outcome of the immune response and any imbalance could have profound pathological effects.

Subsets of CD4+T cells

	T-helper 1 cell	T-helper 2 cell
Function	Cell-mediated immunity	B-cell help
Cytokines:	IFN-γ IL-10 (some) IL-12 TNF-α	IL-4 IL-10

Cytokine profiles in patients with active systemic lupus

Cytokine	Serum level*	Spontaneous	Stimulation *in vitro*
IFN-γ	↑	low	↓
TNF-α	↑ (or normal)	↓ (DR2 DQw1; ↑ reports) ↑ (DR3,4 - negative)	↓
IL-1	n.d.	↑ PGM production	↓ monocyte production
IL-2	↑	low	↓
IL-4	n.d.	low	low
IL-6	↑	↑	↔
IL-10	↑ (or normal)	↑ (or normal)	normal

Table 15 T-cell subsets and cytokines

At the most simplistic level systemic lupus could be expected to be a disease where Th2 cells predominate resulting in too much help for B cells and overproduction of antibodies. In support of this concept, increased levels of IL-10 have been found in patients with lupus; this cytokine suppresses Th1 cells and thus impairs cell-mediated immunity, a characteristic feature of the disease. In mice continuous administration of anti-IL-10 delays the onset of disease.

Tumour necrosis factor (**TNF**-a) is produced by T (Th1) and B lymphocytes, natural killer cells, and mononuclear phagocytic cells. Another cytokine (TNF-b), originally called lymphotoxin, is produced by activated lymphocytes. The genes for both TNF-a and -b are closely linked and located within the major histocompatibility complex (MHC). *In vitro* TNF-a production by mitogen-activated peripheral blood lymphocytes varies according to the HLA class II haplotype of the donor. Elevated production is found in DR3 and DR4 subjects, whereas low production is found in DR2 and DQw1 positive donors. The DR2, DQw1 genotype in patients with lupus is associated with lupus nephritis. DR3 + patients with lupus are not predisposed to lupus nephritis and have elevated levels of TNF-a production. The DR4 haplotype is associated with high TNF-a inducibility and negatively correlated with lupus nephritis (Jacob *et al*. 1990a).

Paradoxically high levels of TNF-a are more frequently associated with rheumatoid arthitis and, as such, therapeutic benefit has been reported using monoclonal antibodies to TNF-a. Two of 20 patients treated with anti-TNF-a antibody developed anti-dsDNA antibodies and other serological abnormalities characteristic of lupus, although neither developed clinical symptoms of the disease. All antibodies disappeared when therapy was stopped (Maini *et al*. 1994).

Both macrophage and natural killer cell-mediated cytotoxicity are frequently impaired in patients with lupus. g-Interferon-induced enhancement of both types of cytotoxicity is also impaired despite normal levels of g-interferon production by lupus Th1 cells. Lupus natural killer cells fail to release soluble factors necessary for killing. Recombinant g-interferon has been used to induce remission in patients with rheumatoid arthitis, but it exacerbated disease in both patients with lupus (Machold and Smolen 1990) and lupus-prone strains of mice (Jacob *et al*. 1990b).

Accessory cells in lupus seem to produce insufficient amounts of IL-1 to provide the necessary activation signal for T cells. This effect cannot be overcome *in vitro* by the addition of exogenous IL-1, suggesting that a defect may exist at the level of the IL-1 receptor on T cells. Alternatively the defect could exist at a distal point in some biochemical pathway.

Both CD4 + and CD8 + T cells have been described to produce either normal or decreased amounts of IL-2 in response to exogenous antigens, mitogens, allo- and autoantigens. Reduced IL-2 generation would have a profound effect on T-cell responses. This impairment is not reversed by the addition of exogenous IL-2. However, *in vitro* treatment of lupus accessory cells with g-interferon plus the addition of IL-2 has been shown to restore T-cell proliferation and the expression of T-suppressor activity. The variability of experimental data in this context raises the question of whether IL-2 functional activity varies between patients rather than merely reflecting levels of IL-2.

Studies *in vitro* on the response to IL-2 of the CD4 + T-cell subset (helper/inducer) indicate that there is altered expression of IL-2 receptors. The CD8 + T-cell subset (suppressor/cytotoxic) has IL-2 receptors but fails to respond as there is no IL-2 signal from CD4 + T cells. This is a secondary defect. The primary disorder relates to the CD4 + subset which expresses IL-2 receptors of low affinity. Functional IL-2 receptors are of high affinity. Variable synthesis of both types of receptors may be governed by altered intracellular synthesis and/or transport.

Apoptosis

The study of programmed cell death in human lupus is now an active area of research. As with experimental models of lupus, an underlying defect of apoptosis prevails in human disease. Decreased apoptosis in systemic lupus has been described in association with an increased production of the soluble form of the *Fas* molecule, which is capable of inhibiting apoptosis after a stimulus to proliferate (Cheng *et al*. 1994).

bcl-2 is a proto-oncogene located at the inner membrane of mitochondria, the endoplasmic reticulum, and the nuclear membrane and exerts a regulatory function during development and maintenance of adult tissue by preventing apoptosis in specific cell types. It is involved in T-cell development and thymic selection and is found in long-lived B lymphocytes within the follicular mantle zone. Over-expression of *bcl-2* in circulating lymphocytes from patients with lupus has been described in both CD4 + and CD8 + T cells but not B cells, and levels of *bcl-2* expression have been shown to correlate with disease activity (Aringer *et al*. 1994). A more recent study of a larger group of patients with lupus has described enhanced *bcl-2* expression in a subpopulation of CD3 + T cells lacking CD4 and CD8 (double-negative T cells). The percentage of these cells expressing *bcl-2* correlated strongly with disease activity (Rose *et al*., unpublished observations). Defects in apoptosis in systemic lupus may therefore involve two genes—either abnormal transcription of the *Fas* gene resulting in a soluble rather than cell membrane-associated product, or increased transcription and expression of the *bcl-2* gene product, both of which are capable of inhibiting apoptosis.

Cell adhesion molecules

Interest in the adhesive mechanisms which facilitate cell–cell and cell–matrix interactions has expanded considerably over the past few years. Adhesion is mediated by several molecules which vary in affinity and structure on the surface of the cells which express them. Adhesion molecules belong to one of three main families: selectins, integrins, or the immunoglobulin superfamily; the important members and their cell distribution, function, and ligands are summarized in Table 16. The mechanism of leucocyte emigration through the vascular endothelium into tissue is simplistically summarized in Fig. 9. As leucocytes roll in the direction of blood flow, random contact is made with the endothelium. Transient adhesion of leucocytes to the endothelium occurs through low affinity binding to E-selectins. The cells

become activated and then bind with higher affinity to vascular integrins. Once cells are firmly attached transendothelial migration occurs and, under the influence of extravascular chemoattractants, subendothelial migration into the extracellular matrix occurs.

Family	Member	Cell	Ligand	Function
Selectins	ELAM-1	Endothelial	Leucocyte Sulfated Carbohydrate	Inflammation, leucocyte extravasation
	CD62	Platelets	?	T-cell recruitment to inflammatory sites
	LECAM-1	T and B cells	?	PMN, lymphocyte binding to endothelium
Integrins	LFA-1 (CD11a/CD18)	Lymphocytes, PMN, monocytes, Macrophages	ICAM-1,2,3	Tc/target cells, MLR, antigen-specific T-cell proliferation, leucocyte adhesion to endothelium
Immunoglobulin supergene	ICAM-1 (CD54)	Monocytes Endothelial cells CD40	LFA-1 (CD11a/18) MAC-1 (CD11b/18)	T-T, T-B, T-APC interactions. Induced by IFN-γ, IL1β, TNF-α
	VCAM-1	Endothelial cells (act), tissue Mo	VLA-4	Leucocyte recruitment

ELAM, endothelial leucocyte adhesion molecule; CD, cluster differentiation; LECAM, leucocyte adhesion molecule; LFA, leucocyte function associated molecule; ICAM, intercellular adhesion molecule; PMN, Lymphocytes, monocytes; Tc, T-cytotoxic cells; MLR, mixed lymphocyte reaction; APC, antigen-presenting cell; VCAM-1, vascular cell adhesion molecule; Mo, macrophages; VLA-4, very late antigen

Table 16 Adhesion molecules

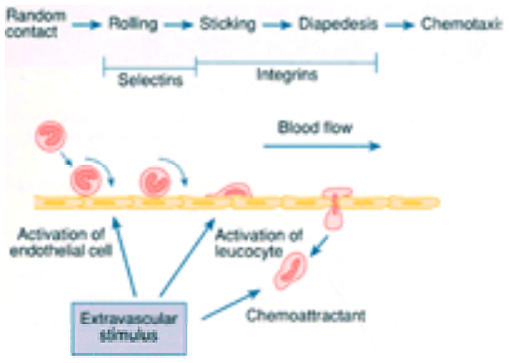


Fig. 9 Adhesive interactions during leucocyte emigration

Up-regulation of the surface expression of three distinct adhesion molecules, E-selectin, VCAM-1, and ICAM-1, has been shown in biopsies of skin, which is non-lesional and has not been exposed to the sun, from 16 patients with lupus. Levels of adhesion molecules were directly correlated with disease activity and in serial biopsy specimens they decreased with clinical improvement (Belmont *et al*. 1994). These findings suggest that excessive complement activation in association with primed endothelial cells induces leucocyte–endothelial cell adhesion and leuco-occlusive vasculopathy. Furthermore, UV irradiation of keratinocytes *in vitro* induces release of epidermal and dermal cytokines and increases ICAM-1 expression, which *in vivo* may lead to vascular activation culminating in the photosensitive lupus syndromes (Norris 1993).

Soluble forms of adhesion molecules are elevated in the circulation of patients with lupus compared with healthy controls. Elevated levels of soluble ICAM-1 (sICAM-1) have been reported to show significant association with skin involvement and disease activity (Sfikakis *et al*. 1994). Lupus patients also have elevated levels of a soluble form of VCAM-1 (Wellicome *et al*. 1993). The importance of these observations may become apparent from studies in murine lupus which have demonstrated up-regulation of ICAM-1 in nephritic MRL-*lpr/lpr* and NZB/W kidneys, particularly in the brush borders of proximal tubules, glomerular mesangium, and endothelium of larger vessels (Wuthrich *et al*. 1990). Similarly VCAM-1 is also up-regulated in MRL-*lpr/lpr* kidneys, not only in the endothelium but also in cortical tubules and glomeruli. Kidney tissue sections from nephritic MRL-*lpr/lpr* mice also display increased adhesiveness for T-cell and macrophage cell lines which can be blocked by monoclonal antibodies to ICAM-1 and VCAM-1 (Wuthrich 1992).

Genetic components

The idea that genetic factors may play a role in susceptibility to lupus stems from the high rate of concordance for disease in monozygotic twins, increased frequency of lupus and immunological abnormalities in relatives of patients with lupus, and the prevalence of lupus among certain ethnic groups. The genes which may influence or predispose to lupus include those which determine sex, colour, complement haplotype, tissue type (HLA), antibody variable regions, and T-cell receptors (Table 17).

Genetic factors	
Complement	C1, C4 deficiencies (~ 75% prevalence, severe disease) C2 deficiency (~ 20% prevalence, severity similar to non-C deficient lupus patients) C4 null-ness, or extreme partially deficient C3(B) deficiencies (acquired rather than genetic)
MHC associations	HLA A1, B8, DR3 – strongest positive DR2 and DR4/5 – some definite contributions DQw1 – most DR contributors Altered C4 – more anti-nuclear DR2 – often Caucasian patients with anti-Sm DR3, DR4/5 (DR4/8, B18) – susceptibility to nephritis DR2, DR3/5 – protective? DR4/8 – early onset, SLE (~ 60 years) Decreased expression of class II on APC Homozygosity may depend upon a single amino acid substitution (tyrosine at position 57 in sequence of DRbeta 1 MHC gene from single SLE patient compared with healthy control)
T-cell receptor	
Human	Little data on human TCR family usage Germline sequence related to TCR-V gene product in SLE Restriction DNA restriction fragment length polymorphism in humans no deletions, insertions in TCRb-J-C chain genes
Mouse	Cry deletions in TCR-alpha/delta and beta-delta-V, assay can reveal TCR (deletion) (in light TCR-Cb-Bb allele) some increases as animals age and as disease and lymphadenopathy develop
Antibody V genes	
Germline encoded	Human monoclonal anti-DNA/Sm germline encoded but use some V gene families as do in non-self. But they can result somatic mutation (germline predisposition to develop anti-Sm etc in SLE patients and not degree mutation measured)
Somatic mutation	V somatic mutation (e.g. like autoantibodies between homologous family) Autoantibody response generally light somatic mutation (increases with immunoglobulin class switch
Idiotype analysis	Preferential use of V gene family with or without autoantibody activity (produced by family members). Specific deposition of anti-DNA pathogenic/nephritogenic idiotypes in kidney
	12/13 14/20 20/21
Percentage SLE	20 75 27
Percentage disease controls	4 4 4
Genetic factors	
Ethnic origin	Blacks, whites, most probably linked to MHC
Sex and hormonal factors	Females > whites, androgens enhances the immune response, testosterone is immunosuppressive

MHC, major histocompatibility complex; APC, antigen-presenting cell; TCR, T-cell receptor

Table 17 Genetic components associated with systemic lupus erythematosus (SLE)

Sex and ethnic background

At the most basic genetic level systemic lupus affects females more than males and shows an ethnic bias in that Afro-Caribbeans are more affected than Orientals, in turn more affected than Caucasians. These basic differences are reflected in other ways. Sex hormones are known to influence disease in both mice and humans with autoimmune disease. Androgens are immunosuppressive and oestrogens are immunoenhancing. This explains the susceptibility of women to autoimmune diseases compared with men (Ansar Ahmed *et al*. 1985).

For example, androgens reduce while oestrogens enhance the spontaneous antibody production in mixed strains of mice (NZB × CBA and NZB × C3H). In MRL-*lpr/lpr* mice, testosterone treatment reduces lupus-like symptoms without affecting lymphoproliferation. In humans where pregnancy occurs during active disease, exacerbations often occur as oestrogen levels rise.

Complement

Lupus is associated with deficiencies of the early classical pathway of complement components C1, C4, and C2. Two alleles are inherited for each complement component thus deficiencies may be partial or complete (homozygous). Congenital deficiencies of C2 and C4 are frequently in linkage disequilibrium with HLA DR3 and DR2, HLA haplotypes associated with lupus. C4 is composed of two distinct but homologous proteins, C4A and C4B. These bind to immune complexes and

prevent precipitation. C4A deficiency is rare in the general population but complete (homozygous) deficiency has been found in 10 to 15 per cent of white patients with lupus. Partial C4A deficiency occurs in 10 to 20 per cent of controls but in 50 to 80 per cent of patients with lupus (Kemp *et al*. 1987). A single null C4A allele increases the relative risk for lupus by 3 and two null alleles by 17. Deficiencies of C4A can arise by two mechanisms; one is a 30 kilobase deletion of the DNA encoding all C4A along with a small portion of C4B and occurs in DR3 + patients (Kemp *et al*. 1987), and the other does not involve a deletion but may be the result of a regulatory gene linked to the MHC causing reduced synthesis of C4A (Fronek *et al*. 1988). The consequence of C4A deficiency in lupus is that immune complex clearance and solubilization is seriously impaired, leading to deposition in lungs and kidneys and subsequent inflammation. C4 is important in drug-induced lupus as the drugs bind to the active site of C4 and prevent covalent binding to immune complexes. C4A is inhibited more than C4B, resulting in a relative deficiency of C4A with consequent defective clearance of immune complexes (Gatenby 1991).

Complement receptors, CR1 and CR2, have been studied in lupus. The CR1 ligands are C3b and C4b bound to immune complexes. The receptor is present on peripheral B lymphocytes, erythrocytes, monocytes, and tissue macrophages and binds, internalizes, processes, and transports immune complexes which have activated complement. Low expression of CR1 on erythrocytes and peripheral blood leucocytes was described in patients with lupus and their healthy family members (Walport and Lachmann 1988). This was originally interpreted as an inherited defect which could predispose to the disease. More recently this defect is thought to be acquired, as normal erythrocytes infused into patients with lupus also showed a decrease in CR1 receptors. Levels of CR1 deficiency in patients with lupus correlate with disease activity. Controversy still abounds in this area as to the precise nature of this defect and is confounded further by the possibility of a functional defect of CR1 receptors on polymorphonuclear neutrophils.

Complement receptor 2, **CR2** (CD21), has recently been described on peripheral T cells in healthy individuals. The ligands for CR2 are C3d, C3g, and Epstein–Barr virus (EBV). The receptor is present on mature, circulating, and lymph node cells, follicular dendritic cells and on 10 to 40 per cent of peripheral blood CD4 + or CD8 + T cells. In patients with active or inactive lupus, B-cell CR2 expression is significantly diminished. This may be a consequence of the activated state of lupus B cells reflecting the loss of CR2 as cells differentiate into immunoglobulin-secreting cells and then plasma cells, or the levels may be modulated by high levels of circulating immune complexes or cytokines. Expression of CR2 on T cells from patients with inactive lupus is similar to that found in healthy individuals, but is increased in some patients with active lupus (90 per cent of CD4 + and CD8 + peripheral T cells expressing increased levels of CR2). In this context CR2 is important in signalling and its increased expression on T cells may play a role in cell adhesion or cytotoxicity (Levy *et al*. 1992).

Major histocompatibility complex

The associations between the MHC and systemic lupus must take into account the ethnic origin of the patient. A recent report highlights this aspect by demonstrating that only in Caucasian patients of English/Irish descent is lupus associated with an MHC extended haplotype (HLA B8, SCO1, DR3) (Schur *et al*. 1990). In black American people DRw52b is positively associated with renal disease and negatively associated with antinuclear RNP antibodies. DR3 (DRw17) and DQw2 are highly associated with the ability to produce anti-Ro/anti-La antibodies. These antibodies are associated with subacute cutaneous lupus, lymphopenia, neonatal lupus and complete congenital heart block. DR4 is associated with the ability to make anti-RNP and with a reduced risk for lupus nephritis. In contrast to the protection conferred by DR4, DR2 confers susceptibility to nephritis. Associations with DQw1 and with DQb1.AZH and DQb2 are more recent findings. Early-onset systemic lupus (before 20 years of age) is associated with DRw8 and the frequency of neuropsychiatric involvement correlates negatively with a DQa fragment (Reveille *et al*. 1989). Compared with the normal ethnic population, Afro-Caribbean patients with antibodies to Sm have an increased frequency of DR2 and a reduced frequency of DR3, regardless of anti-DNA antibody status (Olsen *et al*. 1993).

DQw7 correlates significantly with lupus anticoagulant, although there are patients who have lupus anticoagulant but lack the DQw7 haplotype. However, these patients all have DQw8. Amino acids in position 71 to 77 of the third hypervariable region of DQB1 chains were identical in DQw7, DQw6, and DQw8 individuals, leading to the proposal that this region might constitute the 'epitope' for mediation of this autoimmune response (Arnett *et al*. 1991). For a more complete description of the known associations between MHC class II genes and autoantibody subsets the reader is referred to an excellent review by Arnett and Reveille (1992) .

T-cell receptors (TCR)

Patients with anti-Ro antibodies show association with TCR-b gene products compared with patients with lupus lacking this antibody specificity (Frank *et al*. 1990). The association with this particular gene segment is with the specificity of the autoantibody produced rather than risk factors for the disease. The anti-Ro antibody response is also associated with HLA-DQw1 (Harley *et al*. 1986) which may be important for recognition by the Ro-specific TCR (Scofield *et al*. 1994). In the lupus-prone mouse strains no evidence of unusual or abnormal TCR gene usage has been demonstrated compared with non-autoimmune mice of the same genetic background (Theofilopoulos *et al*. 1989).

The p70 (Ku) autoantigen has been described as a non-histone nuclear protein recognized by antibodies from patients with lupus. It has been shown that the p70 antigen is a DNA-binding protein and specifically binds to the TCR-b-chain gene enhancer thus playing a role in regulation of TCR-b gene expression (Messier *et al*. 1993).

A human non-specific suppressor factor has been isolated and characterized from lupus ascitic fluid and can inhibit proliferation of T and B cells and suppress IgG production *in vitro*. Suppression was inhibited by anti-TCR-a antibodies but not by those directed against the b chain (Xavier *et al*. 1994).

B-cell immunoglobulin receptors and antibody V genes

The most frequently studied antibodies are those directed against DNA, particularly since anti-idiotypic antibodies raised against the variable regions of murine anti-DNA antibodies have been shown to cross-react with V regions on human anti-DNA antibodies. Many idiotypic markers have been defined on antibodies to DNA and the interpretation of the data has led to much confusion. Only the major findings and their significance with respect to lupus will be discussed.

Several groups have described anti-DNA associated idiotypes defined by anti-idiotype antibodies raised either as monoclonal antibodies in mice or as polyclonal antibodies by immunization of rabbits. In most cases, the immunogen, or anti-DNA antibody-bearing idiotypes, is a human monoclonal antibody made by fusing human B lymphocytes with an immortalized human myeloma/lymphoblastoid cell line to form stable hybridomas which secrete the human antibody of interest. The 16/6 idiotype is located on the heavy chain of a human IgM_ monoclonal anti-DNA antibody derived from a human lupus spleen hybridoma. The anti-idiotype to the monoclonal antibody bound not only to the parent molecule but to anti-DNA antibodies from different patients with lupus and to IgG lacking anti-DNA activity from healthy relatives, spouses, and family members of patients. Serum levels of the 16/6 idiotype correlated with disease activity and deposits of 16/6 positive immunoglobulin were found in skin and kidney biopsies from patients with lupus (reviewed by Watts and Isenberg 1990).

Among the other DNA antibody idiotypes which have been described are two idiotypic markers which distinguish between nephritogenic antibodies and non-nephritogenic antibodies (Hahn *et al*. 1990). Neither of these idiotypes correlated with HLA class II haplotypes known to be associated with lupus nephritis. High serum levels of idiotypes associated with lupus nephritis are thought to arise from polyclonal B cell activation rather than from idiotype-specific up-regulation associated with one or more of the class II genes that predispose to nephritis. On the contrary, two lupus-associated idiotypes on a somatically mutated anti-DNA antibody have been described, providing evidence that such an antibody reflects the selection pressure of antigen (Davidson *et al*. 1990). However, new data on hybridomas from non-autoimmune strains of mice (Balb/c) shows that pathogenic IgG antibodies with specificity for DNA can be produced by these mice. These antibodies were derived from germ-line genes and were not the products of somatic mutation normally found on antibodies of the same specificity from autoimmune strains of mice. This suggests that such germ-line encoded antibodies are under extremely strict regulation in normal animals. The defect in autoimmunity appears to lie in the products of somatic mutation, possibly at the level of anti-idiotypic regulation (Shefner and Diamond 1990).

The immunoglobulin receptors on B cells are also able to 'present' antigen to T cells. This occurs by binding to antigen-specific immunoglobulin receptors which are then able to internalize antigen and express processed antigen on the cell surface in the context of MHC class II. This process has been demonstrated in mice using snRNP autoantigens which are presented to autoreactive T cells (Mamula *et al*. 1994).

Summary of immunopathological events in systemic lupus

Systemic lupus erythematosus is a multifactorial disease in which the relative contribution of each factor increases the relative risk of disease susceptibility. A significant contribution arises from genetic components, such as those encoding tissue type (MHC), complement components, cell receptors, cytokines, and their respective receptors, together with environmental elements, such as drugs, toxins, diet, and infectious agents. The acquisition of each gene will increase the relative risk of disease development. Genetic predisposition may be so strong that a relatively minor environmental insult may be sufficient to trigger disease, whereas a

modest genetic predisposition together with a strong environmental stimulus would be sufficient to lead to manifestations of disease. The latter would encompass idiopathic lupus and drug-induced lupus.

Systemic lupus is characterized by abnormal immune function. Increased numbers of hyperactive B cells, together with impaired cell-mediated immunity, suggest the dominance of Th2 cells and their cytokines in mediating hypergammaglobulinaemia and the appearance of IgG antibodies with specificity for both self and non-self antigens. There is evidence that B cells expand both in response to specific antigen and in a polyclonal fashion.

Molecular mimicry may also contribute to the autoantibody response seen in lupus. Do autoantibodies arise by molecular mimicry (e.g. with viral or bacterial proteins such as anti-U1RNP antibodies and influenza B virus (Guldner *et al*. 1990)) or do they become anti-self and/or pathogenic through somatic mutation of evolutionary 'useful' antibodies and thus bind to host structures/determinants (Diamond and Scharff 1984)?

The transfer of peripheral blood lymphocytes from patients with lupus into SCID mice, described above, which showed deposition of mouse complement and human IgG in the kidneys, suggests that all the information required to cause the histological change associated with lupus resides in the peripheral blood lymphocytes (Duchosal *et al*. 1990). The exact nature of this information remains to be elucidated.

The treatment of lupus

It is evident that the diverse effects of lupus require a variety of treatments. These will be divided into pharmacological and other approaches. However, it must be stressed that a number of general measures may be most useful (see Table 18).

1. Rest as appropriate; try to avoid stress.
2. Avoid over-exposure to heat and sunlight. Use sun protection factor 15+ (30+ in USA) if in a sunny country; avoid exposing an arm on an open car window.
3. Try to adhere to a low fat diet and consider adding fish oil derivatives.
4. Vaccination, for foreign travel etc., apart from 'live' vaccines in patients on immunosuppressives, is not contraindicated though the precise nature of the immune response differs from that in healthy individuals.
5. Medium or high oestrogen contraceptive pills should be avoided—progesterone only or the lowest possible oestrogen pill (or other methods of contraception) are advised.
6. The use of hormone replacement in the menopause remains controversial. Many patients do tolerate it without flaring, but not all.

Table 18 Treatment of lupus—general measures

Pharmacological

Patients with lupus are treated with four main groups of drugs, often in combination. Recommendations about precisely when to commence therapy, the initial dose of a given drug, the likely response of a given symptom, and the duration of treatment vary widely. In Table 19 the broad indications for use of these four types of drugs are shown. In Table 20 suggestions are provided as to the initial doses and duration of treatment with the antimalarials, corticosteroids, and cytotoxic drugs. These are intended purely as guidelines and there will undoubtedly be patients who require larger doses for longer periods of time.

	NSAID	Antimalarial	Corticosteroids	Cytoxic agents
Malaise	–	+	+	–
Fever	+	–	+	–
Serositis	+	–	+	+
Arthralgia	+	+	+	+
Arthritis	+	+	+	+
Myalgia	+	+	+	–
Myositis	–	–	+	+
Malar/discoid rash	–	+	+	–
Pneumonitis	–	–	+	+
Carditis	–	–	+	+
Vasculitis	–	–	+	+
CNS disease	–	–	†*	†
Renal	–	–	+	+
Haemolytic anaemia	–	–	+	+
Thrombocytopenia	–	–	+	+
Raynaud's	–	–	†	†
Alopecia	–	–	†	†

+ = usually beneficial; – = not beneficial; †=effective/indicated in
*=safety precautions must always be taken when steroids are prescribed in early pregnancy

Table 19 Drug therapy in systemic lupus

Symptom	Drugs to try	Dose and duration
Arthralgia	Non-steroidal anti-inflammatory drugs	No special recommendations
Myalgia Lethargy Arthralgia	Hydroxychloroquine	Start 400 mg/day for 3–4 months then reduce to 200 mg/day for 3–4 months, then to 200 mg five times week for 3 to 4 months; repeat courses may be necessary and retinal checks every 6–9 months are generally recommended
Arthritis Pleuritis Pericarditis	Prednisolone	20–40 mg per day initially for 3–4 weeks, reducing in 5–10 mg decrements per week. If patient is responding, treatment is likely to be required for several months
Autoimmune haemolytic anaemia/thrombocytopenia	Prednisolone often accompanied by azathioprine	60–80 mg prednisolone for 1–2 weeks reducing in 10 mg increments in response to the blood test results; aim for 2.5–5 mg/kg azathioprine. Treatment will last for several months
Renal	Prednisolone plus azathioprine or cyclophosphamide	Depending upon the severity of the renal lesion, anything from 30–60 mg/day is required; cyclophosphamide can be given by intravenous boluses (750 mg–1 g) monthly for 6 months, then every 3 months for 2 years, some groups prefer prednisolone and azathioprine at 2–3 mg/kg in the first instance. Treatment is likely to be required for several years
Central nervous system	Prednisolone plus an appropriate drug, e.g. an anticonvulsant, antidepressant, etc.	Controversial — but 20–100 mg prednisolone have been prescribed sometimes accompanied by azathioprine or intravenous cyclophosphamide (rubenic). Treatment is likely to be required for months

Table 20 Recommendations for drug usage in lupus

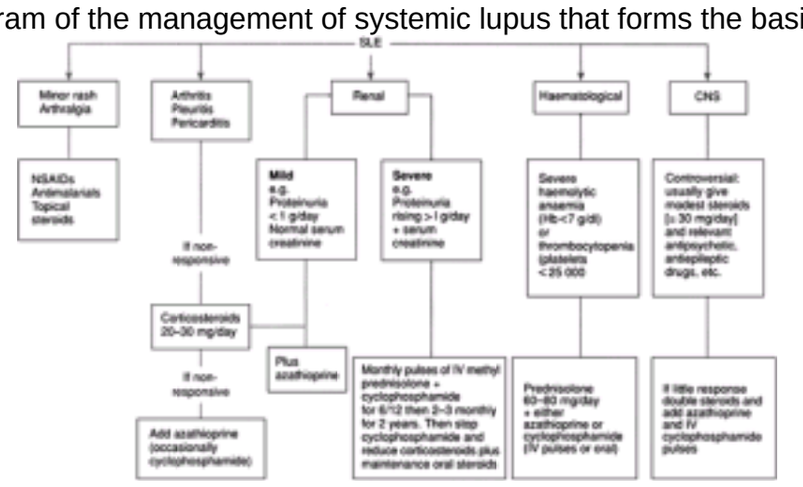
In general the patient with mildly active lupus can be managed with combinations of non-steroidal anti-inflammatory drugs and antimalarials. Patients with lupus are at no lesser risk of gastrointestinal and renal complications of the non-steroidal anti-inflammatory drugs than other patients and thus the usual type of monitoring (clinical history, blood tests) is required. Hydroxychloroquine (Plaquenil) is the antimalarial drug of choice. It is still recommended by some authorities that ophthalmological examinations are undertaken approximately every 9 months to ensure that no retinal damage is occurring.

Corticosteroids in the main are required when non-steroidal anti-inflammatory drugs and antimalarials are insufficient to relieve the patients symptoms. Thus severe arthritis, pleuritis, pericarditis, autoimmune haemolytic anaemia, thrombocytopenia, nephritis, and a wide range of neuropsychiatric problems frequently require treatment with corticosteroids.

Corticosteroids are usually prescribed and taken by mouth but they may also be given intramuscularly and intravenously. Intravenous or pulse therapy has been widely used in the past 15 years. If given over a 15 to 20-min period there is a danger of reactive arthropathy, and in our experience intravenous pulses are best given slowly over a 3 to 4-h period. We use pulse therapy: for example, 1 g on 3 successive days for patients with severe disease which does not seem to be responding to oral corticosteroids. Although some claims have been made about the advantages of this type of approach, the evidence that anything more than a temporary benefit

is obtained is controversial.

Box 1 Flow diagram of the management of systemic lupus that forms the basis of our practice



Some centres have attempted to use alternate-day oral steroid regimes, although the evidence that there is a reduction in the number or the severity of side-effects compared with daily use is lacking. The major side-effects, however corticosteroids are prescribed, are increased risk of infection, osteoporosis, diabetes, hypertension, cushingoid facies, abdominal striae, and insomnia. Corticosteroids are thus no panacea.

Various control trials of cytotoxic drugs in lupus have been reported. The group from the National Institutes of Health at Bethesda has argued strongly that intravenous boluses of cyclophosphamide, monthly for 6 months, subsequently once every 3 months for 2 years, are the treatment of choice in patients with severe renal involvement (Boumpas *et al*. 1991; Boumpas *et al*. 1992; Boumpas *et al*. 1993). The problems of side-effects with this drug (profound nausea, alopecia, infertility especially in patients over 30, and bone marrow suppression) have made others more wary about its routine use. In common with many European groups we prefer to use steroids and maintenance azathioprine in the first instance for patients with mild/moderately active renal disease. Occasionally however, we have seen patients being treated with steroids and azathioprine for other manifestations who develop severe renal disease and require cyclophosphamide urgently. For patients with endstage renal disease, kidney dialysis and kidney transplantation are available. Interestingly, it is rare for a patient with lupus with a transplanted kidney to develop lupus nephritis in the new organ. A flow diagram of the way we use drugs to treat the various aspects of lupus is shown in Box 1.

Other treatments

In the late 1970s and early 1980s there was a great vogue for using plasma exchange. The concept was that the removal of circulating, presumptively pathogenic, immune complexes offered a therapeutic advantage. In practice, it became evident that in some patients a 'rebound' phenomenon occurred in which patients' symptoms and signs dramatically improved but returned within a few days or weeks. This form of treatment requires good venous access, much patience on the part of both physician and patient, and is extremely expensive. A double-blind study using a sham exchange procedure has been performed in mild lupus. The frequency and degree of clinical improvement was the same in both groups (Wei *et al*. 1983). In common with most units we now reserve plasma exchange for those patients seemingly resistant to conventional drug therapy. Attempts are currently being made to achieve a more synchronized deletion of immunoglobulins followed by cytotoxic therapy. Autologous stem cell transplants are also now being considered for lupus.

Lymphoid radiation

Fractionated total lymphoid irradiation, a radiotherapy technique adapted from the method used to treat Hodgkin's disease was shown to improve the survival of NZB/W mice. There have been conflicting opinions as to its value in human patients with lupus. This treatment had not found widespread acceptance. Another form of radiation using ultraviolet A1, has been shown to reduce disease activity in patients in a provisional study (McGrath 1994).

Diet therapy

Many patients with lupus are anxious to know if some form of dietary modification might be of help. It is now well established that diet content can affect the course of disease in NZB/W and MRL-*lpr/lpr* mice. Although some benefits have been demonstrated from total calorie restriction, a restricted amino acid diet, and one in which dietary zinc is reduced, the influence of dietary fat on autoimmunity, certainly in the mouse, appears to be particularly important. In addition, supplementation of the diet by fish oils has been shown to be beneficial. In a recent double-blind, cross-over study in which patients with lupus were put on to low fat diets, those who were concurrently taking 10 g of fish oil per day were shown to have done significantly better over a 3-month period (Walton *et al*. 1991).

Intravenous high-dose gammaglobulins

Intravenous high-dose gammaglobulins may be effective in the treatment of immune thrombocytopenic purpura, immune neutropenia, and myasthenia gravis. It has also been used with moderate success in patients with lupus with low platelet counts. A claim that it was of value in patients with severe renal lupus has not been substantiated. When this approach does not work for patients with thrombocytopenia, splenectomy is beneficial in four or five out of six cases, provided the problem has not been left to become long-term and chronic (Silvestris *et al*. 1994).

Sex hormone therapy

Given the marked predilection of lupus for females, it is not surprising that attempts have been made to treat the condition by manipulating the level of sex hormones.

However, the clinical use of sex hormones in lupus and other autoimmune diseases has neither been extensive nor particularly successful. One drug, danazol, an androgen with reduced virilizing capacity has been used by several groups; as is so often the case with new drugs, the initial optimism has given way to the view that it adds little to the treatment of lupus.

Treatment of experimental lupus

The occurrence of spontaneous lupus in mice has the advantage that manipulation of the disease *in vivo* may be studied. Early studies in the MRL-*lpr/lpr* strain showed that neonatal removal of the thymus delayed the onset of disease to resemble that of the congenic MRL $^{+/+}$ strain. This implicated T cells in the pathogenesis of the disease. Attempts to down-regulate autoantibody production by treatment with anti-idiotype were initially successful in suppressing idiotype-positive anti-DNA antibodies in MRL-*lpr/lpr* mice. Subsequently new clones of idiotype-negative anti-DNA-positive antibodies emerged (Hahn and Ebling 1984). Treatment of both NZB/W F1 and MRL-*lpr/lpr* mice with a monoclonal antibody directed against the CD4 receptor on T helper/inducer cells also improved the clinical status of these animals (Wofsy and Seaman 1985; Santoro *et al*. 1988).

Infection of lupus-prone NZB/W F_1 mice with the parasite *Plasmodium chabaudi* retards the development of their autoimmune disease. Survival was prolonged and high-grade proteinuria and IgG anti-DNA antibodies were delayed for 6 months when parasite inoculation was given either before (3 months) or after (7 months) the onset of the clinical symptoms. Similar beneficial effects, although less pronounced, were obtained when mice were treated with IgG or IgM or cryoglobulin preparations isolated from *P. chabaudi*-infected BALB/c mice, while similarly prepared fractions from uninfected mice had little effect. In surviving mice, levels of anti-DNA antibodies, particularly the IgG1 isotype, were significantly decreased. Flow cytometric analysis of various T-cell subsets showed that the number of T cells expressing Vb 8.1,2, Vb 10 and Vb 14 TCR antigens, which increased with age, were significantly reduced. The mode of therapeutic action is thought to arise from the malarial induction of high levels of natural antibodies bearing the D23 idiotype characteristic of polyreactive natural autoantibodies with enhanced activity against Fab and Fc fragments of IgG. These antibodies have immunoregulatory properties and attempt, at least transitorily, to rescue a natural autoantibody network that is deficient in B/W mice (Hentati *et al*. 1994). These experiments may parallel the findings of human disease treated with intravenous gammaglobulin.

As with human disease, experimental models of lupus have become the focus of much research into cytokine-directed therapeutic intervention. In NZB/W F_1 mice, administration of interferon-g aggravated the autoimmune response, whereas monoclonal antibodies to interferon-g (Jacob *et al*. 1990b) and replacement therapy with recombinant tumour necrosis factor-a (TNF-a) delayed disease development (Jacob and McDevitt 1988). Significantly serum TNF-a levels increased in anti-IL-10

treated NZB/W mice while disease onset was delayed. Simultaneous treatment with a neutralizing antibody to TNF-a at the same time as anti-IL-10 treatment resulted in the rapid onset of lupus and increased mortality supporting the concept that TNF-a is protective in this model. Similarly the role of IL-10 was confirmed to accelerate disease by direct administration of recombinant IL-10 (Ishida *et al*. 1994).

In the same strain of mice, chronic treatment with rat monoclonal antibodies to IL-6 prevented anti-dsDNA antibody production, decreased proteinuria, and prolonged life, provided that mice attained tolerance to rat immunoglobulin by a single injection of anti-CD4 antibodies at the start of therapy (Finck *et al*. 1994). Treatment of MRL-*lpr/lpr* mice for 5 weeks at 15 weeks of age with neutralizing antibodies to the IL-6 receptor showed improvements in glomerular structure and function and a decrease in anti-dsDNA antibodies after 2 weeks of treatment but a gradual increase by week 4 (Kiberd 1993).

A novel 'T-cell vaccination' in MRL-*lpr/lpr* mice has shown a highly significant amelioration of disease parameters. Intravenous administration of irradiated *lpr* cells recovered from hyperplastic lymph nodes of adult diseased animals to young MRL-*lpr/lpr* mice resulted in selective depletion of Vb 8.2 T cells in lymph nodes, in addition to eliciting a surge in peripheral T cells capable of conferring disease protection in adoptive transfer experiments (De Alboran *et al*. 1992).

The major reversal of disease in MRL-*lpr/lpr* mice has been described in the section on experimental models of lupus and refers to the introduction of the *Fas* gene under the control of the CD2 promoter and enhancer (Wu *et al*. 1994).

Treatment of autoimmune mice with drugs used for humans allows not only efficacy but also the mode of action to be tested. For example, dexamethasone given to MRL-*lpr/lpr* mice from 4 weeks of age, prevents lymphadenopathy and renal injury (proteinuria), suppresses a fourfold increase in MHC class II antigen expression (22 weeks), but has no effect on the costimulatory molecules ICAM-1 and TNF-a (Jevnikar *et al*. 1992). Similarly, methylprednisolone down-regulates renal expression of endothelin-1 and its receptors, TGF-b and TNF-a mRNA in NZB/W mice and suppresses development of renal histological lesions (Nakamura *et al*. 1993).

A low fat diet has been reported to delay disease symptoms in MRL-*lpr/lpr* mice (Morrow *et al*. 1986). Since then the nature of marine long-chain fatty acids in the diet has been investigated in female NZB/W mice fed diets containing 10 per cent fish oil, with control mice fed diets containing 10 per cent corn oil. Compared with control mice those maintained on a diet rich in marine oils had an extended life span, later onset of proteinuria, increased proliferative responses to T-cell mitogens, and decreased circulating anti-dsDNA antibodies. Splenocyte analysis compared with controls showed decreased Ig +, higher lymphocyte ECAM-1 expression, elevated mRNA for IL-2, IL-4, and TGF-b, and higher TGF-b, lower c-Myc and c-Ha-Ras proteins. Changes in membrane fatty acid composition may contribute to the altered immune function and gene expression during the development of murine lupus (Fernandes *et al*. 1994).

Prognosis and survival

Studies on duration of disease and overall survival rates have frequently been confounded by the numbers of patients lost to follow-up and inadequate attention paid to the ethnic group, age of onset, and socio-economic status of individual patients. With these possible confounding factors in mind, and the division of patients with lupus into those with overt nephritis and those without, it is reasonable to state that the 5-year survival in lupus is presently 90 per cent or greater, but at 15 years only 60 per cent of those with nephritis will still be alive compared with around 85 per cent of those without nephritis. In the United States, it has been claimed that black patients with lupus, males, those from poorer socio-economic groups, and possibly children, have poorer survival, especially if nephritis is present. It has also been suggested that there exists a bimodal mortality curve. Patients who die within 5 years usually have very active disease, with a requirement for substantial doses of steroids and other immunosuppressives. Those patients dying much later tend to do so from cardiovascular disease and possibly infection. Overall most patients with lupus die from active generalized disease, nephritis, sepsis, and cardiovascular disease. Evidence that patients are more predisposed to malignancy has been discussed earlier in the chapter but this seems to be a relatively minor cause of death in lupus.

Summary

In the introduction to this chapter an analogy was drawn between systemic lupus erythematosus and the Hydra monster of Greek legend. In Fig. 10 this analogy is reiterated to confirm the multiple factors involved in the aetiology and pathogenesis of lupus. It is evident that this remarkable disease presents a wide spectrum of clinical features and is characterized by multiple autoantibodies, although clearly not due to random polyclonal B-cell activation. The treatment and general management of lupus continues to present a challenge. While lupus may for some patients represent a relatively mild set of problems, many others require large doses of immunosuppressive drugs which carry long-term concerns about side-effects.

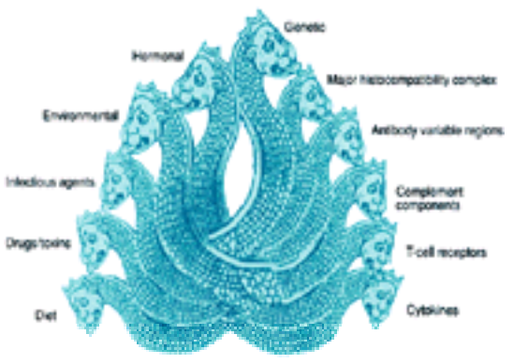


Fig. 10 The Hydra—an analogy of the multiple factors involved in lupus.

Chapter References

Abramson, S., Kramer, S.B., Radin, A., and Holzman, R. (1985). Salmonella bacteremia in systemic lupus erythematosus. Eight year experience at a municipal hospital. *Arthritis and Rheumatism*, **28**, 75–9.

Ansar Ahmed, S., Penhale, W.J., and Talal, N. (1985). Sex hormones, immune responses and autoimmune diseases: mechanisms of sex hormone action. *American Journal of Pathology*, **121**, 531–51.

Aringer, M. *et al*. (1994). High levels of bcl-2 protein in circulating T lymphocytes, but not B lymphocytes, of patients with systemic lupus erythematosus. *Arthritis and Rheumatism*, **37**, 1423–30.

Arnett, F.C. and Reveille, J.D. (1992). Genetics of systemic lupus erythematosus. [Review.] *Rheumatic Disease Clinics of North America*, **18**, 865–92.

Arnett, F.C., Olsen, M., Anderson, K., and Reveille, J.D. (1991). Molecular analysis of major histocompatibility complex alleles associated with the lupus anticoagulant. *Journal of Clinical Investigation*, **87**, 1490–5.

Ashany, D., Hines, J., Gharavi, A., Mouradian, J., and Elkon, K.B. (1992 *a*). Analysis of autoantibody production in SCID-systemic lupus erythematosus (SLE) chimeras. *Clinical and Experimental Immunology*, **88**, 84–90.

Ashany, D., Hines, J.J., Gharavi, A.E., Mouradian, J., Drappa, J., and Elkon, K.B. (1992 *b*). MRL/lpr—severe combined immunodeficiency mouse allografts produce autoantibodies, acute graft-versus-host disease or a wasting syndrome depending on the source of cells. *Clinical and Experimental Immunology*, **90**, 466–75.

Belmont, H.M., Buyon, J., Giorno, R., and Abramson, S. (1994). Up-regulation of endothelial cell adhesion molecules characterizes disease activity in systemic lupus erythematosus. The Shwartzman phenomenon revisited. *Arthritis and Rheumatism*, **37**, 376–83.

Bhalla, R. *et al*. (1993). Systemic lupus erythematosus and Hodgkin's lymphoma. *Journal of Rheumatology*, **20**, 1316–20.

Blanco, F.J., De La Mata, J., Lopez-Fernandez, J.I., and Gomez-Reino, J.J. (1994). Light, immunofluorescence and electron microscopy renal biopsy findings as predictors of mortality in 85 Spanish patients with systemic lupus erythematosus. *British Journal of Rheumatology*, **33**, 260–6.

Bonfa, E. *et al*. (1987). Association between lupus psychosis and anti-ribosomal P protein antibodies. *New England Journal of Medicine*, **317**, 265–71.

Boumpas, D.T. *et al*. (1991). Pulse cyclophosphamide for severe neuropsychiatric lupus. *Quarterly Journal of Medicine*, **81**, 975–84.

Boumpas, D.T. *et al*. (1992). Controlled trial of methyl prednisolone versus two regimens of pulse cyclophosphamide in severe lupus nephritis. *Lancet*, **340**, 741–5.

Boumpas, D.T. *et al*. (1993). Risk for sustained amenorrhea in patients with systemic lupus erythematosus receiving intermittent pulse cyclophosphamide therapy. *Annals of Internal Medicine*, **119**, 366–9.

Callen, J.P. and Klein, J. (1988). Subacute cutaneous lupus erythematosus. *Arthritis and Rheumatism*, **31**, 1007–13.

Cervera, R. *et al*. (1992). Systemic lupus erythematosus: clinical and immunologic patterns of disease expression in a cohort of 1000 patients. *Medicine (Baltimore)*, **72**, 113–24.

Cheng, J. *et al*. (1994). Identification of a soluble form of the fas molecule that protects cells from fas-mediated apoptosis. *Science*, **263**, 1759–61.

Crow, M.K. *et al*. (1994). Autoantigen-specific T cell proliferation induced by the ribosomal P2 protein in patients with systemic lupus erythematosus. *Journal of Clinical Investigation*, **94**, 345–52.

Davidson, A., Manheimer-Lory, A., Aranow, C., Peterson, R., Hannigan, N., and Diamond, B. (1990). Molecular characterisation of a somatically mutated anti-DNA antibody bearing two systemic lupus erythematosus-related idiotypes. *Journal of Clinical Investigation*, **85**, 1401–9.

Deapen, D. *et al*. (1992). A revised estimate of twin concordance in systemic lupus erythematosus. *Arthritis and Rheumatism*, **35**, 311–8.

De Alboran, I.M. *et al*. (1992). lpr T cells vaccinate against lupus in MRL/lpr mice. *European Journal of Immunology*, **22**, 1089–93.

Diamond, B. and Scharff, M. (1984). Somatic mutation of the T15 heavy chain gives rise to an antibody with autoantibody specificity. *Proceedings of the National Academy of Sciences* (*USA*), **81**, 841–4.

Dubois, W.L. and Tuffenelli, D.L. (1964). Clinical manifestations of systemic lupus erythematosus. *Journal of the American Medical Association*, **190**, 104–11.

Duchosal, M.A., McConahey, P.J., Robinson, C.A., and Dixon, F.J. (1990). Transfer of human systemic lupus erythematosus in severe combined immunodeficient (SCID) mice. *Journal of Experimental Medicine*, **172**, 985–8.

Ehrenstein, M.R., Katz, D.R., Griffiths, M.H., Papadeki, L., Winkler, T.H., Kalden, J.R., and Isenberg, D.A. (1995). Human IgG anti-DNA antibodies deposit in kidneys and induce proteinurea in SCID mice. *Kidney International*, **48**, 705–11.

Erkeller-Yuksel, F., Isenberg, D.A., Dhillon, V.B., Latchman, D.S., and Lydyard, P.M. (1992). Surface expression of heat shock protein 90 by blood mononuclear cells from patients with systemic lupus erythematosus. *Journal of Autoimmunity*, **5**, 803–14.

Estes, D. and Christian, C. (1971). The natural history of systemic lupus erythematosus. *Medicine*, **50**, 85–95.

Feinglass, E.J., Arnett, F.C., Dorsch, D.A., Zizic, T.M., and Stevens, M.B. (1976). Neuropsychiatric manifestations of systemic lupus erythematosus: diagnosis, clinical spectrum and relationship to other features of the disease. *Medicine*, **55**, 323–37.

Fernandes, G., Bysani, C., Venkatraman, J.T., Tomar, V., and Zhao, W. (1994). Increased TGF-beta and decreased oncogene expression by omega-3 fatty acids in the spleen delays onset of autoimmune disease in B/W mice. *Journal of Immunology*, **152**, 5979–87.

Fessel, W.J. (1988). Epidemiology of systemic lupus erythematosus. *Rheumatic Disease Clinics of North America*, **14**, 15–23.

Finck, B.K., Chan, B., and Wofsy, D. (1994). Interleukin 6 promotes murine lupus in NZB/NZW F1 mice. *Journal of Clinical Investigation*, **94**, 585–91.

Frank, M.B., McArthur, R., Harley, J.B., and Fujisaku, A. (1990). Anti-Ro (SS-A) autoantibodies are associated with T cell receptor b genes in systemic lupus erythematosus patients. *Journal of Clinical Investigation*, **85**, 33–9.

Fries, J., Porta, J., and Liang, M.H. (1978). Marginal benefit of renal biopsy in systemic lupus erythematosus. *Archives of Internal Medicine*, **138**, 1386–9.

Fronek, Z. *et al*. (1988). Major histocompatibility complex associations with systemic lupus erythematosus. *American Journal of Medicine*, **85**, 42–4.

Furie, R.A. and Chartash, E.K. (1988). Tendon rupture in systemic lupus erythematosus. *Seminars in Arthritis and Rheumatism*, **18**, 127–33.

Galli, M. *et al*. (1990). Anticardiolipin antibodies (ACA) directed not to cardiolipin but to a plasma protein cofactor. *Lancet*, **335**, 1544–7.

Gatenby, P.A. (1991). The role of complement in the aetiopathogenesis of systemic lupus erythematosus. [Review.] *Autoimmunity*, **11**, 61–6.

Gemelig-Meyling, F., Dawisha, S., and Steinberg, A. (1992). Assessment of in vivo frequency of mutated T cells in patients with systemic lupus erythematosus. *Journal of Experimental Medicine*, **175**, 297–300.

Geppert, T.D. and Jasin, H.E. (1990). Immunoglobulin (Ig) and autoantibody production in mice with severe combined immunodeficiency (SCID) injected with lymphocytes from patients with systemic lupus erythematosus (SLE). *Arthritis and Rheumatism*, **33** (Suppl.), 599.

Gladman, D.D. *et al*. (1994). Sensitivity to change of 3 systemic lupus erythematosus disease activity indices: international validation. *Journal of Rheumatology*, **21**, 1468–71.

Gladman, D.D. *et al*. (1996). The development and initial validation of the Systemic Lupus International Collaborating Clinics/American College of Rheumatology damage index for systemic lupus erythematosus. *Arthritis and Rheumatism*, **39**, 363–9.

Goulet, J.R., Mackenzie, T., Lewinton, C., Hayslett, J.P., Campi, A., and Esdaile, J.M. (1993). The long term prognosis of lupus nephritis: the impact of disease activity. *Journal of Rheumatology*, **20**, 59–65.

Grigor, R., Edmonds, J., Lewkonia, R., Bresnihan, B., and Hughes, G.R.V. (1978). Systemic lupus erythematosus. A prospective analysis. *Annals of the Rheumatic Diseases*, **37**, 121–8.

Guldner, H.H., Netter, H.J., Szostecki, C., Jaeger, E., and Will, H. (1990). Human anti-p68 autoantibodies recognise a common epitope of nuclear ribonucleoprotein and influenza B virus. *Journal of Experimental Medicine*, **171**, 819–29.

Hahn, B.H. and Ebling, F.M. (1984). Suppression of murine lupus nephritis by administration of anti-idiotypic antibody to anti-DNA. *Journal of Immunology*, **132**, 187–90.

Hahn, B.H. *et al*. (1990). Idiotypic characteristics of immunoglobulins associated with human systemic lupus erythematosus. *Arthritis and Rheumatism*, **33**, 978–84.

Hanley, J.G. *et al*. (1993). Cognitive impairment and autoantibodies in systemic lupus erythematosus. *British Journal of Rheumatology*, **32**, 291–6.

Harley, J.B., Reichlin, M., Arnett, F.C., Alexander, E.L., Bias, W.B., and Provost, T.T. (1986). Gene interaction at HLA-DQ enhances autoantibody production in primary Sjögren's syndrome. *Science*, **232**, 1145–7.

Haupt, P.M., Moore, W.G., and Hutchins, G.M. (1981). The lung in systemic lupus erythematosus. Analysis of the pathologic changes in 120 patients. *American Journal of Medicine*, **71**, 791–7.

Hay, E.M. and Isenberg, D.A. (1993). Autoantibodies in central nervous system lupus. *British Journal of Rheumatology*, **32**, 329–32.

Hay, E.M. *et al*. (1993). The BILAG index: a reliable and valid instrument for measuring clinical disease activity in systemic lupus erythematosus. *Quarterly Journal of Medicine*, **86**, 447–58.

Hentati, B., Sato, M.N., Payelle-Brogard, B., Avrameas, S., and Ternynck, T. (1994). Beneficial effect of polyclonal immunoglobulins from malaria-infected BALB/c mice on the lupus-like syndrome of (NZB × NZW)F1 mice. *European Journal of Immunology*, **24**, 8–15.

Hishikawa, T. *et al*. (1990). HLA-DP + T cells and deficient interleukin-2 production in patients with systemic lupus erythematosus. *Clinical Immunology and Immunopathology*, **55**, 285–96.

Hoffbrand, B.I. and Beck, E.R. (1965). 'Unexplained' dyspnoea and shrinking lungs in systemic lupus erythematosus. *British Medical Journal*, **1**, 1273–7.

Hopkinson, N.D., Doherty, M., and Powell, R.J. (1993). The prevalence and incidence of systemic lupus erythematosus in Nottingham, UK, 1989–1990. *British Journal of Rheumatology*, **32**, 110–15.

Horsfall, A.C., Rose, L.M., and Maini, R.N. (1989). Autoantibody synthesis in salivary glands of Sjögren's patients. *Journal of Autoimmunity*, **2**, 559–68.

Horwitz, D.A. and Jacob, C.O. (1994). The cytokine network in the pathogenesis of systemic lupus erythematosus and possible therapeutic implications. *Springer Seminars in Immunopathology*, **16**, 181–200.

Isenberg, D.A. (1983). Immunoglobulin deposition in skeletal muscle in primary muscle disease. *Quarterly Journal of Medicine*, **52**, 297–310.

Isenberg, D.A. and Malick, J. (1994). Male lupus—the Loch Ness syndrome revisited. *British Journal of Rheumatology*, **33**, 307–8.

Isenberg, D.A. and Snaith, M.L. (1981). Muscle disease in SLE: a study of its nature, frequency and cause. *Journal of Rheumatology*, **8**, 917–24.

Ishida, H., Muchamuel, T., Sakaguchi, S., Andrade, S., Menon, S., and Howard, M. (1994). Continuous administration of anti-interleukin 10 antibodies delays onset of autoimmunity in NZB/W F1 mice. *Journal of Experimental Medicine*, **179**, 301–10.

Ishikawa, O., Ohnishi, K., Miyachi, Y., and Ishizaka, H. (1994). Cerebral lesions in systemic lupus erythematosus detected by magnetic resonance imaging. Relationship to anti-cardiolipin antibody. *Journal of Rheumatology*, **21**, 87–90.

Izui, S., Elder, J.H., McConahey, P.J., and Dixon, F.J. (1981). Identification of retroviral gp70-anti-gp70 antibodies involved in circulating immune complexes in NZB × NZW mice. *Journal of Experimental Medicine*, **153**, 1151–60.

Jacob, C.O. and McDevitt, H.O. (1988). Tumour necrosis factor-a in murine autoimmune 'lupus' nephritis. *Nature*, **331**, 356–8.

Jacob, C.O., Fronek, Z., Lewis, G.D., Koo, M., Hansen, J.A., and McDevitt, H.O. (1990 *a*). Heritable major histocompatibility complex class II-associated differences in relevance to genetic predisposition to systemic lupus erythematosus. *Proceedings of the National Academy of Sciences* (*USA*), **87**, 1233–37.

Jacob, C.O., van der Meide, P.H., and McDevitt, H.O. (1990 *b*). *In vivo* treatment of (NZB × NZW) F1 lupus-like nephritis with monoclonal antibody to g interferon. *Journal of Experimental Medicine*, **166**, 798–803.

Jevnikar, A.M., Singer, G.G., Brennan, D.C., Xu, H.W., and Kelley, V.R. (1992). Dexamethasone prevents autoimmune nephritis and reduces renal expression of Ia but not costimulatory signals. *American Journal of Pathology*, **141**, 743–51.

Johnson, A.E., Gordon, C., Palmer, R.G., and Bacon, P.A. (1995). The prevalence and incidence of systemic lupus erythematosus (SLE) in Birmingham, UK, related to ethnicity and country of birth. *Arthritis and Rheumatism*, **38**, 551–8.

Kahl, L.E. (1994). Herpes zoster infections in systemic lupus erythematosus: risk factors and outcome. *Journal of Rheumatology*, **21**, 84–6.

Kalsi, J. *et al*. (1996). Functional and structural analysis of the binding of four human monoclonal anti-DNA antibodies. *Molecular Immunology*, **33**, 471–83.

Kalunian, K.C. *et al*. (1989). Idiotypic characteristics of immunoglobulins associated with systemic lupus erythematosus. *Arthritis and Rheumatism*, **32**, 513–22.

Kammer, G.M. and Stein, R.L. (1990). T lymphocyte dysfunctions in systemic lupus erythematosus. *Journal of Laboratory and Clinical Medicine*, **115**, 273–82.

Kanai, Y., Tokano, Y., Tsuda, H., Hashimoto, H., Okumura, K., and Hirose, S. (1993). HLA-DP positive T cells in patients with polymyositis/dermatomyositis. *Journal of Rheumatology*, **20**, 77–9.

Karsh, J., Klippel, J., Balow, J.E., and Decker, J.L. (1979). Mortality in lupus nephritis. *Arthritis and Rheumatism*, **22**, 764–9.

Kaufman, L.D., Heinicke, M.H., Hamberger, M., and Corevic, P.D. (1989). Male lupus: retrospective analysis of the clinical and laboratory features of 52 patients, with a review of the literature. *Seminars in Arthritis and Rheumatism*, **18**, 189–97.

Keeling, D.M. and Isenberg, D.A. (1993). Haematological manifestations of systemic lupus erythematosus. *Blood Reviews*, **7**, 199–207.

Kelley, V.R., Diaz-Gallo, C., Jevnikar, A.M., and Singer, G.G. (1993). Renal tubular epithelial and T cell interactions in autoimmune renal disease. [Review.] *Kidney International*, **39** (Suppl.), S108–15.

Kemp, M.E., Atkinson, J.P., Skanes, V.M., Levine, R.P., and Chaplin, D.D. (1987). Deletion of C4A genes in patients with systemic lupus erythematosus. *Arthritis and Rheumatism*, **30**, 1015–22.

Khamashta, M.A. *et al*. (1990). Association of antibodies against phospholipids with heart valve disease in systemic lupus erythematosus. *Lancet*, **335**, 1541–4.

Kiberd, B.A. (1993). Interleukin-6 receptor blockage ameliorates murine lupus nephritis. *Journal of the American Society of Nephrology*, **4**, 58–61.

Koh, W.H., Fong, K.R., Boeg, N.L., Feng, P.H. (1994). Systemic lupus erythematosus in 61 oriental males. A study of clinical and laboratory manifestations. *British Journal of Rheumatology*, **33**, 339–42.

Krapf, F.E., Herrmann, M., Leitman, W., and Kalden, J.R. (1989). Are retroviruses involved in the pathogenesis of SLE? Evidence demonstrated by molecular analysis of nucleic acids from SLE patients' plasma. *Rheumatology International*, **9**, 115–21.

Lahita, R.G., Chiorazzi, N., Gibofsky, A., Winchester, R.J., and Kunkel, H.G. (1983). Familial systemic lupus erythematosus in males. *Arthritis and Rheumatism*, **26**, 39–44.

Leaker, M., McGregor, A., Griffiths, M., Snaith, M., Neild, G., and Isenberg, D.A. (1991). Insidious loss of renal function in patients with anti-cardiolipin antibodies and absence of overt nephritis. *Annals of the Rheumatic Diseases*, **30**, 422–5.

Lee, S.S., Li, C.S., and Li, P.C.H. (1993). Clinical profile of Chinese patients with systemic lupus erythematosus. *Lupus*, **2**, 105–9.

Leung, W.H., Wong, K.L., Lau, C.P., Wong, C.K., and Liu, H.W. (1990). Association between antiphospholipid antibodies and cardiac abnormalities in patients with systemic lupus erythematosus. *American Journal of Medicine*, **89**, 411–9.

Levy, E., Ambrus, J., Kahl, L., Molina, H., Tung, K., and Holers, V.M. (1992). T lymphocyte expression of complement receptor 2 (CR2/CD21): a role in adhesive cell–cell interactions and dysregulation in a patient with systemic lupus erythematosus (SLE). *Clinical and Experimental Immunology*, **90**, 235–44.

Liang, M.H., Sacher, S.A., Neal Roberts, W., and Esdaile, J.M. (1988). Measurement of systemic lupus erythematosus activity in clinical research. *Arthritis and Rheumatism*, **31**, 817–25.

Libman, E. and Sachs, B. (1924). A hitherto undescribed form of valvular and mural endocarditis. *Archives of Internal Medicine (Chicago)*, **33**, 701–9.

Love, P.E. and Santoro, S.A. (1990). Anti-phospholipid antibodies: anticardiolipin and the lupus anticoagulant in systemic lupus erythematosus (SLE) and in non-SLE disorders. *Annals of Internal Medicine*, **112**, 682–98.

Lynch, A. *et al*. (1994). Anti-phospholipid antibodies in predicting adverse pregnancy outcome—a prospective study. *Annals of Internal Medicine*, **15**, 470–5.

Machold, K.P. and Smolen, J.S. (1990). Interferon-g induced exacerbation of systemic lupus erythematosus. *Journal of Rheumatology*, **17**, 831–2.

Maini, R.N., Elliott, M.J., Charles, P.J., and Feldmann, M. (1994). Immunological intervention reveals reciprocal roles for tumour necrosis factor-a and interleukin-10 in rheumatoid arthritis and systemic lupus erythematosus. *Springer Seminars in Immunopathology*, **16**, 327–36.

Mamula, M.J., Fatenejad, S., and Craft, J. (1994). B cells process and present lupus autoantigens that initiate autoimmune T cell responses. *Journal of Immunology*, **152**, 1453–61.

Mandell, B. (1987). Cardiovascular involvement in systemic lupus erythematosus. *Seminars in Arthritis and Rheumatism*, **17**, 120–141.

Maruyama, N. *et al*. (1983). Genetic studies of autoimmunity in New Zealand mice. IV. Contribution of NZB and NZW genes to the spontaneous occurrence of retroviral gp70 immune complexes in (NZB xNZW)F1 hybrid and the correlation to renal disease. *Journal of Immunology*, **130**, 740–6.

Matsuda, J., Saitoh, N., Gohchi, I., Gotoh, M., and Tsukamoto, M. (1994). Detection of b2-glycoprotein 1-dependent antiphospholipid antibodies and anti-b2-glycoprotein 1 antibody in patients with systemic lupus erythematosus and in patients with syphilis. *International Archives of Allergy and Immunology*, **103**, 239–44.

Matsukawa, Y., Sawada, S., Hayama, T., Usui H., and Horie, T. (1994). Suicide in patients with systemic lupus erythematosus: a clinical analysis of seven suicidal patients. *Lupus*, **3**, 31–5.

McGrath, H., Jr. (1994). Ultraviolet-A1 irradiation decreases clinical disease activity and autoantibodies in patients with systemic lupus erythematosus. *Clinical and Experimental Rheumatology*, **12**, 129–135.

McLaughlin, J.R., Bombardier, C., Farewell, V.T., Gladman, D.A., and Urowitz, M.B. (1994). Kidney biopsy in systemic lupus erythematosus. III. Survival analysis, controlling for clinical and laboratory variables. *Arthritis and Rheumatism*, **4**, 559–67.

Mendlovic, S. *et al*. (1988). Induction of a systemic lupus erythematosus-like disease in mice by a common human anti-DNA idiotype. *Proceedings of the National Academy of Sciences* (*USA*), **85**, 2260–4.

Menon, S., Snaith, M.L., and Isenberg, D.A. (1993). The asociation of malignancy with SLE: an analysis of 150 patients under long term review. *Lupus*, **2**, 177–81.

Messier, H. *et al*. (1993). p70 lupus autoantigen binds the enhancer of the T-cell receptor b-chain gene. *Proceedings of the National Academy of Sciences* (*USA*), **90**, 2685–9.

Miller, M.H., Urowitz, M.B., Gladman, D.D., and Killinger, D.W. (1983). Systemic lupus erythematosus in males. *Medicine*, **G2**, 327–34.

Morrow, W.J., Homsy, J., Swanson, C.A., Ohashi, Y., Estes, J., and Levy, J.A. (1986). Dietary fat influences the expression of autoimmune disease in MRL/lpr/lpr mice. *Immunology*, **59**, 439–43.

Mountz, J., Wu, J., Cheng, W., and Zhou, T. (1994). Autoimmune disease. A problem of defective apoptosis. *Arthritis and Rheumatism*, **37**, 1415–20.

Murakami, M., Kumagai, S., Sugita, M., Iwai, K., and Imura, H. (1992). *In vitro* induction of IgG anti-DNA antibody from high density B cells of systemic lupus erythematosus patients by an HLA DR-restricted T cell clone. *Clinical and Experimental Immunology*, **90**, 245–50.

Nakamura, T. *et al*. (1993). Renal expression of mRNAs for endothelin-1, endothelin-3 and endothelin receptors in NZB/W F1 mice. *Renal Physiology and Biochemistry*, **16**, 233–43.

Naparstek, Y. *et al*. (1990). Binding of anti-DNA antibodies and inhibition of glomerulonephritis in MRL-lpr/lpr mice by heparin. *Arthritis and Rheumatism*, **33**, 1554–9.

Norris, D.A. (1993). Pathomechanisms of photosensitive lupus erythematosus. [Review.] *Journal of Investigative Dermatology*, **100**, 58S–68S.

Ohteki, T., Okamoto, S., Nakamura, M., Nemoto, E., and Kumagai, K. (1993). Elevated production of interleukin 6 by hepatic MNC correlates with ICAM-1 expression on the hepatic sinusoidal endothelial cells in autoimmune MRL/lpr mice. *Immunology Letters*, **36**, 145–52.

Olsen, M.L., Arnett, F.C., and Reveille, J.D. (1993). Contrasting molecule patterns of MHC class II alleles associated with the anti-Sm and anti-RNP precipitin autoantibodies in systemic lupus erythematosus. *Arthritis and Rheumatism*, **36**, 94–104.

Petri, M., Lakatta, C., Magder, L., and Goldman, D. (1944). Effect of prednisone and hydroxychloroquine on coronary artery disease risk factors in systemic luypus erythematosus: a longitudinal data analysis. *American Journal of Medicine*, **96**, 254–9.

Pistiner, M., Wallace, D.J. Nessim, S., Metzger, A.L., and Klinenberg, J.R. (1991). Lupus erythematosus in the 1980s: A survey of 570 patients. *Seminars in Arthritis and Rheumatism*, **21**, 358–63.

Rajagopalan, S., Zordan, T., Tsokos, G.C., and Datta, S.K. (1990). Pathogenic anti-DNA antibody-inducing T helper cell lines from patients with active lupus nephritis: isolation of CD4–8-T helper cell lines that express the g T-cell antigen receptor. *Proceedings of the National Academy of Sciences* (*USA*), **87**, 7020–4.

Rajagopalan, S., Mao, C., and Datta, S.K. (1992). Pathogenic autoantibody-inducing g/d T helper cells from patients with lupus nephritis express unusual T cell receptors. *Clinical Immunology and Immunopathology*, **62**, 344–50.

Reveille, J.D., Schrohenloher, R.E., Acton, R.T., and Barger, B.O. (1989). DNA analysis of HLA-DR and DQ genes in American blacks with systemic lupus erythematosus. *Arthritis and Rheumatism*, **32**, 1243–51.

Rosner, S. *et al*. (1982). A multicenter study of outcome in systemic lupus erythematosus. II. Causes of death. *Arthritis and Rheumatism*, **256**, 612–7.

Santoro, T.J., Portanova, J.P., and Kotzin, B.L. (1988). The contribution of L3T4 + T cells to lymphoproliferation and autoantibody production in MRL-lpr/lpr mice. *Journal of Experimental Medicine*, **167**, 1713–8.

Schur, P.H., Marcus-Bagley, D., Awdeh, Z., Yunis, E.J., and Alper, C.A. (1990). The effect of ethnicity on major histocompatibility complex complement allotypes and extended haplotypes in patients with systemic lupus erythematosus. *Arthritis and Rheumatism*, **33**, 985–92.

Schwartz, M.M., Lan, S., Bernstein, J., Hill, G.S., Holley, K., and Lewis, E.J. (1993). Irreproducability of the activity and chronicity indices limits their utility in the management of lupus nephritis. *American Journal of Kidney Disease*, **21**, 374–7.

Scofield, R.H. *et al*. (1994). Cooperative association of T cell b receptor and HLA-DQ alleles in the production of anti-Ro in systemic lupus erythematosus. *Clinical Immunology and Immunopathology*, **72**, 335–41.

Segal, R., Globerson, A., Zinger, H., and Mozes, E. (1992). Induction of experimental systemic lupus erythematosus (SLE) in mice with severe combined immunodeficiency (SCID). *Clinical and Experimental Immunology*, **89**, 239–43.

Sfikakis, P.P., Charalambopoulos, D., Vayiopoulos, G., Oglesby, R., Sfikakis, P., and Tsokos, G.C. (1994). Increased levels of intercellular adhesion molecule-1 in the serum of patients with systemic lupus erythematosus. *Clinical and Experimental Rheumatology*, **12**, 5–9.

Shefner, R. and Diamond, B. (1990). A novel class of autoantibody. *Arthritis and Rheumatism*, **33**, S27.

Shortall, E., Isenberg, D., and Newman, S.P. (1995). Factors associated with mood and mood disorders in SLE. *Lupus*, **4**, 272–9.

Siegert, C.E., Daha, M.R., Tseng, C.M., Coremans, I.E., van Es, L.A., and Breedveld, F.C. (1993). Predictive value of IgG autoantibodies against C1q for nephritis in systemic lupus erythematosus. *Annals of the Rheumatic Diseases*, **52**, 851–6.

Sieving, R.R., Kauffman, C.A., and Watanakunakorn, C. (1975). Deep fungal infection in systemic lupus erythematosus—three cases reported, literature reviewed. *Journal of Rheumatology*, **2**, 61–72.

Silvestris, F., Cafforio, P., and Dammacco, F. (1994). Pathogenic anti-DNA idiotype-reactive IgG in intravenous immunoglobulin preparations. *Clinical and Experimental Immunology*, **97**, 19–25.

Spronk, P.E., ter Borg, E.J., and Kallenberg, C.G.M. (1992). Patients with systemic lupus erythematosus and Jaccoud's arthropathy: a clinical subset with an increased C reactive protein response. *Annals of Rheumatic Diseases*, **51**, 358–61.

Stamenkovic, I., Favre, H., Doneth, A., Assimacopoulos, A., and Chatelanet, F. (1986). Renal biopsy in systemic lupus erythematosus irrespective of clinical findings. Long term follow up. *Clinical Nephrology*, **26**, 109–15.

Sthoeger, Z.M., Geltner, D., Rider, A., and Bentwich, Z. (1987). Systemic lupus erythematosus in 49 Israeli males—a retrospective study. *Clinical and Experimental Rheumatology*, **5**, 233–40.

Talal, N. (1994). Oncogenes, autogenes and rheumatic diseases. *Arthritis and Rheumatiem*, **37**, 1421–2.

Tan, E.M. *et al*. (1982). The 1982 revised criteria for the classification of systemic lupus erythematosus (SLE). *Arthritis and Rheumatism*, **25**, 1271–7.

Teh, L. and Isenberg, D.A. (1994). Antiribosomal P protein antibodies in systemic lupus erythematosus—a reappraisal. *Arthritis and Rheumatism*, **37**, 307–15.

Termaat, R.-M. *et al*. (1990). Cross-reactivity of monoclonal anti-DNA sulphate is mediated via bound DNA/histone complexes. *Journal of Autoimmunity*, **3**, 531–45.

Theofilopoulos, A.N., Singer, P.A., Kofler, R., Kono, D.H., Duchosal, M.A., and Balderas, R.S. (1989). B and T cell antigen receptor repertoires in lupus/arthritis murine models. *Springer Seminars in Immunopathology*, **11**, 335–68.

Turner-Stokes, L. and Isenberg, D.A. (1988). Immunisation of patients with rheumatoid arthritis and systemic lupus erythematosus. *Annals of the Rheumatic Diseases*, **47**, 529–31.

Urman, J.D. and Rothfield N.F. (1977). Corticosteroid treatment in systemic lupus erythematosus: survival studies. *Journal of the American Medical Association*, **238**, 2272–6.

Urowitz, M.B., Gladman, D.D., Farewell, V.T., Stewart, J., and McDonald, J. (1993). Lupus and pregnancy studies. *Arthritis and Rheumatism*, **36**, 1392–7.

Valeri, A. *et al*. (1994). Intravenous pulse cyclophosphamide treatment of severe lupus nephritis: a prospective five year study. *Clinical Nephrology*, **42**, 71–8.

Vitali, C. *et al*. (1992). Disease activity in SLE: report of the concensus study group of the European Workshop for Rheumatology Research III: development of a computerized clinical chart and its application to the comparison of different indices of disease activity. *Clinical and Experimental Rheumatology*, **10**, 549–54.

Wallace, D.J., Podell, T., Weiner, J., Klinenberg, J.R., Forouzesh, S., and Dubois, E.L. (1981). Systemic lupus erythematosus: survival pattern. Experience with 609 patients. *Journal of the American Medical Association*, **245**, 934–8.

Walport, M.J. and Lachmann, P.J. (1988). Erythrocyte complement receptor type I, immune complexes and the rheumatic diseases. *Arthritis and Rheumatism*, **31**, 153–8.

Walton, A.J.E., Snaith, M.L., Locniskar, M., Cumberland, A.G., Morrow, W.J.W., and Isenberg, D.A. (1991). Dietary fish oil reduces the severity of symptoms in patients with SLE. *Annals of the Rheumatic Diseases*, **33**, 463–6.

Ward, M.M. and Studensky, S. (1990). Systemic lupus erythematosus in men—a multivariate analysis of gender differences in clinical manifestations. *Journal of Rheumatology*, **17**, 220–4.

Watts, R. and Isenberg, D.A. (1989). Pancreatic complications of the autoimmune rheumatic diseases. *Seminars in Arthritis and Rheumatism*, **19**, 158–65.

Watts, R. and Isenberg, D.A. (1990). DNA antibody idiotypes: an analysis of their clinical connections and origins. *International Review of Immunology*, **5**, 279–93.

Wei, N., Klippel, J.H., and Husto, D.P. (1983). Randomised trial of plasma exchange in mild SLE. *Lancet*, **i**, 17–22.

Weinrib, L., Sharma, O.P., and Quismario, F.P. (1990). A long term study of interstitial lung disease in systemic lupus erythematosus. *Seminars in Arthritis and Rheumatism*, **16**, 479–81.

Wellicome, S.M., Kapahi, P., Mason, J.C., Lebranchu, Y., Yarwood, H., and Haskard, D.O. (1993). Detection of a circulating form of vascular cell adhesion molecule-1: raised levels in rheumatoid arthritis and systemic lupus erythematosus. *Clinical and Experimental Immunology*, **92**, 412–18.

Whiting-O'Keefe, Q., Henke, J.E., Sheard, M.A., Hopper J., Jr., Biava, C.G., and Epstein, W.V. (1982). The information content from renal biopsy in systemic lupus erythematosus. Stepwise linear regression analysis. *Annals of Internal Medicine*, **96**, 718–27.

Williams, W. and Isenberg, D.A. (1994). Idiotypes and autologous anti-idiotypes in human autoimmune disease —some theoretical and practical observations. *Autoimmunity*, **17**, 343–52.

Wofsy, D. and Seaman, W.E. (1985). Successful treatment of autoimmunity in NZB/NZW F1 mice with monoclonal antibody to L3T4. *Journal of Experimental Medicine*, **161**, 378–91.

Wolff-Vorbeck, G. *et al*. (1994). Characterization of an HLA-DR4-restricted T cell clone recognizing a protein moiety of small ribonucleoprotein (UsnRNP). *Clinical and Experimental Immunology*, **95**, 378–84.

Worrall, J.G., Snaith, M.L., Batchelor, J.R., and Isenberg, D.A. (1990). SLE: a rheumatological view. Analysis of the clinical features, serology, and immunogenetics of 100 SLE patients during long term follow up. *Quarterly Journal of Medicine*, **74**, 319–30.

Wu, J., Zhou, T., Zhang, J., and Mountz, J. (1994). Correction of autoimmune disease in CD2-fas transgenic mice. *Proceedings of the National Academy of Sciences* (*USA*), **91**, 2344–8.

Wuthrich, R.P. (1992). Vascular cell adhesion molecule-1 (VCAM-1) expression in murine lupus nephritis. *Kidney International*, **42**, 903–14.

Wuthrich, R.P., Jevnikar, A.M., Takei, F., Glimcher, L.H., and Kelley, V.E. (1990). Intercellular adhesion molecule-1 (ICAM-1) expression is upregulated in autoimmune murine lupus nephritis. *American Journal of Pathology*, **136**, 441–50.

Wysenbeek, A.J., Leibovici, L., Weinberger, A., and Guadj, D. (1993). Fatigue in systemic lupus erythematosus: prevalence and relation to disease expression. *British Journal of Rheumatology*, **32**, 633–5.

Xavier, R.M., Nakamura, M., and Tsunematsu, T. (1994). Isolation and characterization of a human nonspecific suppressor factor from ascitic fluid of systemic lupus erythematosus. Evidence for a human counterpart of the monoclonal nonspecific suppressor factor and relationship to the T cell receptor a-chain. *Journal of Immunology*, **152**, 2624–32.

Yamada, A., Minota, S., Nojima, Y., and Yazaki, Y. (1993). Changes in subset specificity of anti-T cell autoantibodies in systemic lupus erythematosus. *Autoimmunity*, **14**, 269–73.

Yoshida, S., Castles, J.J., and Gershwin, M.E. (1990). The pathogenesis of autoimmunity in New Zealand mice. *Seminars in Arthritis and Rheumatism*, **19**, 224–42.

5.7.2 Systemic lupus erythematosus in children

Earl D. Silverman and Alison A. Eddy

Systemic lupus erythematosus (SLE) in children and adolescents has many features in common with adult-onset SLE. This chapter, rather than fully reviewing all the possible presentations and manifestations of SLE will highlight the differences between paediatric SLE and adult-onset lupus. This chapter will not cover many topics such as immunopathogenesis, cytokines, and animal models which are outlined in Chapter 5.7.1, nor will it cover antiphospholipid antibodies in depth. Rather, it should serve as a complementary chapter and will place its emphasis on unique features and provide an overview of general features of paediatric SLE. We suggest reading both chapters on SLE. A separate section of this chapter will cover neonatal lupus erythematosus, a disease caused by maternally-transmitted autoantibodies.

Incidence

Although SLE has been recognized for many decades, most physicians rarely consider it in the paediatric age group. Currently there has been no good epidemiological study focusing on paediatric SLE and very few large population studies give details of its incidence prior to the age of 20. A large population study from New York showed an incidence of 6 cases per 100 000 white females of less than 15 years of age as compared with an overall incidence of 25 cases per 100 000 white females of all ages. The incidence rate rapidly rose to 18.9 cases per 100 000 white females aged between 15 and 25 years. Therefore, the true incidence of SLE beginning prior to age 19 was between 6 and 18.9 cases per 100 000 in white females, and was higher in black (20 to 30 per 100 000) and Puerto Rican females (16 to 36.7 per 100 000) (Siegel and Lee 1973).

Most other series of SLE patients with a large number of paediatric cases state that 20 per cent of all cases of SLE have onset prior to age 18 (Kaufman *et al*. 1986; Reeves and Lahita 1987). Therefore, using the best available data, the incidence of SLE beginning prior to age 18 is 10 to 20 cases per 100 000 children and adolescents, with an overall prevalence of 10 to 20 cases per 10 000 people less than 18 years old. These rates are higher in Hispanic, black and oriental people.

Early reports had suggested that there was a high percentage of male cases prior to puberty, while after puberty the percentage of cases occurring in females rose to make up 85 to 90 per cent of cases (as seen in most adult studies). In our series we had an overall male:female ratio of 1:4.3. When further divided by age, we found a male:female ratio of 1:5.5 between 6 and 10 years (prepubertal); 1:3.4 in the age 11 to 13 group (peripubertal); and 1:4.5 between the ages of 14 and 18 (postpubertal). Our data are consistent with more recent larger reviews (King *et al*. 1977; Lehman *et al*. 1989b). These data may reflect both an increased awareness of the possibility of SLE in males and the larger series size would eliminate any potential false bias of previous small studies. Whatever the reasons, the overall ratio of male:female cases of approximately 1:4.5 suggests that there is a higher percentage of male cases in paediatric SLE than in adult cases of SLE.

In the paediatric series, the average age at diagnosis varied from 11 to 14 years (median 12.2 years). The time from onset of symptoms to diagnosis varied from 8 months to 3.3 years (median 1.2 years). This median time of 1.2 years from symptoms to diagnosis emphasizes the difficulty in diagnosis or the lack of awareness of SLE in this age group. The features of SLE at presentation are shown in Table 1; while features at any time during the course of the disease are shown in Table 2 (Zetterstrom and Berglund 1956; Gribetz and Henley 1959; Cook *et al*. 1960; Jacobs 1963; Robinson and Williams 1967 ; Walravens and Chase 1976; Fish *et al*. 1977; King *et al*. 1977; Abeles *et al*. 1980; Caeiro *et al*. 1981; Yancey *et al*. 1981; Schaller 1982; Glidden *et al*. 1983; Emery 1986; Kaufman *et al*. 1986; Lehman *et al*. 1989a; El-Garf and Salah 1990; Lacks and White 1990). These are shown as the range of each feature in: (i) previously published paediatric series; and (ii) our series of 86 patients. Details are given in the organ-specific sections. However, general systemic symptoms such as fever, malaise and weight loss are common, as is evidence of systemic inflammation shown by lymphadenopathy and/or hepatosplenomegaly—this is both at diagnosis and throughout the course of the disease.

	Other series (%)	Our series (%)
Fever	60–90	55
Arthritis	60–88	78
Skin rash (any)	60–78	79
Malar rash	22–60	36
Renal	20–80	51
Cardiovascular	5–30	14
Pulmonary	18–40	18
Central nervous system	5–30	25
Gastrointestinal	14–30	19
Hepatosplenomegaly	16–42	30
Lymphadenopathy	13–45	34

Table 1 Clinical features: at diagnosis

	Other series (%)	Our series (%)
Fever	80–100	86
Arthritis	60–90	80
Any skin rash	60–90	86
Malar rash	30–60	38
Renal	46–100	69
Cardiovascular	25–60	17
Pulmonary	18–61	18
Central nervous system	20–44	34
Gastrointestinal*	24–40	24
Hepatosplenomegaly	19–43	30
Lymphadenopathy*	13–45	34

*Not included in follow-up data in many paediatric series.

Table 2 Clinical features: any time during the course

Musculoskeletal disease

Arthritis and arthralgia are among the most common symptoms in paediatric SLE occurring in more than 90 per cent of cases. Most patients with arthritis have a symmetric polyarthritis affecting both large and small joints. The arthritis usually responds to the treatment of other major organ involvement or to treatment of the general systemic symptoms. Unlike juvenile chronic arthritis, the arthritis is usually episodic and is easy to treat. Severely painful joints are common and usually the pain is out of proportion to the physical findings. The peripheral joints including the fingers, hands, wrists and knees are the commonest joints involved, with involvement of hips and ankles more unusual. In most cases there is no radiographic evidence of joint destruction. However, in 1 to 2 per cent of cases of lupus arthritis there is a deforming erosive arthritis as seen in seropositive juvenile chronic arthritis. This type of arthritis is usually associated with a positive rheumatoid factor, and at least in adults, has been referred to as 'Rhupus'. A more common cause of joint deformities is secondary to ligamentous laxity and periarticular fibrosis, the so-called Jaccoud's arthritis. This can result in deformity of the hand with multiple joint subluxations but with good preservation of function. Radiographs show only osteoporosis without evidence of erosions or joint space loss. Jaccoud's arthritis is more of a periarticular rather than an articular disease (Martini *et al*. 1987). However, in most patients with arthritis of SLE, the arthritis does not result in deformity. Both 'Rhupus' and Jaccoud's arthritis combined occur in less than 5 per cent of paediatric SLE patients with arthritis. Tenosynovitis is common and patients with knee effusions can develop Baker's cysts. Of particular interest have been the reports of patients with American College of Rheumatology (ACR) criteria for juvenile chronic arthritis who develop SLE years later. We have seen a patient with 'classic' systemic-onset juvenile chronic arthritis, including evanescent rash and spiking fever, who after 2 years in remission for her juvenile chronic arthritis developed SLE, including a photosensitive rash and positive antinuclear antibodies (which previously had been negative). Initial reports had shown a transformation from polyarticular rather than systemic juvenile chronic arthritis to SLE (Ragsdale *et al*. 1980; Saulsbury *et al*. 1982), however a more recent report has also demonstrated transformation from systemic juvenile-onset chronic arthritis to SLE (Citera *et al*. 1993).

Analysis of synovial fluid usually shows only mild inflammation with a total white blood-cell count of generally less than 2000 cells/ml. Synovial biopsies show a mild vasculitis or perivasculitis and the diagnostic histological lesion is the hematoxylin body. This may be the equivalent of the LE cell in the joint. However, synovial biopsies are indicated rarely.

Myalgia, as part of the generalized disease process, occurs in approximately 50 to 60 per cent of paediatric cases, while true myositis with proximal muscle weakness or tenderness occurs in less than 10 per cent of cases. When present there is usually a slow progression of weakness and rarely is there involvement of the intercostal or cricopharyngeal muscles. It is seen more commonly in patients with so-called 'overlap syndromes' or with mixed connective tissue disease. Primary muscle involvement must be differentiated from muscle weakness secondary to steroid treatment. This complication is a much talked about steroid side-effect but in our experience it is uncommon. The treatment of steroid-induced myopathy would be to decrease rather than increase the steroid dose and would result in an improvement in muscle strength.

The other steroid-induced musculoskeletal side-effects include avascular necrosis, osteoporosis with fracture or vertebral body collapse, and even growth failure if prolonged high-dose steroids are required. Growth failure may be alleviated partially by alternate-day dose regimens but even using these it still remains a problem. One hope may be through the use of recombinant growth hormone to promote growth, but no studies have been undertaken yet in this group of patients.

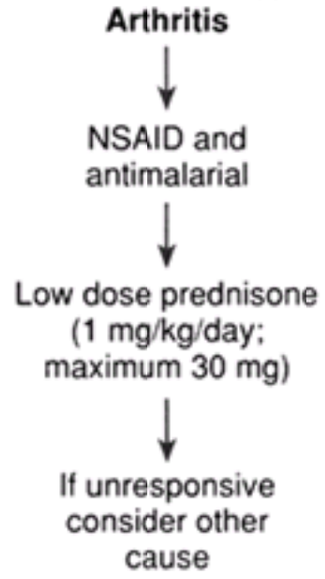
Avascular necrosis occurs in approximately 10 to 15 per cent of paediatric cases and appears to be more common in children than adults (Smith *et al*. 1976). In our experience, avascular necrosis occurs more commonly in SLE than other paediatric autoimmune diseases where prolonged high-dose steroids are used. The reason for this is not readily apparent. Unlike adults, in our experience, paediatric patients with Raynaud's phenomenon and vasculitis do not appear to be at a greater risk of this complication. Usually avascular necrosis is related to the dose and duration of therapy but there have been case reports of it developing prior to or without steroid usage (Abeles *et al*. 1978; Kalla *et al*. 1986). In contrast to avascular necrosis, vertebral collapse is rarer in SLE than in other autoimmune rheumatic diseases and in particular juvenile dermatomyositis. The incidence of vertebral collapse is probably related not only to steroid dose and duration but also to the amount of weight gain and physical activity. The most common lesion involves the thoracolumbar spine with cervical spine lesions occurring more rarely.

The last musculoskeletal side-effect of steroid therapy is increased incidence of fractures secondary to steroid-induced osteoporosis. This appears to be less common in children than adults and may reflect relative differences in bone mineral content between the groups. Treatment or prevention of steroid-induced side-effects is controversial in adults and there has not been a good study in children. The best hope lies with the new generation of 'bone-sparing' steroids and keeping patients active while minimizing weight gain.

Treatment of musculoskeletal involvement

The arthritis of SLE frequently occurs with disease flares elsewhere and usually responds to the treatment for the more serious complication. However, if isolated, the use of a non-steroidal anti-inflammatory drug with an antimalarial drug has a high success rate. The only word of caution is that SLE patients seem to be more susceptible to hepatotoxicity induced by non-steroidal anti-inflammatory drugs, particularly aspirin. We commonly use naproxen (10 to 20 mg/kg per day), tolmetin (20 to 30mg/kg per day) and avoid ibuprofen because of the reports of aseptic meningitis occurring in SLE patients. When the arthritis is more severe or unresponsive we commonly add hydroxychloroquine at a dose of 5 mg/kg per day, although others have advocated the use of chloroquine. Antimalarial therapy appears to be of benefit not only for the arthritis but also for rash and overall disease control. The major side-effects are retinal macular deposition of the drug and gastrointestinal distress. Therefore, we recommend ophthalmological examination every 6 months and taking the drug immediately before bed to lessen gastrointestinal upset. Prednisone at low to moderate dose may be required. For suggested therapy of arthritis see Box 1.

Box 1 Arthritis therapy flow chart



Mucocutaneous involvement

Cutaneous involvement has been reported in 50 to 80 per cent of paediatric SLE patients at the time of diagnosis and in up to 85 per cent of patients during the course of the disease. A malar rash, in the classic 'butterfly distribution', is the most common rash occurring in 30 to 60 per cent of cases at diagnosis and in up to 80 per cent of patients during the course of the disease. The rash may be mild, requiring no treatment, but can be severe and cosmetically unacceptable requiring treatment with steroids, either topically or systemic, and/or antimalarial drugs. The appearance of a malar rash often heralds a disease flare and while usually non-scarring it can be present as a crusted or scaling rash. A truly photosensitive rash only occurs in approximately one-third of paediatric patients. The photosensitive rash can occur not only on the face but in any sun-exposed area, especially the arms and legs (Fig. 1). This photosensitive rash may be maculopapular or papulosquamous and may be associated with anti-Ro and anti-La antibodies. The other anti-Ro/La antibody-associated rash is annular erythema. This rash is

commonly photosensitive and occurs on the face or neck (Deng *et al*. 1984) (Fig. 2). True discoid lupus lesions are seen rarely in patients under the age of 18 years. In patients with a photosensitive rash, sun exposure may not only exacerbate the skin disease but may cause a systemic flare. Therefore, we recommend avoidance of sunbathing along with the use of sun-blocking agents, with high SPF (sun protecting factor) blocking both ultraviolet A and B wavelengths, and protective clothing including long-sleeved shirts and hats. It must also be remembered that unshielded fluorescent lamps may emit ultraviolet B radiation.

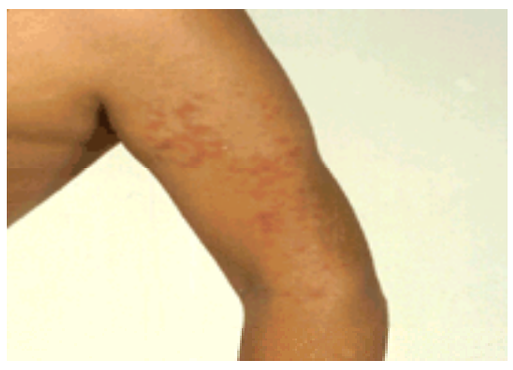


Fig. 1 A patient with known SLE developed a severe, vasculitic rash on arms and legs following sun exposure.

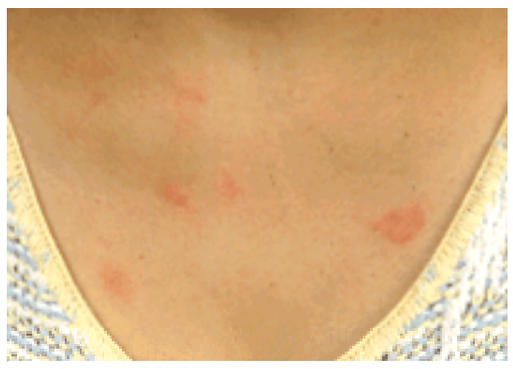


Fig. 2 Annular erythema on the neck of an SLE patient with anti-Ro and anti-La antibodies.

A true vasculitic skin rash has been reported in 10 to 20 per cent of patients. It occurs commonly in fingers or toes and can result in splinter haemorrhages and digital infarcts. Oral or nasal ulcers are caused by the same vasculitic process and they are probably a reflection of alterations seen in other vascular beds. Even when isolated, they may signify active disease. Painless ulcers are more common on the hard palate but may occur on the soft palate. A petechial rash on the hard palate may precede true ulceration while a chronic sore throat may represent mildly active vasculitis. Similarly, hyperaemia of the nasal mucosa usually precedes ulceration or perforation (Fig. 3). Most paediatric series do not comment on the incidence of these lesions but they probably occur in 10 to 20 per cent of patients. For suggested therapy of mucocutaneous disease see Box 2.

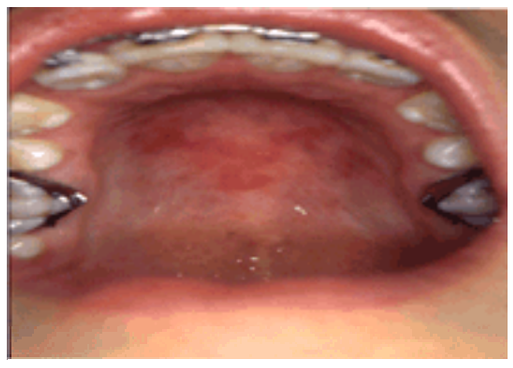


Fig. 3 Hyperaemia and petechiae on hard palate secondary to vasculitis.

Alopecia, no longer part of the revised ACR criteria for SLE, is common and occurs in up to 50 per cent of paediatric patients. Diffuse hair loss during washing or brushing is more common than patchy hair loss. In some patients there is a delay in the hair loss, which may follow a disease flare or increase after the introduction of steroids. Rarely is the hair loss significant enough to cause a cosmetic problem. Scarring is unusual.

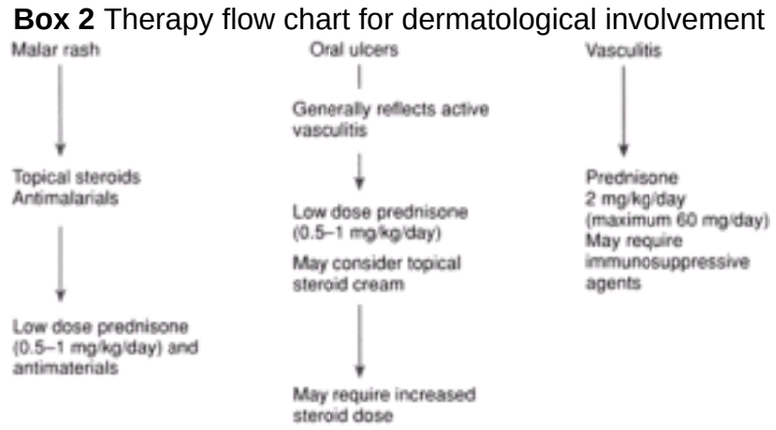
Raynaud's phenomenon appears to be less common in paediatric lupus, occurring in only 10 to 20 per cent of patients. In some patients local therapy including avoidance of cold, use of insulated mittens rather than gloves and hand/feet warmers will suffice. However, many patients will have more severe disease requiring the use of calcium-channel blocking agents or other vasodilating medication, including topical therapy with prostaglandins or nitroglycerin paste. Paediatric patients appear to tolerate calcium-channel blocking agents better than adults do. Patients with Raynaud's phenomenon rarely have normal vascular tone and will often have vasodilation at rest or severe vasoconstriction after only slight provocation. Raynaud's phenomenon may precede overt clinical SLE by months and may initially involve only a few digits and later involve all fingers and toes. Although usually easy to control, Raynaud's may result in distal digital infarction or gangrene.

Central nervous system

Involvement of the central nervous system, which occurs in 20 to 50 per cent of paediatric patients, can be the most interesting yet frustrating target in SLE. The interest stems from the protean manifestations but the frustration stems from the difficulty in diagnosis and assessment of patients with neuropsychiatric SLE (NP-SLE). Central nervous involvement can occur in isolation or it may be associated with disease flares in other organs. The only correlation between central nervous disease and disease elsewhere appears to be the association with vasculitis and thrombocytopenia. Table 3 lists the different manifestations of nervous system involvement seen.

Table 3 Neuropsychiatric lupus

Box 2 Therapy flow chart for dermatological involvement



Neuropsychiatric SLE

The most difficult SLE patients to diagnose and treat are the patients with psychiatric disorders. As a group psychiatric disorders are seen in 10 to 20 per cent of paediatric SLE cases. When a known lupus patient presents with florid psychosis or organic brain syndrome, the diagnosis is easy. However, many patients present with affective or mood disorders. The difficulty is determining whether the abnormalities detected are a direct result of the disease, secondary to steroid treatment, or 'reactive' to the disease or changes in body image. Several different investigations have been proposed to differentiate active disease from other causes of central nervous abnormalities (see below).

It can be very difficult to differentiate neuropsychiatric SLE-induced cognitive impairment and concentration difficulties from other causes of poor school performance. Studies in adults have suggested that cognitive impairment may occur in up to 80 per cent of all SLE patients (Carbotte *et al*. 1986). Defects in cognitive function may reflect active disease or residual defects from previous central nervous involvement (Fisk *et al*. 1993). This is emphasized by a longitudinal study which demonstrated that cognitive functional abnormalities may resolve and therefore they do not necessarily reflect irreversible damage (Hanly *et al*. 1994). To date there has been only one study in paediatric SLE that suggests that cognitive function defects are common (Papero *et al*. 1990). This has also been our experience. Although affective disorders occur less frequently than cognitive impairment, it must be recognized that patients can present with a major affective disorder with very few other signs of lupus.

Seizures occur in approximately 10 to 20 per cent of paediatric SLE cases and may be the presenting sign. Focal seizures are more common than *grand mal* seizures. Seizures are generally easily treated with anticonvulsant medication and usually do not require high-dose steroids for control. Seizures may not be primary but rather secondary to metabolic disturbances caused by uraemia, hypertension, cerebral infarction or central nervous infection. Structural abnormalities should be ruled-out in all patients with new onset or increasing frequency of seizures.

Movement disorders including chorea, cerebellar ataxia, hemiballismus, tremor and Parkinsonian-like movements, occur in 5 to10 per cent of cases. Interestingly chorea appears to be over-represented in the paediatric age group. In one review of 52 cases of SLE-associated chorea, 34 had chorea that developed before the age of 18 years, with the chorea preceding other manifestations of SLE by more than 1 year in 20 per cent of all cases (Bruyn and Padberg 1984). Recently we have had cases of chorea that preceded the development of overt SLE by many years and chorea that developed in patients with long-standing SLE. It has been demonstrated that there is an association between chorea and antiphospholipid antibodies. The exact mechanism which leads to chorea is unknown, although it rarely appears to be the result of an infarction that can be demonstrated on neuroimaging studies. The decline in the incidence of rheumatic fever means that the diagnosis of SLE or antiphospholipid antibody syndrome should be considered in all patients presenting with chorea.

Neuropathies

Both cranial and peripheral neuropathies can occur, with cranial nerve involvement being the more common. Cranial nerve involvement usually affects cranial nerves II, III, IV, and VI resulting in abnormalities of vision, pupils, or extraocular movements. Less frequently facial palsy (VII), trigeminal neuropathy (V) or nystagmus and vertigo (VIII) occur. When peripheral neuropathies occur, a sensory or mixed sensorimotor involvement is more common than an isolated motor neuropathy. However, mononeuritis or mononeuritis multiplex can occur and isolated cases of Guillain–Barré-like syndrome have been seen.

Paresis, while seen in approximately 5 per cent of adult cases, is less common in children. Hemiparesis in association with other major neurological findings, rarely seen in children, is usually secondary to cerebral vascular accident. Hemiparesis and/or cerebral vascular accidents may occur as the result of active SLE but also may occur when the SLE is complicated by the presence of hypertension, antiphospholipid antibodies, or thrombocytopenia. Transverse myelitis may present with acute paraplegia or quadriplegia; in this latter syndrome antiphospholipid antibodies must be sought.

Headache

Headache requires a separate category as this is a common symptom occurring in approximately 20 per cent of cases. The differentiation of the type of headache is important. The severe, unremitting lupus headache is the most serious and reflects active disease or may represent cerebral vein thrombosis (see below). A special problem occurs with migraine headache secondary to SLE, which must be differentiated from non-lupus associated vascular instability, particularly in patients with a family history of migraine headaches (Isenberg *et al*. 1982). A headache may be the presentation of pseudotumour cerebri, which has been described as the presenting diagnosis in paediatric SLE. Migraine headaches may reflect active central nervous SLE (Miguel *et al*. 1994).

Most worrying is the association of severe headache with cerebral vein thrombosis. Although cerebral vein thrombosis has been felt to be rare, we have recently recognized this complication in patients with headache. The headache associated with cerebral vein thrombosis may occur without any other neurological manifestation (Uziel 1995). We recommend neuroimaging studies for patients with persistent or severe headaches (see below).

Investigation

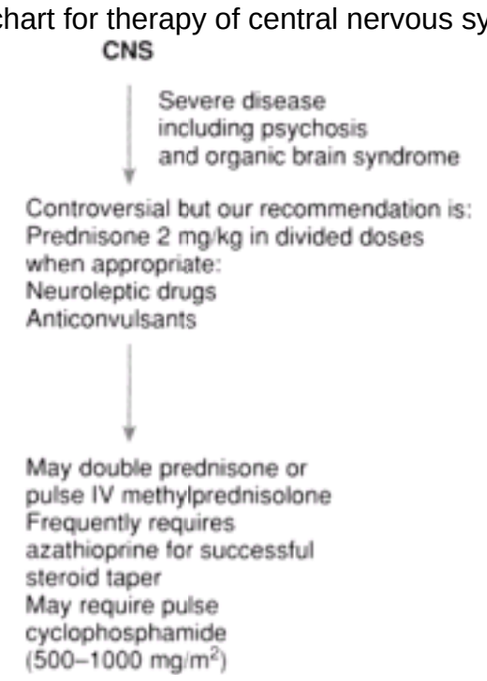
As the differential diagnosis neuropsychiatric SLE is large, so is the list of investigations. Abnormalities of cerebrospinal fluid cell count, protein and complement levels have been described. However, the findings are inconsistent and complement levels cannot be routinely obtained. It has been suggested that measurements of the integrity of the blood–brain barrier and of immunoglobulin synthesis in the cerebrospinal fluid may correlate with neuropsychiatric SLE, but the demonstration of these abnormalities is not specific for SLE (Hirohata *et al*. 1985; McLean *et al*. 1995). Furthermore, large controlled, prospective studies are not available in SLE to test the utility of measuring these parameters. However, the simplest test, examination and culture of cerebrospinal fluid, remains important when the possibility of infection or haemorrhage exists. An elevated cerebrospinal fluid protein and/or white blood-cell count, in the absence of infection, is suggestive of cerebritis.

The correlation of antineuronal antibodies and cognitive function has sparked much interest recently. One prospective adult study suggested that a combination of the

measurement of antineuronal antibodies with a large battery of sophisticated neurophysiological cognitive function tests correlated well with fluctuations of disease (Long *et al*. 1989). Our limited experience has suggested that this combination of testing may be of benefit in diagnosis and disease monitoring. Despite its promise, cognitive function testing requires more time and expertise than is routinely available. Furthermore, longitudinal studies of cognitive function may be required to determine whether a defect is old, permanent or temporary (Hanly *et al*. 1994).

Serial measurement of more routine autoantibodies and routine blood tests may not correlate with neuropsychiatric SLE (Miguel *et al*. 1994). Patients may have normal serology and complement levels despite active SLE of the central nervous system. One exception may be the association of antiribosomal P antibody and depression. Initial studies had suggested that although antiribosomal P antibodies are present in approximately 15 per cent of all lupus patients, these autoantibodies occur in most lupus patients with depression (Bonfa *et al*. 1987). However, other studies have failed to demonstrate this high degree of association with psychosis or depression (Teh *et al*. 1993). Importantly, preliminary studies suggest that the presence of the antiribosomal P antibodies may distinguish depression secondary to SLE from other causes including reactive, steroid-induced or other non-SLE causes of depression (Schneebaum *et al*. 1991). We have recently examined antiribosomal P antibodies in paediatric SLE and found that the antibodies were present only in patients with psychosis secondary to SLE and not in patients with other causes of psychosis. However, many patients with psychosis did not have these antibodies, and many SLE patients with the antibodies did not have psychosis. Therefore, although anti-ribosomal P antibodies were specific for psychosis secondary to SLE, they had a low sensitivity in SLE patients with psychosis and a low specificity for psychosis in our SLE population (Press *et al*. 1996).

Box 3 Flow chart for therapy of central nervous system disease



Radiological investigation of the central nervous system may be helpful in demonstrating specific structural lesions such as infarction, embolus and subdural or intracranial haemorrhage. However, neither computed tomography (CT) nor magnetic resonance imaging (MRI) have been shown to be consistently helpful in measuring overall disease activity in the central nervous system (O'Connor 1988). The use of MRI to determine the presence of diffuse disease activity in the central nervous system remains controversial (Isshi *et al*. 1994; Jarek *et al*. 1994). Although abnormalities have been described, these findings may not correlate with clinical disease.

An SLE patient with persistent or severe headache should be investigated for the presence of intracranial thrombosis and in particular cerebral vein thrombosis. The best investigation is by a combination of CT and MRI scans. A CT scan, with 'wide-windows' should be performed in all patients when this diagnosis is considered (Uziel 1993). This examination will detect the presence of infarction and thrombosis. The use of 'wide-windows' will allow for the appropriate visualization of the cerebral veins. The neuroradiologist should be advised of the possibility of this diagnosis as the use of 'wide-windows' is not part of the routine CT scan. The diagnosis of cerebral vein thrombosis may be confirmed by the absence of flow on an MR venogram. An MR venogram is a very sensitive examination but it may not be readily available and may require sedation, while a CT scan is easier to perform and more readily available.

In most series, the use of nuclear medicine brain scans has not been shown to be of any diagnostic use. However, the use of single-photon-emission computed tomography nuclear brain scans (SPECT) holds promise. These scans may be a functional assessment of brain activity (Holman 1991). SPECT scans appear to be a sensitive test, both in adults and children, to measure active disease in the central nervous system (Rubbert *et al*. 1993; Szer *et al*. 1993; Kodama *et al*. 1995). However, the ability of this test to differentiate active SLE from other causes of abnormalities in the central nervous system remains to be determined. In addition patients without overt central nervous disease may have an abnormal SPECT scan. Whether the perfusion abnormalities in this latter group of SLE patients reflects subclinical dysfunction of the central nervous system or the generalized disease process, remains to be determined. Positron-emission tomography (PET) scans are not universally available and are considered generally as a research tool.

Treatment of central nervous system disease

The therapy of central nervous system disease varies with the manifestation. Currently the treatment of isolated cognitive disorders is controversial, although one adult study has suggested that these defects may be steroid-responsive. Obviously when infection is present the therapy is dictated by the microbiological identification of the organism. However, it must be remembered that these patients are susceptible to encapsulated organisms, Gram-negative organisms and opportunistic infections.

Active psychosis and/or organic brain syndrome are potentially life-threatening complications of central nervous system disease and should be treated aggressively with high-dose corticosteroids. In our experience these patients frequently require immunosuppressive therapy with azathioprine, to allow for the successive tapering of the steroid dose. For resistant central nervous disease, treatment with cyclophosphamide may be required. In addition, the use of psychotropic drugs may be needed for the control of the psychosis. We do not recommend the isolated use of psychotropic medication without the use of steroids. Although reported by others, we have not seen a case of steroid-induced psychosis. For suggested therapy of severe central nervous system disease see Box 3.

Treatment of seizures should be directed at finding their cause, whether they are secondary to infarction, active central nervous system disease or metabolic disturbance. Anti-convulsant medication may be required to control the seizures, but in many cases it is needed only for a short-term if the underlying cause of the seizure can be corrected.

As previously described, headaches can occur for a variety of reasons in these patients. The treatment is therefore dictated by the cause. If the headache is mild and it resolves with the use of routine analgesia then no further investigation or therapy is required. However, persistent headache resistant to analgesia may reflect active SLE or may be secondary to an intracranial thrombosis (in particular cerebral vein thrombosis). The management of 'lupus' headache secondary to active SLE requires better control of the SLE, which may include the use of steroids. We suggest that cerebral vein thrombosis requires anticoagulation, initially at high dose for 3 to 6 months followed by long-term low-dose anticoagulation (Uziel 1995). However, the need for and duration of long-term anticoagulation is controversial.

Renal disease

Renal involvement occurs in a significant number of children with SLE. In a literature review of 540 children followed at centres other than our own, 72 per cent developed nephritis during the course of their disease, similar to our overall incidence of 61 per cent in 138 children with SLE. The high susceptibility of the kidneys to injury in this disease is not entirely understood but it is due in part to haemodynamic factors; to the unique architecture of the glomerular capillary wall that permits direct exposure of blood-borne molecules to the glomerular basement membrane; and to the expression of renal antigens that may react with the autoantibodies present in these patients. The pathogenesis of SLE is reviewed in Chapter 5.7.1.

In most published series of childhood SLE, prognosis is most closely related to the severity of the renal disease. The World Health Organization (WHO) developed a morphological classification of lupus nephritis that was modified by the Pathology Advisory Group for the International Study of Kidney Disease in Children in 1980 (Table 4). Although the latter classification is more complex, it is our impression that it provides more useful information when making patient management decisions.

Unfortunately, neither classification accurately evaluates extraglomerular (tubular, interstitial and vascular) disease. It has been our practice to biopsy any child with clinical or laboratory evidence of renal involvement, because a diagnosis of diffuse proliferative lupus nephritis (DPLN) adversely affects long-term outcome and thus influences our initial patient management; an approach that is supported by the experience of others (Esdaile *et al*. 1994). In the report by Platt *et al*. (Platt *et al*. 1982), 59 per cent (21) of the children with biopsy-proven DPLN were alive without endstage renal failure 10 years after diagnosis compared with an 85 per cent 10-year survival rate for all 70 children with SLE (Fig. 4). It is important to note that the outcome of patients with DPLN did not deviate from that of the entire group for 7 years, emphasizing the need for long-term clinical trials to validate the efficacy of treatment protocols. A more recent paediatric series (McCurdy *et al*. 1992) reported a 62 per cent incidence of renal insufficiency in 24 children who developed DPLN between 1970 and 1983. Hopefully, earlier diagnosis of, and newer therapeutic approaches to, DPLN will improve the long-term outcome.

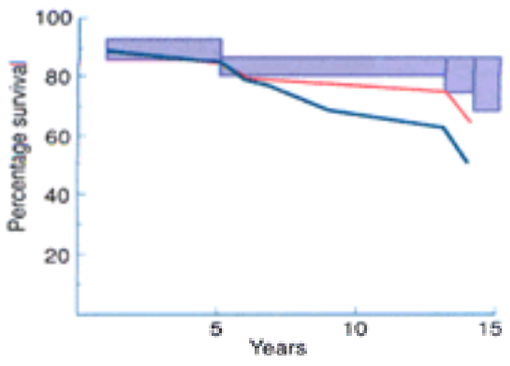


Fig. 4 Long-term survival of 70 children with SLE followed at The University of Minnesota. The patient survival rates for the subset of 21 children with diffuse proliferative lupus nephritis (DPLN) begins to deviate from the entire group (shaded area) at 7 years. Survival of the DPLN group without dialysis or renal transplantation is illustrated by the bolder line (Platt 1982).

I. *Normal*
A. Nil
B. Normal by light microscopy, but deposits present
II. *Pure mesangiopathy*
A. Mild (+)
B. Moderate (++)
III. *Segmental and focal proliferative glomerulonephritis*
A. Active necrotizing
B. Active and sclerosing
C. Sclerosing
IV. *Diffuse proliferative glomerulonephritis*
A. Without segmental necrotizing lesions
B. With segmental necrotizing lesions
C. With segmental active and sclerotic lesions
D. Inactive, sclerotic
V. *Diffuse membranous glomerulonephritis*
A. Pure membranous
B. Associated with lesions II (A or B)
C. Associated with lesions III (A, B, or C)
D. Associated with lesions IV (A, B, C, or D)
VI. *Advanced sclerosing glomerulonephritis*

Pediatr Nephrol (1988)

Table 4 International Study of Kidney Disease in Children: classification of lupus nephritis

Unfortunately, approximately 40 per cent of patients with SLE nephritis have the most severe subtype, DPLN. Since 1979, when electron microscopy was introduced as a routine part of the evaluation of renal biopsies performed at our institution, 39 per cent of 79 children with biopsy-confirmed lupus nephritis had DPLN (Table 5). Three additional patients underwent histological transformation to DPLN from focal proliferative and mesangial lupus nephritis within 9, 14, and 17 months respectively. Although it is relatively easy for the clinician to predict that a patient presenting with SLE and severe clinical manifestations of nephritis will have DPLN confirmed on renal biopsy, we have been impressed that this lesion may be present even in children with mild clinical manifestations. The profile of 29 children at the time of biopsy-confirmed DPLN illustrates this point (Table 6).

	General (%)[a] (n=368)	Seven paediatric series (%)[b] (n=424)	HSC 1979–1995 (%) (n=79)
Mesangial	27	24	19
Focal proliferative	18	24	20
Diffuse proliferative	39	44	39
Membranous	16	8	22

[a]Data obtained from Pollak and Pirani (1993)
[b]Data obtained from Cassidy *et al*. (1977), King *et al*. (1977), Abeles *et al*. (1980), Platt *et al*. (1982), Gladden *et al*. (1983), Yang *et al*. (1994), Cameron (1994).

Table 5 Histological patterns of SLE nephritis

	Mean ± SD	Median	Range	Percentage
Age (y)	13 ± 4	14	4–17	
Female				67
Microhaematuria				100
Proteinuria (mg/kg per day)	46 ± 26	35	3–129	
> 50				38
35–49				10
< 34				52
Serum albumin (g/l)	30 ± 7	30	16–47	
> 35				30
25–34				50
< 25				20
GFR[c] ml/min per 1.73 m²	55 ± 33	102	17–151	
> 100				52
90–99				33
< 50				14
Hypertension				33
C3[c] (g/l)	0.47 ± 0.15	0.49	0.09–0.70	

[a]Diffuse proliferative lupus nephritis diagnosed at our institution since 1979; clinical data available on 29/33 children.
[b]GFR, glomerular filtration rate determined by ^{51}Cr-EDTA or iothalamate clearance.
[c]C3, third component of complement; normal value is 0.81–1.8 g/l in our laboratories.

Table 6 Clinical manifestations at the time of biopsy of children with diffuse proliferative lupus nephritis [a]

Severe proteinuria, even in the nephrotic range, does not always predict the presence of DPLN. Two children in our series with an unequivocal diagnosis of SLE had nephrotic syndrome at presentation and minimal glomerular histological changes consistent with the diagnosis of mesangial lupus nephritis (Fig. 5). Electron microscopy revealed that the foot processes of the glomerular epithelial cells were fused. The nephrotic syndrome responded quickly to high-dose prednisone therapy and both girls were in complete remission within 10 days. We believe that their nephrotic syndrome was more typical of a 'minimal lesion-type disease' similar to previously reported cases (Abuelo 1984; Bakir *et al*. 1989). Since 1979, our incidence of membranous lupus nephritis has been 22 per cent, a frequency that is higher than most published paediatric series (8 per cent of 424 published cases). Whether this is because of an increase in the appreciation of this lesion by ultrastructural studies or due to a changing incidence of epimembranous disease in the past decade is unclear. A previous report of 70 children, who had biopsies taken at this

institution between 1970 and 1985, revealed an 8 per cent incidence of membranous lupus nephritis (Baumal *et al*. 1987). Nephrotic-range proteinuria was present at the time of renal biopsy in 56 per cent of the current series of children with membranous nephropathy, and nephrotic syndrome subsequently developed in another 13 per cent; microscopic haematuria was present in 63 per cent of patients. Importantly, a diagnosis of SLE should be considered in any child with membranous nephropathy, a form of nephropathy that is rare in children and usually not of the idiopathic type.

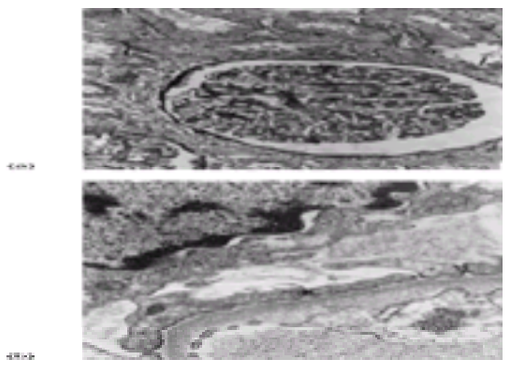


Fig. 5 Renal biopsy of a 14-year-old girl with SLE associated with nephrotic syndrome. (a) A few glomeruli showed a mild increase in mesangial matrix without mesangial cell proliferation. Several immune deposits were demonstrated in mesangial regions by immunofluorescence and electron microscopy. (b) Marked fusion of the foot processes of glomerular epithelial cells (star) was present. The nephrotic syndrome was in complete remission after 10 days of prednisone therapy. These clinical and histological features are typical of a 'minimal lesion-type disease'.

In our experience the presence of focal necrotizing lesions, even in patients with focal proliferative lupus nephritis, is a bad prognostic sign. These lesions were present in 50 per cent of our patients with biopsy-proven focal proliferative disease, all of whom required treatment with a cytotoxic drug (azathioprine, cyclophosphamide or methotrexate) before their disease could be controlled adequately. In a long-term follow-up study of childhood lupus nephritis of all classes, renal insufficiency was reported in 53 per cent of patients who had focal necrotizing glomerular lesions on the initial biopsy (McCurdy *et al*. 1992). Although rare, focal necrotizing glomerulonephritis in the absence of significant glomerular immune deposits (pauci-immune) has been reported in patients with SLE (Akhtar *et al*. 1994).

In addition to providing a histological classification, renal biopsies can be used to evaluate the degree of active and chronic renal damage. Semiquantitative scoring indices have been developed and although some controversy regarding the utility of these scores exists, they are particularly valuable in predicting the long-term outcome in patients with DPLN. A high score for the chronicity index is a bad prognostic sign. Austin and his colleagues from the National Institute of Health (NIH) (Austen *et al*. 1984) reported that patients with DPLN and any evidence of chronic damage on the initial renal biopsy had a much higher incidence of renal failure (Fig. 6). This group has recently reanalysed their experience, following the introduction of intensive treatment regimes that include intravenous pulse cyclophosphamide or methylprednisolone for patients with severe lupus nephritis (Austin *et al*. 1994). By multivariant survival analysis the combination of cellular crescents and interstitial fibrosis was shown to be the strongest histological predictor of a poor long-term prognosis at 80 months (p = 0.0003). In our series of patients followed since 1979, the three patients who have developed endstage renal disease all had advanced chronic disease on the renal biopsy done at the initial presentation and they developed irreversible renal failure within 3 years of diagnosis.

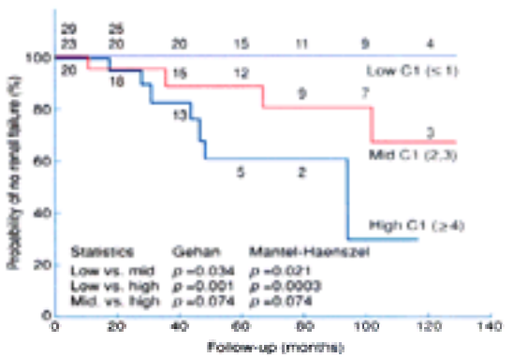


Fig. 6 Cumulative survival curves of 72 lupus patients with diffuse proliferative lupus nephritis (DPLN) demonstrating the negative impact of histological evidence of chronic renal damage on the long-term outcome of renal function. The chronicity index (CI; maximum score of 12 points) adversely affects renal survival rates (reproduced from Austin *et al*. (1984) with permission).

The evaluation of acute tubulointerstitial injury in lupus nephritis has been neglected until recently. Acute interstitial inflammation is commonly observed in patients with DPLN (75 per cent in our series) where it may or may not be associated with tubulointerstitial immune deposits. In our patient population, four patients without DPLN also had significant interstitial inflammation: three with membranous lupus nephritis (Fig. 7) and in one with atypical mesangial lupus nephritis associated with the nephrotic syndrome. Two of the later four patients had foci of immune deposits along tubular basement membranes. A study by Alexopoulos *et al*. (Alexopoulos *et al*. 1990) reported a positive correlation between the number of interstitial mononuclear cells and both renal function and the degree of chronic renal damage. Magil *et al*. reported a poorer outcome in patients, with tubulointerstitial immune deposits present in more than 20 per cent of the renal biopsy specimen (Magil *et al*. 1984). Lupus interstitial nephritis should be considered in the differential diagnosis of older children presenting with renal tubular disorders (Bagga *et al*. 1993). We hope soon to gain a better understanding of the pathogenesis and treatment of the tubulointerstitial disease in lupus nephritis since chronic tubulointerstitial damage predicts a poor long-term prognosis.

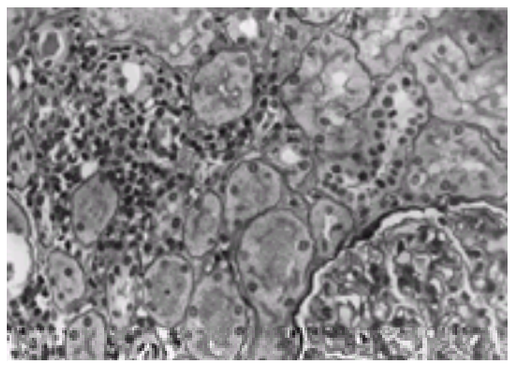


Fig. 7 Light photomicrograph of the renal biopsy of a 14-year-old girl with lupus nephritis illustrating the presence of interstitial inflammation. In this patient immunofluorescence and electron microscopy demonstrated the presence of immune deposits along tubular basement membranes. Interstitial inflammation is a common finding in lupus nephritis and often occurs in the absence of extraglomerular immune complex deposition.

As recently reviewed by Appel *et al*., a variety of renal vascular lesions may be observed in SLE patients, including those associated with vascular immune complex deposits, non-inflammatory necrotizing vasculopathy, thrombotic microangiopathy and true renal vasculitis (Appel *et al*. 1994). Studies of renal vascular lesions in

childhood SLE have not been published. Experience in adults suggests a worse prognosis than in patients without vascular lesions.

Regular evaluation of the urine is important in all SLE patients. Many patients with renal involvement develop urinary abnormalities within 3 years of diagnosis, but a longer lag period can occur (Tucker *et al*. 1995). One of our patients recently developed focal proliferative lupus nephritis 5 years after the initial diagnosis of SLE which was made at 11 years of age. The importance of follow-up is illustrated by the fact that there is now a 72 per cent incidence of renal disease in our patients that were reported in the 1993 edition of this chapter, while only 42 per cent of patients referred to our clinic since 1991 have so far had clinical evidence of nephritis.

Hypertension is a common manifestation in patients with SLE. An elevated blood pressure prior to steroid therapy suggests that the patient has DPLN or renovascular disease. The latter is unusual but has been observed in patients with the antiphospholipid antibodies who have a hypercoaguable state. One patient followed at our institution presented with an infarct to the lower pole of one kidney as the result of a segmental occlusion of the renal vein, in the absence of evidence of glomerulonephritis.

It is rare that patients present with haematuria due to a bleeding diathesis. We treated an 11-year-old girl who had gross haematuria resulting from factor-II deficiency. The bleeding reversed with plasma therapy and factor-II levels were normalized by treatment with steroids (Eberhard *et al*. 1994).

Urological manifestations of SLE are not widely recognized but are clearly documented in the literature. Urinary frequency may occur as a result of autoimmune cystitis. However, it is always important to rule out an infectious cause, particularly in immunosuppressed patients.

Treatment of nephritis

There is probably no topic in medicine more disputed than the management of patients with SLE. Once renal involvement is documented, immunosuppression is indicated and steroids remain the mainstay of therapy. The majority of patients with mesangial and focal proliferative disease are successfully managed with steroids alone, although those with associated focal segmental necrotizing lesions have a poorer prognosis and are likely to require additional treatment including the use of cytotoxic agents. Evaluation of newer treatment protocols for patients with DPLN is impeded by the need to follow adequate numbers of patients for 10 to 15 years in order to observe meaningful differences. Although our patients with DPLN are not managed by a rigid protocol, several principles of therapy have evolved.

1. Initial treatment always includes high-dose steroids (prednisone 2 mg/kg per day; maximum 60 mg/day), divided three times daily for 4 to 8 weeks, consolidated to once daily and then tapered slowly over several months. Most patients are maintained on low-dose prednisone for years in an attempt to minimize the risk of subsequent relapses.
2. Azathioprine (2 to 3 mg/kg per day) is initiated within the first month of diagnosis because we believe that the best information available today suggests that patients given cytotoxic drugs in addition to prednisone have a better long-term outcome. Although preliminary data suggests that cyclophosphamide might be a better cytotoxic drug for these patients (Austin *et al*. 1986), the majority of our patients do well without it. Since the long-term toxicities of cyclophosphamide (especially malignancy and infertility) are correlated with the lifetime accumulated dose, we are reluctant to use this drug initially in these children facing a lifetime of chronic illness. Concern about accelerated atherosclerosis by prolonged steroid use (Rubin *et al*. 1985) further rationalizes our use of azathioprine as a steroid-sparing agent. An interesting alternative that we have not used (Cameron 1994) is a protocol that includes an initial 8- to 12-week course of oral cyclophosphamide (3 mg/kg ideal body weight for height, reduced if renal insufficiency is present) followed by maintenance therapy with azathioprine.
3. Intravenous pulse cyclophosphamide given in a slightly modified version of the NIH protocol is used primarily to treat patients whose renal disease is controlled inadequately by steroids plus azathioprine. To date we have treated seven DPLN patients with intravenous cyclophosphamide. Although cyclophosphamide is an important addition to the therapeutic armamentarium, renal relapses following cessation of therapy clearly occur and cyclophosphamide is not universally effective. We have only seen a sustained remission in one in seven patients treated at our institution. Two patients experienced significant disease reactivation (carotid artery thrombosis and renal failure; recurrence of nephrotic syndrome and deterioration of renal function) after completion of the 6 monthly doses during the 3-month period before the seventh dose was administered. One patient who was also maintained on high-dose prednisone appeared resistant to intravenous cyclophosphamide. After 8 monthly doses she remained hypocomplementaemic, serologically active and had a deterioration in her protein excretion from 0.7 to 2.4 g/24 h. Two additional patients were treated with cyclophosphamide for non-compliance with prednisone and azathioprine. One of these patients had normal renal function after a 2-year course of treatment but in follow-up remained non-compliant. He now has chronic renal failure after 9 years of the disease. The second patient refused to take prednisone during cyclophosphamide therapy. The protocol was discontinued after 7 doses due to non-responsive disease. One year ago this patient consented to treatment with prednisone and azathioprine and is currently in remission. The final patient that we treated had significant chronic renal disease at presentation. Her therapy was discontinued after 4 doses. Although she required dialysis during her initial admission and eventually came off dialysis for 2.5 years, we would no longer recommend cyclophosphamide therapy in this kind of patient (see next point).
4. The chronicity index should be taken into consideration when making decisions about the use of cytotoxic drugs. If advanced chronic disease is present indicating that endstage renal disease is inevitable and severe extrarenal manifestations of SLE are not present, then cytotoxic drugs are not justified. The survival of SLE patients on dialysis and following renal transplantation is very good (Bumgardner 1988; Nossent *et al*. 1991). The risk of recurrent lupus nephritis in renal allografts is very low. Unfortunately the overall incidence of malignancy following solid organ transplantation is approximately 4 per cent (Penn 1988) and we are concerned about increasing this risk, particularly by treating such a patient with cyclophosphamide.
5. High-dose pulse steroids and plasmapheresis may have a role in the management of patients presenting with acute fulminating renal disease. Methotrexate therapy may be useful in selected patients with long-term disease that is difficult to control (Abud-Mendoza *et al*. 1993), but its role has not been established. Methotrexate should be restricted to patients with good renal function to avoid serious drug toxicities.
6. Over the past couple of years we have treated an increasing number of renal flares with intravenous methylprednisolone at a dose of 30 mg/kg (maximum 1000 mg) daily for 3 consecutive days each month for 6 months. Two of the five children recently treated were patients with DPLN unresponsive to intravenous cyclophosphamide. Short-term results have been encouraging but whether a sustained long-term benefit will be achieved is not yet clear.

Close follow-up of these patients is essential in order to monitor disease activity, complications of therapy, the patient's understanding of their disease and compliance with treatment. In the evaluation of the renal response to therapy, follow-up urinalysis and quantification of urinary protein excretion rates are obvious. The presence of either red cell or white cell casts may precede a renal relapse. Persistent proteinuria one year after the initiation of treatment predicts a worse long-term prognosis (Fraenkel *et al*. 1994). We feel that two additional parameters are useful markers of renal disease activity. The first is the serum complement profile, particularly levels for the third component of complement (C3). As demonstrated by Laitman *et al*., patient outcome is improved if CH50 levels are normalized during the initial phase of therapy and if they remain normal during the maintenance phase (Laitman *et al*. 1989). Since SLE relapses tend to be mimetic, nephritis is likely to recur during this period. Although we do not consider serological abnormalities alone (i.e., hypocomplementaemia, increasing anti-DNA antibody levels) an indication to re-treat a patient, clinically active disease frequently ensues and these patients need to be followed closely (Fig. 8).

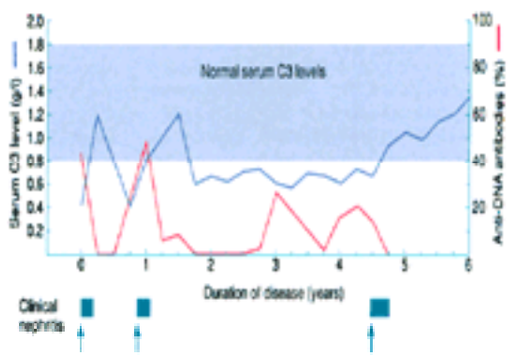


Fig. 8 Schematic summary of the clinical course of a girl with congenital C4 deficiency who developed SLE at 4 years of age associated with diffuse proliferative lupus nephritis. The value of monitoring serum C3 levels (solid line) is highlighted by her clinical course. Initial prednisone therapy (arrow) failed to maintain the C3 level in the normal range. She developed a clinical relapse of nephritis 10 months after the initial presentation, that was treated with prednisone and azathioprine (arrow). One year later the C3 level fell below the normal range. However, she remained clinically well for 3.5 years before findings of recurrent nephritis were documented and high-dose immunosuppression was reinitiated (arrow).

Preservation of renal function is an obvious goal in patients with lupus nephritis. Unfortunately, obtaining an accurate measurement of the glomerular filtration rate is

not easy. Taking inulin clearance as the 'gold standard', Meyers and his colleagues (Shemesh *et al*. 1985) have demonstrated that creatinine clearance determinations overestimate the glomerular filtration rate during the acute phase of lupus nephritis, probably as a result of tubular secretion of creatinine (Table 7). Isotopic tests (e.g. 99Tc-DTPA) appear to provide a more accurate measure of glomerular filtration rate in these patients.

Inulin clearance (ml/min per 1.73m^2)	n	Creatinine /inulin clearance
>80	13	1.20 ± 0.08
40–80	10	1.57 ± 0.11
<40	21	2.21 ± 0.16

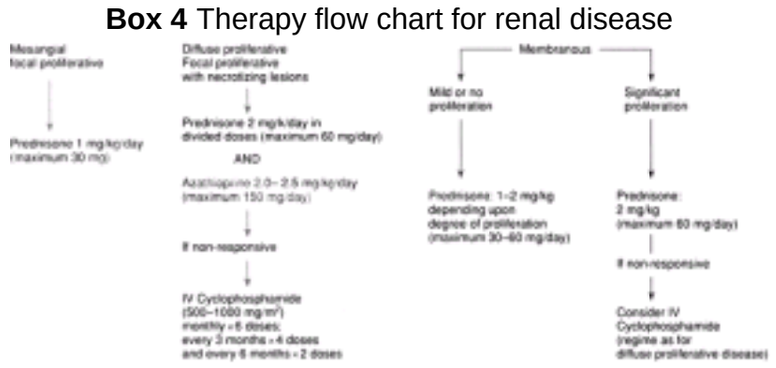
[a]As illustrated by the study of Shemesh *et al*. (1985) the creatinine clearance does not provide an accurate measure of the true glomerular filtration rate (assessed by inulin clearance rates) in patients with lupus nephritis. The more severe the reduction in renal function, the greater is this discrepancy.
C = clearance.

Table 7 Creatinine clearance is an unreliable measure of glomerular filtration rate in lupus nephritis [a]

The optimal treatment of patients with lupus membranous nephropathy is unknown. We agree with Appel *et al*. that the long-term outcome of these patients is less favourable than was suggested originally (Appel *et al*. 1987). A subset of these patients appears destined to develop chronic renal failure and we believe that they should be treated with high-dose steroids. Less clear is the indication for alkylating agents. There is some encouraging preliminary evidence that alkylating agents are beneficial in the treatment of patients with the severe idiopathic form of membranous nephropathy (Imperiale *et al*. 1995) and it is tempting to speculate that alkylating agents may also be useful in patients with membranous nephropathy due to SLE. Unfortunately appropriate clinical trials have not yet addressed these important issues. A recent review by Sloan *et al*. (1996) reported that long-term prognosis in patients with SLE membranous nephritis was determined by the degree of associated glomerular inflammation seen on biopsy. The only patient at our institution with membranous SLE and diffuse endocapillary proliferation developed endstage renal disease.

For an overview of suggested therapy for renal disease see Box 4.

Box 4 Therapy flow chart for renal disease

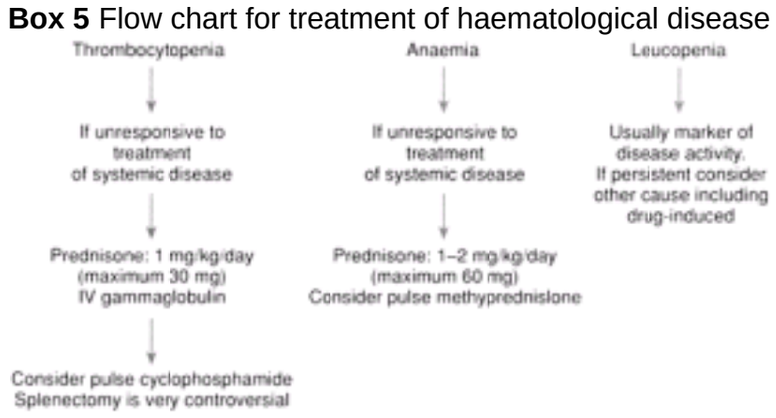


Haematological involvement

Anaemia, thrombocytopenia, and leucopenia are very common laboratory abnormalities seen in 50 to 75 per cent of patients with SLE. The most common anaemia is normochromic normocytic, which when persistent usually becomes a microcytic and hypochromic anaemia. There is both a decrease in the serum iron level and in the iron-binding capacity, with increased iron in macrophages. The serum ferritin is normal or elevated, as ferritin production is increased as part of the acute-phase response. In our experience, in the absence of an immune-mediated haemolytic anaemia (see below), the anaemia is rarely severe.

The Coombs' test is positive in approximately 30 to 40 per cent of our patients but less than 10 per cent of patients have overt haemolysis. Haemolysis is seen generally only in the presence of immunoglobulin on the surface of erythrocytes. The likely explanation for the lack of haemolysis when only complement is present is that red blood cells bind immune complexes through their C3b receptor (Wilson *et al*. 1989). Therefore, the presence of complement components (C3b) on the red blood cell probably represents the presence of immune complexes rather than of autoantibodies directed against erythrocytes. Furthermore, complement activation products, C4d and C3b, may non-specifically coat red blood cells and therefore, can be detected in a Coombs' test. When present, the autoimmune haemolytic anaemia may be a warm or cold. Although rarer than the warm, cold autoimmune haemolytic anaemia is usually associated with Raynaud's phenomenon or cold intolerance and patients more commonly present with haemolysis. Drug-induced haemolytic anaemia in our patients is not a common problem. Haemolysis secondary to autoimmune haemolytic anaemia is usually mild but occasionally it may be severe enough to require treatment with high-dose steroids for prolonged periods of time or even azathioprine or cyclophosphamide. Occasionally transfusions are required, but if the anaemia is secondary to an indirect Coombs' test then the cross-match will be incompatible and type O Rh-negative blood is recommended. For an overview of suggested therapy for anaemia see Box 5.

Box 5 Flow chart for treatment of haematological disease



Thrombocytopenia is present in 15 to 45 per cent and in our experience, it may be the initial presentation in up to 15 per cent of paediatric cases. The thrombocytopenia is secondary to peripheral destruction rather than bone marrow suppression. It can occur early in the disease course and may present as classic idiopathic thrombocytopenic purpura, which would be better referred to as autoimmune thrombocytopenic purpura. We have found that these patients have a positive antinuclear antibody and most patients will have decreased complement levels (uncommon in isolated idiopathic thrombocytopenic purpura). Patients with autoimmune thrombocytopenic purpura may have a transient response to intravenous immunoglobulin treatment. The presentation of isolated idiopathic thrombocytopenic purpura progressing to SLE is more common in children than adults. The development of SLE may take 20 years or more. Most patients with idiopathic thrombocytopenic purpura and autoimmune haemolytic anaemia, so-called Evans syndrome, will probably develop SLE. Interestingly, many patients with thrombocytopenia secondary to SLE will respond to therapy with intravenous immunoglobulin.

Despite the high incidence of thrombocytopenia in paediatric SLE, bleeding is unusual and usually occurs only when the platelet count is less than 10 000. Most patients with thrombocytopenia will respond to steroid therapy. If the patient becomes either steroid-resistant or steroid-dependent then pulse steroid therapy or intravenous immunoglobulin therapy should be tried, with cyclophosphamide reserved for resistant cases (Lipnick *et al*. 1990). Splenectomy is very controversial as some reports have suggested not only a lack of effect but also an unacceptably high increase in infections of these patients postsplenectomy. These findings are not universal and splenectomy may have a role in resistant, persistent life-threatening thrombocytopenia. For an overview of suggested therapy for thrombocytopenia see Box 5.

In addition to thrombocytopenia, acquired abnormalities of platelet function have been described. These include serum inhibitors that decrease aggregation and block

uptake and storage of adenosine diphosphate and serotonin. Collagen-induced aggregation may be absent and adenosine diphosphate and adrenaline-induced aggregation impaired (Decker *et al*. 1979). These defects generally present as purpura while overt bleeding is rare.

Leucopenia is seen in 20 to 40 per cent of cases of paediatric SLE. Both lymphopenia and granulocytopenia can be seen. The lymphopenia may be secondary to the presence of circulating lymphocytotoxic antibodies which may be cold- or warm-reacting; IgM or IgG; and directed against resting or activated cells (Peake *et al*. 1988). In many patients lymphopenia is a sensitive marker of general disease activity and will rarely if ever require specific therapy. A lymphopenia may also be secondary to therapy with azathioprine or cyclophosphamide. Granulocytopenia is usually secondary to a central depression of granulopoiesis or splenic sequestration and more rarely antigranulocyte antibodies. Drugs including prednisone, azathioprine, and cyclophosphamide decrease granulocyte function and/or numbers. All of these problems probably contribute to the increased susceptibility to infection seen in SLE patients.

Other unusual haematological problems in SLE include myelofibrosis, and thrombotic thrombocytopenic purpura. Myelofibrosis, although rarely seen, is more common in the paediatric age group while there have been reports of thrombotic thrombocytopenic purpura progressing to SLE. Thrombotic thrombocytopenic purpura, as opposed to haemolytic-uraemic syndrome, is rarely seen in childhood. When it does occur, the diagnosis of SLE should be considered. An interesting association is the development of SLE in adolescents with sickle-cell anaemia; the SLE usually develops prior to the age of 18 (Katsanis *et al*. 1987).

Following thrombocytopenia, the presence of the lupus anticoagulant is the most common coagulation defect. This abnormality presents with an elevated partial thromboplastin time and occurs in 20 to 30 per cent of paediatric cases. The prothrombin time is usually normal or minimally elevated and if markedly prolonged a second defect is probably present (see below). The lupus anticoagulant reacts with the phospholipid portion of the prothrombin activator complex. This antibody cross-reacts with anticardiolipin antibodies, which are seen commonly in paediatric SLE (Shergy *et al*. 1988), and may be responsible for the false-positive VDRL (Veneral Disease Research Laboratory). Patients with the lupus anticoagulant do not bleed but rather have an increased incidence of deep vein thrombosis, thromboemboli, and less commonly arterial thrombosis. The increased partial thromboplastin time will usually resolve with steroids; however, treatment should not be directed to this laboratory finding alone, but reserved for patients who have clotting problems. When a venous or arterial thrombosis occurs, then patients should be subjected to anticoagulation with heparin followed by warfarin. We recommend that these patients receive full anticoagulant treatment for 3 to 6 months followed by long-term low-dose anticoagulation. We suggest that the INR should be maintained between 1.5 and 2.0. A recent adult study supports the view of long-term anticoagulation although the suggested INR range is higher (Khamashta *et al*. 1995). In many patients the presence of the anticoagulant correlates with disease activity, especially with vasculitis. The anticoagulant may be IgG, IgM or both. The full spectrum of diseases associated with antiphospholipid antibodies and the lupus anticoagulant is fully discussed in Chapter 5.7.1 and Chapter 5.7.3. However, it must be remembered that patients with antiphospholipid antibodies may not only present with venous or arterial events, but chorea may be the only manifestation of these autoantibodies. These patients with chorea are at risk for thrombosis or other manifestations of the antiphospholipid antibody syndrome.

Specific inhibitors of other factors of the coagulation cascade have been described. The most common of these is prothrombin deficiency which is associated usually with the lupus anticoagulant. These patients, unlike those with the anticoagulant only, present with bleeding and a prolonged prothrombin and partial thromboblastin time. This defect is reported in approximately 5 per cent of adult patients with the lupus anticoagulant or in less than 1 per cent of all adult SLE patients. Prothrombin deficiency appears to be more common in paediatric SLE and a review of our patients demonstrated this abnormality in approximately 4 per cent of our SLE patients (Eberhard *et al*. 1994). As expected these patients presented with bleeding rather than clotting. Many patients with prothrombin deficiency have an associated mild thrombocytopenia, but the platelet counts are generally greater than 50 000. The aetiology is unknown and it may result from an acquired production defect or secondary to the presence of an antiprothrombin antibody. We have found that the prothrombin deficiency will rapidly respond to steroid therapy. More rarely a steroid-responsive acquired von Willebrand's defect has been described.

Splenomegaly is quite common occurring in 20 to 30 per cent of paediatric cases. In our series of patients splenomegaly was seen commonly in patients less than 10 years of age and may be the result of the generalized inflammatory state. Functional asplenia has been described and this abnormality of splenic function may increase the incidence of sepsis (Malleson 1989).

The final haematological complication we will discuss is the association of SLE and cancer. Despite the frequent use of immunosuppressive therapy and abnormal immunoregulation, malignancy as a cause of death is rare. There may be a higher incidence of neoplasia than in the general population, but most tumours are of epithelial origin and are easily treated. There have been case reports of lymphoma, including Burkitt's lymphoma, in patients with SLE; from the other perspective, patients with lymphoma may have many features suggestive of SLE, including the presence of a high titre of antinuclear antibodies (Posner *et al*. 1990). We have seen cases of lymphomas presenting as 'SLE' including a positive, high-titre, speckled antinuclear antibody. Therefore, the presence of a high-titred antinuclear antibody does not help in differentiating these two disorders.

Cardiac involvement

Symptomatic pericarditis is the most common cardiac manifestation occurring in approximately 5 to 25 per cent of patients with SLE and is commonly associated with pleurisy (De Inocencio and Lovell 1994). The diagnosis is made by physical examination, chest radiography, electrocardiogram and confirmed by echocardiogram. Cardiac tamponade rarely occurs. Studies in adults have suggested that echocardiographic evidence of pericarditis may occur in up to 75 per cent of patients (Doherty and Siegel 1985). Similar studies have not been performed in children. Most cases of pericarditis will rapidly respond to either non-steroidal anti-inflammatory drugs alone, or low to moderate dose of corticosteroids. We suggest the initial use of indomethacin at a dose of 3 mg/kg per day divided into three doses. Indomethacin appears to work well in patients with SLE and is well tolerated in the paediatric patient. Rarely pulse intravenous methylprednisolone is required and on occasion pericardiocentesis is necessary if tamponade is impending. Antimalarials may be of benefit in the long-term management of pericarditis.

The diagnosis of myocarditis or endocarditis is uncommon, with clinically detectable or significant myocarditis in less than10 per cent of patients (Badui *et al*. 1985). However, autopsy studies in adults have shown evidence of myocarditis in 25 to 50 per cent of patients. There is no similar autopsy data available in children. Occasionally, patients with myocarditis may present with first degree atrioventricular block or arrhythmia. A study from 1985 showed that Libman–Sacks endocarditis as present in 30 to 50 per cent of all hearts from patients with SLE at autopsy, while echocardiographic evidence of small vegetations may occur in 2 to 5 per cent of cases (Doherty and Siegel 1985). These vegetations commonly occur at valvular rings and commissures. However, clinically significant lesions causing aortic or mitral stenosis or regurgitation are very rare.

Unlike Libman–Sacks endocarditis which appears to be declining with steroid usage, atherosclerotic heart disease and myocardial infarction are increasing with greater steroid usage and longevity of patients (Rubin *et al*. 1985). Other risk factors, especially hypertension and hyperlipidaemia further increase the risk of myocardial infarction. The presence of either the lupus anticoagulant or antiphospholipid antibodies may further predispose to thrombosis and myocardial infarction. Hyperlipidaemia can either be the result of a primary hyperlipidaemia of SLE or secondary to treatment with steroids, which alters lipid profiles including cholesterol, triglycerides and both high-density and low-density lipoprotein levels (Ilowite *et al*. 1988). Although myocardial infarction during childhood or adolescence is rare, the incidence of myocardial infarction in young adults (under 25 years of age) with initial onset of SLE in the paediatric age is increasing. When myocardial infarction or chest pain occurs in young patients or shortly following the diagnosis of SLE, the presence of coronary arteritis must be suspected (Friedman *et al*. 1990). Although it occurs in less than1 per cent of cases, aggressive treatment of this potentially life-threatening complication with high-dose steroids and immunosuppressive agents is required. We have found recently that many paediatric patients may have asymptomatic abnormalities of cardiac perfusion, which can be demonstrated by nuclear cardiac imaging. The clinical significance of these abnormalities is unknown but they may predispose these patients to early ischaemic heart disease.

Congenital heart block and neonatal lupus erythematosus will be dealt with in another section.

Pulmonary involvement

Pulmonary involvement is common in paediatric SLE occurring in 25 to 75 per cent of cases. The manifestations range from severe life-threatening pulmonary haemorrhage or infection to asymptomatic abnormalities of pulmonary function tests (see Table 8 for a complete list of pulmonary complications). Decreased carbon monoxide diffusing capacity is the most common abnormality in both paediatric and adults patients; in one paediatric series it was seen in 100 per cent of unselected patients (Delgado *et al*. 1990). The next most common abnormality in the pulmonary function test is a restrictive pattern in 35 to 60 per cent of cases, while an obstructive defect is uncommon. The functional defects are probably secondary to chronic fibrotic changes that in turn are secondary to mild subclinical lupus pneumonitis, as suggested by the diffusing capacity abnormalities and restrictive rather than obstructive defects. These abnormalities rarely require specific therapy.

Pleuritis
Diaphragm involvement including shrinking lungs
Pneumonitis (acute or chronic)
Vasculitis
Pulmonary haemorrhage
Pulmonary emboli
Isolated pulmonary function test abnormalities
Drug-induced changes
Pulmonary hypertension

Table 8 Pulmonary involvement in childhood SLE

Pleural involvement occurs in up to 50 per cent of cases at some time during the disease course and is seen commonly in association with pericarditis. However, isolated pleuritis, presenting as chest pain, can occur during a systemic flare and may be unilateral. Pleural fluid examination reveals decreased levels of C3 and C4, an increased white blood-cell count, normal glucose and can show a positive antinuclear antibody. There have been reports of a positive antinuclear antibody only in the pleural fluid but not the peripheral blood. When the pleuritis is mild, treatment can consist of anti-inflammatory doses of non-steroidals, but usually prednisone, at a low to moderate dose, is required for complete resolution of symptoms. Similar to patients with pericarditis, antimalarials may be of benefit in the long-term management of pleuritis.

Dyspnoea, a common symptom, can be secondary to acute or chronic pneumonitis, pulmonary infection, pulmonary haemorrhage or shrinking lung syndrome. The dramatic presentation of acute pneumonitis with fever, cough, dyspnoea, hypoxia and chest pain occurs in only 3 to10 per cent of patients (Carette *et al*. 1984). Pulmonary haemorrhage has a reported death rate of up to 50 per cent and in many paediatric studies this complication accounts for 10 to 20 per cent of deaths despite its occurrence in less than 5 per cent of patients (Nadorra and Landing 1987). Pulmonary haemorrhage or acute pneumonitis must be differentiated from congestive heart failure, pulmonary infection, and non-cardiogenic pulmonary oedema that are the side-effects of drugs and secondary to uraemia, pancreatitis, or sepsis. The clinical differentiation of acute SLE pneumonitis from pulmonary haemorrhage may be aided by a fall in haemoglobin, the presence of haemoptysis and an increased carbon monoxide diffusing capacity on pulmonary function tests seen in pulmonary haemorrhage. The treatment for both is similar and the acute management may require intubation and ventilation in addition to aggressive immunosuppressive therapy, which may include high-dose prednisone and cyclophosphamide. Pulmonary haemorrhage may be the initial and sole manifestation of paediatric SLE and this disease should be considered in young patients presenting with pulmonary haemorrhage even in the absence of other disease manifestations.

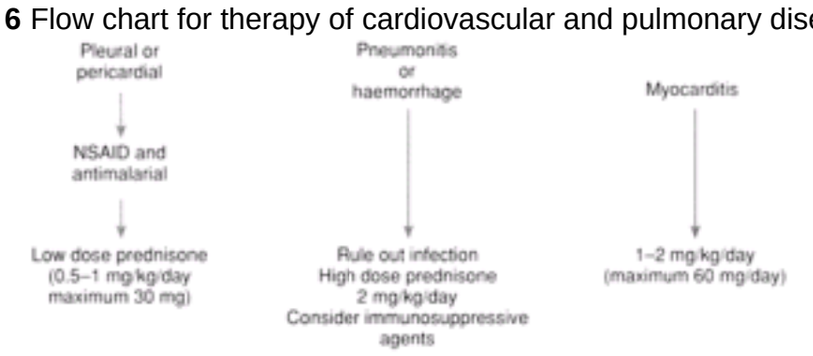
In the acutely ill patient the possibility of pulmonary infection must be considered. There is an increased incidence of infection even prior to steroid or immunosuppressive treatment. In addition to the more common bacterial and viral causes of pneumonia, these patients are at risk for opportunistic infection including fungal, parasitic and protozoan infection. We have seen a patient on prednisone and azathioprine who developed *Pneumocystis carinii* pneumonia and others have described patients with this complication while on methotrexate and prednisone. When patients present with acute respiratory failure and fever, we recommend treatment with broad-spectrum antibiotics and high-dose steroids including pulse therapy. If the above-mentioned opportunistic relevant infections are suspected clinically, then antifungal, antiprotozoan, antilegionella, antipneumocystis and even antinocardia therapy should be considered. If the patient does not improve quickly then a diagnostic open lung biopsy is necessary. The use of bronchial washings obtained via bronchoscopy should be considered early, especially if intubation is required.

In contrast to acute pneumonitis, chronic or interstitial pneumonitis has an insidious onset. Clinically significant disease is uncommon and usually follows long-standing SLE. When severe, most patients have clinically evident pulmonary involvement and present with cough and chest pain leading to dyspnoea and a decreased diffusing capacity. An autopsy study of 26 paediatric patients with SLE showed a mild chronic interstitial pneumonitis in all patients, with additional pulmonary lesions in 18 patients (Nadorra and Landing 1987). Most of the chronic lesions were asymptomatic prior to death.

A more unusual cause of dyspnoea is the shrinking lung syndrome. This disorder is diagnosed by chest radiography that demonstrates an elevated diaphragm in the clinical setting of dyspnoea, which might be mild. A paediatric review demonstrated radiographic evidence of an elevated diaphragm in 12 per cent of patients (Delgado *et al*. 1990). They hypothesized that this syndrome was caused by a combination of inspiratory muscle dysfunction, recurrent pleuritis and atelectasis. The long-time outcome is generally good despite the lack of effective therapy, as it is usually a slowly progressive disease. Pulmonary function tests show abnormal diaphragm function and restrictive airway disease.

One of the most fascinating pulmonary manifestations of SLE is pulmonary hypertension (Asherson and Oakley 1986). Difficult to treat, pulmonary hypertension fortunately is rare and is seen usually only with long-standing disease. These patients present with classic signs and symptoms of primary pulmonary hypertension but it may evolve from one of five processes. The first is best regarded as primary pulmonary hypertension of SLE. There is no known cause or predisposing clinical setting. The second cause is secondary to pulmonary artery thrombosis or emboli and is seen in patients with antiphospholipid antibodies and/or with severe Raynaud's phenomenon. The third cause is secondary to chronic left ventricular failure. The fourth is secondary to chronic hypoxic from diffuse interstitial lung disease, with the hypertension usually out of proportion to the parenchymal disease. The final cause is obstruction of peripheral pulmonary vasculature secondary to vasculitis. Overall, the outcome of pulmonary hypertension is poor but luckily it is a rare manifestation, especially in children. An overview of suggested therapy for cardiovascular and pulmonary disease is given in Box 6.

Box 6 Flow chart for therapy of cardiovascular and pulmonary disease



Gastrointestinal disease

Gastrointestinal involvement occurs in 20 to 40 per cent of paediatric patients. The most common complaint is abdominal pain that can be the result of peritoneal inflammation (serositis), vasculitis, pancreatitis and/or direct bowel wall involvement (enteritis). In many patients, the abdominal pain heralds a disease flare in another system. The presentation of serositis may vary from mild crampy or colicky pain to severe pain with a rigid abdomen. It can be difficult to differentiate peritoneal inflammation as a manifestation of the underlying SLE from infective peritonitis.

Bowel wall inflammation presenting as crampy, abdominal pain and diarrhoea can be caused by a primary enteritis or be secondary to a mesenteric vasculitis or thrombosis (Hoffman and Katz 1980). The latter two are generally more acute and may be accompanied by bloody diarrhoea and nausea or vomiting. These disorders must be differentiated because vasculitis with thrombosis may require urgent surgery owing to the risk of bowel ischaemia, while most cases of non-thrombotic enteritis will respond to steroids. Involvement of other organ systems may aid in the diagnosis as there is an association of severe enteritis with central nervous system disease and lupus cystitis (Orth *et al*. 1983). This disorder requires aggressive immunosuppressive therapy and although this is an unusual association, paediatric cases have been described. We have seen a patient present with cystitis who rapidly developed severe enteritis with bowel perforation, disease of the central nervous system and nephritis. Despite aggressive immunosuppressive therapy, the patient died (Eberhard *et al*. 1991a).

An autopsy study of 26 young patients demonstrated that abdominal pain occurred in 65 per cent of all paediatric patients and gastrointestinal bleeding in 46 per cent. Dysphagia and symptoms of reflux were unusual. In 15 out of 25 autopsies a gastrointestinal 'vasculopathy' was found with ischaemic lesions, and a non-specific

infiltrate was found in 96 per cent. In 20 per cent of the cases, autopsy findings of ischaemic bowel were clinically asymptomatic (Nadorra *et al*. 1987). These authors suggested that chronic lupus enteritis, as defined by the presence of a non-specific infiltrate, was common at autopsy. However, these patients probably had severe SLE, as they died before reaching adulthood.

Pancreatitis, although it must always be considered, is a rare cause of abdominal pain in paediatric patients. There are very few reports in the literature, with an overall incidence of less than 2 per cent. In order to determine the true incidence of pancreatic dysfunction in SLE, we prospectively evaluated pancreatic function and performed pancreatic ultrasonography in 36 patients. We did not find any significant pancreatic pathology over the 2 year study period and could not find any association of pancreatic function abnormalities with steroid or azathioprine use (Eberhard *et al*. 1992). However, others have reported pancreatitis associated with disease flares, some of which responded to increased steroid dose (Nadorra *et al*. 1987). Rarely a pancreatic pseudocyst may be present at diagnosis.

Liver disease

Hepatomegaly occurs in 40 to 50 per cent of paediatric patients, while abnormalities of liver function tests may occur in up to 25 per cent. However, the elevation of liver enzymes is usually mild and transient. When there is marked elevation of liver function tests another cause must be sought (Miller 1977). The clinical features of autoimmune chronic active hepatitis, so-called 'lupoid hepatitis', may mimic SLE. Both these diseases preferentially occur in adolescent females. Many patients with the autoimmune form of chronic active hepatitis have a malar rash, arthritis and autoantibodies including anti-DNA antibodies, but they rarely have renal involvement (Hall *et al*. 1985; Hall *et al*. 1986). The differentiation between SLE and autoimmune hepatitis may require a liver biopsy. When jaundice is a prominent feature in a known SLE patient then a second disease such as obstruction, haemolysis or viral hepatitis is probably the cause.

Liver involvement may be part of the generalized vasculitis and may explain the increased salicylate sensitivity seen in 10 to 20 per cent of SLE patients. However, salicylate toxicity rarely results in significant, permanent pathological changes. A similar percentage of patients with systemic juvenile arthritis, another disease with systemic vasculitis, develop salicylate sensitivity. An increased incidence of hepatotoxicity to other drugs including azathioprine, and even cyclophosphamide, has been reported. Rarely systemic vasculitis has resulted in hepatic rupture secondary to liver infarction. These patients present with an acute abdomen and usually have evidence of active vasculitis and Raynaud's phenomenon elsewhere.

Endocrine involvement

Thyroid involvement is the most common endocrine organ involved in SLE. In a prospective study we found that antithyroid antibodies are present in 45 per cent of paediatric patients and clinical hypothyroidism was present in 15 per cent (Eberhard *et al*. 1991b). The development of hypothyroidism usually preceded or was coincidental with the development of SLE. The association of these two diseases may reflect similar genetic susceptibility or a common inciting agent. Hyperthyroidism, although less common than hypothyroidism, occurs with an increased incidence in SLE patients than the general paediatric population. Steroid-induced diabetes mellitus occurs in up to 10 per cent of patients but a lower percentage require insulin treatment. There does not appear to be an increased incidence of diabetes mellitus in the absence of steroid therapy. Delayed puberty and menstrual abnormalities are seen commonly. However, these abnormalities are probably secondary to chronic illness and active disease, rather than as a direct result of the disease process. Rarely hypoparathyroidism has been reported in association with SLE. This diagnosis should be considered in the presence of unexplained hypocalcaemia.

Autoantibodies

The hallmark of SLE is the production of autoantibodies. There is a long list of antibodies directed against histone, non-histone, RNA-binding, cytoplasmic and nuclear proteins. Many articles have reviewed the structure and function of the autoantigens and their role in autoimmune disease. A good review of autoantibodies is found in Chapter 5.7.3 and therefore, in this section we will focus on the limited literature regarding specific autoantibodies and disease manifestation in paediatric SLE.

The most common autoantibody is antinuclear. Depending on the series and method of detection, a positive antinuclear antibody is seen in 85 to 100 per cent of paediatric patients. As in adults, adolescent patients negative for this antibody have been described. The incidence of true antinuclear antibody-negative SLE has decreased with the use of human cell line substrates and procedures that eliminate the loss of some of the extractable nuclear antigens. Historically, anti-DNA antibodies were present in 78 to 87 per cent of cases depending on the series. The older series tend to have a higher incidence than the more recent series. This probably reflects the increase in awareness of paediatric SLE without anti-DNA antibodies and we have found anti-DNA antibodies in approximately 70 per cent of our cases. Studies in adults have suggested that patients with renal disease have higher avidity antibodies than patients with central nervous system disease, in whom anti-DNA antibodies are less common. Similar studies have not been undertaken in children. A paediatric study demonstrated that a lower percentage of paediatric SLE patients had a lower percentage of patients with IgG anti-DNA antibodies than adults, while conversely IgM anti-DNA antibodies were present in a higher percentage of paediatric patients. The anti-DNA antibodies measured were directed against single-stranded DNA (Shergy *et al*. 1989). Similarly there was an increased percentage of paediatric patients with IgM anti-Sm, anticardiolipin and anti-70 kDa ribonucleoprotein antibodies when compared with adult SLE patients (Ward *et al*. 1990). The authors suggested that this was the result of a time-related maturation of the immune response. Therefore, evidence for not only IgG antibodies but also IgM antibodies should be sought in paediatric patients. In enzyme-linked immunosorbent assays (ELISA), the overall incidence of anti-Sm antibodies was 58 per cent, anticardiolipin antibodies were present in 50 per cent of patients, and anti-70 kDa ribonucleoprotein antibodies were present in more than 90 per cent of patients.

Rheumatoid factor has been seen in 12 to 29 per cent of patients by conventional methods. To date there has been no study of the incidence of rheumatoid factor as detected by ELISA. Anticardiolipin antibodies may be present in up to 30 per cent of patients and the lupus anticoagulant is present in approximately 20 per cent of cases. No studies have examined the incidence of anti-Ro and/or anti-La antibodies in paediatric patients, but similar to adults, the presence of anti-Ro and anti-La antibodies defines a population which is at risk of a photosensitive rash and subacute cutaneous lupus.

As previously stated in the haematology section, the presence of anticardiolipin antibodies, the lupus anticoagulant and antiphospholipid antibodies is associated with an increased risk of thrombosis. These autoantibodies are common in paediatric patients and these patients are at increased risk for thrombosis (Shergy *et al*. 1988; Molta *et al*. 1993; Ravelli *et al*. 1994). In our experience, thrombosis may occur in up to 50 per cent of patients with detectable antiphospholipid antibodies. The thrombosis may be arterial or venous. Furthermore, these patients are at risk for the development of other manifestations of the antiphospholipid antibody syndrome including chorea, avascular necrosis, epilepsy, migraine headache, and livedo reticularis (Ravelli *et al*. 1994). There have been recent reports of neonatal thrombosis as the result of the transplacental passage of maternal antiphospholipid antibodies (see section on neonatal lupus erythematosus).

Antiribosomal P antibodies have been addressed in more detail in the section on central nervous system disease. These antibodies appear to be present in approximately 15 per cent of all patients with SLE and there is an increased incidence of antiribosomal P antibodies in patients with psychosis. In patients with psychosis titres of antiribosomal P antibodies may vary with disease activity in the central nervous system. These autoantibodies may be either IgM or IgG.

Similar to the findings seen in spouses of adults with SLE, autoantibodies have been found to be increased in the sera of relatives of paediatric SLE patients. More recently, the presence and titre of anti-Ro antibodies in the asymptomatic mothers of children with SLE correlated with the onset of SLE prior to age 10 and with the male sex (Lehman *et al*. 1989a). The reasons for these observations are not clear.

Neonatal lupus erythematosus

The neonatal lupus erythematosus syndrome (NLE) is a disease of the newborn defined by the demonstration of maternal autoantibodies and characteristic clinical features in the neonatal period. NLE is assumed to be the result of fetal and/or neonatal damage caused by the transplacental passage of maternal IgG autoantibodies. The major clinical manifestations are cardiac and dermatological, with complete congenital heart block being the most significant lesion (Reed *et al*. 1983; Watson *et al*. 1984; Lee *et al*. 1986). More rarely haemolytic anaemia, thrombocytopenia, urinary abnormalities and liver dysfunction are seen (Watson *et al*. 1984; McCune *et al*. 1987). Most early reports suggested that children with NLE are born to mothers with anti-Ro antibodies (Reed *et al*. 1983; Watson *et al*. 1984; Lee *et al*. 1986). However, in larger studies, we and others have demonstrated that the presence of anti-La antibodies in association with anti-Ro antibodies is a more specific disease maker, especially in patients with congenital heart block, and that the presence of anti-La antibodies was a greater risk factor for the development of congenital heart block than the presence of anti-Ro antibodies alone (Buyon and Winchester 1990; Silverman *et al*. 1991). In addition, antibodies directed against 52 kDa Ro were more specific than antibodies directed against the 60 kDa form in determining the risk of development of NLE, although other studies have suggested the importance of antinative 60 kDa Ro antibodies (Buyon *et al*. 1993; Lee *et al*. 1994; Reichlin *et al*. 1994).

The clinical spectrum

Cardiac lesions

The characteristic cardiac lesion of NLE is isolated congenital heart block although there have been case reports of congenital heart block in association with endomyocardial fibroelastosis, valvular insufficiency and patent ductus arteriosus. Ho *et al*. described the histopathology of eight hearts with complete congenital heart block, seven of which were associated with maternal anti-Ro antibodies (Ho *et al*. 1986). There was no mention of the maternal anti-La antibody status. These seven hearts all showed identical abnormalities in the conducting system, namely, a lack of contact between atrial myocardial tissue and the more distal part of the atrioventricular conduction (Lev *et al*. 1971).

In some cases of congenital heart block, in addition to the above mentioned changes, ventricular endomyocardial fibrosis and small inflammatory infiltrates have been described (Hogg 1957). The histopathology of the lesion suggests either early interference with normal organogenesis, or intrauterine inflammatory lesions resulting in subsequent scarring (Carter *et al*. 1974). We believe the latter hypothesis to be true for the following reasons: (i) bradycardia is usually a later intrauterine event corresponding to the time when maternal autoantibodies begin to cross the placenta; (ii) immunofluorescent studies have shown deposition of IgG, complement components and fibrin indicating at least an initial inflammatory event; and (iii) case reports of mothers with SLE which describe the sudden onset of fetal bradycardia and congestive heart failure at 23 to 24 weeks and the reversal of intrauterine congestive heart failure and myocarditis but not the heart block with treatment of the mother with dexamethasone, with or without plasmapheresis. Taken together these data suggests there is initially an inflammatory lesion in the fetal heart. The inflammatory lesion hypothesis is supported by the pathological demonstration of fibrosis at the chordae tendinae and the ventricular septum (Smith and Ho 1994; Weber and Myers 1994).

Investigators have attempted to use animal models to determine the cause of congenital heart block. Anti-Ro and/or anti-La antibodies may alter membrane action potential of rabbit conduction tissue and these abnormalities may resemble the cardiac conduction disorders seen in NLE (Alexander *et al*. 1992; Garcia *et al*. 1994). However, these conduction changes were not specific for anti-Ro/La sera derived from mothers of children with NLE, and sera obtained from mothers who had delivered unaffected children had similar effects on the rabbit conducting tissue (Garcia *et al*. 1994).

Congenital heart block was described originally as a relatively benign condition. However, studies of large series of patients with this condition are not as optimistic. Of our 60 patients with congenital heart block, approximately 50 per cent have required pacemaker therapy. The pacemaker may not be necessary neonatally but rather in later life. We have had neonatal deaths and later deaths as a result of pacemaker failure. A follow-up report of 14 children with congenital heart block showed that 3 of the children (21 per cent) died neonatally of congestive heart failure secondary to the block and a further 5 required pacemakers (McCune *et al*. 1987). Therefore, although congenital heart block has been felt to be rather benign, most series have demonstrated both intrauterine and neonatal deaths with the potential for further deaths to occur secondary to pacemaker failure.

Skin lesions

The skin lesions of NLE are very similar to those of subacute cutaneous lupus erythematosus in which at least 80 per cent of patients have anti-Ro antibodies (Deng *et al*. 1984). Both Ro and La antigens have been documented in human skin (Harmon *et al*. 1984; Lee *et al*. 1985), and exposure to ultraviolet light increases the expression of Ro on the surface of keratinocytes (LeFeber *et al*. 1984). This feature may explain the photosensitive nature of the skin rash in both NLE and subacute cutaneous lupus erythematosus (Fig. 9). Maternally transmitted factors other than anti-Ro and anti-La antibodies may play a role in the development of cutaneous NLE. Although cutaneous NLE is almost universally associated with anti-Ro and/or anti-La antibodies, there have been reports of the condition with anti-U1 ribonucleoprotein antibodies, instead of anti-Ro or anti-La antibodies. In our prospective study, we demonstrated maternal anti-Ro antibodies in 100 per cent of cutaneous NLE patients, while anti-La antibodies were present in 75 per cent of sera.

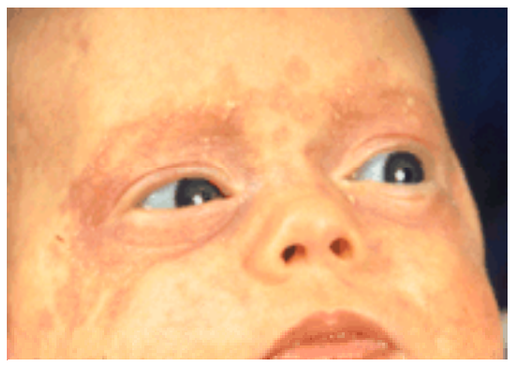


Fig. 9 The face of a 3-month-old baby delivered to a healthy mother with anti-Ro and anti-La antibodies. Note the scaly erythematosus areas around the eyes. This rash healed without scarring following the use of topical steroid cream.

Liver disease

Enlargement of the liver, spleen or both occurs in approximately 30 per cent of cases of NLE. The enlargement of these organs may be secondary to congestive heart failure, but there may also be primary hepatic enlargement. In most cases, the liver disease is associated with other clinical manifestations of NLE. These patients presented with liver enlargement, elevated levels of hepatic enzymes and evidence of cholestasis (Laxer *et al*. 1990). The histological changes seen on biopsy include giant cell transformation, ductal obstruction and extramedullary haematopoiesis. The liver changes generally are not severe and recovery from the liver disease usually occurs despite the presence of residual fibrosis on repeat biopsy. However, we and others have seen cases of severe intrahepatic cholestasis secondary to NLE (Lee *et al*. 1993; Rosh *et al*. 1993).

Haematological problems occasionally occur including thrombocytopenia, rarely leading to bleeding, and a mild haemolytic anaemia. Urinary abnormalities include transient pyuria or urethritis, but to date there have been no cases of nephritis.

Treatment

Cutaneous NLE rarely requires treatment as usually it is self-limited and usually heals without scarring. In the unusual case with more severe or scarring lesions, topical steroids may be required. The liver disease of NLE will usually spontaneously resolve by the age of six months and does not require steroid treatment. However, if the cholestasis is severe, these infants may require a formula high in medium-chain triglycerides.

The treatment or prevention of congenital heart block is much more complex than the treatment of cutaneous NLE. Previously, it had been suggested that all mothers at risk for delivering a child with congenital heart block should be treated with plasmapheresis and dexamethasone throughout the pregnancy (Buyon *et al*. 1988). However, this form of treatment will subject both the mother and the fetus to significant risks during the pregnancy, and up to 90 per cent of mothers will deliver a normal child without congenital heart block, despite having previously delivered a child with this condition (McCune *et al*. 1987; Buyon *et al*. 1988). Therefore, the current recommendation is to monitor the fetus using fetal echocardiography to assess the developing heart, prior to the initiation of potentially toxic treatment. When congenital heart block and fetal hydrops is discovered during gestation, treatment of the mother with dexamethasone and plasmapheresis may reverse the fetal congestive heart failure but not the heart block. However, the efficacy of this therapy in altering the natural history of congenital heart block is controversial and unproven. Children diagnosed with this condition should be delivered in a high-risk neonatal centre which can provide cardiac pacing. If the fetal bradycardia is recognized as congenital heart block rather than fetal distress, then unnecessary early, emergency delivery may be prevented. However, further prospective collaborative studies are required to determine the population at risk for this block and how to prevent it.

Chapter References

Abeles, M., Urman, J.D., and Rothfield, N.F. (1978). Aseptic necrosis of bone in systemic lupus erythematosus. *Archives of Internal Medicine*, **138**, 750–54.

Abeles, M., Urman, J.D., Weinstein, A., Lowenstein, M., and Rothfield, N.F. (1980). Systemic lupus erythematosus in the younger patient: survival studies. *Journal of Rheumatology*, **7**, 515–22.

Abud-Mendoza, C., Sturbaum, A.K., Vazquez-Compean, R., and Gonzalez-Amaro, R. (1993). Methotrexate therapy in childhood systemic lupus erythematosus. *Journal of Rheumatology*, **20**, 731–3.

Abuelo, J.G., Esparza, A.R., and Garella, S. (1984). Steroid-dependent nephrotic syndrome in lupus nephritis. *Archives of Internal Medicine*, **144**, 2411–12.

Akhtar, M., Al-Dalaan, A., and El-Ramahi, K.M. (1994). Pauci-immune necrotizing lupus nephritis: report of two cases. *American Journal of Kidney Diseases*, **23**, 320–25.

Alexander, E., Buyon, J.P., Provost, T.T., and Guarnieri, T. (1992). Anti-Ro/SS-A antibodies in the pathophysiology of congenital hearth block in neonatal lupus syndrome, an experimental model. *Arthritis and Rheumatism*, **35**, 176–89.

Alexopoulos, E., Seron, D., Hartley, R.B., and Cameron, J.S. (1990). Lupus nephritis: correlation of interstitial cells with glomerular function. *Kidney International*, **37**, 100–109.

Appel, G.B., Cohen, D.J., Pirani, C.L., Meltzer, J.I., and Estes D. (1987). Long-term follow-up of patients with lupus nephritis. A study based on the classification of the World Health Organization. *American Journal of Medicine*, **83**, 877–85.

Appel, G.B., Pirani, C.L., and D'Agati, V. (1994). Renal vascular complications of systemic lupus erythematosus. *Journal of the American Society of Nephrology*, **4**, 1499–515.

Asherson, R.A. and Oakley, C.M. (1986). Pulmonary hypertension and systemic lupus erythematosus. *Journal of Rheumatology*, **13**, 1–5.

Austin III, H.A., Muenz, L.R., Joyce, K.M., Antonovych, T.T., and Balow, J.E. (1984). Diffuse proliferative lupus nephritis: identification of specific pathologic features affecting renal outcome. *Kidney International*, **25**, 689–95.

Austin III, H.A. *et al*. (1986). Therapy of lupus nephritis. Controlled trial of prednisone and cytotoxic drugs. *New England Journal of Medicine*, **314**, 614–19.

Austin III, H.A., Boumpas, D.T., Vaughan, E.M., and Balow, J.E. (1994). Predicting renal outcomes in severe lupus nephritis: contributions of clinical and histologic data. *Kidney International*, **45**, 544–50.

Badui, E. *et al*. (1985). Cardiovascular manifestations in systemic lupus erythematosus. prospective study of 100 patients. *Angiology—Journal of Vascular Diseases*, **36**, 431–41.

Bagga, A., Jain, Y., Srivastava, R.N., and Bhuyan, U.N. (1993). Renal tubular acidosis preceding systemic lupus erythematosus. *Pediatric Nephrology*, **7**, 735–6.

Bakir, A.A., Rhee, H.L., Ainis, H., and Dunea, G. (1989). Nephrotic syndrome, hematuria, and hypocomplementemia in a case of mesangial lupus nephritis evolving later to a membranous lesion. *American Journal of Medicine*, **86**, 609–11.

Baumal, R., Farine, M., and Poucell, S. (1987). Clinical significance of renal biopsies showing mixed mesangial and global proliferative lupus nephritis. *American Journal of Kidney Disease*, **10**, 236–40.

Bonfa, E. *et al*. (1987). Association between lupus psychosis and anti-ribosomal P protein antibodies. *New England Journal of Medicine*, **317**, 265–71.

Bruyn, G.W. and Padberg, G. (1984). Chorea and systemic lupus erythematosus. *European Neurology*, **23**, 435–48.

Bumgardner, G.L. *et al*. (1988). Single-center 1–15-year results of renal transplantation with systemic lupus erythematosus. *Transplantation*, **46**, 703–9.

Buyon, J.P., and Winchester, R. (1990). Congenital complete heart block. *Arthritis and Rheumatism*, **33**, 609–14.

Buyon, J.P. *et al*. (1988). Complete congenital heart block: risk of occurrence and therapeutic approach to prevention. *Journal of Rheumatology*, **15**, 1104–8.

Buyon, J.P. *et al*. (1993). Identification of mothers at risk for congenital heart block and other neonatal lupus syndromes in their children. *Arthritis and Rheumatism*, **36**, 1263–73.

Caeiro, E., Michielson, F.M.C., Bernstein, R., Hughes, G.R., and Ansell, B.M. (1981). Systemic lupus erythematosus in childhood. *Annals of the Rheumatic Diseases*, **40**, 325–31.

Cameron, J.S. (1994). Lupus nephritis in childhood and adolescence. *Pediatic Nephrology*, **8**, 230–49.

Carbotte, R.M., Denburg, S.D., and Denburg, J.A. (1986). Prevalence of cognitive impairment in systemic lupus erythematosus. *Journal of Nervous and Mental Disease*, **174**, 357–64.

Carette, S., Macher, A.M., Nussbaum, A., and Plotz, P.H. (1984). Severe, acute pulmonary disease in patients with systemic lupus erythematosus: ten years of experience at the National Institutes of Health. *Seminars in Arthritis and Rheumatism*, **14**, 52–9.

Carter, J.B., Blieden, L.C., and Edwards, J.E. (1974). Congenital Heart Block. *Archives of Pathology*, **97**, 51–7.

Cassidy, J.T., Sullivan, D.B., Petty, R.E., and Ragsdale, C. (1977). Lupus nephritis and encephalopathy. Prognosis in 58 children. *Arthritis and Rheumatism*, **20**, 315–22.

Citera, G., Espada, G., and Maldonado Cocco, J.A. (1993). Sequential development of 2 connective tissue diseases in juvenile patients. *Journal of Rheumatology*, **20**, 2149–52.

Cook, C.D., Wedgwood, J.P., Craig, J.M., Hartmann, J.R., and Janeway, C.A. (1960). Systemic lupus erythematosus: description of 37 cases in children and a discussion of endocrine therapy in 32 of the cases. *Pediatrics*, **26**, 570–85.

Decker, J.L., Steinberg, A.D., Reinertsen, J.L., Plotz, P.H., Balow, J.E., and Klippel, J.H. (1979). Systemic lupus erythematosus: evolving concepts. *Annals of Internal Medicine*, **91**, 587–604.

De Inocencio, J. and Lovell, D.J. (1994). Cardiac function in systemic lupus erythematosus. *Journal of Rheumatology*, **21**, 2147–56.

Delgado, E.A., Malleson, P.N., Pirie, G.E., and Petty, R.E. (1990). The pulmonary manifestations of childhood onset systemic lupus erythematosus. *Seminars in Arthritis and Rheumatism*, **19**, 285–93.

Deng, J-S., Sontheimer, R.D., and Gilliam, J.N. (1984). Relationships between antinuclear antibodies and anti-Ro/SSA antibodies in subacute cutaneous lupus erythematosus. *Journal of the America Academy of Dermatology*, **11**, 494–9.

Doherty, N.E. and Siegel, R.J. (1985). Cardiovascular manifestations of systemic lupus erythematosus. *American Heart Journal*, **110**, 1257–65.

Eberhard, A., Shore, A., Silverman, E., and Laxer R. (1991 *a*). Bowel perforation and interstitial cystitis in childhood systemic lupus erythematosus: a case report. *Journal of Rheumatology*, **18**, 746–7.

Eberhard, A., Laxer, R., Eddy, A., and Silverman, E. (1991 *b*). Presence of thyroid abnormalities in children with systemic lupus erythematosus. *Journal of Pediatrics*, **119**, 277–9.

Eberhard, A., Couper, R., Durie, P., and Silverman, E. (1992). Exocrine pancreatic function in children with systemic lupus erythematosus. *Journal of Rheumatology*, **19**, 964–7.

Eberhard, A., Sparling, C., Sudbury, S., Ford, P., Laxer, R., and Silverman, E. (1994). Hypoprothrombinemia in childhood systemic lupus erythematosus. *Seminars in Arthritis and Rheumatism*, **24**, 12–18.

El-Garf, A. and Salah, S. (1990). Juvenile systemic lupus erythematosus among Egyptian children. *Journal of Rheumatology*, **17**, 1168–70.

Emery, H. (1986). Clinical aspects of systemic lupus erythematosus in childhood. *Pediatric Clinics of North America*, **33**, 1177–90.

Esdaile, J.M., Joseph, L., Mackenzie, T., Kashgarian, M., and Hayslett, J.P. (1994). The benefit of early treatment with immunosuppressive agents in lupus nephritis. *Journal of Rheumatology*, **21**, 2046–51.

Fish, A.J., Blau, E.B., Westberg, N.G., Burke, B.A., Vernier, R.L., and Michael, A.F. (1977). Systemic lupus erythematosus within the first two decades of life. *American Journal of Medicine*, **62**, 99–117.

Fisk, J.D., Eastwood, B., Sherwood, G., and Hanly, J.G. (1993). Patterns of cognitive impairment in patients with systemic lupus erythematosus. *British Journal of Rheumatology*, **32**, 458–62.

Fraenkel, L., Mackenzie, T., Joseph, L. Kashgarian, M., Hayslett, J.P., and Esdaile, J.M. (1994). Response to treatment as a predictor of longterm outcome in patients with lupus nephritis. *Journal of Rheumatology*, **21**, 2052–7.

Friedman, D.M., Lazarus, H.M., and Fierman, A.H. (1990). Acute myocardial infarction in pediatric systemic lupus erythematosus. *Journal of Pediatrics*, **117**, 263–6.

Garcia, S. *et al*.(1994) Cellular mechanism of the conduction abnormalities induced by serum from anti-Ro/SSA-positive patients in rabbit hearts. *Journal of Clinical Investigation*, **93**, 718–24.

Glidden, R.S., Mantzouranis, E.C., and Borel, Y. (1983). Systemic lupus erythematosus in childhood: clinical manifestations and improved survival in fifty-five patients. *Clinical Immunology and Immunopathology*, **29**, 196–210.

Gribetz, D. and Henley, W.L. (1959). Systemic lupus erythematosus in childhood. *Mount Sinai Journal of Medicine*, **26**, 289–96.

Hall, S., McCormick, Jr., J.L., Greipp, P.R., Michet, Jr., C.J., and McKenna, C.H. (1985). Splenectomy does not cure the thrombocytopenia of systemic lupus erythematosus. *Annals of Internal Medicine*, **102**, 325–8.

Hall, S., Czaja, A.J., Kaufman, D.K., Markowitz, H., and Ginsburg, W.W. (1986). How lupoid is lupoid hepatitis? *Journal of Rheumatology*, **13**, 95–8.

Hanly, J.G., Fisk, J.D., Sherwood, G., and Eastwood, B. (1994). Clinical course of cognitive dysfunction in systemic lupus erythematosus. *Journal of Rheumatology*, **21**, 1825–31.

Harmon, C.E., Deng, J-S., Peebles, C.L., and Tan, E.M. (1984). The importance of tissue substrate in the SS-A/Ro antigen-antibody system. *Arthritis and Rheumatism*, **27**, 166–73.

Hirohata, S., Hirose, S., and Miyamoto, T. (1985). Cerebrospinal fluid IgM, IgA, and IgG indexes in systemic lupus erythematosus. *Archives of Internal Medicine*, **145**, 1843–6.

Ho, Y.S., Esscher, E., Anderson, R.H., and Michaelsson M. (1986). Anatomy of congenital complete heart block and relation to maternal anti-Ro antibodies. *American Journal of Cardiology*, **58**, 291–4.

Hoffman, B.I. and Katz, W.A. (1980). The gastrointestinal manifestations of systemic lupus erythematosus: a review of the literature. *Seminars in Arthritis and Rheumatism*, **9**, 237–47.

Hogg, G.R. (1957). Congenital acute lupus erythematosus associated with subendocardial fibroelastosis. *American Journal of Clinical Pathology*, **28**, 648–53.

Holman, B.L. (1993). Functional imaging in systemic lupus erythematosus: an accurate indicator of central nervous system involvement? *Arthritis and Rheumatism*, **36**, 1193–95.

Ilowite, N.T., Samuel, P., Ginzler, E., and Jacobson, M.S. (1988). Dyslipoproteinemia in pediatric systemic lupus erythematosus. *Arthritis and Rheumatism*, **31**, 859–63.

Imperiale, T.F., Goldfarb, S., and Berns, J.S. (1995). Are cytotoxic agents beneficial in idiopathic membranous nephropathy? A meta-analysis of the controlled trials. *Journal of the American Society of Nephrology*, **5**, 1553–8.

Isenberg, D.A., Meyrick-Thomas, D., Snaith, M.L., McKeran, R.O., and Royston, J.P. (1982). A study of migraine in systemic lupus erythematosus. *Annals of the Rheumatic Diseases*, **41**, 30–32.

Isshi, K., Hirohata, S., Hashimoto, T., and Miyashita, H. (1994). Systemic lupus erythematosus presenting with diffuse low density lesions in the cerebral white matter on computed axial tomography scans: its implication in the pathogenesis of diffuse central nervous system lupus. *Journal of Rheumatology*, **21**, 1758–62.

Jacobs, J.C. (1963). Systemic lupus erythematosus in childhood: report of thirty-five cases, with discussion of seven apparently induced by anticonlvulsant medication, and of prognosis and treatment. *Pediatrics*, **32**, 257–64.

Jarek, M.J., West, S., Baker, M.R., and Rak, K.M. (1994). Magnetic resonance imaging in systemic lupus erythematosus patients without a history of neuropsychiatric lupus erythematosus. *Arthritis and Rheumatism*, **37**, 1609–13.

Kalla, A.A., Learmonth, I.D., and Klemp, P. (1986). Early treatment of avascular necrosis in systemic lupus erythematosus. *Annals of the Rheumatic Diseases*, **45**, 649–52.

Katsanis, E., Hsu, E., Luke, K.-H., and McKee, J.A. (1987). Systemic lupus erythematosus and sickle hemoglobinopathies: a report of two cases and review of the literature. *American Journal of Hematology*, **25**, 211–14.

Kaufman, D.B., Laxer, R.M., Silverman, E.D., and Stein, L. (1986). Systemic lupus erythematosus in childhood and adolescence—the problem, epidemiology, incidence, susceptibility, genetics, and prognosis. *Current Problems in Pediatrics*, **16**, 555–624.

Khamashta, M.A., Cuadrado, M.J., Mujic, F., Taub, N.A., Hunt, B.J., and Hughes, G.R.V. (1995). The management of thrombosis in the antiphospholipid-antibody syndrome. *New England Jourrnal of Medicine*, **332**, 993–7.

King, K.K., Kornreich, H.K., Bernstein, B.H., Singsen, B.H., and Hanson, V. (1977). The clinical spectrum of systemic lupus erythematosus in childhood. *Arthritis and Rheumatism*, **20** (Suppl.), 287–94.

Kodama, K. *et al*. (1995). Single photon emission computed tomography in systemic lupus erythematosus with psychiatric symptoms. *Journal of Neurology, Neurosurgery, and Psychiatry*, **58**, 307–11.

Kunkel, H.G. (1980). The immunopathy of systemic lupus erythematosus. *Hospital Practice*, **15**, 47.

Lacks, S. and White, P. (1990). Morbidity associated with childhood systemic lupus erythematosus. *Journal of Rheumatology*, **17**, 941–5.

Laitman, R.S., Glicklich, D., Sablay, L.B., Grayzel, A.I., Barland, P., and Bank, N. (1989). Effect of long-term normalization of serum complement levels on the course of lupus nephritis. *American Journal of Medicine*, **87**, 132–8.

Laxer, R.M. *et al*. (1990). Liver disease in neonatal lupus. *Journal of Pediatrics*, **116**, 238–42.

Lee, L.A., Harmon, C.E., Huff, J.C., Norris, D.A., and Weston, W.L. (1985). The demonstration of SS-A/Ro antigen in human fetal tissues and in neonatal and adult skin. *Journal of Investigative Dermatology*, **85**, 143–6.

Lee, L.A., Norris, D.A., and Weston, W.L. (1986). Neonatal lupus and the pathogenesis of cutaneous lupus. *Pediatric Dermatology*, **3**, 491–7.

Lee, L.A., Reichlin, M., Ruyle, S.Z., and Weston, W.L. (1993). Neonatal lupus liver disease. *Lupus*, **2**, 333–8.

Lee, L.A., Frank, M.B., McCubbin, V.R., and Reichlin, M. (1994). Autoantibodies of neonatal lupus erythematosus. *Journal of Investigative Dermatology*, **102**, 963–6.

LeFeber, W.P. *et al*. (1984). Ultraviolet light induces binding of antibodies to selected nuclear antigens on cultured keratinocytes. *Journal of Clinical Investigation*, **74**, 1545–51.

Lehman, T.J.A. *et al*. (1989*a*). Intermittent intravenous cyclophosphamide therapy for lupus nephritis. *Journal of Pediatrics*, **144**, 1055–60.

Lehman, T.J.A., McCurdy, D.K., Bernstein, B.H., King, K.K., and Hanson, V. (1989 *b*). Systemic lupus erythematosus in the first decade of life. *Pediatrics*, **83**, 235–9.

Lev, M., Silverman, J., Fitzmaurice, F.M., Paul, M.H., Cassels, D.E., and Miller, R.A. (1971). Lack of connection between the atria and the more peripheral conduction system in congenital atrioventricular block. *American Journal of Cardiology*, **27**, 481–90.

Lipnick, R.N., Tsokos, G.C., Bray, G.L., and White, P.H. (1990). Autoimmune thrombocytopenia in pediatric systemic lupus erythematosus: alternative therapeutic modalities. *Clinical and Experimental Rheumatology*, **8**, 315–19.

Long, A.A., Denburg, S.D., Carbotte, R.M., Singal, D.P., and Denburg, J.A. (1989). Serum lymphocytotoxic antibodies and neurocognitive function in systemic lupus erythematosus. *Annals of the Rheumatic Diseases*, **49**, 249–53.

Magil, A.B., Ballon, H.S., Chan, V., Lirenman, D.S., and Rae, A. (1984). Diffuse proliferative lupus glomerulonephritis. Determination of prognostic significance of clinical, laboratory, and pathologic factors. *Medicine*, **63**, 210–20.

Malleson, P.N. (1989). The role of the renal biopsy in childhood onset systemic lupus erythematosus: a viewpoint. *Clinical and Experimental Rheumatology*, **7**, 563–6.

Martini, A., Ravelli, A., Viola, S., and Burgio, R.G. (1987). Systemic lupus erythematosus with Jaccoud's arthropathy mimicking juvenile rheumatoid arthritis. *Arthritis and Rheumatism*, **30**, 1062–4.

McCune, A.B., Weston, W.L., and Lee, L.A. (1987). Maternal and fetal outcome in neonatal lupus erythematosus. *Annals of Internal Medicine*, **106**, 518–23.

McCurdy, D.K. *et al*. (1992). Lupus nephritis: prognostic factors in children. *Pediatrics*, **89**, 240–46.

McLean, B.N., Miller, D., and Thompson, E.J. (1995). Oligoclonal banding of IgG in CSF, blood–brain barrier function, and MRI findings in patients with sarcoidosis, systemic lupus erythematosus, and Behçet's disease involving the nervous system. *Journal of Neurology, Neurosurgery, and Psychiatry*, **58**, 548–54.

Miguel, E.C. *et al*.(1994). Psychiatric manifestations of systemic lupus erythematosus: clinical features, symptoms, and signs of central nervous system activity in 43 patients. *Medicine*, **73**, 224–32.

Miller, J.J.I. (1977). Drug-induced lupus-like syndrome in children. *Arthritis and Rheumatism*, **20** (Suppl.), 308–11.

Molta, C. *et al*. (1993). Childhood-onset systemic lupus erythematosus: antiphospholipid antibodies in 37 patients and their first-degree relatives. *Pediatrics*, **92**, 849–53.

Nadorra, R.L. and Landing, B.H. (1987). Pulmonary lesions in childhood onset systemic lupus erythematosus: analysis of 26 cases, and summary of literature. *Pediatric Pathology*, **7**, 1–18.

Nadorra, R.L., Nakazato, Y., and Landing, B.H. (1987). Pathologic features of gastrointestinal tract lesions in childhood-onset systemic lupus erythematosus: study of 26 patients, with review of the literature. *Pediatric Pathology*, **7**, 245–59.

Nossent, H.C., Swaak, T.J.G., and Berden, J.H.M. (1991). Systemic lupus erythematosus after renal transplantation: patient and graft survival and disease activity. *Annals of Internal Medicine*, **114**, 183–8.

O'Connor, P. (1988). Diagnosis of central nervous system lupus. *Canadian Journal of Neurological Sciences*, **15**, 257–60.

Orth, R.W., Weisman, M.H., Cohen, A.H., Talner, L.B., Nachtsheim, D., and Zvaifler, N.J. (1983). Lupus cystitis: primary bladder manifestations of systemic lupus erythematosus. *Annals of Internal Medicine*, **98**, 323–6.

Papero, P.H., Bluestein, H.G., White, P., and Lipnick, R.N. (1990). Neuropsychologic deficits and antineuronal antibodies in pediatric systemic lupus erythematosus. *Clinical and Experimental Rheumatology*, **8**, 417–24.

Peake, P.W., Greenstein, J.D., Timmermans, V., Gavrilovic, L., and Charlesworth, J.A. (1988). Lymphocytotoxic antibodies in systemic lupus erythematosus: studies of their temperature dependence, binding characteristics, and specificty in vitro. *Annals of the Rheumatic Diseases*, **47**, 725–32.

Penn, I. (1988). *Development of new tumors after transplantation*. J.B. Lippincott Co., Philadelphia.

Platt, J.L., Burke, B.A., Fish, A.J., Kim, Y., and Michael, A.F. (1982). Systemic lupus erythematosus in the first two decades of life. *American Journal of Kidney Diseases*, **11**, 212–22.

Pollak, V.E. and Pirani, C.L. (1993). In *Dubois' lupus erythematosus*, (4th edn)(ed. D.J. Wallace and B.H. Hahn), pp. 525–41. Lea and Febiger, Philadelphia.

Posner, M.A., Gloster, E.S., Bonagura, V.R., Valacer, D.J., and Ilowite, N.T. (1990). Burkitt's lymphoma in a patient with systemic lupus erythematosus. *Journal of Rheumatology*, **17**, 380–82.

Press, J. *et al*. (1996). Antiribosomal P antibodies in pediatric patients with systemic lupus erythematosus and psychosis. *Arthritis and Rheumatism*, **39**, 671–6.

Ragsdale, C.G., Petty, R.E., Cassidy, J.T., and Sullivan, D.B. (1980). The clinical progression of apparent juvenile rheumatoid arthritis to systemic lupus erythematosus. *Journal of Rheumatology*, **7**, 50–55.

Ravelli, A., Martini, A., and Burgio, G.R. (1994). Antiphospholipid antibodies in paediatrics. *European Journal of Pediatrics*, **153**, 472–9.

Reed, B. *et al*. (1983). Autoantibodies to SS-A/Ro in infants with congenital heart block. *Journal of Pediatrics*, **103**, 889–92.

Reeves, W.H. and Lahita, R.G. (1987). Clinical presentation of systemic lupus erythematosus in the adult. In *Systemic lupus erythematosus*, pp.355–75. John Wiley, New York..

Reichlin, M. *et al*. (1994). Concentration of autoantibodies to native 60-kd Ro/SS-A and denatured 52-kd Ro/Ss-A in eluates from the heart of a child who died with congenital complete heart block. *Arthritis and Rheumatism*, **37**, 1698–703.

Robinson, M.J., and Williams, A.L. (1967). Systemic lupus erythematosus in childhood. *Australian Paediatric Journal*, **3**, 36–47.

Rosh, J.R., Silverman, E.D., Groisman, G., Dolgin, S., and LeLeiko, N.S. (1993). Intrahepatic cholestasis in neonatal lupus erythematosus. *Journal of Pediatric Gastroenterology and Nutrition*, **17**, 310–12.

Rubbert, A. *et al*. (1993). Single-photon-emission computed tomography analysis of cerebral blood flow in the evaluation of central nervous system involvement in patients with systemic lupus erythematosus. *Arthritis and Rheumatism*, **36**, 1253–62.

Rubin, L.A., Urowitz, M.B., and Gladman, D.D. (1985). Mortality in systemic lupus erythematosus: the bimodal pattern revisited. *Quarterly Journal of Medicine*, **55**, 87–98.

Saulsbury, F.T., Kesler, R.W., Kennaugh, J.M., Barber, J.C., and Chevalier, R.L. (1982). Overlap syndrome of juvenile rheumatoid arthritis and systemic lupus erythematosus. *Journal of Rheumatology*, **9**, 610–12.

Schaller, J. (1982). Lupus in childhood. *Clinics in the Rheumatic Diseases*, **8**, 219–28.

Schneebaum, A.B. *et al*. (1991). Association of psychiatric manifestations with antibodies to ribosomal P proteins in systemic lupus erythematosus. *American Journal of Medicine*, **90**, 54–62.

Shemesh, O., Golbetz, H., Kriss, J.P., and Myers, B.D. (1985). Limitations of creatinine as a filtration marker in glomerulopathic patients. *Kidney International*, **28**, 830–38.

Shergy, W.J., Kredich, D.W., and Pisetsky, D.S. (1988). The relationship of anticardiolipin antibodies to disease manifestations in pediatric systemic lupus erythematosus. *Journal of Rheumatology*, **15**, 1389–94.

Shergy, W.J., Kredich, D.W., and Pisetsky, D.S. (1989). Patterns of autoantibody expression in pediatric and adult systemic lupus erythematosus. *Journal of Rheumatology*, **16**, 1329–34.

Siegel, M. and Lee, S.L. (1973). The epidemiology of systemic lupus erythematosus. *Seminars in Arthritis and Rheumatism*, **3**, 1–54.

Silverman, E.D., Mamula, M., Hardin, J.A., and Laxer R.M. (1991). Importance of the immune response to the Ro/La particle in the development of congenital heart block and neonatal lupus erythematosus. *Journal of Rheumatology*, **18**, 120–24.

Sloan, R.P., Schwartz, M.M., Korbet, S.M., Borok, R.Z., and the Lupus Nephritis Collaborative Study Group (1996). Long-term outcome in systemic lupus erythematosus membranous glomerulonephritis. *Journal of the American Society of Nephrology*, **7**, 299–305.

Smith, F.E., Sweet, D.E., Brunner, C.M., and Davis IV, J.S. (1976). Avascular necrosis in SLE. An apparent predilection for young patients. *Annals of the Rheumatic Diseases*, **35**, 227–32.

Smith, N.M. and Ho, S.Y. (1994). Heart block and sudden death associated with fibrosis of the conduction system at the margin of a ventricular septal defect. *Pediatric Cardiology*, **15**, 139–42.

Szer, I.S., Miller, J.H., Rawlings, D., Shaham, B., and Bernstein, B. (1993). Cerebral perfusion abnormalities in children with central nervous system manifestations of lupus detected by single photon emission computed tomography. *Journal of Rheumatology*, **20**, 2143–8.

Teh, L.S. *et al*. (1993). Anti-P antibodies are associated with psychiatric and focal cerebral disorders in patients with systemic lupus erythematosus. *British Journal of Rheumatology*, **32**, 287–90.

Tucker, L.B., Menon, S., Schaller, J.G., and Isenberg, D.A. (1995). Adult- and childhood-onsent systemic lupus erythematosus: a comparison os sonet, clinical features, serology and outcome. *British Journal of Rheumatology*, **34**, 866–72.

Uziel, Y., Laxer, R.M., Blaser, S., Andrew, M., Scneider, R., and Silverman, E.D. (1995). Cerebral vein thrombosis in childhood systemic lupus erythematosus. *Journal of Pediatrics*, **126**, 722–7.

Walravens, P. A. and Chase, H.P. (1976). The prognosis of childhood systemic lupus erythematosus. *American Journal of Diseases of Children*, **130**, 929–33.

Ward, M.M., Dawson, D.V., Kredich, D.W., and Pisetsky, D.S. (1990). Expression of IgM and IgG autoantibodies in pediatric and adult systemic lupus erythematosus. *Clinical Immunology and Immunopathology*, **55**, 273–84.

Watson, R.M., Lane, A.T., Barnett, N.K., Bias, W.B., Arnett, F.C., and Provost, T.T. (1984). Neonatal lupus erythematosus. A clinical, serological and immunologenic study with review of the literature.

Medicine Baltimore, **63**,362–78.

Weber, H.S. and Myers, J.L. (1994). Maternal collagen vascular disease associated with fetal heart block and degenerative changes of the atrioventricular valves. *Pediatric Cardiology*, **15**, 204–6.

Wilson, W.A., Armatis, P.E., and Perez, M.C. (1989). C4 concentrations and C4 deficiency alleles in systemic lupus erythematosus. *Annals of Rheumatic Diseases*, **48**, 600–604.

Yancey, C.L., Doughty, R.A., and Athreya, B.H. (1981). Central nervous system involvement in childhood systemic lupus erythematosus. *Arthritis and Rheumatism*, **24**, 1389–95.

Yang, L-Y, Chen, W-P., and Lin, C-Y. (1994). Lupus nephritis in children—a review of 167 patients. *Pediatrics*, **94**, 335–40.

Zetterstrom, R. and Berglund, G. (1956). Systemic lupus erythematosus in childhood a clinical study. *Acta Paediatrica Hungarica*, **45**, 189–204.

5.7.3 The antiphospholipid antibody syndrome

Munther A. Khamashta and G. R. V. Hughes

Introduction

In 1983, few rheumatologists would have predicted the interest that the introduction of the anticardiolipin test would generate in subsequent years (Harris *et al*. 1983). New autoantibodies turn up frequently in patients with systemic lupus erythematosus and there may have seemed little reason why anticardiolipin antibodies should have merited any more than passing interest. However, these particular autoantibodies had unusual characteristics that attracted the attention of investigators from a variety of disciplines (Fig. 1). One reason may have been the novelty of phospholipid molecules as antigens, given that most attention up to that time had focused on protein, DNA, and even polysaccharide antigens (Harris *et al*. 1994). In addition, these antibodies exhibited an easily detectable biological effect, as evidenced by their ability to prolong clotting times of plasma. More intriguing was their association with an unusual combination of clinical complications that included venous and arterial thrombosis, pregnancy loss, and thrombocytopenia (Hughes 1983; Hughes 1993).

Fig. 1 Antiphospholipid antibodies—linking many specialties (reproduced with permission of the Editor of *Clinical and Experimental Rheumatology*).

During the last decade, considerable progress has been made in understanding antiphospholipid antibodies and the disorder with which they are associated, but many questions remain unanswered, particularly those of pathogenesis and optimal treatment (Khamashta and Asherson 1995). A chronology of the major developments in the unfolding of the antiphospholipid syndrome story is listed in Table 1.

1906	Wassermann reaction (reagin)
1941	Reagin binds cardiolipin
1952	False positive test for syphilis
1959	Cofactor requirements for lupus anticoagulant activity
1960s	Lupus anticoagulant: association with thrombosis
1970s	Lupus anticoagulant is due to immunoglobulin
1975	Lupus anticoagulant: association with recurrent abortions
1983	Anticardiolipin antibodies: detection by radioimmunoassay
1985	Anticardiolipin antibodies: detection by ELISA
1980s	Clinical description of the antiphospholipid syndrome
1987	Diagnostic criteria for the antiphospholipid syndrome
1989	Lupus anticoagulant and anticardiolipin: separate antibody subgroups
1990	Cofactor requirements for anticardiolipin antibody binding
1990	Anticardiolipin cofactor: β_2-glycoprotein I
1991	Lupus anticoagulant cofactor: prothrombin
1991	Animal models: passive immunization
1992	Animal models: active immunization
1992	Lupus anticoagulant cofactor: β_2-glycoprotein I
1992	Anti-β_2-glycoprotein I: association with thrombosis
1994	Phospholipid binding site: fifth domain of β_2-glycoprotein I

ELISA, enzyme-linked immunosorbent assay.

Table 1 Antiphospholipid antibodies and the antiphospholipid syndrome—history

Detection of antiphospholipid antibodies

Antiphospholipid antibodies are detected by a variety of laboratory tests, the most useful for identifying patients with the antiphospholipid syndrome being the lupus anticoagulant and the anticardiolipin antibody tests. These antibodies are distinct and separable immunoglobulins present alone or in combination in the plasma of people with the antiphospholipid syndrome. The autoantibodies sometimes bind phospholipids utilized in the Venereal Disease Research Laboratories (**VDRL**) test; hence, some patients may have a false-positive test for syphilis. However, the VDRL test is not positive frequently enough to make it valuable in diagnosing the antiphospholipid syndrome.

The lupus anticoagulant is a functional assay measuring the ability of antiphospholipid antibodies to prolong clotting via their inhibition of the conversion of prothrombin to thrombin or the activation of factor X (both reactions are catalysed by phospholipids). Tests for the lupus anticoagulant have been difficult to standardize, and no single test appears to be adequate (Permpikul *et al*. 1994; Triplett 1994). The test begins with an attempt to demonstrate an abnormal coagulation screening test, such as a prolonged activated partial thromboplastin time, dilute Russell viper venom time, or kaolin clotting time. If any of these is positive, the test is repeated, using a sample in which the patient's plasma has been mixed with normal plasma. If the patient's disorder is a clotting deficiency, the test should become normal. If, on the other hand, lupus anticoagulant or some other clotting inhibitor is present, the clotting time will remain prolonged. The presence of lupus anticoagulant is confirmed by the return to normal of the clotting test after addition of freeze–thawed platelets or excess phospholipids, either of which bind the antibodies. The lupus anticoagulant test must be performed on platelet-poor plasma. The test cannot be performed reliably if the patient is receiving heparin or oral

anticoagulants.

The most sensitive test for antiphospholipid antibodies is the anticardiolipin antibody test, introduced in 1983 and extensively improved since that time (Gharavi *et al*. 1987; Harris *et al*. 1987; Harris *et al*. 1990; Khamashta and Hughes 1993). This test uses enzyme-linked immunosorbent assay to determine antibody binding to solid plates coated either with cardiolipin or other phospholipids. Serum or plasma samples may be used for the anticardiolipin assay. The availability of reference sera which are isotype specific (IgG and IgM) has greatly improved interlaboratory testing and quantification of anticardiolipin antibodies (Harris *et al*. 1990). IgG and IgM isotype concentrations are expressed as GPL and MPL units, respectively. One unit represents the binding activity of 1 μg/ml of affinity purified anticardiolipin antibody. Results are expressed as low, medium, and high positive according to levels below 20 units, between 20 and 80 units, and above 80 units, respectively. IgA anticardiolipin reference sera are now also available. Many laboratories currently measure all three isotypes and sensitive kits are commercially available. Flow cytometry has also been used to test for anticardiolipin antibodies (Stewart *et al*. 1993). This sytem allows the simultaneous measurement of antiphospholipid antibody isotypes with different phospholipid specificity. The routine detection of other phospholipids, such as phosphatidylserine, phosphatidylinositol, and phosphatidic acid, gives little additional information.

In our experience, although a strong association between the anticardiolipin and the lupus anticoagulant results was found, these tests were discordant in 40 per cent of cases. The majority of patients with positive anticardiolipin antibodies were found to be positive for the IgG isotype (93 per cent). IgM anticardiolipin antibodies were demonstrated in only a quarter of the patients (Khamashta *et al*. 1995). The unrelated behaviour of lupus anticoagulant and anticardiolipin antibodies in the course of disease and in individual patients indicates that both assays are required if most cases with the antiphospholipid syndrome are to be identified.

Clinical features

The antiphospholipid syndrome is a thrombophilic disorder in which patients may develop both venous and arterial occlusion. The clinical ramifications are extensive (Table 2). Table 3 illustrates the frequency of the major manifestations of the antiphospholipid syndrome in our patients.

Table 2 Clinical features of the antiphospholipid syndrome

Clinical features	No. of patients	Percentage
Thrombosis	147/171	86
Venous:	80/147	54
Lower limbs DVT	54/80	67
Pulmonary embolism	19/80	24
Other	7/80	9
Arterial:	67/147	46
Stroke	32/67	48
TIA	24/67	36
Other	11/67	16
Thrombocytopenia	40/171	23
Pregnancy loss	48/147 (female)	33
First trimester	61/109 (pregnancies)	56
Second trimester	78/109 (pregnancies)	46
Third trimester	30/109 (pregnancies)	18

"St. Thomas' Hospital Lupus Clinic database (1985–1994).
DVT, deep venous thrombosis; TIA, transient ischaemic attack.

Table 3 Clinical manifestations of 171 patients with antiphospholipid syndrome [a]

Thrombosis

Although some antiphospholipid antibodies prolong *in vitro* clotting tests, haemorrhage is rare in patients with these antibodies. When haemorrhage does occur, other causes such as severe thrombocytopenia or clotting factor inhibitors should be excluded (Harris *et al*. 1993).

Instead of being associated with haemorrhage, antiphospholipid antibodies are paradoxically associated with thrombosis. Thrombosis can occur anywhere in the venous or arterial circulation. Vessels of all sizes may be affected, and the vascular pathological appearance has consistently been of bland occlusion without inflammatory infiltrate (Fig. 2) (Lie 1994). Thus, it is unlikely that thrombotic occlusion of blood vessels in patients with the antiphospholipid syndrome is caused by vasculitis, which may be seen in other patients with autoimmune rheumatic disorders. The distinction is important not only for discovering the pathogenesis of the vascular lesions, but also for the choice of treatment.

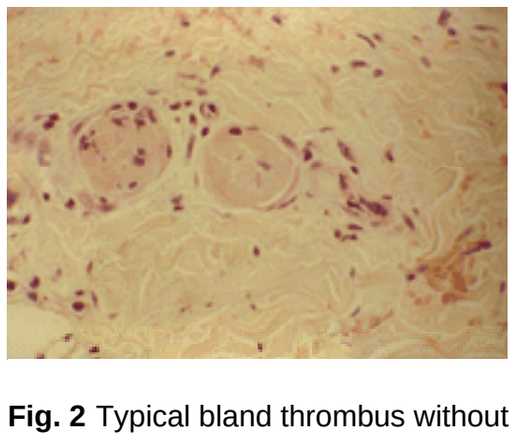


Fig. 2 Typical bland thrombus without inflammatory infiltrate in the vessel of a patient with primary antiphospholipid syndrome.

In the venous circulation, thrombosis of the deep and superficial veins of the lower extremities has been reported most frequently (occasionally after the use of oral contraceptive pills containing oestrogen). It is often recurrent and may be accompanied by pulmonary embolism. It has been estimated that up to 19 per cent of patients with deep vein thrombosis and/or pulmonary thromboembolism are suffering from antiphospholipid coagulopathy and may demonstrate a positive lupus anticoagulant test, antibodies to cardiolipin, or both. Some patients with antiphospholipid antibodies also have pulmonary hypertension, perhaps caused by recurrent pulmonary emboli or intravascular thromboses (Asherson *et al*. 1990). Other reported venous sites of thrombosis include the axillary, ocular, renal, and hepatic veins and the inferior vena cava. The antiphospholipid syndrome is now considered one of the most frequent causes of the Budd–Chiari syndrome (Fig. 3) (Pelletier *et al*. 1994). Antiphospholipid antibodies have recently been implicated in the development of adrenal vein thrombosis leading to Addison's disease (Asherson and Hughes

1991). Interestingly, most documented cases with adrenal insufficiency associated with antiphospholipid antibodies were in patients with history of previous venous thromboses.

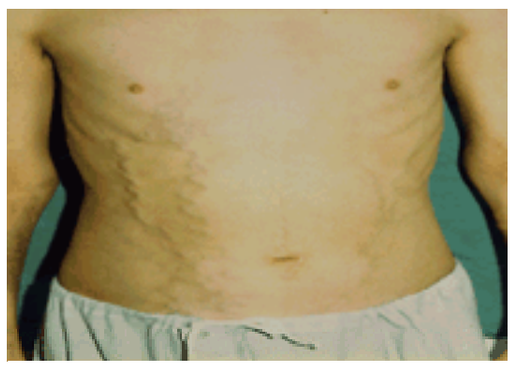


Fig. 3 Budd–Chiari syndrome in a 20-year-old patient with primary antiphospholipid syndrome (by courtesy of Dr L. Pallares, Servicio de Medicina Internal, Hospital Son Dureta, Palma de Mallorca, Spain).

Unlike other known clotting disorders, arterial thromboses are a major feature of the antiphospholipid syndrome. Occlusion of the intracranial arteries has been reported most frequently, with the majority of patients presenting with stroke and transient ischaemic attacks. Magnetic resonance imaging scans show changes that vary from single lesions to multiple widely-scattered infarcts (Fig. 4) (Asherson *et al*. 1989a; Stimmler *et al*. 1993). In some patients, untreated recurrent cerebral thrombosis has led to multi-infarct dementia and psychiatric features have been prominent in the presentation of some of our patients with the antiphospholipid syndrome. Antiphospholipid antibodies are now internationally recognized as an important aetiological factor and may be present in 7 per cent of all patients who have suffered a stroke (Montalban *et al*. 1991). They should be sought especially in young patients with strokes, where they may account for up to 18 per cent (Nencini *et al*. 1992). The prevalence of myocardial infarction in patients with antiphospholipid antibodies has yet to be established. A figure of 4 per cent was derived from our studies assessing patients with lupus and related disorders (Asherson *et al*. 1989b). Another study, found that one-fifth of all young patients with myocardial infarction had antiphospholipid antibodies (Hamsten *et al*. 1986). Other arterial thromboses have involved the retina, mesenteric and peripheral arteries. Malignant hypertension with renal insufficiency secondary to thrombosis of the renal glomeruli and renal thrombotic microangiopathy (without classical lupus nephritis) has also been associated with the presence of antiphospholipid antibodies (Fig. 5) (Amigo *et al*. 1992; Asherson *et al*. 1993a; Piette *et al*. 1994).

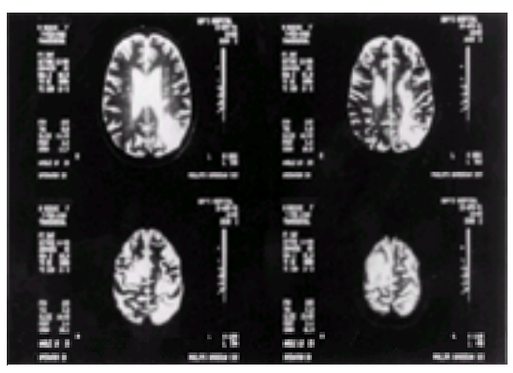


Fig. 4 Cerebral magnetic resonance imaging scan showing multiple widely-scattered infarcts in a patient with primary antiphospholipid syndrome.

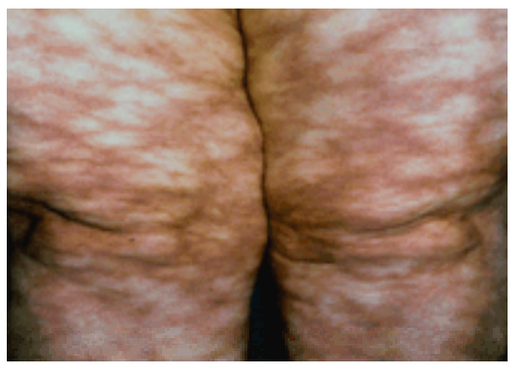


Fig. 5 Extensive livedo reticularis in a patient with primary antiphospholipid syndrome.

Occasionally, patients positive for antiphospholipid antibody develop acute medical collapse with severe thrombocytopenia, adult respiratory distress syndrome, and multiple organ (notably cerebral and renal) failures. The aetiology of this usually fatal condition is unknown, though the limited reports available suggest widespread thrombosis as the pathogenesis. The syndrome often appears to have had a trigger, such as preceding viral infection. We have chosen to call this rare but life-threatening presentation the 'catastrophic antiphospholipid syndrome' (Asherson 1992a).

Fetal loss

Recurrent spontaneous pregnancy losses are one of the most consistent complications of the antiphospholipid syndrome. Losses can occur at any stage of pregnancy, although miscarriages associated with antiphospholipid antibody are strikingly frequent (about 50 per cent of cases) during the second and third trimester. This differs from the pattern of pregnancy loss in the normal population, which usually occurs during the first trimester and is most often due to non-immunological factors, i.e. morphological or chromosomal abnormalities (Branch 1994). The rate of miscarriage in patients positive for antiphospholipid antibody is still uncertain, although the epidemiology is being studied and, increasingly, testing for this antibody is becoming a routine investigation in women with recurrent miscarriages. Fewer than 2 per cent of apparently normal pregnant women have either anticardiolipin antibody or lupus anticoagulant in any titre and less than 0.2 per cent have high titre antibody (Lockwood *et al*. 1989; Harris and Spinnato 1991). Hence, screening normal pregnant women has little value. Previous pregnancy history is of importance in determining the significance of a positive laboratory test for antiphospholipid antibodies. It has been estimated that if a patient with lupus has a positive lupus anticoagulant, or at least moderate levels of IgG anticardiolipin antibodies, the risk of spontaneous abortion during the first pregnancy is 30 per cent and if she has a history of at least two spontaneous abortions, the risk is 70 per cent during the following pregnancy (Lockshin *et al*. 1987). In our series, the presence of antiphospholipid antibodies was an important predictor of poor fetal outcome, as was a poor obstetric history (Buchanan *et al*. 1992; Lima *et al*. 1995). These findings are in keeping with the results of other authors (Branch I>et al. 1992).

The mechanism of pregnancy loss associated with antiphospholipid antibody remains uncertain. Progressive thrombosis of the microvasculature of the placenta and subsequent infarction resulting in placental insufficiency, fetal growth retardation and, ultimately, fetal loss, is a plausible explanation. Not all placentas examined, however, have shown areas of thrombosis or infarction and other mechanisms may be operative in these patients (Out *et al*. 1991). The aborted fetus is often normal except for evidence of growth retardation. It seems, therefore, that placental disease rather than fetal abnormality is responsible for fetal deaths related to antiphospholipid antibody. Pre-eclampsia is common in pregnant patients with the antiphospholipid syndrome and may provide clues to the pathogenesis of

pregnancy loss related to this antibody (Branch *et al*. 1992).

Thrombocytopenia

Thrombocytopenia is common in patients with antiphospholipid antibodies, though not severe enough to cause haemorrhage (Khamashta and Machin 1991). The platelet count often remains stable for many years; then, for reasons that are often obscure, the count drops, sometimes catastrophically. A survey of sera from patients presenting with idiopathic thrombocytopenic purpura, found anticardiolipin antibodies in 30 per cent of the cases (Harris *et al*. 1985). This finding raises the possibility that patients with the antiphospholipid syndrome may present only with severe thrombocytopenia and will later develop pregnancy loss or thrombosis. This form of presentation was observed in a very small number of our patients with the syndrome. Some patients with antiphospholipid antibodies and thrombocytopenia also develop haemolytic anaemia with positive direct Coombs' test. This is widely known as Evans' syndrome. In a study of 12 patients with Evans' syndrome and systemic lupus, 10 patients had evidence of antiphospholipid antibodies (Deleze *et al*. 1989). Similarly, in 70 patients with the primary antiphospholipid syndrome, 10 per cent were described as having Evans' syndrome (Asherson *et al*. 1989c).

Other manifestations

Epilepsy and chorea are less frequent manifestations of the antiphospholipid syndrome and have, intriguingly, been seen to improve in some patients treated with anticoagulants (Asherson *et al*. 1987; Herranz *et al*. 1994). Transverse myelopathy, though rare, is strongly associated with the presence of antiphospholipid antibodies (Alarcon-Segovia *et al*. 1989). Occasionally, in some patients with bizarre, transient/recurrent neurologial signs (resembling multiple sclerosis), antiphospholipid antibodies have been detected in the absence of other immunological abnormalities. Its recognition is important as anticoagulation therapy may be effective in these patients. Migraine is a common finding in patients with the antiphospholipid syndrome, and often pre-dates the diagnosis by many years. However, recent prospective studies have not demonstrated a significant statistical association between migraine headaches and the presence of antiphospholipid antibodies (Montalban *et al*. 1992; Tsakiris *et al*. 1993).

Heart valve disease, particularly mitral valve involvement, is strikingly associated with antiphospholipid antibodies (Khamashta *et al*. 1990; Cervera *et al*. 1991; Galve *et al*. 1992). In some cases this is due to a combination of valvular thrombosis and degeneration. In our prospective echocardiographic studies, the valves were involved in more than one-third of the patients with lupus or primary antiphospholipid syndrome. Most patients with heart valve disease associated with antiphospholipid antibodies are asymptomatic, though heart insufficiency requiring surgical valve replacement has been reported. Emboli from sterile valvular vegetations can cause multiple cerebral lesions. Large intracardiac thrombosis associated with antiphospholipid antibodies can mimic atrial myxoma.

One of the most striking physical signs in patients positive for antiphospholipid antibody is livedo reticularis (Fig. 5), sometimes widespread, sometimes subtle, e.g. confined to a small area on the back of the wrist. In one prospective study of patients with livedo reticularis, 43 per cent had anticardiolipin antibodies (Asherson *et al*. 1989d). We found that surprisingly few patients with Sneddon's syndrome (a triad of livedo reticularis, ischaemic cerebrovascular disease, and hypertension) were positive for antiphospholipid antibody, suggesting either that this syndrome results from many coagulopathies or that antiphospholipid antibodies are indeed important but that in our study the antibodies had long since disappeared. More dramatic skin manifestations associated with vascular thrombosis include widespread skin ulceration, notably in the lower extremities. Clinically, some patients with antiphospholipid antibodies may develop nail splinter haemorrhages (Fig. 6) and clubbing, posing major diagnostic difficulties, in those with heart valve disease, in differentiating from bacterial endocarditis (Mujic *et al*. 1995a).

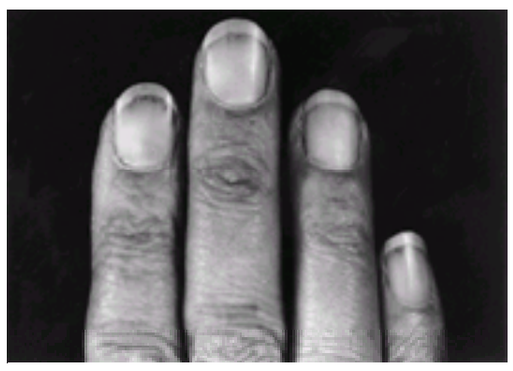


Fig. 6 Subungual splinter haemorrhages in a patient with primary antiphospholipid syndrome.

Avascular necrosis of bone is an uncommon complication in lupus patients and clearly associated with high steroid dosage. We have noted an increased risk of avascular necrosis in individuals positive for antiphospholipid antibody, possibly as a result of small arterial occlusions, notably of the head of the femur (Asherson *et al*. 1993b).

Many patients with the antiphospholipid syndrome seem to develop widespread arteriopathy. The systemic narrowing of major arteries is similar in many respects to the widespread endarterial disease seen in some patients after heart–lung transplantation. Thus, antiphospholipid antibodies might be associated with accelerated vascular disease, including atherosclerosis (Lahita *et al*. 1993; Vaarala *et al*. 1993; Vaarala *et al*. 1996).

Epidemiology

The epidemiology of antiphospholipid antibodies is still being investigated worldwide. Efforts are being made in clinics throughout the world to assess the importance of this factor in recurrent abortion, stroke, myocardial infarction, epilepsy, and so on. In cardiovascular disease, Hamsten *et al*. (1986) found that 21 per cent of young patients in Sweden with myocardial infarction were antiphospholipid positive—this figure has not been reached in other studies. A collaborative United Kingdom/Spanish study found a prevalence of antiphospholipid positivity of 6.8 per cent in a large cohort of patients with stroke (Montalban *et al*. 1991). More recently, Nencini *et al*. (1992) in a study of 'young' stroke patients found that 18 per cent were positive for antiphospholipid antibody. In many of the other specialties, there is not enough data for adequate analysis. In obstetrics, the associations are now more clearly defined, though still with wide disparity between series, due possibly more to variations in test standardization than to clinical selection. Lynch *et al*. (1994) in a recent, large, prospective cohort study of 389 nulliparous mothers assessed at study entry and delivery, showed that 95 (24 per cent) were positive for antiphospholipid antibody of which 15.8 per cent had pregnancy loss, compared with 6.5 per cent of the women who were negative for antiphospholipid antibody.

Antiphospholipid antibodies are detected in patients with a variety of autoimmune, infectious, malignant, and drug-induced disorders, as well as in some apparently healthy individuals. Other than patients with the antiphospholipid syndrome, the single disorder in which these antibodies have been reported most frequently is systemic lupus erythematosus, in which lupus anticoagulants are reported in 10 to 20 per cent of cases, and anticardiolipin antibodies in 20 to 40 per cent of cases (Love and Santoro 1990; Cervera *et al*. 1993).

The specificities of antiphospholipid antibodies probably differ in various disorders. Large retrospective studies of patients with thrombotic complications suggest that those with high concentrations of IgG anticardiolipin antibodies appear to be at greatest risk for thrombosis, whereas the risk of clotting appears to be much lower in patients with infection-related or drug-induced antiphospholipid antibodies. Families positive for antiphospholipid antibody exist, and HLA studies have suggested associations with DR7, DR4, DRw53, DQw7, and C4 null alleles (Wilson *et al*. 1988; Arnett *et al*. 1991; Asherson *et al*. 1992; Wilson *et al*. 1995).

Diagnosis

Tests for anticardiolipin and lupus anticoagulant are essential to a diagnosis of the antiphospholipid syndrome. This syndrome is best defined as the occurrence of venous or arterial thrombosis, pregnancy losses, or thrombocytopenia associated with persistently positive tests for anticardiolipin or lupus anticoagulant. Patients may have one, two, or all three clinical features present, but they must also be positive for at least one of the two laboratory tests (Table 4). Many of the patients reported to have the syndrome have lupus and can be regarded as having secondary antiphospholipid syndrome. Some patients do not have any underlying systemic disease. These patients may be regarded as having primary antiphospholipid syndrome (Alarcon-Segovia and Sanchez-Guerrero 1989a; Asherson *et al*. 1989c;

Mackworth-Young *et al*. 1989). For research and classification purposes, the term primary is useful, although there appear to be few differences in complications related to antiphospholipid antibody or in antibody specificity in the presence or absence of systemic lupus erythematosus (Vianna *et al*. 1994). Although some patients with primary antiphospholipid syndrome progress to systemic lupus, most do not show such progression (Mujic *et al*. 1995b).

Clinical	Laboratory
Venous thrombosis Arterial thrombosis Recurrent fetal loss Thrombocytopenia	IgG anticardiolipin antibodies (moderate/high titres) IgM anticardiolipin antibodies (moderate/high titres) Positive lupus anticoagulant test

[a]Patients with the syndrome should have at least one clinical plus one laboratory finding during their disease. Antiphospholipid antibody test must be positive on at least two occasions more than 3 months apart.

Table 4 Criteria for the classification of the antiphospholipid syndrome [a]

Apart from lupus and primary antiphospholipid syndrome, antiphospholipid antibodies can occur in a wide variety of rheumatic, inflammatory, drug-induced, malignant, or infectious disorders. Usually, they are of low titre, of the IgM isotype, and unassociated with thrombotic events (Asherson 1992b). Antiphospholipid antibodies, using standardized techniques, are detected in less than 1 per cent of apparently normal persons and in up to 3 per cent of the elderly population without clinical manifestations of the antiphospholipid syndrome.

Differential diagnosis

Careful family and personal history and physical examination of patients with unexplained thromboses are of the utmost importance as thrombotic events and, notably, venous thromboses, often have explanations other than the antiphospholipid syndrome. It is advisable that thrombophilia testing, as suggested in Table 5, be performed in all patients, before initiating anticoagulant treatment.

Full blood count including platelets and blood microscopy (e.g. schistocytes)
Protein C
Activated protein C resistance (ratio)
Protein S (free and total)
Antithrombin III
Activated partial thromboplastin time
Prothrombin time
Reptilase time
Fibrinogen concentration
Fibrin split products
Plasminogen concentration
Serum and urine homocysteine
Antinuclear antibodies
Complement profile (C3, C4 and CH50)
Lupus anticoagulant
Anticardiolipin antibodies (IgG and IgM)
Syphilis serology
Others[a] Anticardiolipin antibody (IgA) Anti-β_2-glycoprotein I antibody Antibodies against other phospholipids: antiphosphatidylserine, antiphosphatidic acid, antiphosphatidylinositol, antiphosphatidylethanolamine, antiphosphatidylcholine

[a]For patients with a clinical picture highly suggestive of the antiphospholipid syndrome but negative for standard tests for anticardiolipin antibodies and lupus anticoagulant.

Table 5 Laboratory investigation of unexplained thromboses

Deficiencies in protein C, protein S, or antithrombin III are usually associated with recurrent venous thromboses. Echocardiographic studies should be performed in patients with cerebral infarcts as arterial emboli originating from cardiac valvular lesions are particularly common.

The most striking feature of the antiphospholipid syndrome is the frequent observation of life-threatening thrombosis in the setting of thrombocytopenia. A number of other conditions can result in thrombocytopenia and thrombosis, including heparin-induced thrombocytopenia and thrombotic thrombocytopenic purpura.

Heparin-induced thrombocytopenia develops in 1 to 5 per cent of patients receiving standard heparin and somewhat less frequently in those who receive preparations of low molecular weight. Approximately 10 to 20 per cent of patients with substantial heparin-induced thrombocytopenia have venous or arterial thrombosis, including pulmonary, cardiac, and cerebral thrombosis. Thrombocytopenia typically develops approximately a week after exposure in persons not previously treated with heparin and sooner in those with previous exposure.

Thrombotic thrombocytopenic purpura is associated with neurological syndromes, but it is chiefly a microvascular disorder and confusion, seizures, or changes in the level of consciousness are more frequent than isolated, cerebral, large vessel thrombosis. Microangiopathic haemolysis with evidence of schistocytes in peripheral blood is usually a prominent finding. Elevation in the serum level of lactate dehydrogenase is found in nearly all cases and is a sensitive marker of the severity of the disorder.

We suggest that the tests for IgA anticardiolipin, anti-b $_2$-glycoprotein I, and antibodies against other phospholipids be ordered only for patients strongly suspected of having the antiphospholipid syndrome despite negative assays for IgG and IgM anticardiolipin and for lupus anticoagulant (Viard *et al*. 1992; Amengual *et al*. 1996).

Pathogenesis

Precisely how antiphospholipid antibodies relate to thrombosis is unknown. Several mechanisms have been proposed to explain the prothrombotic nature of the antiphospholipid syndrome. The range of possible mechanisms include effects of antiphospholipid antibodies on platelet membranes, on endothelial cells, and on clotting components such as antithrombin III, protein C, and protein S (Table 6) (Khamashta *et al*. 1989; Roubey 1994).

Decreased prostacyclin production and/or release by endothelial cells
Inhibition of factor XII/prekallikrein-mediated fibrinolytic activity
Decreased protein C activation
Decreased free protein S levels
Interference with thrombomodulin function on endothelial cells
Decreased antithrombin III activation
Increased platelet activation and aggregation
Interference with the function of plasma β_2-glycoprotein I
Increased tissue factor expression
Decreased function of annexin V (placental anticoagulant protein I)
Direct vascular damage
Complement activation
Vascular activation and release of von Willebrand factor multimer

Table 6 Proposed mechanisms of thrombosis mediated by antiphospholipid antibody

In 1990, three groups independently reported that the anticardiolipin antibodies detected by enzyme-linked immunosorbent assay are not directed against cardiolipin alone, because purified IgG from anticardiolipin-positive patients did not bind to cardiolipin unless a plasma protein cofactor was present (Galli *et al*. 1990; Matsuura *et al*. 1990; McNeil *et al*. 1990). This protein was b$_2$-glycoprotein I. It has since become clear that anticardiolipin antibody in patients with the syndrome is dependent on both cardiolipin and b$_2$-glycoprotein I for optimum binding, though the relative importance of the two molecules, or their combination, is uncertain. Recent data indicate that certain lupus anticoagulants also require a plasma protein cofactor. This cofactor has been identified as prothrombin or b $_2$-glycoprotein I (Bevers *et al*. 1991; Oosting *et al*. 1992).

b$_2$-Glycoprotein I (apolipoprotein H) is a 50-kDa protein present at approximately 200 µg/ml in normal plasma. Although its physiological role is not known, *in vitro* data suggest that b$_2$-glycoprotein I may play a role in coagulation. b $_2$-Glycoprotein I binds to anionic phospholipids and inhibits the contact phase of intrinsic blood coagulation, platelet aggregation dependent on adenosine diphospate, and the prothrombinase activity of platelets. Although these data suggest an anticoagulant role for b$_2$-glycoprotein I, deficiency of this protein is not a clear risk factor for thrombosis. Patients with antiphospholipid antibodies have normal or somewhat elevated levels of b$_2$-glycoprotein I (Roubey 1994).

b$_2$-Glycoprotein I has its critical role in the recognition of cardiolipin by anticardiolipin antibodies in patients with autoimmune disease, but not from patients with infection (Hunt *et al*. 1992; Hunt *et al*. 1994). Anticardiolipin antibodies associated with syphilis bind to cardiolipin in the absence of b $_2$-glycoprotein I but this binding is inhibited by its presence, presumably because the antibodies and b $_2$-glycoprotein I bind to similar phospholipid structures. This difference in antigenic specificity may explain why the autoimmune type of anticardiolipin antibody is associated with the lupus anticoagulant and clinical complications such as thrombosis and recurrent fetal loss, whereas anticardiolipin antibodies associated with infection are not (Roubey 1994).

One hypothetical model for the mechanism of thrombosis proposes persistent endothelial cell damage and/or platelet activation, resulting in increased exposure of anionic phospholipid surfaces. Plasma proteins such as b $_2$-glycoprotein I and prothrombin are relatively abundant and may bind to the lipid surface, exposing novel epitopes. The immune response is directed to modified plasma proteins, therefore, rather than to lipids (Fig. 7) (Comfurius *et al*. 1995).

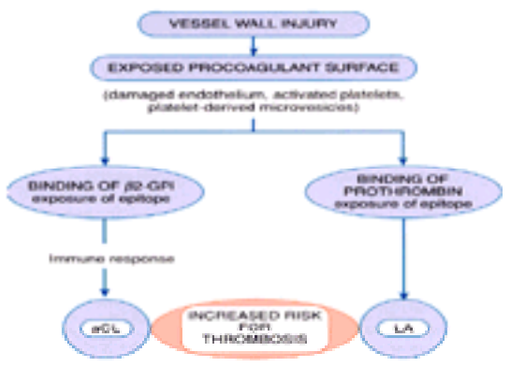


Fig. 7 Hypothesis explaining the generation of antiphospholipid antibodies as a normal response of the body to a potentially dangerous situation. This hypothesis predicts the possible existence of antibodies directed against all lipid-binding plasma proteins. As a consequence, it also explains the coexistence of anticardiolipin antibodies and lupus anticoagulant that is frequently observed in patients with the antiphospholipid syndrome. (Reproduced with the permission of the Editor of *Lupus* and from Dr P. Comfurius.)

We have recently established five, monoclonal, IgManticardiolipin antibodies from three patients with the antiphospholipid syndrome and showed that binding to anionic phospholipids was absolutely dependent on the presence of b $_2$-glycoprotein I (Ichikawa *et al*. 1994). Furthermore, a mixture of b$_2$-glycoprotein I and cardiolipin inhibited the binding of monoclonal anticardiolipin antibodies to b $_2$-glycoprotein I, but cardiolipin or b $_2$-glycoprotein I alone did not. These results suggest that anticardiolipin antibodies may recognize a cryptic epitope, which appears as a result of b $_2$-glycoprotein I binding to anionic phospholipids. Recent studies by Matsuura *et al*. (1994) confirmed that anticardiolipin antibodies bind to an epitope on b $_2$-glycoprotein I that is expressed when b$_2$-glycoprotein I binds to g-irradiated microtitre plates. Interestingly, immunization of healthy mice and rabbits with b $_2$-glycoprotein I resulted in high titres of anti-b $_2$-glycoprotein I and anticardilipin antibodies, whereas cardiolipin alone was not immunogenic (Gharavi *et al*. 1992).

It has recently been demonstrated that the fifth C-terminal domain of b $_2$-glycoprotein I contains the major phospholipid binding site, a region critical for binding anticardiolipin antibodies (Hunt *et al*. 1993). Using synthetic peptides spanning the fifth domain, it has been shown that the major phospholipid binding site is restricted to the sequence Cys^{281}–Lys–Asn–Lys–Glu–Lys–Lys–Cys^{288} (Hunt and Krilis 1994). A preparation of b$_2$-glycoprotein I clipped in the C-terminus between Lys^{317} and Thr^{318} completely loses both its ability to bind negatively-charged phospholipid and to act as a cofactor. In addition, anticardiolipin antibodies purified from patients with autoimmune disease could bind directly to b $_2$-glycoprotein I but not to wells coated with the preparation of b $_2$-glycoprotein I cleaved between Lys^{317} and Thr^{318}. These results were confirmed recently using monoclonal, anticardiolipin antibodies derived from patients with the antiphospholipid syndrome (Wang *et al*. 1995).

Animal models are providing useful clues to the pathogenesis of the antiphospholipid syndrome (Blank *et al*. 1991; Bakimer *et al*. 1992). Passive transfer and active immunization of BALB/c mice with human or mouse anticardiolipin monoclonal antibodies induced features of the antiphospholipid syndrome. The experimental antiphospholipid syndrome was characterized by serological markers (a panel of antiphospholipid antibodies, prolonged activated partial thromboplastin time, indicating the presence of the lupus anticoagulant), haematological findings (thrombocytopenia), and clinical manifestations (recurrent fetal resorption, the equivalent of human fetal loss). A role for antiphospholipid antibodies in causing thrombosis has been more difficult to demonstrate. Pierangeli and Harris (1994) provided recently the most persuasive evidence yet for a cause-and-effect role for antiphospholipid antibodies in thrombosis. In an *in vivo* mouse model, the femoral vein was exposed and damaged by a standardized 'pinch' injury. A fibreoptic transilluminator was placed underneath the vessel and the clotting process observed through a stereoscopic operating microscope. Infusion of anticardiolipin antibodies dramatically enhanced both the size of the thrombus formed locally as well as the time of persistence of the clot.

The basis of the heterogeneity of the clinical features of the antiphospholipid syndrome is not well understood. If one hypothesizes that these antibodies are pathogenic on the basis of their antigenic specificity, it is not clear why certain patients have venous thrombosis as opposed to arterial thrombosis or thrombocytopenia or recurrent fetal loss, or some combination of these manifestations (Roubey 1994). It is also not clear why only about 30 per cent of individuals with antiphospholipid antibodies will develop the clinical syndrome, although greater risk is associated with higher antibody titres (Harris *et al*. 1986; Ginsburg *et al*. 1992).

Treatment

One of the reasons for the widespread interest in the antiphospholipid syndrome has been its effect on approaches to therapy. Before its recognition, most features of systemic lupus erythematosus were attributed to inflammatory phenomena, requiring anti-inflammatory measures such as corticosteroids. Now it is recognized that features as diverse as fits, miscarriage, endocardial disease, and hypertension may all be the result of a thrombotic process. This concept has spread beyond the confines of systemic lupus, and pinpointing antiphospholipid-associated thrombosis and taking appropriate anticoagulation measures have become important considerations in specialties as diverse as neurology and obstetrics (Hughes and Khamashta 1994). Table 7 shows our preferred treatment for the different clinical features associated with antiphospholipid antibodies.

Clinical situation	Suggested treatment
1. Asymptomatic individuals	Observations ± low-dose aspirin
2. Thrombotic events	
Deep vein thrombosis ± pulmonary embolism	Life-long oral anticoagulants (INR = 3)
Large vessel arterial occlusion (e.g. stroke)	Life-long oral anticoagulants (INR = 3) ± low-dose aspirin
Transient ischaemic attack	Low-dose aspirin
Catastrophic antiphospholipid syndrome	Oral anticoagulants (INR = 3) ± plasmapheresis ± corticosteroids or immunosuppressants
3. Pregnancy	
No previous history of thrombosis or abortions	Observation ± low-dose aspirin
History of 1st trimester abortions	Low-dose aspirin
History of 2nd or 3rd trimester fetal loss	Low-dose aspirin ± subcutaneous heparin
Previous history of thrombosis ± pregnancy loss	Low-dose aspirin+subcutaneous heparin
4. Thrombocytopenia	
Mild (100 000–150 000)	Observation
Moderate (50 000–100 000)	Observation ± low-dose aspirin
Severe (< 50 000)	Corticosteroids ± intravenous gammaglobulin

INR, international normalized ratio

Table 7 Preferred treatment for the different clinical manifestations associated with antiphospholipid antibodies

Identification and treatment of additional risk factors for thrombosis

In treating patients with the antiphospholipid syndrome, attention should first be given to removal or reduction of other risk factors that might predispose to thrombosis. Treatment of hypertension and hypercholesterolaemia is required, along with advice to stop smoking and, for female patients, counselling against use of oral contraceptives containing oestrogen (Levine *et al*. 1992; Vianna *et al*. 1994). Prophylaxis with heparin administered subcutaneously at the time of surgery should be considered for all patients.

Prophylactic treatment of asymptomatic patients with antiphospholipid antibodies

Antiphospholipid antibodies persist for many years, possibly a lifetime. Thus, one of the unresolved key clinical questions is what additional factors lead to the sudden development of thrombosis, which occurs only in a minority of these patients (Hughes 1993).

The controversy concerning whether or not prophylactic treatment is indicated for patients with antiphospholipid antibodies who have no history of thrombosis remains unresolved. Most clinicians do not consider this sufficient reason for prophylactic anticoagulation, though prophylaxis should certainly be given to cover high-risk situations such as surgery. In our daily practice, we recommend that patients with a persistently positive test for lupus anticoagulant and/or with moderate to high levels of IgG anticardiolipin antibody and no history of thrombosis, be treated with low-dose aspirin (75 mg/day). These patients should be monitored carefully because they have the greatest risk of thrombosis (Khamashta and Wallington 1991; Hunt and Khamashta 1996).

In pregnancy, inappropriate treatment of women who do not have the antiphospholipid syndrome must be avoided. Anticardiolipin tests are positive at low titres in up to 2 per cent of the normal obstetric population, and such findings do not appear to be associated with an adverse outcome (Lockwood *et al*. 1989; Harris and Spinnato 1991). The risk is relatively low for primiparas, even if they have high titres of antiphospholipid antibody, and therefore their recommended prophylactic treatment is either close observation only or observation plus low-dose aspirin (Lockshin 1993).

Prevention of recurrent thrombosis

Optimal management of patients with thrombotic features associated with antiphospholipid antibodies remains a problem. Controlled treatment trials have proved difficult to perform in these patients because of the limited number of eligible patients with antiphospholipid syndrome available for study at a single centre and the necessity of long-term follow-up. There is now good evidence that those with thrombosis will be subject to recurrences, but there is still no consensus regarding the duration and extent of prophylactic antithrombotic treatments in these patients (Khamashta 1996). Many patients with the antiphospholipid syndrome in whom anticoagulation has been stopped have had major recurrent thromboses (Asherson *et al*. 1985; Rosove and Brewer 1992; Derksen *et al*. 1993). We have recently assessed our experience of the management of antiphospholipid-associated thrombosis in 147 patients over a 10-year period (Khamashta *et al*. 1995). Our study showed that long-term and intensive oral anticoagulation (international normalized ratio 3) was the most effective therapeutic option in the secondary prevention of venous and arterial thrombosis in these patients. Moreover, our study has shown that the first 6 months after stopping warfarin therapy were associated with the highest rate of recurrences (1.30 thrombotic events per year). This high probability of recurrent thromboses in patients with antiphospholipid antibodies and previous thromboembolic disease without oral anticoagulation, suggests that these patients require indefinite warfarin therapy. The benefits of long-term anticoagulation should, however, be balanced by the risks of bleeding. In our series, bleeding complications occurred in 29 patients (0.071 occasions per patient–year) and in 7 (0.017 occasions per patient–year) they were severe, suggesting that the benefits of warfarin in the antiphospholipid syndrome are greater than the risks. There was no evidence that low-dose aspirin (75 mg/day) alone prevented further thrombotic events in our patients. A similar finding was observed in the study of Rosove and Brewer (1992). It should be emphasized, however, that some patients with transient ischaemic attacks may respond to low-dose aspirin or dipyridamole therapy.

A significant number of patients require high doses of warfarin (up to 20 mg/day) to maintain the international normalized ratio in therapeutic range (3.0 to 4.0). In our experience, most of these patients were receiving other drugs and, notably, azathioprine at the same time as warfarin therapy. An important drug interaction has been pointed out recently between azathioprine and warfarin (Singleton and Conyers 1992; Rivier *et al*. 1993). When azathioprine is reduced or discontinued, anticoagulation may increase with the potential for bleeding if the international normalized ratio is not carefully monitored in these patients.

The role of steroids and immunosuppressive drugs in treatment of patients with antiphospholipid antibodies and thrombosis is uncertain. Such drugs have severe side-effects when given for prolonged periods and we, and others, have found that antiphospholipid antibodies are not always suppressed by these agents (Out *et al*. 1989; Harris *et al*. 1993). Furthermore, in our series of patients with the antiphospholipid syndrome, corticosteroids and immunosuppressive therapy, prescribed in some patients to control lupus activity, did not prevent further thrombotic events (Asherson *et al*. 1991; Khamashta *et al*. 1995). The use of these drugs is probably justified only in patients with repeated episodes of thrombosis despite adequate anticoagulant therapy, i.e. catastrophic antiphospholipid syndrome. In this rare but life-threatening condition, plasmapheresis has also been used (Asherson 1992; Asherson and Piette 1996).

It is of interest to note that our studies have shown that there were no broad significant differences in the levels of anticardiolipin antibodies between those patients who developed recurrent thromboembolic events and those who did not (Asherson *et al*. 1991; Khamashta *et al*. 1995). These findings further support the recommendation that treatment aimed at suppressing antibody formation is not warranted in patients with thrombosis associated with antiphospholipid antibodies.

The use of intra-arterial fibrinolysis has been described to be of benefit in patients with acute myocardial infarction associated with antiphospholipid antibodies. Prostacyclin analogues (iloprost) were also successfully used in patients with severe ischaemic necrotic toes associated with antiphospholipid syndrome (Zahavi *et al*. 1993). Elective pulmonary thromboendarterectomy can be very effective and lifesaving in selected patients with chronic large-vessel thromboembolic pulmonary hypertension (Sandoval *et al*. 1996).

Prevention of fetal losses

The treatment of patients with recurrent pregnany loss associated with antiphospholipid antibodies remains controversial and there are no clear guidelines for management. However, consensus is being reached on some aspects. An appropriate approach would be to assemble a multidisciplinary team (rheumatologist, obstetrician, and clinical haematologist), exclude causes of pregnancy loss other than the antiphospholipid syndrome, and select a drug regimen which has some reported efficacy and which will cause the least harm for the mother and fetus. Most importantly, the woman should be followed carefully throughout her pregnancy with frequent ultrasound evaluation for early detection of intrauterine growth retardation, and Doppler studies to monitor the flow-velocity waveforms of umbilical and uterine arteries (Harris 1990; Kerslake *et al*. 1992; Buchanan *et al*. 1993; Kamashta and Hughes 1996).

Over the past 7 years there has been a re-evaluation of the significance of antiphospholipid antibodies in pregnancy and, whilst an increased rate of fetal loss is seen in this group, experience suggests that with close monitoring this can be significantly reduced (Trudinger *et al*. 1988; McHugh *et al*. 1989). Recent data from our pregnancy clinic show that the greatest impact on fetal outcome in our series of patients with systemic lupus erythematous and antiphospholipid syndrome has been the regularity and quality of obstetric care, judicious monitoring, and timely intervention (Buchanan *et al*. 1992; Lima *et al*. 1995). In view of these findings, many of the high-dose (40 to 60 mg/day) prednisolone strategies, which were the treatments of choice in the early 1980s for patients with recurrent abortions associated with

antiphospholipid antibodies, are being abandoned. These doses regularly produce severe cushingoid effects, hypertension, and diabetic manifestations. Furthermore, controlled clinical trials now favour low-dose aspirin (75 mg/day) and subcutaneous heparin (10 000 to 15 000 IU/day) in those with a history of thrombosis, over corticosteroid therapy (Lockshin *et al*. 1989; Cowchock *et al*. 1992). In our recently reported study of pregnant women with the antiphospholipid syndrome, pregnancy outcome improved from 19 to 70 per cent using low-dose aspirin in all patients and subcutaneous heparin in those with previous thrombosis (Lima *et al*. 1996).

The presence of antiphospholipid antibodies in pregnant women in the absence of a previous history of thrombosis or fetal loss is not an indication for treatment. If, however, the patient does have a history of previous thrombosis and is being given warfarin, careful management is needed. Warfarin may be teratogenic, even early in the first trimester. The risks of continuing warfarin and also the problems with using heparin should be carefully explained to the patient. For the benefit of the fetus, warfarin should ideally be converted to subcutaneous heparin prior to conception, as heparin does not cross the placenta. Some women prefer to continue with warfarin until the first missed period, and then convert at this time.

Standard heparin or heparins of low molecular weight can be used during pregnancy, although the latter are not licensed for this purpose. There is no theoretical advantage of either. However, as one can monitor the levels of heparin with a low molecular weight very accurately with anti-Xa activity, we prefer to use them (Hunt *et al*. 1997). Their other practical advantage is that the patient can be started on one injection daily, in contrast to the necessary two or even three injections a day with standard subcutaneous heparin. Heparin doses need to be increased during pregnancy as the plasma volume expands. The patients can be converted back to warfarin after delivery.

Despite widespread use of heparin in pregnancy, there are relatively few reports of obstetric complications. The risks associated with heparin treatment include dose-related problems such as bleeding or osteopenia and idiosyncratic reactions such as immune thrombocytopenia, alopecia, and local reactions. The possibility of maternal osteoporosis is a limiting factor in the use of heparin. Heparin of low molecular weight, which has a long half-life and can be given once daily, may have less effect on bone. The combined use of corticosteroids and heparin may result in severe osteopenia.

Aspirin has been an integral part of treatment for women with the antiphospholipid syndrome. It was used on the basis of several reports suggesting that prostacyclin synthesis or release by endothelial cells was decreased *in vitro* by antiphospholipid antibodies. This would lead to an imbalance in the thromboxane/prostacyclin ratio and to platelet aggregation and vasospasm. Aspirin given in doses of under 150 mg/day preferentially blocks arterial thromboxane synthesis. Such dosages are associated with few (if any) maternal or fetal risks.

For patients who continue to have pregnancy loss despite a heparin and low-dose aspirin regimen, high-dose intravenous gammaglobulin (0.4 g/kg on 4 days each month) might be considered. The treatment should be initiated as soon as pregnancy is ascertained and continued throughout the pregnancy; it has proved very effective but is expensive. Fluid overload and hypertension are the major complications.

Clearly, optimal management of patients with recurrent pregnancy loss associated with antiphospholipid antibodies remains a problem, but the recent establishment of animal models for the antiphospholipid syndrome provides an excellent opportunity to develop rational and more-targeted therapies. A recent example is the demonstration by Fishman *et al*. (1993) that mice with experimental antiphospholipid syndrome are deficient in interleukin-3 and that fetal losses in these mice could be prevented by administration of recombinant interleukin-3. This novel approach has not yet been used in patients with antiphospholipid syndrome.

Treatment of thrombocytopenia

Mild thrombocytopenia with platelet counts between 100 000 and 150 000/mm^3 are common in patients with antiphospholipid antibody and usually does not require intervention (Khamashta and Machin 1991). In a minority of cases it can be severe. In these cases, corticosteroid therapy should always be the treatment of choice. However, it should be noted that there have been several case reports of peripheral thrombocytopenia unresponsive to steroid therapy, with platelet counts returning to normal following low-dose aspirin therapy (Alarcon-Segovia and Sanchez-Guerrero 1989b). Aspirin, by reducing the degree of spontaneous platelet activation, also presumably reduces binding of antiphospholipid antibody and immune-type platelet destruction (Khamashta *et al*. 1988). Intravenous gammaglobulin infusion, danazol, and dapsone have been given successfully to some patients with severe thrombocytopenia associated with antiphospholipid antibody. There is no published experience reported with splenectomy in these patients, though we have one patient (a 33-year-old male) with severe thrombocytopenia and antiphospholipid antibodies in whom splenectomy failed to improve the platelet count. This procedure should be considered with caution in view of an increased thromboembolic risk related to post-splenectomy thrombocytosis.

Antiphospholipid syndrome in childhood

There have been relatively few reports of clinical associations of antiphospholipid antibodies in children and the spectrum of clinical findings remains at present unknown. Antiphospholipid antibodies in childhood-onset systemic lupus erythematosus have been described in several small clinical reports, occurring in one-third of the patients. The clinical manifestations are similar to those encountered in adults, particularly recurrent deep-vein thrombosis, strokes, and chorea. Devastating thrombotic complications of the antiphospholipid syndrome in children have been reported including digital ischaemia and myocardial infarction (Tucker 1994). The risk of maternal transmission of antiphospholipid antibodies to infants during pregnancy is unknown, though there have been several case reports of thrombotic events in neonates of mothers with the antiphospholipid syndrome (Silver *et al*. 1992).

Prognosis

Although there are a number of studies of prognostic factors in systemic lupus erythematosus, only a few have addressed the possible role of the antiphospholipid syndrome in the mortality rates of these patients. Drenkard *et al*. (1994) have recently analysed the influence of the antiphospholipid syndrome in the survival of their series of 667 patients with systemic lupus erythematosus. Thrombocytopenia and arterial occlusions were the manifestations related to antiphospholipid antibody that were associated with decreased survival. The syndrome itself was also associated with increased mortality rate, independently of other variables. A negative impact of positivity for IgM anticardiolipin antibody on the probability of survival of lupus patients also was found in another study (Gulko *et al*. 1993).

Chapter References

Alarcon-Segovia, D. and Sanchez-Guerrero, J. (1989*a*). Primary antiphospholipid syndrome. *Journal of Rheumatology*, **16**, 482–8.

Alarcon-Segovia, D. and Sanchez-Guerrero, J. (1989*b*). Correction of thrombocytopenia with small dose aspirin in the primary antiphospholipid syndrome. *Journal of Rheumatology*, **16**, 1421–3.

Alarcon-Segovia, D., Deleze, M., Oria, C.V., Sanchez-Guerrero, J., and Gomez-Pacheco, L. (1989). Antiphospholipid antibodies and the antiphospholipid syndrome in systemic lupus erythematosus: a prospective analysis of 500 consecutive patients. *Medicine* (*Baltimore*), **68**, 353–65.

Amengual, O., Atsumi, T., Khamashta, M.A., Koike, T., and Hughes, G.R.V. (1996). Specificity of ELISA for antibody to b2-glycoprotein I in patients with antiphospholipid syndrome. *British Journal of Rheumatology*, **35**, 1239–43.

Amigo, M.C. *et al*. (1992). Renal involvement in the primary antiphospholipid syndrome. *Journal of Rheumatology*, **18**, 1181–5.

Arnett, F.C., Olsen, M.L., Anderson, K.L., and Reveille, J.D. (1991). Molecular analysis of major histocompatibility complex alleles associated with the lupus anticoagulant. *Journal of Clinical Investigation*, **87**, 1490–5.

Asherson, R.A. (1992*a*). The catastrophic antiphospholipid syndrome. *Journal of Rheumatology*, **19**, 508–12.

Asherson, R.A. (1992*b*). Antiphospholipid antibodies and syndromes. In *Systemic lupus erythematosus* (2nd edn) (ed. R.G. Lahita), pp. 587–635. Churchill Livingstone Inc., New York.

Asherson, R.A. and Hughes, G.R.V. (1991). Hypoadrenalism, Addison's disease and antiphospholipid antibodies. *Journal of Rheumatology*, **18**, 1–3.

Asherson, R.A. and Piette, J.C. (1996). The catastrophic antiphospholipid syndrome: acute multi-organ failure associated with antiphospholipid antibodies: a review of 31 patients. *Lupus*, **5**, 414–17.

Asherson, R.A., Chan, J.K.H., Harris, E.N., Gharavi, A.E., and Hughes, G.R.V. (1985). Anticardiolipin antibody, recurrent thrombosis and warfarin withdrawal. *Annals of the Rheumatic Diseases*, **44**, 823–5.

Asherson, R.A., Derksen, R.W.H.M., Harris, E.N., and Hughes, G.R.V. (1987). Chorea in systemic lupus and lupus-like disease: association with antiphospholipid antibodies. *Seminars in Arthritis and*

Rheumatism, **16**, 253–9.

Asherson, R.A. *et al*. (1989*a*). Cerebrovascular disease and antiphospholipid antibodies in systemic lupus erythematosus, lupus-like disease and the primary antiphospholipid syndrome. *American Journal of Medicine*, **86**, 391–9.

Asherson, R.A., Khamashta, M.A., Baguley, E., Oakley, C., Rowell, N.R., and Hughes, G.R.V. (1989 *b*). Myocardial infarction and antiphospholipid antibodies. *Quarterly Journal of Medicine*, **73**, 1103–15.

Asherson, R.A. *et al*. (1989*c*). The primary antiphospholipid syndrome: major clinical and serological features. *Medicine* (*Baltimore*), **68**, 366–74.

Asherson, R.A., Mayou, S.C., Merry, P., Black, M.M., and Hughes, G.R.V. (1989 *d*). The spectrum of livedo reticularis and anticardiolipin antibodies. *British Journal of Dermatology*, **120**, 215–21.

Asherson, R.A., Higenbottam, T.W., Dinh Xuan, A.T., Khamashta, M.A., and Hughes, G.R.V. (1990). Pulmonary hypertension in a lupus clinic: experience with twenty-four patients. *Journal of Rheumatology*, **17**, 1292–8.

Asherson, R.A., Baguley, E., Pal, C., and Hughes, G.R.V. (1991). Antiphospholipid syndrome: five year follow up. *Annals of the Rheumatic Diseases*, **50**, 805–10.

Asherson, R.A., Doherty, D.G., Vergani, D., Khamashta, M.A., and Hughes, G.R.V. (1992). Major histocompatibility complex associations with primary antiphospholipid syndrome. *Arthritis and Rheumatism*, **35**, 124–5.

Asherson, R.A., Khamashta, M.A., and Hughes, G.R.V. (1993 *a*). Hypertension and the antiphospholipid antibodies. *Clinical and Experimental Rheumatology*, **11**, 465–7.

Asherson, R.A. *et al*. (1993*b*). Avascular necrosis of bone and antiphospholipid antibodies in systemic lupus erythematosus. *Journal of Rheumatology*, **20**, 284–8.

Bakimer, R., Fishman, P., Blank, M., Sredni, B., Djaldetti, M., and Shoenfeld, Y. (1992). Induction of experimental antiphospholipid syndrome in mice by immunization with human monoclonal anticardiolipin antibody (H-3). *Journal of Clinial Investigation*, **89**, 1558–63.

Bevers, E.M., Galli, M., Barbui, T., Comfurius, P., and Zwaal, R.F.A. (1991). Lupus anticoagulant IgG's (LA) are not directed to phospholipids only, but to a complex of lipid-bound human prothrombin. *Thrombosis and Haemostasis*, **66**, 623–32.

Blank, M., Cohen, J., Toder, V., and Shoenfeld, Y. (1991). Induction of antiphospholipid syndrome in naive mice with mouse lupus monoclonal and human polyclonal anticardiolipin antibodies. *Proceedings of the National Academy of Science* (*USA*), **88**, 3069–73.

Branch, D.W. (1994). Thoughts on the mechanism of pregnancy loss associated with the antiphospholipid syndrome. *Lupus*, **3**, 275–80.

Branch, D.W., Silver, R.M., Blackwell, J.L., Reading, J.C., and Scott, J.R. (1992). Outcome of treated pregnancies in women with antiphospholipid syndrome: an update of the Utah experience. *Obstetrics and Gynecology*, **80**, 614–20.

Buchanan, N.M.M., Khamashta, M.A., Morton, K.E., Kerslake, S., Baguley, E.A., and Hughes, G.R.V. (1992). A study of 100 high risk lupus pregnancies. *American Journal of Reproductive Immunology*, **28**, 192–4.

Buchanan, N.M.M., Khamashta, M., Kerslake, S., Hunt, B.J., and Hughes, G.R.V. (1993). Practical management of pregnancy in systemic lupus erythematosus. *Fetal and Maternal Medicine Review*, **5**, 223–30.

Cervera, R. *et al*. (1991). High prevalence of significant heart valve lesions in patients with the primary antiphospholipid syndrome. *Lupus*, **1**, 43–7.

Cervera, R. *et al*. (1993). Systemic lupus erythematosus: clinical and immunologic patterns of disease expression in a cohort of 1000 patients. *Medicine* (*Baltimore*), **72**, 113–24.

Comfurius, P., Bevers, E.M., Galli, M., and Zwaal, R.F.A. (1995). Regulation of phospholipid asymmetry and induction of antiphospholipid antibodies. *Lupus*, **4** (Suppl. 1), 19–22.

Cowchock, F.S., Reece, E.A., Balaban, D., Branch, D.W., and Plouffe, L. (1992). Repeated fetal losses associated with antiphospholipid antibodies. A collaborative randomized trial comparing prednisolone with low-dose heparin treatment. *American Journal of Obstetrics and Gynecology*, **166**, 1318–23.

Deleze, M. *et al*. (1989). Hemocytopenia in systemic lupus erythematosus. Relationship to antiphospholipid antibodies. *Journal of Rheumatology*, **16**, 926–30.

Derksen, R.H.W.M., de Groot, P.H.G., Kater, L., and Nieuwenhuis, H.K. (1993). Patients with antiphospholipid antibodies and venous thrombosis should receive long term anticoagulant treatment. *Annals of the Rheumatic Diseases*, **52**, 689–92.

Drenkard, C., Villa, A.R., Alarcon-Segovia, D., and Perez-Vazquez, M.E. (1994). Influence of the antiphospholipid syndrome in the survival of patients with systemic lupus erythematosus. *Journal of Rheumatology*, **21**, 1067–72.

Fishman, P. *et al*. (1993). Prevention of fetal loss in experimental antiphospholipid syndrome by *in vivo* administration of recombinant interleukin-3. *Journal of Clinical Investigation*, **91**, 1834–7.

Galli, M. *et al*. (1990). Anticardiolipin antibodies (ACA) directed not to cardiolipin but to a plasma protein cofactor. *Lancet*, **336**, 1544–7.

Galve, E. *et al*. (1992). Valvular heart disease in primary antiphospholipid syndrome. *Annals of Internal Medicine*, **116**, 293–8.

Gharavi, A.E., Harris, E.N., Asherson, R.A., and Hughes, G.R.V. (1987). Anticardiolipin antibodies: isotype distribution and phospholipid specificity. *Annals of the Rheumatic Diseases*, **46**, 1–6.

Gharavi, A.E., Sammaritano, L.R., Wen, J., and Elkon, K.B. (1992). Induction of antiphospholipid autoantibodies by immunization with b $_2$-glycoprotein I. *Journal of Clinical Investigation*, **90**, 1105–9.

Ginsburg, K.S. *et al*. (1992). Anticardiolipin antibodies and the risk for ischemic stroke and venous thrombosis. *Annals of Internal Medicine*, **117**, 997–1002.

Gulko, P.S., Reveille, J.D., Koopman, W.J., Burgard, S.L., Bartolucci, A.A., and Alarcon, G.S. (1993). Anticardiolipin antibodies in systemic lupus erythematosus: clinical correlates, HLA associations and impact on survival. *Journal of Rheumatology*, **20**, 1684–93.

Hamsten, A., Norberg, R., Bjorkholm, M., de Faire, U., and Holm, G. (1986). Antibodies to cardiolipin in young survivors of myocardial infarction: an association with recurrent cardiovascular events. *Lancet*, **i**, 113–16.

Harris, E.N. (1990). Maternal antibodies and pregnancy: the antiphospholipid antibody syndrome. *Baillière's Clinical Rheumatology*, **4**, 53–68.

Harris, E.N. and Spinnato, J.A. (1991). Should anticardiolipin tests be performed in otherwise healthy pregnant women? *American Journal of Obstetrics and Gynecology*, **165**, 1272–7.

Harris, E.N. *et al*. (1983). Anticardiolipin antibodies: detection by radioimmunossay and association with thrombosis in systemic lupus erythematosus. *Lancet*, **ii**, 1211–14.

Harris, E.N. *et al*. (1985). Anticardiolipin antibodies in autoimmune thrombocytopenic purpura. *British Journal of Haematology*, **59**, 231–4.

Harris, E.N., Chan, J.K.H., Asherson, R.A., Aber, V.R., Gharavi, A.E., and Hughes, G.R.V. (1986). Thrombosis, recurrent fetal loss and thrombocytopenia. Predictive value of the anticardiolipin test. *Archives of Internal Medicine*, **146**, 2153–6.

Harris, E.N., Gharavi, A.E., Patel, S.P., and Hughes, G.R.V. (1987). Evaluation of the anticardiolipin antibody test: report of an international workshop held 4th April 1986. *Clinical and Experimental Immunology*, **68**, 215–22.

Harris, E.N. and the Kingston Anti-phospholipid Groups (KAPS) (1990). The second international anticardiolipin standardisation workshop/The Kingston Anti-phospholipid Antibody Study (KAPS) Group. *American Journal of Clinical Pathology*, **94**, 476–84.

Harris, E.N., Khamashta, M.A., and Hughes, G.R.V. (1993). Antiphospholipid antibody syndrome. In *Arthritis and allied conditions* (12th edn) (ed. D.J. McCarty and W.J. Koopman), pp. 1201–12. Lea and Febiger, Philadelphia.

Harris, E.N., Gharavi, A.E., Asherson, R.A., Khamashta, M.A., and Hughes, G.R.V. (1994). Antiphospholipid antibodies—middle aged but robust. *Journal of Rheumatology*, **21**, 978–81.

Herranz, M.T., Rivier, G., Khamashta, M.A., Blaser, K.U., and Hughes, G.R.V. (1994). Association between antiphospholipid antibodies and epilepsy in patients with systemic lupus erythematosus. *Arthritis and Rheumatism*, **37**, 568–71.

Hughes, G.R.V. (1983). Thrombosis, abortion, cerebral disease and lupus anticoagulant. *British Medical Journal*, **287**, 1088–9.

Hughes, G.R.V. (1993). The antiphospholipid syndrome: ten years on. *Lancet*, **342**, 341–4.

Hughes, G.R.V. and Khamashta, M.A. (1994). The antiphospholipid syndrome. *Journal of the Royal College of Phyicians of London*, **28**, 301–4.

Hughes, G.R.V. and Khamashta, M.A. (1996). *Management of the Hughes syndrome. Clinical and Experimental Rheumatology*, **14**, 115–17.

Hunt, B.J. *et al*. (1997). Thromboprophylaxis with low molecular weight heparin (Fragmin) in high risk pregnancies. *Thrombosis and Haemostasis*, **77**, 39–43.

Hunt, J. and Krilis, S. (1994). The fifth domain of b $_2$-glycoprotein I contains a phospholipid binding site (Cys^{281}–Cys^{288}), and a region recognized by anticardiolipin antibodies. *Journal of Immunology*, **152**, 653–9

Hunt, J.E., McNeil, H.P., Morgan, G.J., Crameri, R.M., and Krilis, S.A. (1992). A phospholipid-b $_2$-glycoprotein I complex is an antigen for anticardiolipin antibodies occurring in autoimmune disease but not with infection. *Lupus*, **1**, 83–90.

Hunt, J.E., Simpson, R.J., and Krilis, S.A. (1993). Identification of a region of b $_2$-glycoprotein I critical for lipid binding and anticardiolipin cofactor activity. *Proceedings of the National Academy of Sciences* (*USA*), **90**, 2141–5.

Hunt, J.E., Adelstein, S., and Krilis, S.A. (1994). New basic aspects of the antiphospholipid syndrome. *Clinical and Experimental Rheumatology*, **12**, 661–8.

Ichikawa, K., Khamashta, M.A., Koike, T., Matsuura, E., and Hughes, G.R.V. (1994). b $_2$-Glycoprotein I reactivity of monoclonal anticardiolipin antibodies from patients with the antiphospholipid syndrome. *Arthritis and Rheumatism*, **37**, 1453–61.

Kerslake, S. *et al*. (1992). Early Doppler studies in lupus pregnancy. *American Journal of Reproductive Immunology*, **28**, 172–5.

Khamashta, M.A. (1996). Management of thrombosis in the antiphospholipid syndrome. *Lupus*, **5**, 463–6.

Khamashta, M.A. and Asherson, R.A. (1995). Hughes syndrome—antiphospholipid antibodies move closer to thrombosis in 1994. *British Journal of Rheumatology*, **34**, 493–4.

Khamashta, M.A. and Hughes, G.R.V. (1993). Detection and importance of anticardiolipin antibodies. *Journal of Clinical Pathology*, **46**, 104–7.

Khamashta, M.A. and Hughes, G.R.V. (1996). Pregnancy in systemic lupus erythematosus. *Current Opinion in Rheumatology*, **8**, 424–9.

Khamashta, M.A. and Machin, S.J. (1991). Hematological immune cytopenias and antiphospholipid antibodies. In *Phospholipid binding antibodies* (ed. E.N. Harris, T. Exner, G.R.V. Hughes, and R.A. Asherson), pp. 247–54. CRC Press, Boca Raton, Florida.

Khamashta, M.A. and Wallington, T. (1991). Management of the antiphospholipid syndrome. *Annals of the Rheumatic Diseases*, **50**, 959–62.

Khamashta, M.A. *et al*. (1988). Immune mediated mechanism for thrombosis: antiphospholipid antibody binding to platelet membranes. *Annals of the Rheumatic Diseases*, **47**, 849–54.

Khamashta, M.A., Asherson, R.A., and Hughes, G.R.V. (1989). Possible mechanisms of action of the antiphospholipid binding antibodies. *Clinical and Experimental Rheumatology*, **7** (Suppl. 3), 85–9.

Khamashta, M.A. *et al*. (1990). Association of antiphospholipid antibodies with heart valve disease in systemic lupus erythematosus. *Lancet*, **335**, 1541–4.

Khamashta, M.A., Cuadrado, M.J., Mujic, F., Taub, N.A., Hunt, B.J., and Hughes, G.R.V. (1995). The Management of thrombosis in the antiphospholipid-antibody syndrome. *New England Journal of Medicine*, **332**, 993–7.

Lahita, R.G., Rivkin, E., Cavanagh, I., and Romano, P. (1993). Low levels of total cholesterol, high density lipoprotein, and apolipoprotein A1 in association with anticardiolipin antibodies in patients with systemic lupus erythematosus. *Arthritis and Rheumatism*, **36**, 1566–74.

Levine, S.R., Brey, R.L., Joseph, C.L.M., and Havstad, S. (1992). Risk of recurrent thromboembolic events in patients with focal cerebral ischemia and antiphospholipid antibodies. *Stroke*, **23** (Suppl. 1), 29–32.

Lie, J.T. (1994). Vasculitis in the antiphospholpid syndrome: culprit or consort? *Journal of Rheumatology*, **21**, 397–9.

Lima, F., Buchanan, N.M.M., Khamashta, M.A., Kerslake, S., and Hughes, G.R.V (1995). Obstetric outcome in systemic lupus erythematosus. *Seminars in Arthritis and Rheumatism*, **25**, 184–92.

Lima, F., Khamashta, M.A., Buchanan, N.M.M., Kerslake, S., Hunt, B.J., and Hughes, G.R.V (1996). A study of sixty pregnancies in patients with the antiphospholipid syndrome. *Clinical and Experimental Rheumatology*, **14**, 131–6.

Lockshin, M.D. (1993). Which patients with antiphospholipid antibody should be treated and how? *Rheumatic Disease Clinics of North America*, **19**, 235–47.

Lockshin, M.D., Qamar, T., Druzin, M.L., and Goei, S. (1987). Antibody to cardiolipin, lupus anticoagulant, and fetal death. *Journal of Rheumatology*, **14**, 259–62.

Lockshin, M.D., Druzin, M.L., and Qamar, T. (1989). Prednisolone does not prevent recurrent fetal death in women with antiphospholipid antibody. *American Journal of Obstetrics and Gynecology*, **160**, 439–43.

Lockwood, C.J. *et al*. (1989). The prevalence and biologic significance of lupus anticoagulant and anticardiolipin antibodies in a general obstetric population. *American Journal of Obstetrics and Gynecology*, **161**, 369–73.

Love, P.E. and Santoro, S.A. (1990). Antiphospholipid antibodies: anticardiolipin and the lupus anticoagulant in systemic lupus erythematosus (SLE) and non-SLE disorders: prevalence and clinical significance. *Annals of Internal Medicine*, **112**, 682–98.

Lynch, A. *et al*. (1994). Antiphospholipid antibodies in predicting adverse pregnancy outcome. *Annals of Internal Medicine*, **120**, 470–5.

Mackworth-Young, C.G., Loizou, S., and Walport, M.J. (1989). Primary antiphospholipid syndrome: features of patients with raised anticardiolipin antibodies and no other disorder. *Annals of the Rheumatic Diseases*, **48**, 362–7.

Matsuura, E., Igarashi, Y., Fujimoto, M., Ichikawa, K., and Koike, T. (1990). Anticardiolipin co-factor(s) and differential diagnosis of autoimmune disease. *Lancet*, **336**, 177–8.

Matsuura, E., Igarashi, Y., Yasuda, T., Triplett, D.A., and Koike, T. (1994). Anticardiolipin antibodies recognize b $_2$-glycoprotein I structure altered by interacting with an oxygen modified solid phase surface. *Journal of Experimental Medicine*, **179**, 457–62.

McHugh, N., Reilly, P.A., and McHugh, L.A. (1989). Pregnancy outcome and autoantibodies in connective tissue disease. *Journal of Rheumatology*, **16**, 42–6.

McNeil, H.P., Simpson, R.J., Chesterman, C.N., and Krilis, S. (1990). Antiphospholipid antibodies are directed against a complex antigen that includes a lipid-binding inhibitor of coagulation: b_2-glycoprotein I (apolipoprotein H). *Proceedings of the National Academy of Sciences* (*USA*), **87**, 4120–4.

Montalban, J., Codina, A., Ordi, J., Vilardell, M., Khamashta, M.A., and Hughes, G.R.V. (1991). Antiphospholipid antibodies in cerebral ischemia. *Stroke*, **22**, 750–3.

Montalban, J. *et al*. (1992). Lack of association between anticardiolipin antibodies and migraine in systemic lupus erythematosus. *Neurology*, **42**, 681–2.

Mujic, F., Lloyd, M., Cuadrado, M.J., Khamashta, M.A., and Hughes, G.R.V. (1995 *a*). Prevalence and clinical significance of subungual splinter haemorrhages in patients with the antiphospholipid syndrome. *Clinical and Experimental Rheumatology*, **13**, 327–31.

Mujic, F., Cuadrado, M.J., Lloyd, M., Khamashta, M.A., Page, G., and Hughes, G.R.V. (1995 *b*). Primary antiphospholipid syndrome evolving into systemic lupus erythematosus. *Journal of Rheumatology*, **22**, 1589–92

Nencini, P., Baruffi, M.C., Abbati, R., Massai, G., Amaducci, L., and Inzitari, P. (1992). Lupus anticoagulant and anticardiolipin antibodies in young adults with cerebral ischemia. *Stroke*, **23**, 189–93.

Oosting, J.D., Derksen, R.H.W.M., Entjes, H.T.I., Bouma, B.N., and DeGroot, P.G. (1992). Lupus anticoagulan activity is frequently dependent on the presence of b $_2$-glycoprotein I. *Thrombosis and Haemostasis*, **67**, 499–502.

Out, H.J., DeGroot, P.G., Hasselaar, P., van Vliet, M., and Derksen, R.H.W.M. (1989). Fluctuations of anticardiolipin antibody levels in patients with systemic lupus erythymatosus: a prospective

study. *Annals of the Rheumatic Diseases*, **48**, 1023–8.

Out, H.J., Kooijman, C.D., Bruinse, H.W., and Derksen, R.H.W.M. (1991). Histopathological findings in placentae from patients with intra-uterine fetal death and antiphospholopid antibodies. *European Journal of Obstetrics and Gynaecology and Reproductive Biology*, **341**, 179–86.

Pelletier, S. *et al*. (1994). Antiphospholipid syndrome as the second cause of non-tumorous Budd–Chiari syndrome. *Journal of Hepatology*, **21**, 76–80.

Permpikul, P., Mohan Rao, V., and Rapaport, S.I. (1994). Functional and binding studies of the roles of prothrombin and b $_2$-glycoprotein I in the expression of lupus anticoagulant activity. *Blood*, **83**, 2878–92.

Pierangeli, S.S. and Harris, E.N. (1994). Antiphospholipid antibodies in an *in vivo* thrombosis model in mice. *Lupus*, **3**, 247–51.

Piette, J.C., Cacoub, P., and Wechsler, B. (1994). Renal manifestations of the antiphospholipid syndrome. *Seminars in Arthritis and Rheumatism*, **23**, 357–66.

Rivier, G., Khamashta, M.A., and Hughes, G.R.V. (1993). Warfarin and azathioprine: a drug interaction does exist. *American Journal of Medicine*, **95**, 342.

Rosove, M.H. and Brewer, P.M.C. (1992). Antiphospholipid thrombosis: clinical course after the first thrombotic event in 70 patients. *Annals of Internal Medicine*, **117**, 303–8.

Roubey, R.A.S. (1994). Autoantibodies to phospholipid-binding plasma proteins: a new view of lupus anticoagulants and other 'antiphospholipid' antibodies. *Blood*, **84**, 2854–67.

Sandoval, J., *et al*. (1996). Primary antiphospholipid syndrome presenting as thromboembolic pulmonary hypertension. Treatment with thromboendartectomy. *Journal of Rheumatology*, **23**, 772–5.

Silver, R.K., MacGregor, S.N., Pasternak, J.F., and Neely, S.E. (1992). Fetal stroke asociated with elevated maternal anticardiolipin antibodies. *Obstetrics and Gynecology*, **80**, 497–9.

Singleton, J.D. and Conyers, L. (1992). Warfarin and azathioprine: an important drug interaction. *American Journal of Medicine*, **92**, 217.

Stewart, M.W. *et al*. (1993). Detection of antiphospholipid antibodies by flow cytometry: rapid detection of antibody isotype and phospholpid specificity. *Thrombosis and Haemostasis*, **70**, 603–7.

Stimmler, M.M., Coletti, P.M., and Quismorio, F.P. (1993). Magnetic resonance imaging of the brain in neuropsychiatric systemic lupus erythematosus. *Seminars in Arthritis and Rheumatism*, **22**, 335–49.

Triplett, D.A. (1994). Assays for detection of antiphospholipid antibodies. *Lupus*, **3**, 281–7.

Trudinger, B.J., Stewart, G.J., Cook, C.M., Connelly, A., and Exner, T. (1988). Monitoring lupus anticoagulant-positive pregnancies with umbilical artery flow velocity waveforms. *Obstetrics and Gynecology*, **72**, 215–8.

Tsakiris, D.A. *et al*. (1993). Lack of association between antiphospholipid antibodies and migraine. *Thombosis and Haemostasis*, **69**, 415–7.

Tucker, L.B. (1994). Antiphospholipid syndrome in childhood: the great unknown. *Lupus*, **3**, 367–9.

Vaarala, O. (1996). Antiphospholipid antibodies and atherosclerosis. *Lupus*, **5**, 442–7.

Vaarala, O., Alfthan, G., Jauhiainen, M., Leirisalo-Repo, M., Aho, K., and Palosuo, T. (1993). Cross-reaction between antibodies to oxidised low-density lipoprotein and to cardiolipin in systemic lupus erythematosus. *Lancet*, **341**, 923–5.

Vianna, J.L. *et al*. (1994). Comparison of the primary and secondary antiphospholipid syndrome: a European multicenter study of 114 patients. *American Journal of Medicine*, **96**, 3–9.

Viard, J.P., Amoura, Z., and Bach, J.P. (1992). Association of anti-b $_2$ glycoprotein I antibodies with lupus-type circulating anticoagulant and thrombosis in systemic lupus erythematosus. *American Journal of Medicine*, **93**, 181–6.

Wang, M.X. *et al*. (1995). Epitope specificity of monoclonal anti-b $_2$ glycoprotein I antibodies derived from patients with the antiphospholipid syndrome. *Journal of Immunology*, **155**, 1629–36.

Wilson, W.A., Perez, M.C., Michalski, J.P., and Armatis, P.E. (1988). Cardiolipin antibodies and null alleles of C4 in black Americans with systemic lupus erythematosus. *Journal of Rheumatology*, **15**, 1768–72.

Wilson, W.A. *et al*. (1995). Familial anticardiolipin antibodies and C4 deficiency genotypes that co-exist with MHC DQB1 risk factors. *Journal of Rheumatology*, **22**, 227–35.

Zahavi, J., Charach, G., Schafer, R., Toeg, A., and Zahavi, M. (1993). Ischaemic necrotic toes associated with antiphospholipid syndrome and treated with iloprost. *Lancet*, **342**, 862.

5.8 Scleroderma and related disorders in adults and children

Carol M. Black and Christopher P. Denton

'Scleroderma', a word meaning hard skin, is a part of many syndromes including localized, limited, and generalized scleroderma. Related to these are environmentally induced 'scleroderma-like' diseases, overlap syndromes, undifferentiated autoimmune rheumatic disease, and localized fibroses (Table 1). Most of these disorders in the scleroderma spectrum can occur at any stage of life, although the pattern of scleroderma occurring in childhood is different from that in the adult (see below). The disease we call scleroderma or systemic sclerosis has been confused with other diseases that have cutaneous features resembling it, for example scleromyxoedema, scleroedema, and primary amyloidosis (Table 2).

Raynaud's phenomenon	Primary Raynaud's
	Secondary Raynaud's
Scleroderma — localized	Morphoea (plaque, guttate, generalized)
	Linear
	En coup de sabre
Scleroderma — systemic	Limited cutaneous systemic sclerosis
	Diffuse cutaneous systemic sclerosis
	Scleroderma sine scleroderma
Chemically induced*	Environmental/occupational
	Drugs
Scleroderma-like diseases	Metabolic
	Immunological/inflammatory
	Localized systemic sclerosis and visceral diseases

*See Table 2

Table 1 Spectrum of scleroderma and scleroderma-like syndromes

Metabolic/inherited	Scleredema of Buschke Scleromyxoedema } with or without paraproteinaemia
	Insulin-dependent diabetes mellitus (digital sclerosis)
	Carcinoid syndrome
	Acromegaly
	Phenylketonuria
	Porphyrias (congenital and porphyria cutanea tarda)
	Lichen sclerosis et atrophicus
	Acrodermatitis chronica atrophicans
	Amyloidosis
	Inherited premature ageing syndromes — progeria, Rothmund's, Werner's
Immunological/inflammatory	Chronic graft-versus-host disease
	Eosinophilic fasciitis
	Overlap syndromes (systemic sclerosis with rheumatoid arthritis, systemic lupus, Sjögren's, polymyositis/dermatomyositis)
	Undifferentiated autoimmune rheumatic disease
Localized systemic sclerosis and visceral diseases	Idiopathic pulmonary fibrosis
	Amyloidosis
	Sarcoidosis
	Infiltrating carcinomas
	Infiltrating panniculopathy
	Oesophageal and intestinal hypomotility syndromes

Table 2 Scleroderma-like syndromes

The milestones in the history of scleroderma (see Table 3) bear testimony to the gradual realization of the heterogeneity of the disorder. The last 50 years have seen an explosion of interest in the disease and its mechanisms, which has fostered collaboration between many branches of basic science and clinical medicine. For a more detailed discussion of the fascinating history of this disease the reader is referred to the excellent historical review by Rodnan, the 'father' of modern-day clinical scleroderma (Rodnan and Benedek 1962).

1753	Curtis	Description of young woman of Naples with 'excessive hardness of skin — possibly scleroderma, but probably scleredema of Buschke'
1847	Grimaud	First use of name 'Sclerodermie'
1847	Forget	First description of pelvic involvement in scleroderma
1854	Addison	First description of linear scleroderma
1862	Raynaud	Description of 'local asphyxia and symmetrical gangrene of the extremities'
1878	Weber	Coexistence of scleroderma and calcinosis noted
1892	Osler	Tendency for scleroderma patients to die of pulmonary or renal disease noted
1903	Hutchinson	Association of scleroderma and Raynaud's phenomenon noted
1903	Ehrmann	Association of scleroderma and dysphagia recorded
1910	Thibierge and Weissenbach	'Rediscovery' of coexistence of scleroderma and calcinosis
1924	Rosow	First clear description of visceral involvement, with sclerosis of lungs, gastrointestinal tract and kidneys
1943	Weiss	Clear description of myocardial involvement in scleroderma
1945	Goetz	Coined the term 'progressive systemic sclerosis'
1964	Winterbauer	Described the CREST subset (calcinosis, Raynaud's, oesophagitis, sclerodactyly, and telangiectasia)

Table 3 History of scleroderma

The group of syndromes called localized scleroderma includes morphoea (limited and guttate), linear scleroderma with the *en coup de sabre* variety, and generalized morphoea. These syndromes are almost never associated with systemic involvement but may demonstrate abnormal autoimmune serology and inflammatory

histological changes in skin (Falanga 1989). These forms of scleroderma predominate in childhood.

Raynaud's phenomenon now best classified as either primary or secondary (the terms Raynaud's disease and Raynaud's syndrome should be abandoned and replaced by these more accurate, simple descriptive terms) and can be a forerunner one of the autoimmune rheumatic diseases such as generalized scleroderma (see below).

When generalized scleroderma is called systemic sclerosis the term is preferable to progressive systemic sclerosis because not all systemic sclerosis is progressive.

Systemic sclerosis is a multisystem disease, predominantly affecting females and of unknown cause. It is characterized pathologically by the overproduction of several elements of the connective tissue, notably collagen (Mauch *et al*. 1993), widespread vascular damage with the development of microvascular obliteration, and tissue infiltration of mononuclear inflammatory cells often in a perivascular distribution (Prescott *et al*. 1992). It embraces a clinical spectrum ranging from widespread skin thickening (diffuse scleroderma) to skin thickening either limited to the face and distal extremities (limited scleroderma) or absent (systemic sclerosis *sine* scleroderma).

There is no single diagnostic test for systemic sclerosis and, for the purposes of separating it from other autoimmune rheumatic diseases and identifying cases to permit comparison of reported series, preliminary criteria were developed and published in 1980 (Subcommittee for Scleroderma Criteria of the American Rheumatism Association Diagnostic and Therapeutic Criteria Committee 1980; Pope and Bellamy 1993). These criteria had a 97 per cent sensitivity and 98 per cent specificity for definite systemic sclerosis, but they are less sensitive for limited cutaneous disease, which is the largest subset, and they fail to identify at least 10 per cent of such patients. In the future, criteria taking full account of the limited-disease subset, and also including clinical, pathological, and immunological data (such as recent-onset Raynaud's phenomenon, capillary abnormalities, and circulating antibodies) should be developed. There is a readily identifiable group of presclerotic patients who should be included in diagnostic criteria, classification, and subsetting.

Classification of systemic sclerosis is difficult (Masi 1988), and a number of different systems have been proposed (LeRoy *et al*. 1988). Central to all of them is the extent of skin involvement, and whether a two-subset or three-subset model is superior is still debated. Currently the most widely used classification defines two subsets, based on the extent of skin involvement together with a number of reliable clinical laboratory and natural history associations. The two-subset model divides the disease into limited cutaneous systemic sclerosis and diffuse cutaneous systemic sclerosis (Table 4) (LeRoy *et al*. 1988). Over 60 per cent of patients with systemic sclerosis fall into the limited cutaneous subset, where visceral involvement is a late event and tends to occur 10 to 30 years after the onset of Raynaud's. The term limited cutaneous systemic sclerosis is preferable to CREST (**C** alcinosis, **R** aynaud's, (O)**E** sophageal dysphagia, **S** clerodactyly, **T** elangiectasia syndrome), because cutaneous manifestations often extend beyond sclerodactyly and calcinosis may be present only late or radiologically. Diffuse cutaneous systemic sclerosis, the more serious form of the disease, is much more rapid in onset, with organ failure often present within 5 years of the first symptoms.

Table 4 Classification of the subsets of systemic sclerosis (SSc)

This model will almost certainly be changed and developed as knowledge of the pathogenesis of the disorder advances, and the ever-increasing numbers of autoantibodies are matched to clinical subsets. Within each subset, there is great variability in the pace of the disease: for example, some patients with diffuse cutaneous systemic sclerosis develop extensive internal-organ complications within 2 to 4 years; others have widespread skin disease but only grumbling interstitial lung disease. Some patients with limited cutaneous systemic sclerosis may never develop clinically apparent pulmonary hypertension or mid-gut disease whereas other individuals develop this complication, usually late on, but occasionally as early as 5 to 7 years after diagnosis. Thus, not only is there disease heterogeneity, but differential progression within a subset. Notwithstanding these problems a useful and practical scheme is to divide the disease into early and late stages (see Table 5). This is a useful concept because it permits the doctor, patient, and patient's family to anticipate certain developments, which will concentrate attention on early detection and correct management of complications.

Table 5 Characteristic findings in the early and late stages of systemic sclerosis subsets

Epidemiological considerations

Scleroderma is an uncommon disorder and virtually all of the descriptive epidemiology is derived from retrospective or prospective reviews of patients attending hospitals or institutions serving a defined denominator population: there is only one true population-based study (Maricq *et al*. 1989). A summary of salient epidemiological data is given in Table 6 (Silman 1995). The demographic conclusions that may be drawn from these data are that the disease is rare in childhood, and that its incidence increases steadily with age amongst adults. Its rarity suggests that the genetic and/or environmental exposures necessary for disease susceptibility occur infrequently in the population. Scleroderma is a disorder with a female excess (4F:1M overall, and in the child-bearing years 15F:1M. The disease occurs most frequently and severely in young black women, but overall there are no prominent racial, seasonal, or socioeconomic differences. Of great interest is the possibility that the disease clusters in certain areas. A clustering has been described close to international airports in the United Kingdom (Silman 1995), a cluster of patients with a variety of autoimmune rheumatic diseases is reported from The Republic of Georgia (Freni-Titulaer *et al*. 1989), and a clearer cluster of systemic sclerosis in a region of Italy close to Rome (Valesini *et al*. 1993). Unlike some other autoimmune rheumatic diseases, the genetic contribution to the disease would appear to be relatively weak, although it is stronger for individual organ involvement (e.g. the lung). Finally, again in contrast to other autoimmune rheumatic diseases, a growing number of environmental agents have been implicated in systemic sclerosis (see Table 7).

Mortality	USA 1968–77	1.5 males and 3.5 females per million population
	UK 1974–85	1.9 males and 4.0 females per million population
Incidence (new cases detected/ population at risk/time period)	Rates variable Before 1980 report	2–10 per million population per year
	Pittsburgh study (1963–82)	10.9 per million population per year
	UK 1980–85	3.7 per million population per year
Prevalence (number of cases living at particular time or during time interval)	Estimates 1947–52 1977–79 1985 1986 1986 1988	4 per million (hospital studies) 128 per million (hospital studies) 253 per million (community studies) 292 per million (point prevalence) 36.8 per million (hospital and community) 93 per million (hospital study)
Survival		
Cumulative survival rates using life-table methods	Overall	65–70% @ 5 years 40–59% @ 10 years
	By subset localized cutaneous diffuse cutaneous	80% @ 5 years 50% @ 12 years 30% @ 5 years 10% @ 12 years
	Older patients: reduced survival	

Factors associated with decreased survival: increasing skin involvement, organ involvement etc.

Table 6 Epidemiological facts in systemic sclerosis

Table 7 Environmental agents implicated in the development of scleroderma

Aetiology and pathogenesis

Although the basic aetiology of systemic sclerosis is unknown, it is almost certainly multifactorial with genetic and environmental factors playing a part (Briggs *et al*. 1990). As mentioned earlier, systemic sclerosis occurs mostly in females and sex may be the strongest genetic marker. There are, however, several lines of evidence that indicate familial or genetic predisposition to systemic sclerosis. First, although rare, there are familial clusters of systemic sclerosis and related diseases, particularly Raynaud's phenomenon. Autoantibodies associated with systemic sclerosis are found in high frequency in blood relatives of patients with systemic sclerosis, although the incidence of antinuclear antibodies in spouses suggests an environmental component. Thirty-six per cent of relatives with antinuclear antibodies had clinical features of autoimmune rheumatic disease not observed in spouses (Briggs *et al*. 1993). The presence of autoantibodies in the blood of patients' spouses in the United Kingdom study was not confirmed in an Australian study. Thirdly, many centres worldwide have observed abnormal frequencies of the major histocompatibility complex (**MHC**) antigens associated with systemic sclerosis. The association is complex (see Table 8) and the strongest link is between HLA-DRw52a and patients with systemic sclerosis who have lung fibrosis (relative risk 16.7).

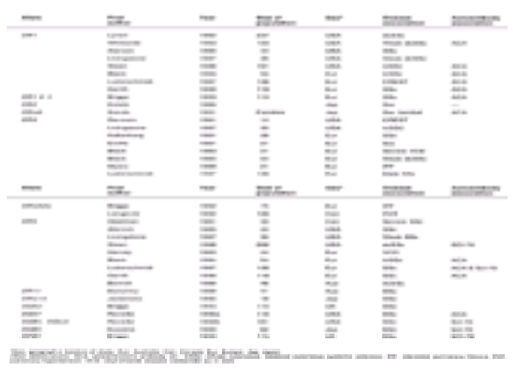


Table 8 Immunogenic associations of scleroderma

As one of the primary roles of MHC class II molecules is the presentation of processed antigen to the T-cell receptor on helper T lymphocytes resulting in an antigen-specific immune response, autoantibody subsets in scleroderma might be expected to show correlations with class II MHC polymorphism and indeed they do, although again, they appear to be complex. In the initial reports, anticentromere antibody was found to be associated with HLA-DR5 (-DR11), -DR4 (-D13 subtypes), -DR1, and -DR8. These findings appear to reflect linkage disequilibrium of *HLA-DR5* (*-DR11*) and many *HLA-DR4* (*-D13* subtypes) with *HLA-DQ7*, and *-DR1* with *-DQ5*. These HLA-DR specificities share no unique amino-acid sequences, which raised the possibility that another linked gene might be more highly correlated with this antibody response. More recently, in a study by Reveille *et al*. (1992a), oligotyping showed that the anticentromere antibody response was most closely associated with *HLA-DQB1* alleles in linkage disequilibrium with *HLA-DR1, -DR4, -DR5* (*DR11*), and *-DR8*. These *HLA-DQB1* alleles had in common a polar tyrosine or a glycine at position 26 of the outermost domain of the HLA-DQB molecule, as opposed to a hydrophobic leucine residue. In a British study (Briggs *et al*. 1993) implication of the *HLA-DQB1* locus was inferred, as virtually all of the anticentromere antibody-positive patients had either HLA-DR1 or -DR4. However, McHugh *et al*. (1994) have indicated that, although at least one *HLA-DQB1* allele not coding at position 26 of the first domain appears necessary, it may not be sufficient for the generation of anticentromere antibody. Reveille and colleagues have also extended the known associations of antitopoisomerase antibodies with HLA-DR5, -DR2, and -DR52a to include four *HLA-DQB1* alleles (Reveille *et al*. 1993b). Japanese workers have found a similar allele association (Kuwana *et al*. 1993). Localization of 'susceptible epitopes', however, has been less definitive. Suggestions include an American population in which a tyrosine at position 30 or the TRAELDT sequence spanning positions 71 to 77 of the *HLA-DQB1* outermost domain is seen (Reveille *et al*. 1992), and a Japanese population with systemic sclerosis in which a tyrosine at position 26 of the *HLA-DQB1* outermost domain is present (Kuwana *et al*. 1993). In a British study an HLA-DPB1 association was suggested, with the presence of an acidic amino-acid residue at position 69 in the third hypervariable region of the outermost domain. Autoantibodies to the anti-Pm/Scl antigen (see below) are nearly 100 per cent correlated with the presence of the *HLA DR3-DQw2* haplotype (Reveille *et al*. 1992).

Important racial and ethnic differences in the frequencies of these autoantibodies have been described, with Western Europeans and North American whites having a significantly higher frequency of anticentromere autoantibodies and a lower frequency of antitopoisomerase-1 than American blacks, Choctaw native Americans, Thai, and Italians (Kallenberg 1994). Particular forms of the disease may not be so strongly influenced by sex. For example, certain chemically induced systemic sclerosis-like disorders tend to be associated with males, partly due to an occupational bias. The range of known agents that can induce a systemic sclerosis-like disease is large and growing (Table 7), although it must be noted that in formal epidemiological studies, no excess was found. It is almost certain that sporadic cases can follow certain occupational exposures, but that both the absolute and attributable risks are low. There are some MHC associations with the environmentally induced cases, for example toxic oil syndrome is characterized by a raised incidence of HLA-DR4, while vinyl chloride disease is primarily associated with HLA-DR5, HLA-DR3 being a marker of severity. There is (see Table 8) inconsistency in the frequencies for genetic markers of susceptibility between patients from different study centres. This may reflect at least in part different agents present at these various geographical locations. The newest arrivals to the group of environmentally induced, systemic sclerosis-like syndromes are the eosinophilic–myalgic syndrome associated with the oral ingestion of the essential amino acid l-tryptophan (discussed elsewhere) and a renewed interest in the possible association between disease development and silicone breast implants. Much of the work in this area has been generated by the need to obtain accurate data because of considerable litigation, particularly in the United States. Hochberg (1994) has carefully reviewed the epidemiological aspects of the literature and has reasonably concluded that none of the available studies has demonstrated a statistical association between

augmentation mammoplasty with silicone gel-filled prostheses and scleroderma or other autoimmune rheumatic disorders.

Autoimmune considerations

An increasing number of immune abnormalities are being reported in systemic sclerosis and the designation of systemic sclerosis as an autoimmune disease now has widespread support. Some of the clinical features of scleroderma bear similarities to other autoimmune disorders such as systemic lupus erythematosus, dermatomyositis, and rheumatoid arthritis, and there are also patients who have overlap syndromes or who have sequential development of more than one autoimmune rheumatic disease. Cases of familial systemic sclerosis and familial associations of systemic sclerosis with other autoimmune diseases such as rheumatoid arthritis and systemic lupus erythematosus occur and have already been mentioned.

There is considerable evidence that abnormalities in both humoral and cell-mediated immunity occur in systemic sclerosis, although the precise importance of these immunological events in the pathogenesis remains uncertain. Some of this evidence is summarized in Table 9, and discussed below. The lack of a generalized immune dysfunction in systemic sclerosis suggests that the derangement of immune-cell dysfunction may be *specific* to certain antigens or cell types (Padula *et al*. 1986; Lupoli *et al*. 1990).

Immunogenetic associations (see Table 8)
Autoantibody associations (see Table 10)
Increased circulating levels of soluble CD4, soluble IL2 receptor and lymphocyte-derived cytokines (IL2, IL4, IL6, IL8) in some patients
Circulating T cells that are reactive to laminin, collagen (type I), and show increased adhesion to endothelial cells and fibroblasts
Elevated levels of IL5 and IL8 in bronchoalveolar lavage fluid
Perivascular T-cell infiltrates (DR+, 1, 2 integrin expression) in lesional skin
Increased tissue expression of lymphocyte-binding cell-surface adhesion molecules in lesional and prelesional skin and lung tissue
Clinical and pathological similarities with chronic graft-versus-host disease

IL, interleukin

Table 9 Summary of evidence for immune-system involvement in systemic sclerosis

The association of systemic sclerosis with particular major HLA antigens and the close association of certain HLA alleles with scleroderma-specific antibodies (see Table 8) is indirect evidence for T-cell involvement in systemic sclerosis. There is considerable evidence for T-cell activation in systemic sclerosis, including an increased ratio of circulating CD4+:CD8+ cells (Degiannis *et al*. 1990; White 1994), reflecting an increased number of CD4+ and/or a reduced number of CD8+ lymphocytes. A particular role for gd T-cells has been suggested (White 1994) and others have reported increased numbers of lymphokine-activated killer and natural killer cells (Kantor *et al*. 1992) in blood samples from patients with systemic sclerosis. Furthermore, several studies have found increased soluble interleukin-2 receptor in scleroderma, sometimes appearing to correlate with disease activity (Kahaleh 1991). Support for the possibility that activated T-cells are important in pathogenesis is provided by the presence of infiltrates of CD3+, CD4+, CD450+, interleukin 2-producing, HLA-DR-positive+, leucocyte function-associated antigen-1-positive, a/b^{+} T cells in lesional tissues (Prescott *et al*. 1992). Also, chronic graft-versus-host disease in humans shows several histological and clinical similarities with systemic sclerosis, and is known to be a T-cell-mediated process (Chosidow *et al*. 1992). Humoral abnormalities in systemic sclerosis are most clearly reflected by the presence of autoantibodies with well-defined target epitopes; mapping of the precise binding sites for some of these is currently being undertaken in several centres (Bona and Rothfield 1994). Although circulating immune complexes have been reported in systemic sclerosis, most studies have not found functional complement abnormalities (Seibold *et al*. 1982), probably because most systemic sclerosis-associated autoantibodies do not activate the complement cascade (White 1994).

Nevertheless, autoantibody production is an early and almost universal feature of systemic sclerosis. The number of autoantibody targets identified in systemic sclerosis continues to grow (see Table 10) but the major ones are topoisomerase-1, centromeric proteins, and RNA polymerases I, II, and III. About 97 per cent of patients have detectable antinuclear antibodies when HEp-2 lines are used as the detection tissue. Characteristic staining patterns for antinuclear antibodies within the nuclear and subnuclear structures are relatively specific (Tan 1989) and can be confirmed by more sophisticated tests. A diffusely grainy pattern of staining is associated with the presence of antibodies to topoisomerase-1 (Scl-70), a nuclear enzyme important in the unwinding of DNA for replication and RNA transcription. Antibodies to Scl-70 occur in up to 40 per cent of patients with diffuse systemic sclerosis and 15 per cent of those with limited disease. The occurrence of this antibody varies considerably between laboratories and studies have shown that the immunoblot technique is more sensitive than indirect immunofluorescence and should be the 'gold standard' for this test. An anticentromere staining pattern occurs in up to 80 per cent of patients with the limited form of systemic sclerosis. Antigens recognized by positive sera have been identified as CENP-A, CENP-B, and CENP-C, with molecular weights of 19, 80, and 140 kDa, respectively (Earnshaw *et al*. 1986). A correlation has been shown (Jabs *et al*. 1993) between anticentromere antibodies and aneuploidy in patients with systemic sclerosis, and it is possible that anticentromere antibodies could disrupt centromere function and allow chromosomes to segregate inappropriately during mitosis, leading to a high rate of chromosomal breakage and sister chromatid exchange, although to date no correlation has been found between the presence of anticentromere antibody and chromosomal changes. Anti-RNA polymerase (**RNAP**) antibodies are the latest systemic sclerosis-specific antibodies to be described. They occur mainly in patients with diffuse disease, and antibodies against RNAPI, -II, and -III have been described (Bona and Rothfield 1994). The RNAPs are multiprotein complexes and are components of the transcription complex (Reeves *et al*. 1994). Each RNAP is composed of collections of smaller proteins shared by other RNAPs and two large distinct proteins. RNAPI synthesizes ribosomal RNA precursors in the nucleoli, whereas RNAPs II and III are found in the nuclei. RNAPII synthesises most of the small nuclear RNAs found in ribonucleoprotein particles that mediate pre-mRNA splicing and synthesize precursors of mRNA, and RNAPIII synthesizes small RNAs including single-strand ribosomal RNA and transfer RNA. Anti-RNAP antibodies target both the smaller shared subunits and the larger distinct proteins, which explains antibody reactivity against several RNAPs in one serum sample.

Antigen	Molecular identity	Immunofluorescence pattern	Disease subtype and frequency
Scl 70 (topoisomerase I)	100-kDa protein degrades to 70 kDa	Nuclear (diffuse fine speckles)	Up to 40% with diffuse cutaneous SSc[a] 10–15% limited
Kinetochore-centromere	17-, 80-, 140-kDa proteins at inner and outer kinetochore plates	Centromeric	75–80% in limited cutaneous SSc 5–20% with primary biliary cirrhosis
RNA polymerase I, II, III	Complex of 12 proteins 12.5–210 kDa	Nucleolar (punctate)	20% — especially diffuse; high prevalence of internal organ involvement
PM-Scl	Complex of 11 proteins 20–110 kDa	Nucleolar (homogeneous)	3% — high prevalence of myositis, more renal involvement
To or Th	65-kDa protein associated with 7.2 & 8.2 kDa RNAs	Nucleolar (homogeneous)	Rare — localized cutaneous SSc
Mitochondrial M2	70-kDa protein — dihydrolipoamide acyltransferase	Cytoplasmic (rod like)	25% of CREST[a] (95% primary biliary cirrhosis)

[a]Eds: systemic sclerosis; [a]CREST (see Table 8); RNAP, ribonucleoprotein

Table 10 Autoantibodies in scleroderma

The relation between HLA status and autoantibody production is also of increasing interest in scleroderma. It would appear that certain of these antibodies are closely related to particular HLA alleles, for example it has recently been shown that class II MHC haplotype is an important factor determining *in vitro* responsiveness to topoisomerase antigen, both in patients with systemic sclerosis and in healthy control individuals (Kuwana *et al*. 1995). It is important to consider that there may be racial differences in HLA associations for the various autoantibodies (see Table 8). These antibodies not only mark out certain subsets of patients with systemic sclerosis but are of increasing importance in defining a subgroup of Raynaud's patients likely to develop scleroderma.

The antinuclear antibody profile is not as clear-cut in juveniles as in the adult form, although some trends are emerging. Serum antinuclear antibodies have been reported in 25 to 55 per cent of juveniles with localized scleroderma, the association being most marked in the linear group and in patients with extensive cutaneous

lesions. Antibodies to single-stranded DNA also appear to be correlated with the extent of localized disease, whereas antibodies to double-stranded DNA are rarely found. It is interesting that in the generalized form of childhood scleroderma no anticentromere antibodies have been reported, even in those children with disease identical to that found in the adult.

A pathogenetic role for autoantibodies in systemic sclerosis has long been sought. Defined epitopes seem to be targets for the autoantibodies and there has been recent work showing homology between target autoantigens in systemic sclerosis and retroviral proteins, suggesting molecular mimicry, which may have significance in disease pathogenesis. There are reports that some of these antibodies are also able to enter intracellular compartments (Levine *et al*. 1991; Ma *et al*. 1991) and thereby to mediate intracellular events, such as the reported ability of anticentromeric antibodies to disrupt the centromere. In addition, autoantibodies might be able to activate cells that bear the target autoantigens; for example, patients with systemic sclerosis produce antibodies that bind *Fc* gRI (CD64), II (CD32), and III (CD16) (Boros *et al*. 1993). It has been suggested that some of the autoantibodies in serum from systemic sclerosis may mediate antibody-dependent cytotoxicity, and potential effector cells have been found in the skin of some patients (White 1994). Another speculation is that these antibodies might contribute to the pathology if they mediated complement-dependent cellular lysis or phagocytosis. However, these ideas must be kept in perspective balanced against the lack of correlation of antibody titre with disease duration or activity. Circulating antibodies to the extracellular matrix proteins, collagens I, III, IV and VI, and laminin have also been found in systemic sclerosis, but their role is undetermined (Mackel *et al*. 1982).

Macrophages, mast cells, eosinophils, and basophils are found in increased numbers and in an activated state in tissues of patients with systemic sclerosis. These cells are capable of producing soluble mediators and can thereby modify endothelial and fibroblast function; for example, mast cells produce histamine, which stimulates both proliferation and matrix synthesis by fibroblasts and causes retraction of endothelial cells.

The initiating stimulus in idiopathic scleroderma is unknown, although the identification of chemical precipitants for environmentally induced systemic sclerosis as discussed above (e.g. vinyl chloride and epoxy resin) may provide some clues to the processes involved, particularly in view of the similar immunogenetic associations for both idiopathic and chemically induced disease (Black *et al*. 1983).

The most obvious major targets for the immune response in systemic sclerosis are endothelial cells and the fibroblasts. Stimulation of collagen synthesis could involve an increasing number of cytokines known to modulate the properties of fibroblasts. It is possible that cascades of such cytokines or autocrine/paracrine loops stimulate or maintain the disease process. It is now appreciated that the repertoire of mediators and cytokines produced by immune cells, fibroblasts, and endothelial cells is large. It is possible that the aberrant properties of connective tissue cells (e.g. excess synthesis of collagen, fibronectin, and glycosaminoglycans) and the endothelial-cell damage and vasculopathy, are consequences of the immunological events in systemic sclerosis.

To date, the only well-established animal model of an immune, scleroderma-like disorder is that of the chronic graft-versus-host disease in mice (Bocchieri and Jimenez 1990). There is also a human counterpart following bone marrow transplantation in which patients develop Raynaud's, dermal sclerosis, and vasculopathy (Roumm *et al*. 1984; Chosidow *et al*. 1992). Although these models are interesting, they should be viewed with caution. Graft-versus-host disease is certainly induced by T cells but the final lesion is damage to endothelium, epithelium, or both, related to the actions of natural killer cells, lymphokines (e.g. interleukin 1) and tumour necrosis factor. This damage can occur in many organs, for example lung, gut and skin, and can therefore mimic the end-point of many damaging processes. Such mimicry may be totally irrelevant to the initiating process or even the damaging events.

Pathogenesis

Although the pathogenesis of systemic sclerosis is still uncertain, it is likely that the development of both the *fibrous* and the *vascular* lesion is complicated and involves events that may occur simultaneously or in sequence (Fig. 1).

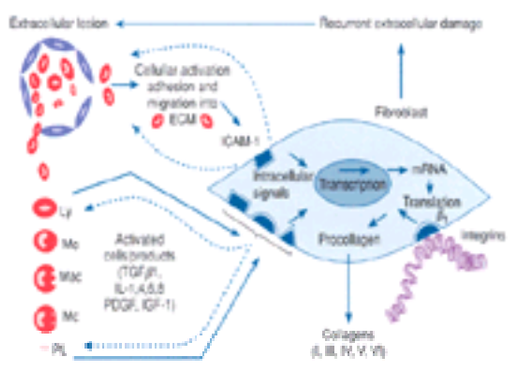


Fig. 1 Hypothesis for the pathogenesis of the scleroderma lesion. Complex interactions between cells of the immune system, vasculature, and connective tissue matrix lead to the development of an activated fibrogenic fibroblast phenotype.

Fibrosis is a hallmark of a number of diseases that includes scleroderma, pulmonary fibrosis, atherosclerosis, liver cirrhosis, and keloids. It is important to remember that the formation of fibrous tissue can be a normal physiological response, for example in wound healing. However, in fibrotic disorders the regulation of this normal response is altered, and in systemic sclerosis it is a widespread, non-organ-specific phenomenon. Excess deposition of collagen and extracellular matrix protein in the skin and internal organs of patients with systemic sclerosis was first demonstrated histopathologically many years ago, and this was confirmed by physical and biochemical means. Subsequently, techniques for culturing fibroblasts have provided valuable insight into the mechanisms involved in the synthesis of extracellular matrix components. Skin fibroblasts from scleroderma, or at least a subset of them, synthesize increased quantities of fibronectin, proteoglycan core proteins, and particularly collagens types I and III and to a lesser degree IV and VI.

Fibroblasts grown from areas of dermal sclerosis continue to synthesize increased amounts of collagen for several passages *in vitro* (LeRoy 1974; Uitto *et al*. 1979). It has also been demonstrated that the amounts of mRNA for these matrix proteins are increased and have been localized predominantly to areas surrounding dermal blood vessels. Nucleic-acid hybridization techniques have since confirmed that not all fibroblasts are activated to produce more normal collagen but rather a group of high-collagen producers is responsible. The increase in collagen RNA could arise through increased an transcription rate or by a reduction in the breakdown of the mRNA, and there is evidence that both mechanisms may operate. The transcriptional rate of genes encoding pro-a2(I) collagen (*COL1A2*) is increased in fibroblasts from systemic sclerosis, suggesting a change in regulatory transcription factors in these cells.

The mechanism for transcriptional regulation of collagen synthesis is still unproven but De Crombrugghe *et al*. (1990) have demonstrated a pathway that could have a direct bearing on the transcriptional defect in scleroderma. Following stimulation by transforming growth factor-b, fibroblasts containing the promoter–enhancer regions of the a2(I) collagen gene linked to a chloramphenicol acetyl transferase reporter show a marked enhancement of collagen gene expression. This appears to be through the interaction of a transacting DNA-binding protein, nuclear factor 1, and the specific promoter of the collagen gene. Such elegant experiments provide a prototype for the study of the activation of a pleomorphic genetic response such as fibrosis, which is not, of course, restricted to systemic sclerosis. Neither is the cytokine stimulus likely to be transforming growth factor-b alone. The activation and perpetuation of systemic sclerosis is a complex affair and other mediators such as platelet-derived growth factor, interleukins 1 and 4, and fibroblast growth factors may be involved. The cytokines possibly act in a cascade, and the timing and site of a biopsy might be critical for its correct interpretation.

One aspect of control that seems to be normal in fibroblasts from systemic sclerosis is that provided by negative feedback from propeptides. In addition, it should be noted that fibroblasts in scleroderma have changes in their proliferative properties. Scleroderma fibroblasts, unlike those from normal skin, display persistent proliferation in serum-free medium, but they are not transformed or immortalized cells, and they have similar doubling times, life-span and monolayer culture patterns to those of normal fibroblasts. What is, however, without doubt is that they are dysregulated with respect to the synthesis of extracellular matrix; whether this abnormality is wholly acquired or whether it is partially inherited and then activated by an inciting event is unknown.

Vascular lesions

Vascular injury is critical to the pathogenesis of systemic sclerosis and may be the primary event (Campbell and LeRoy 1975). The damage in systemic sclerosis is widespread and can be recognized as:

1. Vasomotor instability or Raynaud's phenomenon with repeated 'transient' interruption of tissue perfusion in the digits and internal organs (systemic Raynaud's), which is often an early event in disease development.
2. Microvascular abnormalities with structural changes characterized by proliferative intimal arterial lesions and obliteration of the vessels, leading to chronic ischaemia. Vascular damage can be visualized in the nailfold capillaries but it is also present in the small blood vessels of virtually all the viscera, muscle, subcutaneous tissues, and skin.
3. Intravascular pathology that is manifest by decreased red-cell deformability, increased platelet activity, and enhanced thrombus formation.

Many factors may be important in the vascular damage but it is the endothelial cell that is thought to have a pivotal role. The endothelium is now known to produce numerous molecules (see Table 11) and to regulate many aspects of vascular stability including control of vascular tone, permeability, thrombotic potential, and leucocyte trafficking (Pearson 1990; Kahaleh and Mattuci-Cerinic 1995).

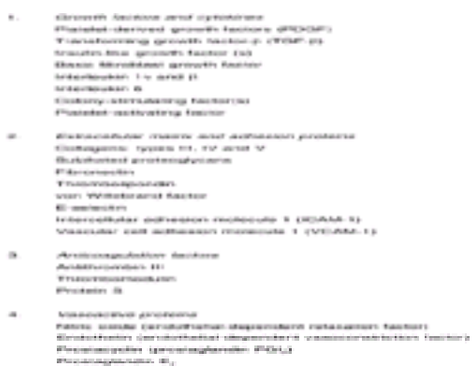


Table 11 Molecules synthesised by the vascular endothelium

Evidence for endothelial-cell dysfunction

The considerable evidence for endothelial-cell dysfunction in scleroderma is summarized below:

1. direct observation of abnormal nailfold capillaries;
2. increased capillary permeability to tracer molecules, with a slowing of flow and increased periods of stasis;
3. changes in circulating levels of endothelial-cell products such as von Willebrand factor, endothelin 1, plasminogen activator, angiotensin-converting enzyme, and prostacyclin/thromboxane metabolites;
4. the presence of a circulating cytotoxic factor identified as granzyme 1, a serine protease present in granules of activated T cells;
5. the presence of autoantibodies that bind to endothelial cells—these autoantibodies are distinct from other circulating antibodies characteristic of scleroderma.
6. increased endothelial cell-surface expression *in vivo* and elevated circulating levels of the adhesion molecules, intracellular adhesion molecule 1 (**ICAM-1**), vascular cell adhesion molecule 1 (**VCAM-1**), and E-selectin.

The mechanism for vasospasm in systemic sclerosis is likely to be complex. Interactions between extracellular matrix products such as nitric oxide, endothelin 1 and prostacyclins, platelet-released products (serotonin and b-thromboglobulin) and neuropeptides (calcitonin gene-related peptide and vasoactive intestinal polypeptide) may all contribute to the abnormal vascular tone (Kahaleh and Mattuci-Cerinic 1995).

The mechanism for the development of injury in the extracellular matrix is unknown but both immune (cell-mediated and humoral) and non-immune cytotoxicity have been implicated, and recently, several reports have shown that extracellular matrix can be damaged before there is an obvious inflammatory infiltrate. Prescott *et al*. (1992) studied sequentially the pathological changes in the perivascular spaces in the skin of patients with scleroderma and normal controls. Functional and structural endothelial change with subendothelial oedema was recognized as the first defect, followed by platelet aggregation and lymphocyte migration of both CD+ and CD8+ T cells. Tissue fibrosis occurred after the inflammation had subsided. Supporting evidence for this finding comes from the work of Harrison *et al*. (1991) in which damage to endothelial/epithelial surfaces is shown to be the first ultrastructural change to occur in lung biopsies from patients with systemic sclerosis. Lung sections appearing normal under light microscopy, with no evidence of infiltrating inflammatory cells, were, on electron microscopy, shown to contain widespread endothelial/epithelial damage. The nature of the primary vascular trigger to immune stimulation is critical to our understanding of the disease: environmental agents or an endothelium-seeking virus are possible candidates. In addition, intense vasospasm, which occurs both in the extremities and in internal organs in patients with systemic sclerosis, could lead by reperfusion injury and free-radical damage to structural and functional change in the endothelium and subsequent immune activation (Blann *et al*. 1993).

Following injury to, and activation of, the vascular endothelium adhesion molecules such as E-selectin, VCAM-1 and ICAM-1 are up-regulated in response to cytokines and other factors. These endothelial adhesion molecules bind to specific ligands on T and B lymphocytes, platelets, neutrophils, monocytes and natural killer cells, facilitating their adhesion to vascular endothelium and subsequent migration through what have now become 'leaky vessels' into the extracellular matrix, ultimately with the potential for fibroblast activation. Therefore if endothelial change could be detected, and stabilized at an early stage, it would almost certainly influence the clinical expression and progression of the disease.

The clinical picture

Scleroderma, as discussed earlier, is not one condition but a spectrum of heterogeneous conditions occurring at any age and including localized and systemic forms. Within each subtype the rate of progression and extent of damage varies. Occasionally, localized scleroderma may become systemic and sometimes localized scleroderma and eosinophilic fasciitis merge or overlap, either appearing together or sequentially. In addition, there are sclerotic conditions induced by a variety of occupational, environmental, and metabolic stimuli.

Juveniles may develop any form of scleroderma, but fortunately, there is a predilection for the localized forms in which the skin, subcutaneous fascia, muscle, and bone are the main organs to be attacked. In the localized form the skin at first shows a marked inflammatory reaction, followed by matrix deposition, fibrosis, and ultimately atrophy. The linear lesions are most associated with subcutaneous involvement and growth defects, but in general there are no systemic features (Hanson 1976).

Childhood-onset scleroderma is rare in comparison to the adult disease and to juvenile chronic arthritis. Fewer than 3 per cent of all cases of scleroderma are childhood-onset (Dabich *et al*. 1974), and such children comprised fewer than 3 per cent of all patients seen in a paediatric rheumatology clinic (Hanson 1976). Ascertainment of childhood scleroderma may be biased by referral patterns and subspecialty orientation. In a paediatric rheumatology centre, systemic sclerosis is seen much less frequently than localized scleroderma, approximately one case to every 15 of localized disease. In a busy dermatology clinic, morphoea is more commonly seen and scleroderma *en coup de sabre* may be misdiagnosed as alopecia areata: consequentially, the real prevalence of these conditions is unknown. Like its adult counterpart, childhood-onset scleroderma occurs in all races, with a female predominance. There appears to be no significant familial incidence. HLA studies with sufficient numbers of children in each group are only now being undertaken, and preliminary information presented would suggest that the HLA associations in the childhood disease are quite different from those found in adult scleroderma.

Localized scleroderma

This is separated from systemic sclerosis not only by the absence of vasospasm, structural vascular damage and involvement of internal organs, but also by the

distribution of the dermal lesions, which may, depending on the subtype, follow a dermatomal pattern. The varied clinical features have led to the separation of three main varieties of localize scleroderma, morphea, linear, and *en coup de sabre*.

Morphoea

This may be circumscribed or generalized. In circumscribed morphoea (Fig. 2) there may be just one or two lesions with no generalized spread. The changes often begin with small, violaceous or erythematous skin lesions, which enlarge and progress to firm 'hidebound' skin with variable degrees of hypo- or hyperpigmentation. These lesions eventually settle into a waxy, white appearance with subsequent atrophy. Pruritis is often a problem with the early lesion. Lesions vary in diameter between 1 and 10 cm. The condition generally resolves within 3 to 5 years, although sometimes a patch may persist for over 25 years. In generalized morphoea there are many patches covering a large surface area. The acral parts are usually spared, but the trunk and legs are often involved. Generalized morphoea can be disfiguring and may continue to extend, resulting in contractures, disability, and troublesome ulceration that may occasionally become malignant. In guttate morphoea there are multiple small, hypopigmented and pigmented papules 2 to 10 mm in diameter, with minimal sclerosis, and the lesions closely resemble those of lichen sclerosus et atrophicus. These lesions usually localize to the neck, shoulders, and anterior chest wall.

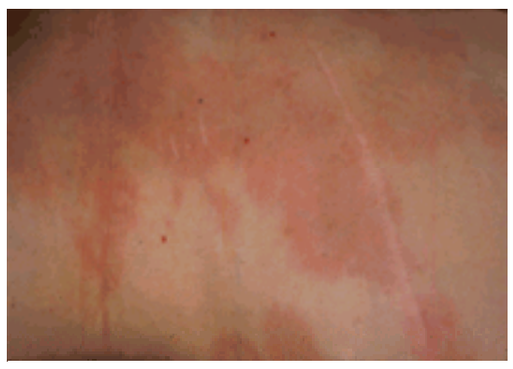


Fig. 2 Localized morphoea showing discrete lesions with central depigmentation and circumferential inflammation.

Linear scleroderma

The sclerotic areas occur in a linear, band-like pattern, often in a dermatomal distribution (Fig. 3). They often cross joint lines, and are associated with atrophy of the soft tissue, muscle, periosteum, bone, and occasionally synovium; they can lead to extensive growth defects in a limb or a part thereof, which can be extremely disfiguring. Fixed valgus or various deformities also occur, and scoliotic changes in the spine can develop as a result of inequalities in limb length. If the toes or fingers are involved, 'hammer toes' or 'claw hand' may develop. All of these changes are much more noticeable in a growing child and most cases of linear systemic sclerosis tend to occur in childhood, as do most cases of *en coup de sabre*, a specialized form of linear disease.

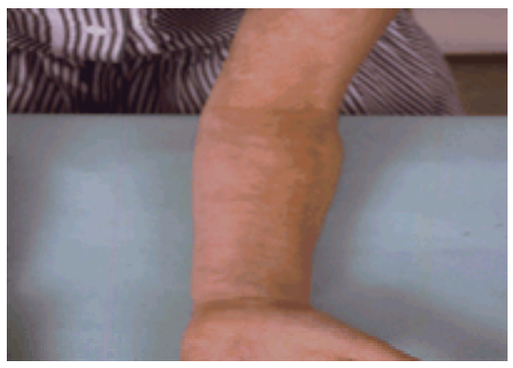


Fig. 3 Linear scleroderma occurring in childhood. Defective growth of the involved limb if a major clinical problem in childhood-onset disease.

En coup de sabre

Linear scleroderma occurring on the face or scalp may assume a depressed, ivory appearance. The lesion was considered originally reminiscent of the scar from a sabre wound so it was termed '*coup de sabre*'. The linear lesion is often associated with hemiatrophy of the face on the same side. It may also be associated with vascular abnormalities of the brain (David *et al*. 1991), and also with morphoea lesions elsewhere. It is also not uncommon for patients to present with morphoea and then later develop linear lesions (Falanga *et al*. 1986). This evolution should be anticipated extremely carefully, as the linear lesions tend to have much greater morbidity than the circumscribed patches of morphoea. The linear lesions may be quietly progressive for a long period, and lengthy follow-up is important.

In addition, in children there has been described a small group with morphoea and/or linear lesions who also have a synovitis, which can be demonstrated by infrared thermography (Allen *et al*. 1987). These patients have a raised erythrocyte sedimentation rate, rheumatoid factor, and circulating autoantibodies. Such cases are unusual, but they have an accelerated course with rapid development of contractures. An additional intermediate and interesting group of juveniles was described by Ansell *et al*. (1976). These children's disease is often mistaken for polyarticular juvenile chronic arthritis, since they may present with extensor and/or flexor tendon nodules in the hands, nodules at the elbows, knees or ankle joints, stiffness and a limitation of joint movement, but with little evidence of synovitis. At the time of presentation, there may only be a small area of localized or linear scleroderma, distant from the joint symptoms. The erythrocyte sedimentation rate and rheumatoid factor are usually normal, but autoantibodies are often present. There is both a clinical and a biochemical association, the nature of which is unclear, between localized scleroderma and eosinophilic fasciitis, in which large sclerotic patches may also occur.

Evaluation of all forms of scleroderma is difficult. In the localized forms, charting of the involved areas is often cumbersome and imprecise. However, the size of the lesion can be recorded, leg length, limb circumferences, and posture can be monitored, and muscle function and neurological status assessed. Charting of new lesions is also essential. In addition, thermography can be used to assess the activity of localized disease (Birdi *et al*. 1992).

Raynaud's phenomenon and the presclerotic state

It is fortunate that almost all cases of generalized systemic sclerosis arise in patients with Raynaud's phenomenon. The overall prevalence of this phenomenon has been variably assessed as between 3 and 10 per cent of adults worldwide, although it may affect as many as 20 per cent of young women. The prevalence varies somewhat depending on climate, skin colour, ethnic background, and occupational exposure to vibrating machines (Belch 1989).

The clinical syndrome was first described by Maurice Raynaud in 1862 as episodic digital ischaemia provoked by cold and emotion. It is classically manifest by episodic pallor of the digits followed by cyanosis, suffusion, and/or pain and tingling. The blanching reflects vasospasm in the digital vessels, the cyanosis the deoxygenation of static venous blood, and the redness reactive hyperaemia following the return of blood flow. Continuous blanching, blueness or pain is not Raynaud's phenomenon and to have implications for autoimmune rheumatic disease the phenomenon must be biphasic and episodic.

The episodic vasospasm can also be observed in the tip of the nose, the nipples, the mouth, and the ear lobes. More recently, and of great interest, there are reports of a systemic 'vasospasm', which may affect the cerebral, coronary, pulmonary, renal or upper gastrointestinal blood supply. With the published evidence of vasospasm in many organs, it is interesting to speculate that abnormalities of the vasculature may exist throughout the whole Raynaud's patient (Belch 1990). A

careful history is still the best way to diagnose the phenomenon, although this can be supported in the laboratory with measurements of digital and organ blood flow.

Important clues to secondary Raynaud's on clinical evaluation are: the development of Raynaud's either in very young children or after the age of 45 years; severe symptoms occurring all year round; digital ulcerations, which rarely, if ever occur in primary Raynaud's; asymmetry of symptoms; and the reoccurrence of chilblains in an adult (Lally 1992). Two simple, inexpensive, non-invasive procedures have high predictive power for detecting patients in the Raynaud's group who will have systemic sclerosis in the future; serum autoantibody determination and nailfold capillary microscopy. These tests should be performed in all Raynaud's patients. The antinuclear antibodies were discussed earlier in the chapter and the presence of disease-specific autoantibodies plus abnormal nailfold capillaries is a powerful predictive tool.

Autoantibodies

The number of autoantibodies present in the serum of patients with scleroderma is varied (Table 9) (Pollard *et al*. 1989; Bona and Rothfield 1994). However, three of these antibodies have particular importance: anticentromere, antitopoisomerase-1 (Scl-70), and antinucleolar antibodies. Anticentromere antibodies react with the centromere portion of chromosomes and particularly with the kinetochore plates that attach the centromere to the spindle during mitosis; they appear to be a constant finding over time, appear early in the disease, and have a predilection for the subset of limited cutaneous systemic sclerosis. In contrast, antitopoisomerase 1 is an antibody found in up to 60 per cent of patients with diffuse cutaneous systemic sclerosis. This proportion tends to vary considerably between laboratories, and retrospective analysis and patient selection may have influenced results. The specificity of antinuclear antibodies as determined by immunoblots was also found to be predictive of the subtypes of systemic sclerosis. A prospective study of primary Raynaud's and undifferentiated autoimmune rheumatic disease (Kallenberg *et al*. 1988), using the immunoblot method, found that the presence of antinuclear antibodies at the time of entry into the study was associated with the evolution of an autoimmune rheumatic disease, usually scleroderma. Furthermore, in those who initially presented with Raynaud's disease alone, anticentromere antibody had a predictive value for the development of limited cutaneous systemic sclerosis (sensitivity 60 per cent, specificity 98 per cent) and Scl-70 for diffuse cutaneous systemic sclerosis (sensitivity 38 per cent, specificity 100 per cent). Antinucleolar antibodies now being characterized biochemically are present in many patients with scleroderma and further definition may improve their diagnostic specificity. Thus it seems that the presence of antinuclear antibodies, particularly anticentromere antibody and Scl-70, in a patient with apparent Raynaud's disease may mark the probability of later progression to one of the subsets of systemic sclerosis (Wollersheim *et al*. 1989).

Nailfold capillaroscopy

The most distal parts of the skin and its appendages receive their nutrient blood supply from capillary loops that arise from and return to a vascular plexus deeper in the skin. These capillary loops can be seen in the skinfold of the finger nail, where the capillary is visible over its long axis (Fig. 4). Direct observation of the nailfold capillary bed dates back almost 70 years and was introduced by German investigators. Recent refinements have permitted permanent photographic recording of a row of horizontal capillary loops at the nailfold, just proximal to the cuticle (Carpentier and Maricq 1990). The characteristic patterns seen in patients destined to develop autoimmune rheumatic disease are:

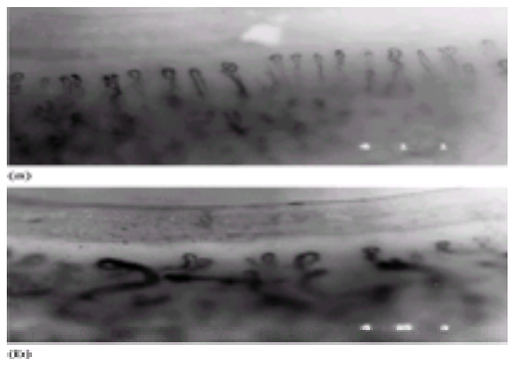


Fig. 4 (a) Photographs if normal nailfold capillaries showing normal spacing, orientation, and indentations. (b) Photograph of nailfold capillaries of a patient with scleroderma showing avascular areas and capillaries that area dilated and irregular in shape and distribution with distributed orientation. Original magnification ×65. (By courtesy of Dr Frances Lefford, Department of Anatomy and Developmental Biology, University College London.)

1. the enlargement of all three portions of the capillary loop—arterial, apical and venular;
2. the loss of capillaries either diffusely or in localized areas often adjacent to enlarged capillaries.

Dilatation without avascular areas is reported to be characteristic of limited cutaneous systemic sclerosis and dilatation with avascular areas characteristic of diffuse cutaneous systemic sclerosis. The capillary patterns are present early in the disease and are remarkably constant over time.

Comment

Tests for autoantibodies and nailfold capillaroscopy together detect more than 90 per cent of patients destined to have generalized systemic sclerosis. Of the 3 to 10 per cent of the population with Raynaud's up to 15 per cent are positive for one or both procedures. These recent observations are now leading to a shrinkage in the number of patients with true Raynaud's disease and an expansion in the number of patients who potentially have autoimmune rheumatic disease.

There is no evidence that symptomatic treatment of Raynaud's phenomenon in any way influences the evolution of scleroderma. However, once the mechanisms of fibrosis and vascular damage are better understood, the predictive power of these two tests, possibly coupled with other serological or genetic markers, ought to allow for a more preventive approach.

Scleroderma once established is a multisystem, multistage disease and each target organ progresses through stages of inflammation, fibrosis, atrophy, not necessarily at the same time or with the same speed. The eventual effect of these pathological processes is to impair organ function. In the following sections some organs, particularly those whose involvement may be fatal, will be discussed in detail; others will be summarized and/or tabulated.

Cutaneous involvement (Fig. 5 and Fig. 6)

The changes in the skin usually proceed through three phases: early, classic, late. The early stage can be difficult to diagnose and a high level of suspicion is needed in the oedematous phase when the only feature may be puffiness of the hands and feet, most marked in the mornings. The face may feel slightly taut at this stage and Raynaud's may be present. On examination there is a non-pitting oedema with intact epidermal and dermal appendages. The subsequent, often sudden, development of firm, taut, hidebound skin proximal to the metacarpophalangeal joints, adherent to deeper structures such as tendons and joints, causing limitation of their movement and subsequent contractures, permits a definitive diagnosis in over 90 per cent of patients. The skin may be coarse, pigmented, and dry at this stage. The epidermis thins, hair growth ceases, sweating is impaired, and skin creases disappear.

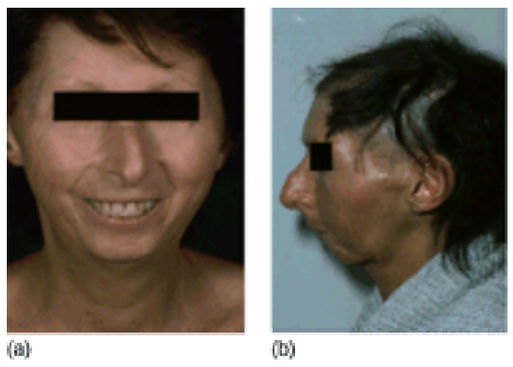


Fig. 5 (a,b) Typical facial appearance of diffuse cutaneous systemic sclerosis, with tight shiny skin, contrasting with the widespread facial telangiectasis often present in advanced, limited cutaneous systemic sclerosis

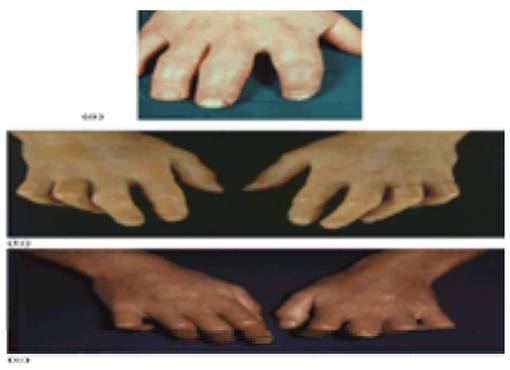


Fig. 6 The three phases of skin involvement in the hands of a patient with systemic sclerosis; (a) initial puffiness of the skin; (b) tight shiny skin with iduration, loss of finger pulp, and contractures; (c) late changes of contractures with atrophic skin and ulceration.

Changes limited to the fingers alone (sclerodactyly) do not carry the same implication. The classical changes, once fully developed, can remain static for many years. Careful mapping of the degree and extent of skin involvement is the single best clinical technique for detecting the patient at risk for life-threatening involvement of internal organs. A number of scoring systems to quantify skin sclerosis have been developed. Most are two-dimensional instruments that attempt to summate the severity of skin involvement at a number of different sites in the body. A modified version of the original Rodnan skin score, consisting of a 0 to 3 grading at 17 skin sites (maximum score 51), has been shown to be better in terms of greater observer agreement and lower bias than a method in which the observer attempts to shade on a mannequin the full extent of skin involvement. The interobserver variability in the use of the modified Rodnan skin score in studies from both the United Kingdom and United States is similar (Clements *et al*. 1990). An overall within-patient variability in scoring (derived from multiple examinations) is about five skin-thickness units, which is similar to the variability found in scoring joint tenderness in rheumatoid arthritis (Ritchie score), justifying its use in clinical trials to follow-up the outcome of skin involvement.

Taut hypo- or hyperpigmented skin proximal to the elbows, knees or clavicles qualifies a patient as having diffuse cutaneous systemic sclerosis and such patients require more frequent multisystem evaluation. Patients with diffuse cutaneous systemic sclerosis have a preponderance of visceral involvement in the first 5 years of symptoms. The exact beginning of the late phase is usually impossible to define, but at some point, and this may be 2 to 15 years after the appearance of the classical changes, the pattern moves into a final phase. The taut truncal arm and leg skin softens, and in some patients, but for the pigmentation, it would be difficult to know that they had ever had systemic sclerosis. Nevertheless, the hands in diffuse cutaneous systemic sclerosis nearly always show the ravages of the early, actively fibrotic period, and tautness and contractures usually remain even after resolution elsewhere. Other skin manifestations include digital pitting scars, loss of fingerpad tissue, ulcers, telangiectasias, and calcinosis. Skin biopsy is usually no more sensitive than the experienced touch in diagnosing the full-blown disease, but may provide useful suggestive information in the early oedematous phase.

In the early phase there are collections of mononuclear cells in the dermis, particularly around blood vessels. The soluble products of these monocytes and lymphocytes may have pathogenetic significance in the disease process. In the classic phase, fibrosis replaces the cellular infiltrate and may extend deep into the connective tissue to surround tendons, nerves, muscle bundles, and joint capsules. In the final stage, the fibrosis may be less evident, with epidermal thinning and loss of appendages the major findings.

Systemic features of disease

General manifestations

The patient with systemic sclerosis must cope with a complex set of symptoms that range from features common to chronic diseases through to complaints attributable to specific visceral involvement; fatigue and lethargy are common throughout the illness, although usually more pronounced in its early phases. Weight loss is almost universal in the diffuse cutaneous form and is less common in the limited variety. Fever is uncommon and if present, other causes such as infection or underlying malignancy should be excluded. Reactive depression is a frequent accompaniment to this often relentless and disfiguring disorder. Patients often feel isolated and support groups provide an invaluable service.

Gastrointestinal tract

The gastrointestinal tract is probably the most commonly involved internal organ system in systemic sclerosis (Shorrock and Rees 1988; Silver 1990). Over 90 per cent of patients with limited cutaneous and diffuse cutaneous systemic sclerosis have oesophageal hypomotility and serious gastrointestinal disease has been estimated to occur in 50 per cent of patients with limited cutaneous systemic sclerosis. It is probable that when systemic sclerosis affects an area of the gastrointestinal tract it does so in a sequential manner with progressive dysfunction (Cohen *et al*. 1980; Greydanus and Camilleri 1989). This concept is important when designing therapeutic regimens. The earliest lesion is neural dysfunction. The basis for this lesion is uncertain, although in the oesophagus there is both physiological and anatomical evidence that it is due to arteriolar changes in the vasa nervorum (D'Angelo *et al*. 1969; Russell *et al*. 1982; Greydanus and Camilleri 1989). An alternative explanation would be compression of nerve fibres by fibrous tissue (Dessein *et al*. 1992). This produces functional change before the next step, which is impairment of muscle contractility. These functional changes may remain asymptomatic for a long period but, if looked for, they usually respond well to prokinetic drugs. Once smooth-muscle atrophy is established, symptoms usually appear. The muscle can respond partially to prokinetic drugs but the response is weak. The final lesion, as with all other organs in systemic sclerosis, is muscle fibrosis superimposed on neural dysfunction and atrophy. At this stage, restoration of function is not possible.

The earliest clinical symptoms may be quite subtle. Patients can often recall a specific event when there was difficulty in swallowing a pill or bolus of hard food. They may also experience retrosternal discomfort or even overt pain, which can be nocturnal. There are patients (forming part of the continuum of systemic sclerosis and who should be included in the group with limited cutaneous systemic sclerosis) who demonstrate Raynaud's phenomenon, abnormal nailfold capillaries, anticentromere antibodies, and decreased motility or at least decreased lower oesophageal-sphincter pressure, but no skin sclerosis. Measurement of lower oesophageal pressure is frequently unacceptable to the patient and therefore in clinical practice the oesophageal transit time (quantitative oesophageal scintigraphy) is usually the preferred screening test. It is non-invasive, cost-effective, and is highly acceptable to patients. In those who have an abnormal scan and those who have frank dysphagia or heartburn, barium studies and/or direct oesophagoscopy may be required to identify structural divisions such as hiatus hernia and oesophageal strictures. Fortunately, Barrett's metaplasia and oesophageal stricture are rare. All patients with systemic sclerosis should be recommended to raise their bed ends with blocks, take small frequent meals, and not eat late at night. The therapies available for oesophageal disease are numerous and their place in management is

summarized in Table 12.

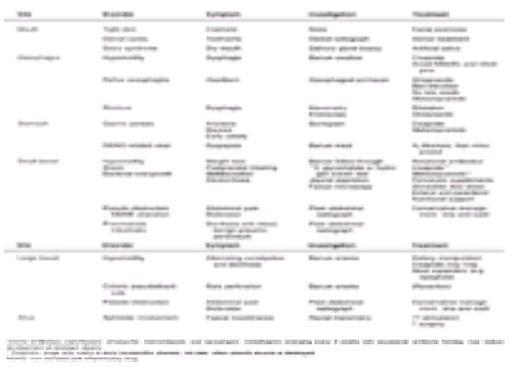


Table 12 Gastrointestinal-tract pathology in systemic sclerosis

Recent studies have also emphasized that oesophageal disease is not predicted by disease subset or duration and may be relatively asymptomatic. Often investigations are not undertaken in the non-complaining patient, but perhaps we should pursue studies early in the disease so that distal hypomotility can be detected and treated aggressively. Such action may delay or prevent irreversible, pathological change.

Small-bowel disease with hypomotility is a major problem in scleroderma and can lead to weight loss, cachexia, malabsorption, and death. The classical symptoms are of a change in bowel pattern, with loose, frequent, floating, foul-smelling stools, and abdominal distension, but a patient may also present with weight loss (otherwise unexplained) or a nutritional anaemia. There are numerous possible ways of investigating the small bowel (some as yet research tools), and each test has its exponent. Details are given in Table 12.

Once the disease is established, bacterial overgrowth with its associated malabsorption is a recurring problem, often punctuated by abrupt episodes of distension and adynamic ileus or pseudo-obstruction as it is now called. The management of such patients is difficult and includes the rotational use of antibiotics, attempts to stimulate the bowel directly with cisapride, and ultimately total parenteral nutrition administered by the patient at home. There are greater than 50 per cent 5- and 10-year survival rates for patients on total parenteral nutrition, which is a distinct improvement over survival rates before total parenteral nutrition was available.

Atony and hypomotility of the rectum and sigmoid colon is frequent and occurs early. It is often missed clinically because patients are reluctant to discuss symptoms such as anal incontinence, a problem for which there is no relief. Constipation is usually manageable with the use of dietary manipulation and stool volume expanders. Codeine can cause constipation and should be avoided.

Surgery to the large bowel or any other part of the gastrointestinal tract must be viewed with great caution. Careful manometry and radiographic localization of affected segments of stomach, small intestine, and colon may allow judicious surgical resection or venting procedures, but these are not without risk and are not always successful.

Pancreatic exocrine function is frequently reduced, but rarely to an extent that is clinically important. Primary biliary cirrhosis may occur and it is associated with the subgroup of limited cutaneous systemic sclerosis. As the gastrointestinal manifestations of systemic sclerosis are frequent, and debilitating if not life-threatening, the goal in this area must be early detection, support, and control, thus permitting as active a life as possible.

Cardiac disease

Vascular, microvascular, and a vasospastic phenomena are, as in all other organs, a feature of cardiac scleroderma (Follansbee 1986). The clinical symptoms are diverse and sometimes difficult to separate from pulmonary or renal disease. They are largely non-specific, including dyspnoea, orthopnoea, paroxysmal nocturnal dyspnoea, oedema, palpitations, and atypical chest pain. Angina pectoris is uncommon and such a history should be regarded as a potential indicator of an underlying complicating factor such as coronary atherosclerosis. No definite predisposition to coronary atherosclerosis has been demonstrated in this disease, although recently the occurrence of large-vessel disease in association with systemic sclerosis has been emphasized. Youssef *et al*. (1993) found a threefold increase in evidence of large-vessel disease among patients with systemic sclerosis compared with the predicted population prevalence of 2.3 per cent for symptomatic vascular disease. Large-vessel disease seems to be associated with limited cutaneous systemic sclerosis and anticentromere antibody-positive patients. Cardiovascular manifestations of systemic sclerosis include myocardial fibrosis, pericarditis, a variety of arrhythmias, conduction abnormalities and pulmonary hypertension, and rarely myocarditis. Cardiac involvement has a poor prognosis, which is probably most directly related to the extent of myocardial fibrosis present. There are many diagnostic techniques useful in evaluating cardiac scleroderma, but the investigation must be carefully matched to the clinical feature (see Table 13).

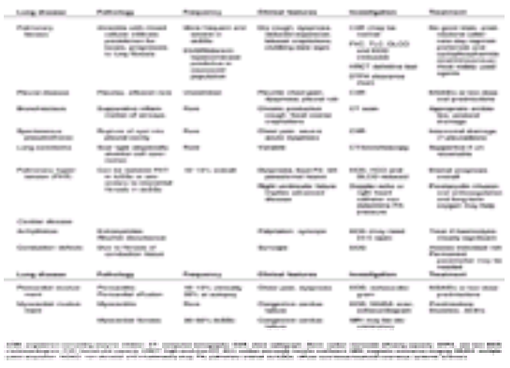


Table 13 Cardiopulmonary manifestations of systemic sclerosis

Two features have been singled out for discussion, myocardial fibrosis because it is the hallmark of cardiac scleroderma, and pulmonary hypertension, with its sequelae including right heart failure, which often presents abruptly and which has such a devastating course. The nature and distribution of the myocardial fibrosis has a number of distinctive features. It is distributed randomly throughout both ventricles and develops throughout the entire thickness of the ventricular wall. It bears no relation to the extramural vascular supply and in some areas is so confluent that it gives the appearance of a circumscribed myocardial infarction. An additional distinctive histological feature of the myocardial fibrosis is the coexistence of contraction-band necrosis. This is typically seen in circumstances of myocardial ischaemia followed by reperfusion. In patients with scleroderma this suggests that vasoconstriction of the small, intramyocardial coronary arteries (a myocardial Raynaud's phenomenon) might be contributing to the pathogenesis of myocardial scleroderma. It would appear from work by Follansbee *et al*. (1990) that patients who develop left ventricular dysfunction in the absence of other apparent underlying causes commonly have advanced myocardial fibrosis associated with contraction-band necrosis, whilst patients who develop ventricular dysfunction in the clinical context of altered myocardial loading conditions (that is, systemic hypertension, pulmonary fibrosis, pulmonary hypertension) tend to have less myocarditis.

Pulmonary hypertension occurs in systemic sclerosis, usually in patients with the limited cutaneous type. It is attributable directly to pulmonary vascular disease, which can be suspected if on testing pulmonary function there is an isolated marked decrease in diffusing capacity for carbon monoxide (<50 per cent of predicted normal) in the absence of significant restrictive ventilatory abnormalities. Pathologically, pulmonary arteries of all sizes show marked intimal and medial hyperplasia; of great interest is the finding by Follansbee *et al*. (1990) that, although the clinical syndrome seems confined to the group with limited cutaneous systemic sclerosis, intimal thickening and narrowing, albeit to a lesser degree, occurred in patients with diffuse cutaneous systemic sclerosis. In addition to the obstructive vascular

lesions, the pulmonary vasculature appears to be abnormally reactive, with significant pulmonary vasoconstriction occurring on exposure to cold, again analogous to a peripheral Raynaud's phenomenon. That systemic sclerosis can be an overwhelmingly vascular disease is perhaps nowhere more convincingly demonstrated than in the subset of patients with severe pulmonary hypertension. It has an extraordinarily poor prognosis; death is usually due to rapidly progressive respiratory insufficiency accompanied by severe right-ventricular hypertrophy and failure.

The treatment of the cardiac manifestations of systemic sclerosis is primarily supportive, empirical, and of moderate value (see Table 13). Caution is necessary when treating patients with large pericardial effusions in diffuse scleroderma to avoid intravascular volume depletion, because of their predisposition to develop renal failure. Patients with systemic sclerosis and coexisting coronary arterial disease are often not good candidates for bypass surgery because of the distal vascular and interstitial disease that will reduce flow despite the presence of a satisfactory graft.

Pulmonary involvement

Pulmonary involvement ranks only second to oesophageal in frequency of visceral disease, and with considerable improvements in the management of renal disease it is now the major cause of death in scleroderma.

A study (Altmann *et al*. 1990) on patients with diffuse skin disease, pulmonary involvement but no cardiac or renal disease found a median survival of 78 months with 60 per cent dead at 5 years. Early diagnosis enabling the institution of effective therapy to halt disease progression is therefore a critical aim in the management of the patient with systemic sclerosis. Fortunately, this is becoming possible with the aid of modern techniques such as high-resolution computed tomography (Wells *et al*. 1993b; Wells *et al*. 1994), [$^{99}Tc^{m}$]diethylene triamine pentacetate (**DTPA**) scanning and analysis of bronchoalveolar lavage fluid (Wells *et al*. 1993a). There is also evidence that the genetic markers HLA-DR3/DR52a (Briggs *et al*. 1990; Langevitz *et al*. 1992) and specific autoantibodies Scl-70, anti-U3 ribonucleoprotein, and antihistone antibodies may help separate this group at presentation.

The two major clinical manifestations of lung involvement are fibrosing alveolitis and pulmonary vascular disease. Other potential complications include aspiration pneumonia, pleural disease, spontaneous pneumothorax, drug-induced pneumonitis, associated pneumoconiosis, and neoplasm. Pulmonary fibrosis occurs in more than three-quarters of the patients with systemic sclerosis and pulmonary vascular disease in approx. 50 per cent. Autopsy studies have always yielded higher percentages than clinical studies.

In contrast to pulmonary hypertension, as discussed above, with other cardiac manifestations, interstitial lung disease often develops insidiously but established fibrosis is irreversible with present-day therapy. Early diagnosis is therefore vital and in the future genetic markers may help to identify this group at presentation. Pulmonary manifestations of systemic sclerosis are listed in Table 13.

The most common symptoms of respiratory involvement are breathlessness, especially on exertion, and a dry cough. Chest pain is infrequent and haemoptysis rare, and if either are present then the presence of additional pathology should be sought. On physical examination the most frequent finding is of bilateral basal crepitations (Alton and Turner-Warwick 1988).

The classical radiographic features consist of reticulonodular shadowing, usually symmetrical and most marked at the lung bases. However, the chest radiograph is an insensitive indicator of fibrosing alveolitis, and should be used only as an initial screen or to exclude infection or aspiration secondary to oesophageal abnormalities. There are many symptomatic patients (often mildly so) with normal chest radiographs despite interstitial lung disease, and lung function tests can be discriminatory. The single-breath diffusion test (*DL* CO) is abnormal in over 70 per cent of patients with diffuse cutaneous systemic sclerosis (including asymptomatic patients with no complaints and an unremarkable chest radiograph). A reduction in *DL* CO is the earliest proven abnormality in patients with systemic sclerosis who develop interstitial lung disease; lung function tests that show normal volumes but reduced transfer of gases in the face of normal imaging are suggestive of pure pulmonary vascular disease (see above). Measurement of the alveolar–arterial oxygen difference during exercise also appears to be a sensitive indicator of lung disease in systemic sclerosis.

Over the past 5 years the application of thin (3 mm)-section, high-resolution CT scanning of the lungs has revolutionized the approach to diffuse lung diseases and has revealed the character and distribution of fine structural abnormalities not visible on chest radiographs (Harrison *et al*. 1989) (Fig. 7). Using this technique, the earliest detectable abnormality is usually a narrow, often ill-defined, subpleural crescent of increased density in the posterior segments of the lower lobes. When more extensive, the shadowing often takes on a more characteristic reticulonodular appearance yet frequently retaining a subpleural distribution. It also becomes associated with fine, honeycomb air spaces and ultimately larger, cystic air spaces—an appearance that mirrors the macroscopic appearance. In a semiquantitative comparison of the predictive value of these CT appearances to mirror the biopsy evidence of an inflammatory alveolitis, a 'ground glass' pattern of opacification on CT was associated with a predominantly cellular biopsy whereas a reticular pattern of abnormality was found in patients whose subsequent lung biopsy confirmed a particularly fibrotic disease process (Muller and Miller 1990; Wells *et al*. 1992). CT scans also confirm the presence of pleural disease or mediastinal lymphadenopathy, which is commonly present. It is important to perform prone as well as supine scans, particularly in more subtle cases, to exclude the contribution of gravity to the radiographic appearances from vascular and interstitial pooling in the dependent areas. In addition to identifying early disease, high-resolution CT scanning can identify a pattern of disease that predicts a better response to therapy and a better prognosis (Wells *et al*. 1993). Furthermore, the extent of disease present, as defined by CT within the lavaged lobe, correlates with the predominant type of inflammatory cell obtained by bronchoalveolar lavage of that same lobe: lymphocytes are present in excess before CT identifies disease; eosinophils appear as the lung becomes abnormal; neutrophils are found in most abundance when at least 50 per cent of the lavaged lobe is involved in the disease process. In other words, the predominant type of inflammatory-cell traffic into the lungs depends upon extent of disease, and this would suggest that different inflammatory cells are involved in different stages of the pathogenesis.

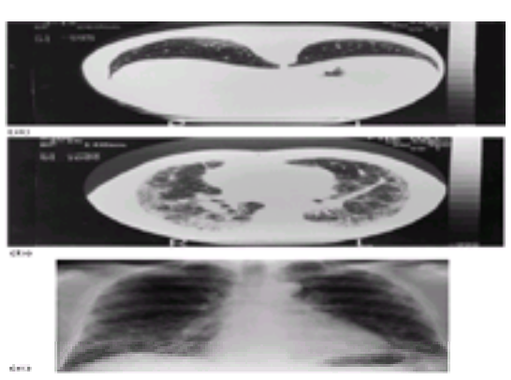


Fig. 7 Thin-section CT scan images illustrating; (a) ground-glass appearance of early pulmonary involvement posteriorly, associated with a normal chest radiograph; (b) extensive honeycomb shadowing and cystic air spaces involving both lower lobes, with corresponding chest radiographic appearances of advanced interstitial lung disease (bibasilar reticulonodular shadowing) (c). (With grateful acknowledgement to Drs A Wells, R du Bois and B Strickland, Departments of Respiratory Medicine and Radiology, Royal Brompton Hospital, London.)

In predicting the histological pattern CT, although useful, has not replaced lung biopsy as the 'gold standard' investigation (Fig. 8). As yet, patients who appear to have early changes on CT should still be considered for a thorascopic biopsy for staging of the disease. In the evaluation of diffuse lung disease, bronchoalveolar lavage has now been used for almost 20 years to sample cells and non-cellular material from the lower respiratory tract. The presence of abnormal numbers of granulocytes, particularly neutrophils and eosinophils (Harrison *et al*. 1989; Miller *et al*. 1990), is typical for a patient with fibrosing alveolitis occurring alone or in the context of systemic sclerosis. Excess lymphocytes are found in some individuals. In a typical patient with fibrosing alveolitis, bronchoalveolar lavage would produce an increase in total cell returns of three- to sixfold (up to 6×10^{5}/ml of fluid return); of these, up to 20 per cent may be neutrophils or eosinophils. Excess lymphocytes may be found (up to 20 per cent of the total cells) and an increase in mast cells may be observed in a small percentage of patients.

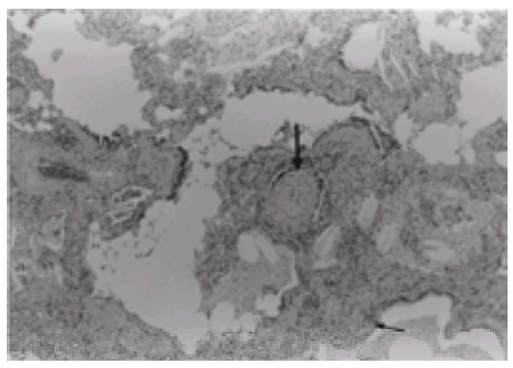


Fig. 8 Open lung biopsy of patient with systemic sclerosis showing interstitial chronic inflammation and fibrosis (small arrow) and thick-walled arteriole showing intimal fibrosis (large arrow). Haematoxylin and eosin, ×300. (Courtesy of dr Mary N Sheppard, Department of Lung pathology, Royal Brompton Hospital, London.)

The prognostic value of bronchoalveolar lavage has been demonstrated in several studies (Harrison *et al*. 1989; Silver *et al*. 1990). Silver *et al*. (1990) reported that patients with scleroderma with persistent alveolitis have greater deterioration in their pulmonary function than alveolitis-negative patients with systemic sclerosis.

The use of DTPA clearance in the management of systemic sclerosis has been the subject of extensive study and has been shown to be of value. It identifies early disease and also identifies a group of patients whose disease will run a more stable, non-progressive course; that is, those with normal clearance (Wells *et al*. 1993). The speed of clearance of the isotope is dependent upon the integrity of the epithelial barrier and therefore anything that disrupts this, either inflammation or fibrosis, will increase the rate of clearance (Barrowcliffe and Jones 1987). DTPA clearance is highly sensitive: cigarette smoking will produce increased clearance rates and the test is, therefore, only of value in non-smokers or those who have given up smoking for at least 1 month before the assessment (Mason *et al*. 1983).

In systemic sclerosis, clearance of DTPA may be abnormal even when chest radiography and pulmonary function tests are normal (Harrison *et al*. 1989). In established disease, clearance is enhanced in comparison with normal individuals. Furthermore, the speed of, and change in, clearance can predict subsequent changes in pulmonary function tests. Patients whose clearance is persistently abnormal are more likely to have a deterioration in pulmonary function tests at follow-up subsequent to the DTPA measurements. In contrast, persistently normal DTPA clearance predicts stable disease and therefore provides a good prognostic index; a study showed significant improvement in pulmonary function tests in 75 per cent of patients whose clearance returned to the normal range whereas similar improvements were not seen in those whose clearance remained normal or abnormally fast (Wells *et al*. 1993).

Considerable space has been devoted here to the lung because it is the organ that represents a major challenge in systemic sclerosis for the next decade, and because the available diagnostic tests have improved. The frequency of reassessments depends on the subset and length of disease. Those with diffuse cutaneous disease should be studied yearly or more frequently if necessary, whilst those with limited cutaneous disease have less need for frequent follow-up in the first 5 years of their illness and could be assessed at 2 years, but with increasing time must be watched on an annual basis for the development of pulmonary complications, either fibrosis or vascular disease.

Renal disease

Although renal disease has now been superseded by lung involvement as the major cause of systemic sclerosis-related death, it remains one of the most important complications of scleroderma and, despite its life-threatening nature, it is amenable to treatment, although the prognosis is much better if appropriate management is instituted early. Both post- and ante-mortem studies suggest that epithelial and endothelial renal lesions occur before there is clinical evidence of renal disease in systemic sclerosis (Kovalchik *et al*. 1978), and certainly precede any histological evidence of fibrosis. This supports the view that epithelial, and particularly endothelial, damage are important early events in the pathogenesis of scleroderma (Prescott *et al*. 1992). The best characterized pattern of renal involvement in systemic sclerosis is an acute or subacute renal hypertensive crisis. This generally occurs in patients with diffuse systemic sclerosis within 5 years of disease onset. The overall incidence of scleroderma renal crisis is uncertain, with differences in the reported frequency even in series from the same unit. This variation probably reflects differences in incidence in the various subsets of systemic sclerosis. In high-risk patients the incidence may be as great as 20 per cent but overall is probably less than 10 per cent (Steen 1984). Traub (1994) proposed the following criteria to diagnose scleroderma renal crisis: abrupt onset of arterial hypertension greater than 160/90 mmHg; hypertensive retinopathy of at least grade III severity; rapid deterioration of renal function and elevated plasma renin activity. Other typical features include the presence of a microangiopathic haemolytic blood film and hypertensive encephalopathy, often complicated by generalized convulsions. It is generally considered important to perform a renal biopsy, once hypertension has been adequately controlled, especially if renal replacement therapy is being contemplated. This allows histological confirmation of the diagnosis and the exclusion of other causes for renal failure of abrupt onset, such as glomerulonephritis or the haemolytic–uraemic syndrome. Histologically, systemic sclerosis renal crisis typically shows fibrinoid necrosis, mucoid or fibromucoid proliferative intimal lesions (when extensive, termed onion skinning) in renal arteries, particularly the arcuate and interlobular vessels; glomerular thrombi occur and ultimately glomerulosclerosis (Fig. 9). The extent of the glomerular lesion can be useful in predicting the eventual degree of functional recovery. Patients usually present with the clinical features of severe hypertension, including headaches, visual disturbances, hypertensive encephalopathy (especially seizures), and pulmonary oedema. Occasionally a similar pattern of renal dysfunction occurs without hypertension (normotensive renal crisis), suggesting that the pathological features are not simply the end-organ consequences of raised arterial pressure (Helfrich *et al*. 1989). A more insidious pattern of renal involvement in systemic sclerosis is also reported in which there is a slow reduction in glomerular filtration rate accompanied by proteinuria. This is believed to reflect a more benign vascular and fibrotic process than scleroderma renal crisis.

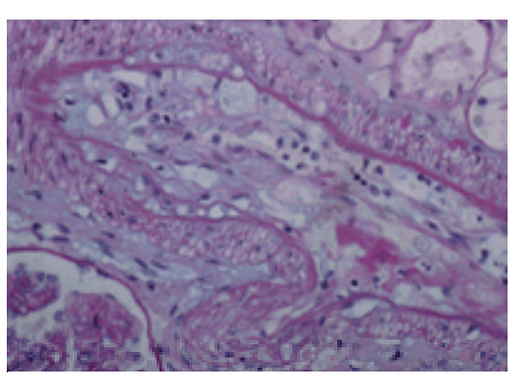


Fig. 9 An artery in the corticomedullary region. This section has been stained by the Alcian-blue/diastase-periodic acid-Sciff method. The lumen of the blood vessel has been entirely effaced by proliferation and swelling of the intima. Intimal cells can be seen to have developed an unusual bubbly, blue (mucoid) cytoplasm that is characteristic of scleroderma-related acute renal arteriopathy. There is scattered small-lymphocyic infilltrate as well. (By courtesy of Dr A. P. Dhillon, Department of Histopathology, Royal Free Hostpital Medical School, London.)

Management of renal systemic sclerosis requires a high index of suspicion to enable early diagnosis and treatment of the renal crisis. Creatinine clearance or isotope glomerular filtration rate should be checked twice yearly in diffuse cutaneous systemic sclerosis for the first 5 years and annually thereafter. In limited cutaneous systemic sclerosis there is much less risk and a less frequent measurement of the glomerular filtration rate is sufficient. Blood pressure should be well controlled (often antihypertensive treatments also help Raynaud's symptoms) and in diffuse cutaneous systemic sclerosis the use of angiotensin-converting enzyme inhibitors is particularly appropriate since there is some anecdotal evidence that they protect from hypertensive crisis (Steen *et al*. 1990). High-dose corticosteroids have now been formally demonstrated in a case–control study to increase the risk of renal crisis in diffuse cutaneous systemic sclerosis and doses above 20 mg prednisolone equivalent daily should be avoided (Steen *et al*. 1994).

Once diagnosed, an acute renal crisis in systemic sclerosis must be treated as a medical emergency. The patient should be admitted immediately and reasonable control of the blood pressure is a priority. Extreme caution must be taken, however; to avoid a precipitous or excessive drop in arterial pressure and also to prevent relative or actual hypovolaemia associated with vasodilatation of constricted vascular beds, both of which can further diminish renal perfusion and compound the renal lesion of systemic sclerosis with acute tubular necrosis. For this reason, powerful parenteral antihypertensives (e.g. intravenous nitroprusside or labetolol) should be avoided; an internal jugular or subclavian venous cannula should be inserted to monitor central venous filling pressure; and an indwelling arterial cannula for pressure monitoring should be considered. Hypertension should be treated with angiotensin-converting enzyme inhibitors (captopril or enalapril up to maximum dose) and calcium-channel blockers (starting with long-acting nifedipine initially), aiming to reduce both diastolic and systolic pressure by 20 mmHg in the first 48 h and ultimately maintaining diastolic pressure at 80 to 90 mmHg. Intravenous prostacyclin, which is believed to help the microvascular lesion without precipitating hypotension, is often administered from diagnosis. Fish-oil capsules are sometimes prescribed in view of their unproven but theoretically beneficial properties (McCarthy and Kenny 1992). Renal function should be closely monitored by twice-weekly creatinine clearance and daily serum creatinine estimations. Regular full blood counts, clotting screening, and estimations of fibrin degradation product are important to monitor the degree of microangiopathic haemolytic anaemia, which often reflects the activity of the disease process. Short-term haemodialysis should be given if necessary and peritoneal dialysis often works well if long-term renal replacement therapy is needed. Interestingly, it has been observed that after a renal crisis, skin sclerosis and other features of systemic sclerosis improve (Denton *et al*. 1994), particularly if a patient is undergoing maintenance dialysis. The reason for this is unknown; it may result from the removal or inactivation of circulating mediators or simply reflect the natural history of the disease. It should be remembered that there is also often considerable recovery in renal function after an acute crisis, sometimes allowing dialysis to be discontinued, and improvement can continue for up to 2 years. Therefore any decisions regarding renal transplantation should not be made before this time.

Musculoskeletal system

Muscle

Skeletal muscle is often involved in scleroderma (Russell 1988). In many instances the weakness and atrophy results from disuse secondary to joint contractures or chronic disease. However, about 20 per cent have a primary myopathy (Medsger 1979), which is a subtle process distinctive for the disease. This chronic myopathy is characterized by mild weakness and atrophy of muscles, minimal elevation of creatine phosphokinase, few or no changes on electromyography, and subtle histological features showing focal replacement of myofibrils with collagen and perimysial and epimysial fibrosis without inflammatory change. This form of myopathy, which can last for many years, is non-progressive, does not warrant intervention, and is often unresponsive to anti-inflammatory mediation. A minority of patients exhibit an inflammatory myositis, indistinguishable from polymyositis; caution must be observed if this occurs in the context of early diffuse disease, when treatment with high-dose steroids might precipitate renal failure, and an alternative treatment should be considered. An atypical inflammatory myositis that requires special histochemical stains to demonstrate the differences in fibre size and composition has been reported in association with myocarditis in a few cases of systemic sclerosis. This form of myopathy, which can last for many years, is often unresponsive to non-steroidal or glucocorticoid anti-inflammatory medications.

Joints

A symmetrical polyarthritis, usually seronegative, anodular and non-erosive, is the presenting feature in a small number of patients destined ultimately to develop systemic sclerosis. By 2 years, frequently much earlier, the synovitis has subsided and classic cutaneous systemic sclerosis is present, often developing abruptly over 1 to 3 months.

The fibrosis characteristic of the classical disease affects the tendons (causing tendon friction rubs), the ligaments, and joint capsules, restricting movement; fibrosis is also found in the synovium. The synovium in systemic sclerosis is often covered by an excessive amount of fibrin; the reason for this is unknown.

Joint destruction is unusual. Management of soft tissue and joint problems is closely linked to skin care and to overall skeletal mobility. True bone changes in the form of distal tufts are usually a late change occurring in the second and third decade of the disease, and are thought to be due to a lack of a vascular supply adequate enough to preserve viable bone. This can occur in patients with long-standing Raynaud's phenomenon without connective tissue disease. Other sites of bone reabsorption, for example the mandible and ribs, have been recorded late in the disease.

Other organ involvement

Other organs involved in systemic sclerosis are listed in Table 14.

Organ	Effect	Pathological process	Frequency
Thyroid	Spectrum of autoimmune thyroid disorders — especially hypothyroid	Thyroid antibodies present in > 50%	20–40%
Liver	Primary biliary cirrhosis (PBC)	Antimitochondrial antibodies in 25% patients with systemic sclerosis. Anticentromere antibodies in 5–20% patients with PBC	3% CREST
	Abdominal pain, bile peritonitis	Inspissation of gallbladder	Rare
	Calcification of liver		Very rare
	Extrahepatic obstructive jaundice	Mucosal ulceration	Very rare
	Nodular regenerative hyperplasia of the liver		Very rare
Nervous system	Trigeminal neuralgia	Collagen deposition in sphenotrium	Most common neurological change
	Carpal tunnel syndrome		3%
	Sensorimotor peripheral neuropathy	Vascular damage of vasa nervorum	
	Autonomic neuropathy		Possibly up to 60% — under-recognized
	Prolonged action of local anaesthetics		
	Subacute combined degeneration of the cord	Malabsorption of vitamin B_{12} (secondary to small-bowel involvement)	Rare
Impotence	Erectile failure (libido intact)	Vascular fibrosis of cavernosal arteries	25–50% males

Table 14 Other organ involvement in scleroderma

Systemic sclerosis and pregnancy

The greater incidence of scleroderma in women has focused interest on the potential interrelations between scleroderma, hormones, and pregnancy. It is of interest that women with Raynaud's phenomenon have an increased likelihood of low birthweight babies and fertility problems both before and after the onset of disease. It is unknown whether these findings are a reflection of the vasospasm affecting pregnancy or whether they reflect a common aetiological link.

The question of pregnancy in systemic sclerosis can be approached from two angles: the effect of scleroderma on pregnancy and its outcome, and the effect of pregnancy on the development and course of systemic sclerosis (Black 1990b).

Case-controlled studies have provided conflicting evidence on the outcome of pregnancy. British and Italian studies have shown an increase in the spontaneous abortion rate in women destined to develop systemic sclerosis. The British workers also found a higher rate of infertility, habitual abortion, and a higher probability that the pregnancy would end in stillbirth or neonatal death. An American study, however, showed only an increase in intrauterine growth retardation and prematurity, but not in miscarriage or fetal death.

The outcome of the pregnancy for the mother is also a subject of discussion. The American workers found no increase in maternal morbidity or mortality. However, the 23 case reports since 1932 (which may well select more interesting and thus severe cases) gave nine deaths, eight in patients with diffuse cutaneous systemic sclerosis, and at least five due to renal failure. Disease progressed in twelve of these patients, regressed in two, regressed during pregnancy but progressed afterwards in two, and developed during the pregnancy in one. In reports of larger series, totalling 103 pregnancies (mainly in women with limited disease), the disease developed during pregnancy in 9, progressed in 32, remitted in 11, and was stable in 35. There were 24 spontaneous abortions, 5 perinatal deaths, 6 cases of toxaemia, and 2 maternal deaths. These reports reinforce the worse prognosis in diffuse disease. Of some encouragement are three more recent case reports. One describes a successful pregnancy in a patient with renal involvement, controlled with angiotensin-converting enzyme inhibitors throughout. The second patient, who had a renal crisis 6 years before treated with angiotensin-converting enzyme inhibitors for 5 years, managed without therapy to reach 38 weeks' gestation before becoming hypertensive and needing caesarean section for delivery of a healthy baby. The third case described a patient with limited cutaneous systemic sclerosis,

livedoid vasculitis with foot ulceration, and a positive ribonucleoprotein autoantibody, and three previous spontaneous abortions, one neonatal death (28 weeks' gestation) and a 4-year period of secondary infertility. She was delivered of a healthy baby at 33 weeks, taking nifedipine before and throughout pregnancy for the foot ulceration.

The most feared complication in pregnancy is renal disease, which usually presents with hypertension especially in the third trimester, and is thus difficult to distinguish from toxaemia. Renal scleroderma should be considered in all but the most typical cases of pre-eclamptic toxaemia. In contrast to ordinary pre-eclamptic toxaemia, patients with systemic sclerosis can be vulnerable in the postpartum period, and must be watched very carefully.

Advice to patients with systemic sclerosis at pregnancy is difficult, since studies are limited and retrospective. Patients with diffuse skin involvement, especially with lung, renal or cardiac involvement, tend to have a worse prognosis, and should be offered therapeutic abortion. The patient with limited disease should be told of the possible development of complications, and the variable and unpredictable outcome, and may then request abortion. Very close monitoring of any patient with systemic sclerosis is then necessary throughout pregnancy.

Drug therapy

Although there is currently no treatment that can induce complete remission of the disease there are therapies available that can offer partial relief, control end-organ damage, and improve quality of life for the patient with scleroderma. The choice and evaluation of any treatment regimen are not easy. This is because (a) the disease is complex and the relation between immune dysfunction, vascular damage, and fibrosis speculative; (b) the disorder is heterogeneous and its extent, severity, and rate of progression are highly variable—therapy must therefore be closely tailored to the individual patient systems involved; (c) there is a tendency towards spontaneous stabilization and/or regression after a few years, particularly within the more benign and numerically larger subset of limited cutaneous systemic sclerosis; and (d) there is a paucity of both clinical and laboratory features for ascertaining improvement (or deterioration) in the disease, especially with respect to visceral change. Therapeutic trials of disease-modifying drugs are essential but their design is critical in this disease. There is growing acceptance that any trial of disease-modifying drugs must be controlled (preferably placebo-controlled), that patients should have early diffuse disease (less than 3 years' duration), and that studies must be of sufficient duration (1 year minimum).

Such is the frustration of the condition that numerous vitamins, hormones, 'alternative' medicines, acupuncture, and surgical procedures have been used. Most have been heralded with great enthusiasm, only to be abandoned once critically assessed.

There are no truly effective antifibrotic therapies. Some of the newer, putative antifibrotic agents that have been used in scleroderma are summarized in Table 15. D-Penicillamine and colchicine have been used in the treatment of scleroderma for many years. Neither drug stops the fibrotic process, but consensus opinion is that D-penicillamine is the more useful. To derive maximum benefit the drug must be used correctly. The most suitable subjects for treatment are those with diffuse active skin disease. Long-term treatment in a dose between 500 and 750 mg is needed. D-Penicillamine should be used throughout the active phase into the stable period and the dose maintained until there is no further improvement in skin thickening. The drug may then be reduced, but low-dose treatment should be maintained for many years—some recommend 10 years.

Therapy	Mechanism	Comments
Anticytokine antibodies	Block mediators of fibroblast activation (e.g. TGF-β)	Toxicity likely to be problematic
Enzyme antagonists	Lysyl and prolyl hydroxylase inhibitors to block post-translational modification of collagen	No data yet
Pretranslational inhibition	Block collagen gene transcription or mRNA translation using antisense nucleotides or gene therapy	Hypothetical and dependent upon progress in understanding of collagen gene regulation

Table 15 Novel potential antibiotic therapies for systemic sclerosis

The ideal group to target and treat would be the 'at risk' patients, those in the presclerotic state. Many such patients can now be identified by circulating antibodies, cytokine production, and nailfold changes. Unfortunately, adequate preventative therapy is still wanting. An extension of this idea and one that can, for example, be applied to the lung is the earliest possible diagnosis of internal organ involvement so that containment therapy may be attempted.

As the immune system may be involved early in the disease, immunomodulatory and immunosuppressive agents have been employed, particularly for early diffuse systemic sclerosis. Table 16 summarizes their use and evaluation. Both antimetabolites and alkylating agents have been used. Chlorambucil, a hopeful therapy in 1985, failed in a placebo-controlled trial. The same fate befell 5-fluorouracil. Cyclophosphamide still has an undetermined place; as a single agent its efficacy is unproven. In combination with steroids or plasma exchange it may also have a role—unfortunately as yet there are no controlled data. The use of the antimetabolites 6-thioguanine and azathioprine has been reported but again all the data are anecdotal. Methotrexate is now being approached in a more rational manner. Pilot studies have been hopeful and controlled trials are being undertaken. Attempts to target the immune system by lymphoplasmapheresis and total lymphoid irradiation again failed and currently under investigation are the use of rabbit antithymocyte globulin, and photophoresis. Monoclonal antibody therapy may hold hope for the future.

Treatment	Mechanism of action	Comments	Efficacy in clinical trials
1. Selective immunosuppression			
Cyclosporin	Inhibits T helper cell activity by reducing IL-2 release	Reported beneficial effects on skin sclerosis may be secondarily increased into decline of renal crisis	Two open study (Kahrilas et al 1985) Two open study (Caldwell et al 1986)
Antithymocyte globulin Antilymphocyte globulin	Temporary suppression of cell-mediated immunity	Possible benefit; confirmation of pilot work is essential, especially because of non-selective therapy once caused morbidity	Two open study (Takeuchi et al 1988) Possible benefit in placebo-controlled pilot study (Sinclair et al 1992, unpublished communication)
Photopheresis	8-Methoxypsoralen photoactivated & methoxypsoralen exhibits activated T cells	Benefit has been reported but good placebo-controlled clinical trial is needed	No-blind study (Scheinfeld et al 1988) Two controlled trial comparing with cyclophosphamide therapy, Williams et al. 1990) Multicentre placebo controlled trial currently in progress
Pooled human gammaglobulin	Possible inhibition of lymphocyte function via regulatory idiotype cross-reactivity	Expensive results but unsubstantiated	Not formally evaluated
2. Non-selective immunosuppression			
Methotrexate	Folic acid antagonist	Currently under formal study (RAID 21 0009)	Controlled trial in progress
Cyclophosphamide Chlorambucil Azathioprine	Alkylating agents suppressing production of immunocompetent lymphocytes	Anecdotal benefit reported in open studies but controlled trial of chlorambucil failed to show superiority over placebo Cyclophosphamide often combined with corticosteroids	Two cyclophosphamide in SSc lung disease – open study (Medsger et al. 1990) No 5-fluorouracil controlled trial (Clements et al. 1984)

Table 16 Immunology therapies for systemic sclerosis

Cyclosporin has been used in a few patients with positive results but high doses of the drug were followed by reports of hypertension and renal failure, both of which might be attributed to either the drug or systemic sclerosis (Denton *et al*. 1994). Physicians must be mindful of the fact that trough blood levels are multiplied two- to fourfold in patients taking calcium-channel blockers. Therefore a lower dose (5 mg/kg body wt per day) has been tried with encouraging results, but awaits confirmation in a controlled study.

The interferons a and g are currently being investigated in controlled trials after initial pilot studies. Recombinant interferon-g has immunoregulatory activities and is a potent inhibitor of collagen production by normal and scleroderma fibroblasts *in vitro*. Interferon-a, although theoretically less potent than interferon-g as an inhibitor of collagen synthesis, does not activate the class II cell-surface antigens—this may be a distinct advantage in a disease that is HLA-linked. None of these drugs should

be used for late-stage, stable, diffuse or limited cutaneous disease.

Long-term high-dose steroids have no place in the management of systemic sclerosis; indeed they are potentially toxic (Steen *et al*. 1994) and may be implicated in 'normotensive renal crisis'. Steroid usage has been restricted to patients with myositis, symptomatic serositis, the early oedematous phase of the skin disease, and refractory arthritis and tenosynovitis. The lowest possible but therapeutically effective dose should be sought.

Raynaud's, as described above, is a prominent feature of systemic sclerosis and occurs in over 97 per cent of cases. It is an aspect of the disease that can be relieved, though the response is variable. This idiosyncratic response may reflect the stage of the disease. If structural damage is present and severe, the patient will respond rather poorly to vasodilators alone. In uncomplicated cases, simple measures may suffice. As the attacks become more frequent, prolonged oral drug therapy, possibly on an intermittent basis, may be needed. Intravenous therapy and limited surgery are restricted to the most severe cases (see Table 17).

Treatment	Examples	Comments
1. *Simple measures*		
Non-drug	Hand warmers Protective clothing	Universally helpful; also useful to minimize cold exposure and ambient temperature changes in work environment
Pharmacological	Evening primrose oil	Evening primrose oil has been shown effective in controlled clinical trial
	Fish oil capsules	
2. *Oral vasodilators*		
Calcium-channel blockers	Nifedipine Nicardipine Felodipine Amlodipine	Relaxant equipotent, other dihydropyridines therefore best to try nonretard drugs in relation to first most effective
5-Hydroxytryptamine antagonist	Ketanserin	Only available on named-patient basis but shown effective in clinical trials
ACE inhibitors	Captopril Enalapril	In Raynaud's secondary to SSDs may protect from hypertensive renal crisis
3. *Topical vasodilators*	GTN patches	Shown effective in short-term use but often cause headaches
4. *Parenteral vasodilators*	Carbaprostacyclin (tromboxane prostaglandin E$_1$	Effective and reasonably tolerated Given for severe frequent attacks of Raynaud's, digital ulceration or gangrene, and before digital surgery
5. *Antibiotics*	Flucloxacillin Erythromycin	Important in secondary infection in Raynaud's; prolonged oral administration may be necessary
6. *Surgical procedures*		
Lumbar sympathectomy	Chemical or operative	For severe Raynaud's of lower limbs
Digital sympathectomy (radical denervation)		Useful treatment for isolated ischaemic of 1 or 2 digits
Debridement amputation	Surgical or auto	Such surgery should be conservative to allow maximum possibility of spontaneous healing

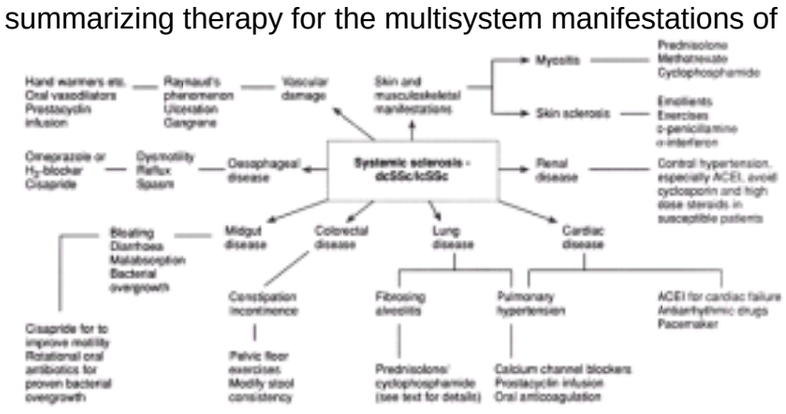
ACE, angiotensin-converting enzyme; GTN, glyceryl trinitrate; SSDs, diffuse connective tissue diseases.

Table 17 Treatment options for Raynaud's phenomenon

Because of the widespread vascular damage there has been a search to find drugs to protect injured endothelial cells and to prevent platelet aggregation and subsequent release of platelet-derived mediators. This search has been disappointing to date. Ketanserin, a serotonin antagonist, although useful in Raynaud's phenomenon, does not improve structural vascular disease. Dipyridamole and aspirin, although reducing the circulating plasma concentrations of b-thromboglobulin or circulating platelet aggregates, were not clinically effective in a randomized double-blind trial. Captopril, the angiotensin-converting enzyme inhibitor that has been so successful in the treatment of renal crisis, has been considered for the primary and possibly prophylactic treatment of vascular disease.

Our own approach to the management for some of the important subgroups of patients within the scleroderma spectrum of disorders is summarized below, and in Box 1. The management of localized scleroderma in adults and children is summarized in Table 18.

Box 1 Algorithm summarizing therapy for the multisystem manifestations of systemic sclerosis



Pattern of disease	Clinical features	Treatment	Prognosis
Localized morphoea	One or a few circumscribed sclerotic plaques with hyper- or hypopigmentation and often inflamed violaceous border	Often unnecessary Serial measurement to assess progress	Good prognosis; lesions less active within 3 years but pigmentary changes often persist
Generalized morphoea	Widespread pruritic lesions, often symmetrical and following the distribution of superficial veins	Suppress inflammatory component using corticosteroids: in children oral doses up to ½ mg/day have been recommended; infusion i.v. may be better D-Penicillamine is often used: in children 3 mg/kg day initially (maximum 13 mg/kg/day), adult daily maintenance dose at least 500 mg Methotrexate and systemic or intralesional interferon may be effective, and cyclosporin has been used in refractory cases	Internal organ pathology very rare. Raynaud's seldom associated Generally improves within 3 years of onset, although textural and pigmentary changes can remain
Linear scleroderma	Sclerotic areas occurring in linear distribution often on limbs and asymmetrical; in childhood can lead to serious growth defect of affected limbs; careful serial measurement of muscle bulk and limb length essential	Suppression of inflammation with oral or i.v. steroids (as above) Maintenance treatment with D-penicillamine or methotrexate Physiotherapy and appropriate regular exercise to minimize growth defect in childhood-onset foci	Long-term effects of childhood-onset foci minimized by effective suppression of inflammatory process, and by good physiotherapy Ultimately the disease tends to resolve, but in part remain active for many years
En coup de sabre	Linear scleroderma affecting the face of scalp, involving underlying subcutaneous tissues, muscles, periosteum and bone; underlying cerebral abnormalities have been reported	Therapeutic options as for linear scleroderma; systemic treatment only for active inflammatory lesions	Scarring, growth deficits, and alopecia persist but inflammatory component usually resolves

Table 18 Management of localized scleroderma in adults and children

Our approach to management of Raynaud's and scleroderma

1. Raynaud's phenomenon without other features of autoimmune rheumatic disease

At presentation

Detailed history and examination to look for evidence of asymptomatic autoimmune rheumatic disease. Thoracic-inlet radiographic imaging to exclude simple structural lesion (e.g. cervical rib). Baseline nailfold capillaroscopy and autoantibody profile (antinuclear antibodies, extractable nuclear antigen, anticentromere antibody) are important to determine whether the patient has primary Raynaud's phenomenon or is likely to develop features of an autoimmune rheumatic disease in the future. Infrared thermography with a standard cold challenge can be used to confirm diagnosis in equivocal cases and assess severity of vasospasm.

Follow-up

If capillaroscopy and autoantibodies are normal/negative, then simple approaches with advice on non-pharmacological relief (hand-warmers, thermal gloves), supplemented as necessary by vasodilator drugs. If the Raynaud's is of short duration (i.e., less than 2 years) the patient is followed annually until the fourth year after onset of symptoms. If capillaroscopy or autoantibody studies are positive, then give the same advice and treatment but follow up at 6 months, and then yearly for 5 years, with repeat capillaroscopy and autoimmune profile. Discharge from regular follow-up at 5 years if no other disease features have developed, and Raynaud's has not required parenteral vasodilator drugs.

2. Diffuse cutaneous systemic sclerosis

At presentation

History and physical examination generally establish the diagnosis. Assess the extent of visceral disease by baseline hand and chest radiography, lung function tests,

oesophageal motility study (scintigraphy or barium swallow), creatinine kinase, creatinine clearance and urinary protein excretion, Doppler echocardiography (with estimation of pulmonary arterial systolic pressure), and electrocardiogram. High-resolution CT lung scan, DTPA clearance, and bronchoscopy with studies of bronchoalveolar lavage fluid are usually performed in highly specialized units. Right-heart catheter studies and open lung biopsies are occasionally necessary to supplement the information obtained from other less invasive investigations. Once diagnosis is confirmed, tissue typing and autoantibody profiles may be useful to identify groups with poor prognosis. Raynaud's symptoms, if prominent, can respond well to intravenous prostacyclin infusions. Immunosuppressive therapy (e.g. antithymocyte globulin) may be considered for the most severe cases within 3 years of onset. Antifibrotic therapy is clearly the most appropriate approach for established diffuse cutaneous systemic sclerosis but unfortunately agents of proven efficacy are lacking. In our unit, for patients with progressive and extensive diffuse cutaneous systemic sclerosis, interferon-a is the first-line agent, monitored by serial skin sclerosis score. The results of formal controlled trials in diffuse cutaneous systemic sclerosis are keenly awaited. Lung fibrosis is treated by combination therapy with prednisolone and cyclophosphamide, as outlined in detail above.

Follow-up

This is especially important during the first 5 years from disease onset. Vigilant monitoring for renal involvement should include regular checks of blood pressure, 6-monthly 24-h urine collection for protein excretion and creatinine clearance. Six-monthly lung function tests and electrocardiograms, with yearly oesophageal motility and echocardiographic studies are generally performed in our own unit.

3. Limited cutaneous systemic sclerosis

At presentation

By the time of presentation the history is usually of some years duration and the physical signs, even if minimal, are generally diagnostic. Diligent attention to the assessment of internal organs is necessary, especially as the disease duration lengthens. The main risks in this scleroderma subset are of pulmonary hypertension, often in the absence of significant interstitial fibrosis, and involvement of the small and large bowel.

Treatment in this group is largely symptomatic and as yet no satisfactory drugs to halt the underlying processes are available. It may, however, be important to treat these patients at an asymptomatic stage in the hope of preventing or slowing down the chronic vascular damage. Current treatment in this group is directed mainly towards the vascular features (Raynaud's and pulmonary vascular disease, when present) and gastrointestinal complications. Reflux oesophagitis is almost universal but fortunately responds well to proton-pump inhibitors such as omeprazole, which we prefer to use rather than H_2-blockers, although the latter also often give good symptomatic relief. Oesophageal spasm may respond to cisapride. In contrast, although midgut disease and anorectal complications are less frequent, they are far more difficult to manage.

Follow-up

Annual follow-up is generally undertaken, with assessment of visceral disease, including renal function by creatinine clearance. These patients often require long-term vasodilator therapy, usually intensified during the winter months. Later in the disease (especially after 10 years of established limited cutaneous systemic sclerosis), pulmonary hypertension more likely. Patients carrying antitopoisomerase 1 (Scl-70) antibody are particularly susceptible to lung fibrosis and their lung function should be carefully monitored.

Education and support forms a particularly important component of the management of patients with systemic sclerosis, who should be made aware of the various support groups and the importance of non-pharmacological aspects of care including appropriate exercises, skin care, and the importance of being in a stable, warm ambient temperature.

Conclusion

Currently therefore, although there is no curative treatment for scleroderma, a careful consideration of the subset and stage of disease of the individual patient can maximize the use of the drugs currently available. It is hoped, more importantly, that the level and degree of activity of research into the cause and pathogenesis of the condition may eventually result in early rational effective treatment.

Acknowledgement

The authors are grateful to Dr Aine Burns, Consultant Nephrologist at the Royal Free Hospital, London, for her help with the section on renal disease in scleroderma.

Chapter References

Alarcon, G. S., Philips, R. M., and Wasner, C. K. (1985). DR antigens in systemic sclerosis: lack of clinical correlations. *Tissue Antigens*, **26**, 156–8.

Allen, R. G., Ansell, B. M., Clark, R. P., Goff, M. R., Waller, R., and Williamson, S. (1987). Localised scleroderma: treatment response measured by infrared thermography. *Thermology*, **2**, 550–3.

Altmann, R. D., Medsger, T. A., Jr, Bloch, D. A., and Michel, B. A (1990). Predictors of survival in systemic sclerosis (scleroderma). *Arthritis and Rheumatism*, **34**, 403–13.

Alton, E. and Turner-Warwick, M. (1988). Lung involvement in scleroderma. In *Systemic sclerosis (scleroderma)* (ed. C. M. Black and M. I. V. Jayson). Wiley, Chichester.

Ansell, B. M., Nasseh, G. A., and Bywaters, E. G. L. (1976). Scleroderma in childhood. *Annals of the Rheumatic Diseases*, **35**, 189–97.

Barnett, A.J., Tait, B.D., Barnett, M.A., and Toh, B.H. (1989). T lymphocyte subset abnormalities and HLA antigens in scleroderma. *Clinical and Experimental Immunology*, **76**, 24–9.

Barrowcliffe, M. P. and Jones, L. G. (1987). Solute permeability of the alveolar capillary barrier. *Thorax*, **42**, 1–10.

Belch, J. J. F. (1989). Raynaud's phenomenon. *Current Opinions in Rheumatology*, **1**, 490–8.

Belch, J. J. F. (1990). Raynaud's phenomenon. *Current Opinions in Rheumatology*, **2**, 937–41.

Birdi, N. *et al*. (1992). Childhood linear scleroderma: a possible role of thermography for evaluation. *Journal of Rheumatology*, **19**, 968–72.

Black, C. M. (1990*a*). Scleroderma—systemic sclerosis. In *Oxford textbook of clinical nephrology* (ed. J. S. Cameron, A. M. Davidson, J. P. Grunfeld, D. N. S. Kerr, and E. Ritz). Oxford University Press.

Black, C. M. (1990*b*). Systemic sclerosis and pregnancy. *Baillière's Clinical Rheumatology*, **4**, 105–24.

Black, C. M. (1990*c*). Systemic sclerosis: is there a treatment yet? I>Annals of the Rheumatic Diseases, **49**, 735–7.

Black, C. M. and Denton, C. P. (1995). The management of systemic sclerosis. *British Journal of Rheumatology*, **34**, 3–7.

Black, C. M. *et al*. (1983). Genetic susceptibility to scleroderma-like syndrome induced by vinyl chloride. *Lancet*, **i**, 53–5.

Black, C. M. *et al*. (1984). HLA antigens, autoantibodies and clinical subsets of scleroderma. *British Journal of Rheumatology*, **23**, 267–71.

Blann, A. D., Illingworth, K., and Jayson, M. I. V. (1993). Mechanisms of endothelial cell damage in systemic sclerosis and Raynaud's phenomenon. *Journal of Rheumatology*, **20**, 1325–30.

Bocchieri, M. H. and Jimenez, S. A. (1990). Animal models of fibrosis. *Rheumatic Disease Clinics of North America*, **16**, 153–67.

Bona, C. and Rothfield, N. (1994). Autoantibodies in scleroderma and tightskin mice. *Current Opinions in Immunology*, **6**, 931–7.

Boros, P., Muryoi, T., Spiera, H., Bona, C., and Unkeless, J. C. (1993). Auto-antibodies directed against different classes of FcR are found in sera of autoimmune patients. *Journal of Immunology*, **50**,

2018–24.

Briggs, D., Black, C. M., and Welsh, K. I. (1990). Genetic factors in scleroderma. *Rheumatic Disease Clinics of North America*, **16**, 31–51.

Briggs, D., Stephens, C., Vaughan, R., Welsh, K. I., and Black, C. M. (1993). A molecular and serologic analysis of the major histocompatibility complex and complement component C4 in systemic sclerosis. *Arthritis and Rheumatism*, **36**, 943–54.

Campbell, P. M. and LeRoy, E. C. (1975). Pathogenesis of systemic sclerosis: a vascular hypothesis. *Seminars in Arthritis and Rheumatism*, **4**, 351–68.

Carpentier, P. H. and Maricq, H. R. (1990). Microvasculature in systemic sclerosis. *Rheumatic Disease Clinics of North America*, **16**, 75–91.

Chosidow, O. *et al*. (1992). Sclerodermatous chronic graft-versus-host disease. Analysis of seven cases. *Journal of the American Academy of Dermatology*, **26**, 49–55.

Clements. P. J., Lachenbruch, P. A., Ng, S. C., Simmons, M., Sterz, M., and Furst, D. E. (1990). Skinscore: a semi-quantitative measure of cutaneous involvement that improves prediction of prognosis in systemic sclerosis. *Arthritis and Rheumatism*, **33**, 1256–63.

Clements, P.J., Lachenbruch, P.A., Sterz, M.A., *et al*. (1994). Cyclosporine in systemic sclerosis. Results of a forty-eight week open safety study in ten patients. *Arthritis and Rheumatism*, **36**, 75–83.

Cohen, S., Laufer, I., Snape, W. J., Shiau, Y-F., Levine, G. M., and Jimenez, S. A. (1980). The gastrointestinal manifestations of scleroderma: pathogenesis and management. *Gastroenterology*, **79**, 155–66.

Dabich, L., Sullivan, D. B., and Cassidy, J. T. (1974). Scleroderma in the child. *Journal of Paediatrics*, **85**, 770–5.

D'Angelo, W. A., Fries, J. F., Masi, A. T., and Shulman, L. E. (1969). Pathologic observations in systemic sclerosis (scleroderma). *American Journal of Medicine*, **46**, 428–40.

David, J., Wilson, J., and Woo, P. (1991). Scleroderma *en coup de sabre*. *Annals of the Rheumatic Diseases*, **50**, 260–2.

De Crombrugghe, B., Vuorio, T., and Karsenty, G. (1990). Control of type I collagen genes in scleroderma and normal fibroblasts. *Rheumatic Disease Clinics of North America*, **16**, 109–23.

Degiannis, D., Seibold, J., Czarnecki, M., Raskova. J., and Raska, K. (1990). Soluble and cellular markers of immune activation in patients with systemic sclerosis. *Clinical Immunology and Immunopathology*, **56**, 259–70.

Denton, C. P., Abdullah, A., Sweny, P., and Black, C. M. (1994). Acute renal failure occurring in scleroderma treated with cyclosporin A—a report of three cases. *British Journal of Rheumatology*, **33**, 90–2.

Dessein, P. H., Joffe, B. I., Metz, R. M., Millar, D. L., Lawson, M., and Stanwix, A. E. (1992). Autonomic dysfunction in systemic sclerosis: sympathetic overactivity and instability. *American Journal of Medicine*, **93**, 143–50.

Dunckley, H., Jazwinska, E. C., Gatenby, P. A., and Serjeatson, S. W. (1989). DNA DR-typing shows HL-DRw11 RFLPs are increased in frequency in both progressive systemic sclerosis and CREST variants of scleroderma. *Tissue Antigens*, **33**, 418–20.

Earnshaw, W. C., Bordwell, B., Marino, C., and Rothfield, N. (1986). Three human chromosomal autoantigens are recognised by sera from patients with anticentromere antibodies. *Journal of Clinical Investigation*, **77**, 426–30.

Ercilla, M. G. *et al*. (1981). HLA-antigens and scleroderma. *Archives of Dermatological Research*, **271**, 381–5.

Falanga, V. (1989). Localised scleroderma. *Medical Clinics of North America*, **73**, 1143–56.

Falanga, V., Medsger, T. A., Jr, Reichlin, M., and Rodnan, G. P. (1986). Linear scleroderma. Clinical spectrum and laboratory abnormalities. *Annals of Internal Medicine*, **104**, 849–57.

Follansbee, W. P. (1986). The cardiovascular manifestations of systemic sclerosis (scleroderma). *Current Problems in Cardiology*, **11**, 242–98.

Follansbee, W. P. *et al*. (1990). Myocardial fibrosis in systemic sclerosis (scleroderma): a case controlled clinicopathologic study. *Journal of Rheumatology*, **17**, 656–62.

Freni-Titulaer, L. J. W. *et al*. (1989). Connective tissue disease in south eastern Georgia: a case-control study. *American Journal of Epidemiology*, **129**, 404–9.

Genth, E. *et al*. (1990). Immunogenetic associations of scleroderma associated autoantibodies. *Arthritis and Rheumatism*, **33**, 657–65.

Germain, B. F. *et al*. (1981). Increased prevalence of DRw3 in the CREST syndrome. *Arthritis and Rheumatism*, **24**, 857–9.

Gladman, D. D. *et al*. (1981). Increased frequency of DR5 in scleroderma. *Arthritis and Rheumatism*, **24**, 854–6.

Greydanus, M. P. and Camilleri, M. (1989). Abnormal post-cibal gastric and small bowel motility due to neuropathy or myopathy in systemic sclerosis. *Gastroenterology*, **96**, 110–15.

Hanson, V. (1976). Dermatomyositis, scleroderma and polyarteritis nodosa. *Clinical Rheumatic Diseases*, **2**, 445–67.

Harrison, N. K. *et al*. (1989). Pulmonary involvement in systemic sclerosis: the detection of early changes by thin section CT scan, bronchoalveolar lavage and ^{99m}Tc-DTPA clearance. *Respiratory Medicine*, **83**, 403–14.

Harrison, N. K. *et al*. (1991). Structural features of interstitial lung disease in systemic sclerosis. *American Review of Respiratory Diseases*, **144**, 706–13.

Harvey, W. (1983). Submucous fibrosis. In *Proceedings of the scleroderma symposium.* Smith Kline and French Laboratories Ltd, Welwyn Garden City.

Helfrich, D. J., Banner, B., Steen, V. D., and Medsger, T. A., Jr (1989). Renal failure in normotensive patients with systemic sclerosis. *Arthritis and Rheumatism*, **32**, 1128–34.

Hochberg, M. C. (1994). Silicone breast implants and rheumatic disease. *British Journal of Rheumatology*, **33**, 601–2.

Jabs, E. W., Tuck-Muller, I. C. M., Anhalt, G. J., Earnshaw, W. C, Wise, R. W., and Wigley, F. (1993). Cytogenetic survey in systemic sclerosis and association with autoantibodies to RNA polymerases I and III. *Journal of Clinical Investigation*, **91**, 1399–404.

Jazwinska, E. C. *et al*. (1990). HLA-DRw15 is increased in frequency in Japanese scleroderma patients. *Disease Markers*, **8**, 323–6.

Kahaleh, M. B. (1991). Soluble immunologic products in scleroderma sera. *Clinical Immunology and Immunopathology*, **58**, 139–44.

Kahaleh, M. B. and Mattuci-Cerinic, M. (1995). Raynaud's phenomenon and scleroderma: dysregulated neuroendothelial control of vascular tone. *Arthritis and Rheumatism*, **38**, 1–4.

Kallenberg, C. G. (1994). Antitopoisomerase and anticentromere antibodies in the sclerodermatous complex. *Clinical Reviews in Allergy*, **12**, 221–35.

Kallenberg, C. G. M., Wonda. A. A., Hoet, M. H., and Van Venrooij, W. J. (1988). Development of connective tissue disease in patients presenting with Raynaud's phenomenon: a six-year follow-up with emphasis on the predictive value of antinuclear antibodies as detected by immunoblotting. *Annals of the Rheumatic Diseases*, **47**, 634–41.

Kantor, T. V., Whiteside, T. L., Friberg, D., Buckingham, R. B., and Medsger T. A., Jr (1992). Lymphokine activated killer cells and natural killer cell activities in patients with systemic sclerosis. *Arthritis and Rheumatism*, **35**, 694–9.

Kondo, H. *et al*. (1985). Histocompatibility antigens in progressive systemic sclerosis. In *Systemic sclerosis (scleroderma)* (ed. D. M. Black and A. R. Myers). Gower Medical, New York.

Kovalchik, M. T., Guggenheim, S. J., Robertson, J. S., and Steigerwald, J. C. (1978). The kidney in progressive systemic sclerosis: a prospective study. *Annals of Internal Medicine*, **89**, 881–7.

Kuwana, M., Kaburaki, J., Okano, Y., Inoko, H., and Tsuji, K. (1993). The *HLA-DR* and *DQ* genes control the autoimmune response to DNA topoisomerase 1 in systemic sclerosis (scleroderma). *Journal of Clinical Investigation*, **92**, 1296–301.

Kuwana, M., Medsger, T.A., and Wright, T.M. (1995). T cell proliferative response induced by DNA topoisomerase 1 in patients with systemic sclerosis and healthy donors. *Journal of Clinical Investigation*, **96**, 586–96.

Lally, E. V. (1992). Raynaud's phenomenon. *Current Opinions in Rheumatology*, **4**, 825–36.

Langevitz, P., Buskila, D., Gladman, D. D., Darlington, G. A., Farewell, V. T., and Lee, P. (1992). HLA alleles in systemic sclerosis: associations with pulmonary hypertension and outcome. *British Journal of Rheumatology*, **31**, 609–13.

LeRoy, E. C. (1974). Increased collagen synthesis by scleroderma skin fibroblasts *in vitro*. A possible defect in regulation or activation of scleroderma fibroblasts. *Journal of Clinical Investigation*, **54**, 880–9.

LeRoy, E. C. *et al*. (1988). Scleroderma (systemic sclerosis): classification, subsets, and pathogenesis. *Journal of Rheumatology*, **15**, 202–5.

Levine, B., Hardwick, J. M., Trapp, B. D., Crawford, T. O., Bollinger, R. C., and Griffin, D. E. (1991). Antibody mediated clearance of alpha virus infection in neurones. *Science*, **254**, 865–70.

Livingstone, J. Z. *et al*. (1987). Systemic sclerosis (scleroderma): clinical, genetic and serological subsets. *Journal of Rheumatology*, **14**, 512–18.

Luderschmidt, C. *et al*. (1987). Associations of progressive systemic scleroderma to several HLA-B and HLA-DR alleles. *Archives of Dermatology*, **123**, 1188–91.

Lupoli, S., Amlot, P., and Black, C. M. (1990). Normal immune responses in systemic sclerosis. *Journal of Rheumatology*, **17**, 323–37.

Lynch, C. J. *et al*. (1982). Histocompatibility antigens in progressive systemic sclerosis (scleroderma). *Journal of Clinical Immunology*, **2**, 314–18.

Ma, J., Chapman, G. V., Chen, S. L., Melick, G., Penny, R., and Briet, S. N. (1991). Antibody penetration of viable human cells: I. Increased penetration of human lymphocytes by anti-RNP IgG. *Clinical and Experimental Immunology*, **84**, 83–91.

McCarthy, G. M. and Kenny, D. (1992). Dietary fish oil in rheumatic diseases. *Seminars in Arthritis and Rheumatism*, **21**, 318–75.

McHugh, N. J. *et al*. (1994). Anti-centromere antibodies (ACA) in systemic sclerosis patients and their relatives: a serological and HLA study. *Clinical and Experimental Immunology*, **96**, 267–74.

Mackel, A. M., De Lustro, F., Harper, F. E., and LeRoy, E. C. (1982) Antibodies to collagen in scleroderma. *Arthritis and Rheumatism*, **25**, 522–31.

Maricq, H. R. *et al*. (1989). Prevalence of scleroderma spectrum disorders in the general population of South Carolina. *Arthritis and Rheumatism*, **32**, 998–1006.

Masi, A. T. (1988). Classification of systemic sclerosis (scleroderma): relationship of cutaneous subgroups in early disease to outcome and serologic reactivity. *Journal of Rheumatology*, **15**, 894–8.

Mason, G. R., Uzzler, J. M., Effros, R. M., and Reid, E. (1983). Rapidly reversible alterations of pulmonary epithelial permeability induced by smoking. *Chest*, **83**, 6–11.

Mauch, C., Eckes, B., Hunzelman, N., Oono, T., Kozlowska, E., and Krieg, T. (1993). Control of fibrosis in systemic scleroderma. *Journal of Investigative Dermatology*, **100**, 92–6S.

Medsger, T. A., Jr (1979). Progressive systemic sclerosis—skeletal muscle involvement. *Clinical Rheumatic Diseases*, **5**, 103–13.

Miller, K. S., Smith, E. A., Kinsella, M., Schabel, S. I., and Silver, R. M. (1990). Lung disease associated with progressive systemic sclerosis. Assessment of interlobar variation by bronchoalveolar lavage and comparison with non-invasive evaluation of disease activity. *American Review of Respiratory Disease*, **141**, 301–6.

Muller, N. L. and Miller, R. R. (1990). Computed tomography of chronic diffuse infiltrative lung disease. *American Review of Respiratory Disease*, **142**, 1206–15; 1440–8.

Myers, A. R. *et al*. (1989). Class II major histocompatibility complex antigens and pulmonary fibrosis in systemic sclerosis. (Abstract). *Arthritis and Rheumatism*, **32** (Suppl.), S77.

Padula, S. J., Clark, R. B., and Korn, J. H. (1986). Cell mediated immunity in rheumatic diseases. *Human Pathology*, **17**, 254–63.

Pearson, J. D. (1990). The endothelium: its role in systemic sclerosis. *Annals of the Rheumatic Diseases*, **50**, 866–71.

Pollard, K. M., Reimer, G., and Tan, E. M. (1989). Autoantibodies in scleroderma. *Clinical and Experimental Rheumatology*, **7** (Suppl. 3), 57–62.

Pope, J. E. and Bellamy, N. (1993). Outcome measurements in scleroderma clinical trials. *Seminars in Arthritis and Rheumatism*, **23**, 22–33.

Prescott, R. J., Freemont, A. J., Jones, C. J., Hoyland, J., and Fielding, P. (1992). Sequential dermal microvascular and perivascular changes in the development of scleroderma. *Journal of Pathology*, **166**, 255–63.

Reeves, W. H., Satih, M., Wang, J., Chou, C-H, and Ajmani, A. K. (1994). Antibodies to DNA, DNA binding proteins, and histones. *Rheumatic Disease Clinics of North America*, **20**, 1–28.

Reveille, J. D., Owerbach, D., Goldstein, R., Moreda, R., Isern, R. A., and Arnett, F. C. (1992 *a*). Association of polar amino acids at position 26 of the *HLA-DQB1* first domain with the anticentromere autoantibody response in systemic sclerosis (scleroderma). *Journal of Clinical Investigation*, **89**, 1208–13.

Reveille, J. D., Durban, E., Macleod, St Clair, *et al*. (1992*b*). Association of amino acid sequences in the *HLA-DQB1* first domain with the anti-topoisomerase 1 autoantibody response in scleroderma. *Journal of Clinical Investigation*, **90**, 973–80.

Rodnan, G. P. and Benedek, T. G. (1962). An historical account of the study of progressive systemic sclerosis (diffuse scleroderma). *Annals of Internal Medicine*, **57**, 305–19.

Roumm, A. D., Whiteside, T. L., Medsger, T. A., Jr, and Rodnan, G. P. (1984). Lymphocytes in the skin of patients with progressive systemic sclerosis. *Arthritis and Rheumatism*, **29**, 645–53.

Russell, M. L. (1988). Muscle and nerve in systemic sclerosis (scleroderma). In *Systemic sclerosis: scleroderma* (ed. C. M. Black and M. I. V. Jayson). Wiley, Chichester.

Russell, M. L., Friesen, D., Henderson, R. D., and Hanna, W. M. (1982). Ultrastructure of the oesophagus in scleroderma. *Arthritis and Rheumatism*, **25**, 1117–23.

Sasaki, T., Denpo, K., Ono, H., and Nakjima, J. (1991). HLA in systemic sclerosis (PSS) and familial scleroderma. *Journal of Dermatology*, **18**, 18–24.

Seibold, J. R., Medsger, T. A., Jr, Winkelstein, A., Kelly, R. H., and Rodnan, G. P. (1982). Immune complexes an progressive systemic sclerosis (scleroderma). *Arthritis and Rheumatism*, 25, 1167–73.

Shorrock, C. J. and Rees, W. D. W. (1988). Gastrointestinal manifestations of systemic sclerosis. In *Systemic sclerosis: scleroderma* (ed. C. M. Black and M. I. V. Jayson). Wiley, Chichester.

Silman, A. J. (1995). Scleroderma. *Ballière's Clinical Rheumatology*, **9**, 471–8.

Silver, R. M. (1990). Clinical aspects of systemic sclerosis (scleroderma). *Annals of the Rheumatic Diseases*, **50** (Suppl. 4), 854–61.

Silver, R. M., Miller, K. S., Kinsella, M. B., Smith E. A., and Schabel, S. I. (1990). Evaluation and management of scleroderma lung disease using bronchoalveolar lavage. *American Journal of Medicine*, **88**, 470–6

Steen, V. D. (1984). Factors predicting the development of renal involvement in progressive systemic sclerosis. *American Journal of Medicine*, **76**, 799–86.

Steen, V. D., Powell, D. L., and Medsger, T. A., Jr (1988). Clinical correlations and prognosis based on serum autoantibodies in patients with systemic sclerosis. *Arthritis and Rheumatism*, **31**, 196–203.

Steen, V. D., Constantino, J. P., Shapiro, A. P., and Medsger, T. A., Jr (1990). Outcome of renal crisis in systemic sclerosis: relation to the availability of converting enzyme inhibitors (ACE). *Annals of Internal Medicine*, **113**, 352–7.

Steen, V. D., Conte, C., and Medsger, T. A., Jr (1994). Case-control study of corticosteroid use prior to scleroderma renal crisis. (Abstract). *Arthritis and Rheumatism*, **37** (Suppl.), S360.

Subcommittee for Scleroderma Criteria of the American Rheumatism Association Diagnostic and Therapeutic Criteria Committee (1980). Preliminary criteria for the classification of systemic sclerosis (scleroderma). *Arthritis and Rheumatism*, **23**, 581–90.

Tan, E. M. (1989). Antinuclear antibodies: diagnostic markers for autoimmune diseases and probes for cell biology. *Advances in Immunology*, **44**, 93–151.

Tarkowski, A. and Lindgren, I. (1994). Beneficial effects of anti-thymocyte globulin in severe cases of progressive systemic sclerosis. *Transplantation Proceedings*, **26**,3197–9.

Traub, Y. M. *et al*. (1984). Hypertension and renal failure (scleroderma renal crisis) in progressive systemic sclerosis. Report of a 25 year experience with 68 cases. *Medicine*, **62**, 335–52.

Uitto, J., Bauer, E., and Eisen, A. Z. (1979). Scleroderma: increased biosynthesis of triple-helical type I and type III procollagens associated with unaltered collagenase by skin fibroblasts in culture.

Journal of Clinical Investigation, **64**, 921–30.

Valesini, G. *et al*. (1993). Geographical clustering of scleroderma in a rural area in the province of Rome. *Clinical and Experimental Rheumatology*, **11**, 41–7.

Warrick, J. H., Bhalla, M., Schabel, S. I., and Silver, R. M. (1991). High resolution computed tomography in early scleroderma lung disease. *Journal of Rheumatology*, **18**, 1520–8.

Wells, A. U., Hansell, D. M., Harrison, N. K., Lawrence, R., Black, C. M., and du Bois, R. M. (1993 *a*). Clearance of inhaled 99m-Tc DTPA predicts the clinical course of fibrosing alveolitis. *European Respiratory Journal*, **6**, 797–802.

Wells, A. U., Hansell, D. M., Rubens, M. B., Cullinan, P., Black, C. M., and du Bois, R. M. (1993 *b*). The predictive value of appearances on thin section computed tomography in fibrosing alveolitis, *American Review of Respiratory Disease*, **148**, 1076–82.

Wells, A. U., Hansell, D.M., Rubens, M.B., *et al*. (1994). Fibrosing alveolitis associated with progressive systemic sclerosis: the relationship between bronchoalveolar lavage cellularity and computed tomographic appearances. *American Journal of Respiratory and Critical Care Medicine*, **150**, 462–8.

White, B. (1994). Immunologic aspects of scleroderma. *Current Opinion in Rheumatology*, **6**, 612–15.

Whiteside, T.L., Medsger, T.A. Jr., and Rodnan, G.P. (1983). HLA-DR antigens in progressive systemic sclerosis. *Journal of Rheumatology*, **10**, 128–31.

Williams, W.V., Rook, A.H., Freundlich, B.F., *et al*. (1995). T-cell receptors in scleroderma skin. The effect of photopheresis. *Annals of the New York Academy of Sciences*, **756**, 424–7.

Wollersheim, H., Thien, T., Hoet, M. H., and Van Venrooy, W. J. (1989). The diagnostic value of several immunological tests for antinuclear antibody in predicting the development of connective tissue disease in patients presenting with Raynaud's phenomenon. *European Journal of Clinical Investigation*, **19**, 535–41.

Youssef, P., Englert, H., and Bertouch, J. (1993). Large vessel occlusive disease associated with CREST syndrome and scleroderma. *Annals of the Rheumatic Diseases*, **52**, 564–9.

Zacheriae, H., Halkier-Sorensen, L., Heickendorff, L., Zacheriae, E., and Hansen, H.E. (1990). Cyclosporin A treatment of systemic sclerosis. *British Journal of Dermatology*, **122**, 677–81.

5.9.1 Polymyositis and dermatomyositis in adults

Ira N. Targoff

Introduction

Polymyositis and dermatomyositis are the most common forms of 'idiopathic inflammatory myopathy', a category that also encompasses inclusion body myositis, and several rare forms. Inflammatory myopathy may also be induced by drugs or infections. Dermatomyositis is distinguished from polymyositis by the presence of a characteristic rash. Weakness affecting skeletal muscle is the major clinical manifestation of most patients, although the skin rash or other extramuscular features may predominate in some.

Wagner first described a case as polymyositis in 1863, while Unverricht first used the term dermatomyositis in 1887. Early studies often included cases that would not fit our present concept. The wide acceptance of the diagnostic criteria of Bohan and Peter (1975) has served to standardize subsequent studies and promote recognition.

Epidemiology

Incidence and prevalence

Estimates of the annual incidence of polymyositis/dermatomyositis have ranged from 1 to 9 cases per million per year, and prevalence from 2.4 to 10.7 cases per 100 000 (Cronin and Plotz 1990; Sigurgeirsson *et al*. 1992; Ahlstrom *et al*. 1993). For example, Oddis *et al*. (1990a) in Pittsburgh, Pennsylvania, found 5.5 per million per year, with an increased incidence over time (2.5 cases per million per year in the first decade, 8.9 per million per year in the second) that could reflect better detection and recognition, or a true increase.

Polymyositis is more common than dermatomyositis in most studies of adults (Bohan *et al*. 1977; Tymms and Webb 1985), but the ratio (less than 2:1) is low and may vary with the population, referral patterns, or criteria for dermatomyositis. Some studies found dermatomyositis to be more common (Ramirez *et al*. 1990; Love *et al*. 1991; Koh *et al*. 1993). Overlap syndromes with other autoimmune rheumatic diseases occur in 15 to 20 per cent.

Risk factors

Polymyositis/dermatomyositis is fourfold more common in the United States in black patients than white patients (Cronin and Plotz 1990). The female:male ratio is 2:1 overall, but lower in myositis with malignancy, and higher during the childbearing years (5:1) and in patients with associated autoimmune rheumatic diseases (Bohan *et al*. 1977). Polymyositis and dermatomyositis may begin at almost any age, but the peak in adults is usually from 40 to 60 years; it tends to develop later in white than in black women (Oddis *et al*. 1990a), and later with cancer (Love *et al*. 1991).

Cases of adult polymyositis and dermatomyositis among family members are very rare, but have been observed (Gurley *et al*. 1992; Garlepp 1993). A family history of myopathy should make one question the diagnosis of polymyositis/dermatomyositis. Other autoimmune disease in relatives of polymyositis and dermatomyositis patients is not unusual (Mbauya *et al*. 1993).

Temporal and geographic factors

Most studies have not found variation in the incidence of adult-onset polymyositis /dermatomyositis with time of year, although one study found a higher rate from March to May (Manta *et al*. 1989). However, certain autoantibody-defined subgroups show significant seasonal differences in the onset of myositis: anti-Jo-1-associated myositis begins more often in the spring, and anti- **SRP**-(signal recognition particle)-myositis more often in the autumn (Leff *et al*. 1991).

Polymyositis and dermatomyositis occur in all parts of the world, but in developing countries infectious myositis is more common. Occasional local clusters of myositis have been reported, which may have atypical features (Nagaraja *et al*. 1992). Certain autoantibodies appeared to differ in prevalence in different regions (Love *et al*. 1992).

Clinical picture

Classification

Several classification systems have been proposed to define subgroups of patients with idiopathic inflammatory myopathies that are more clinically homogeneous (Dalakas 1988a; Miller 1994; Medsger and Oddis 1995), and able to predict course and responsiveness. Disagreement remains, however, due to our lack of knowledge of aetiology.

Most recent studies have used the Bohan and Peter (1975) classification (Table 1), or modifications. Pure polymyositis is separated from dermatomyositis, but not in the presence of malignancy, juvenile onset, or an associated autoimmune rheumatic disease. Patients are usually considered to have dermatomyositis rather than polymyositis only when the rash is observed, but some also include patients without overt skin lesions who have characteristic muscle pathology (Byrne and Dennett 1993). Inclusion-body myositis is a clinically and histologically distinct entity, and should probably be classified separately even when associated with malignancy or autoimmune rheumatic diseases (see later section). Several rare but distinctive forms of idiopathic myositis have also been defined (Table 1).

I. Primary idiopathic polymyositis
II. Primary idiopathic dermatomyositis
III. Polymyositis or dermatomyositis with malignancy
IV. Juvenile dermatomyositis (or polymyositis)
V. Overlap syndrome of polymyositis or dermatomyositis with another autoimmune rheumatic disease
VI. Inclusion-body myositis[b]
VII. Rare forms of idiopathic myositis (a) Granulomatous myositis (b) Eosinophilic myositis (c) Focal myositis (d) Orbital myositis

[a]Classes I to V correspond to Bohan and Peter's original classification (Bohan and Peter (1975).
[b]Patients with inclusion body myositis in association with malignancy or other autoimmune rheumatic diseases should be considered to be class VI.

Table 1 Classification of idiopathic inflammatory myopathies [a]

Certain specific autoantibodies can define subgroups of patients that differ from the overall myositis population with regard to clinical features, response to therapy, prognosis, and HLA type (Love *et al*. 1991). These autoantibodies can be used to classify patients in a way that is complementary to clinical classification.

Clinical features

Myositis

Weakness

Muscle weakness is the main clinical feature of both polymyositis and dermatomyositis, occurring in almost all patients. It usually develops insidiously over weeks to months, generally more slowly in polymyositis than dermatomyositis (Casademont *et al*. 1993). More indolent cases may be seen, progressing over years (more in polymyositis than dermatomyositis), that must be distinguished from inclusion-body myositis. More rapid onset occurs occasionally.

The weakness is typically symmetrical, affecting the large proximal muscles around the shoulders, hips, thighs, trunk and neck. The lower extremities are often involved first, but in most patients both upper and lower extremity involvement occurs (Plotz *et al*. 1989; Henriksson and Lindvall 1990). Patients often have impairment in performance of daily activities such as standing from a chair, getting out of a car, climbing stairs, reaching (e.g. into cabinets), working overhead (hanging clothes, etc.), or combing their hair. The gait may be affected. Getting out of bed, sitting from a supine position, or raising the head off the pillow may become difficult. Weakness of distal muscles is uncommon in polymyositis/dermatomyositis, but may occur late in the course (in about 10 per cent of caes, Love *et al*. 1991). Patients may note impairment of chewing or dysphagia (see below). Involvement of the face is unusual, and involvement of extraocular muscles is rare, and should suggest other diagnoses (Dalakas 1991). Within regions of weakness, there is usually diffuse involvement, unlike some myopathies in which weakness and atrophy may be highly selective for specific muscles.

Other muscle manifestations

Myalgia and muscle tenderness occur in about half of patients, usually as mild aching or soreness. They are usually not predominant, but may be more severe when myositis develops acutely, and occasionally can lead to confusion with polymyalgia rheumatica (Hopkinson *et al*. 1991). There may be atrophy in chronic disease, seen in 9 per cent by Love *et al*. (1991), more commonly in polymyositis than dermatomyositis. Contractures may occur with disease of long duration.

Examination

Muscle strength can be assessed as part of initial evaluation and later monitoring of progress by observing activities and by direct muscle testing. Walking, standing from a squatting position or low chair without using the arms, sitting up or raising the head from a supine position, and raising the arms overhead, should be tested, or multiple repetitions timed (Csuka and McCarty 1985; Moxley 1994). Deltoids, biceps, iliopsoas, quadriceps, and other proximal muscles should be manually tested directly, and graded by systems such as the Medical Research Council scale. In order to overcome insensitivity and subjectivity, and standardize and quantify such testing, there is continuing interest in the use of biomechanical measures of muscle strength with machines that can measure applied force (Moxley 1994). Fafalak *et al*. (1994) found that improved strength by a mechanical measure correlated well with a functional assessment scale. All types of muscle strength testing depend on effort, and may be complicated by fatigue or pain of arthritis or myalgia.

Cutaneous manifestations

Rash

The dermatomyositis rash, found in about 30 to 40 per cent of adults with myositis, most commonly precedes the weakness by weeks to several months, or even longer (Hochberg *et al*. 1986; Rockerbie *et al*. 1989). At presentation, 93 per cent of adult dermatomyositis patients of Bohan *et al*. (1977) had a rash but only 53 per cent weakness. The activity of the rash may parallel that of the weakness or may be independent, and can persist after the myositis resolves.

Gottron's lesions

Erythematous or violaceous, sometimes scaly, papules or plaques (Gottron's papules) or macular patches (Gottron's sign) may occur over the metacarpophalangeal and proximal (and less often distal) interphalangeal joints (Fig. 1) (Franks 1988), and also over the extensor surfaces of the knees, wrists, elbows or medial malleoli

(Fig. 2). These lesions may be found in 70 to 80 per cent of dermatomyositis patients, and are considered pathognomonic (Euwer and Sontheimer 1994). Telangiectasia and atrophy can occur. The erythematous finger rash of systemic lupus erythematosus differs in that it usually occurs between the knuckles.

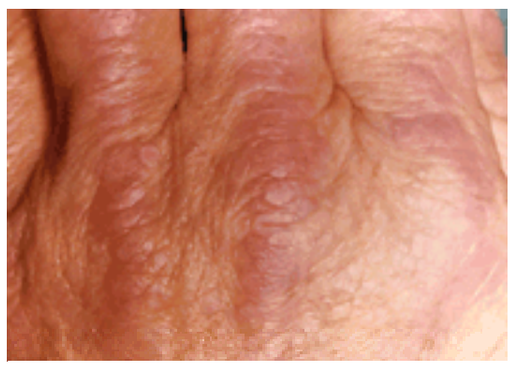


Fig. 1 Skin over the metacarpophalangeal joints of a patient with dermatomyositis shows the characteristic erythematous lesions of Gottron's sign. (By courtesy of the Department of Dermatology, University of Oklahoma Health Sciences Center.)

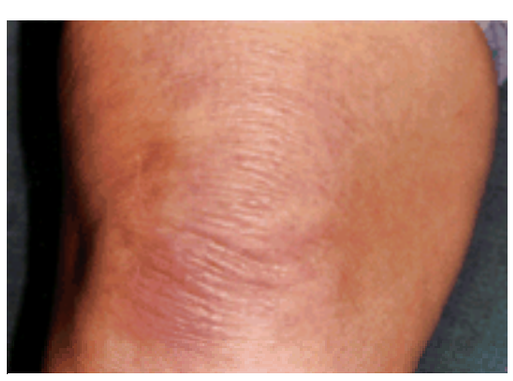


Fig. 2 Skin over the knee of a patient with dermatomyositis shows a characteristic erythematous, violaceous lesion. (By courtesy of the Department of Dermatology, University of Oklahoma Health Sciences Center.)

Erythematous and/or poikilodermatous rash

A macular erythematous or violaceous eruption may involve the upper chest, neck, shoulders, extremities, hands, scalp and face. It may develop into poikiloderma, varied hyper- and hypopigmentation with atrophy and fine telangiectasias. Typical of dermatomyositis are the 'V' sign (involvement at the anterior base,'V', of the neck; 36 per cent of cases), and the 'shawl' sign (back of neck, upper torso, and shoulders in a shawl-like pattern; 22 per cent) (Love *et al*. 1991). Erythema may extend from the joints along the course of the extensor tendons (linear extensor erythema, Franks 1988; Fig. 3). Kasteler and Callen (1994) noted a high frequency of scalp involvement (82 per cent), marked by erythema, atrophy, scale, and sometimes alopecia that can be misdiagnosed as psoriasis or seborrheic dermatitis. Rash on the malar areas, forehead and chin may lead to confusion with systemic lupus erythematosus, although involvement of the nasolabial folds may be a clue (Plotz *et al*. 1989).

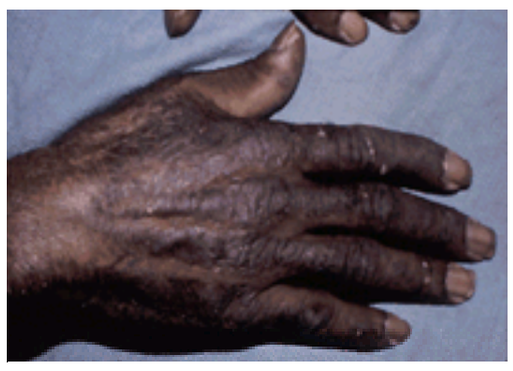


Fig. 3 Hand of a patient with dermatomyositis shows Gottron's papules over the proximal interphalangeal and metacarpophalangeal joints, with marked linear extensor erythema extending from the joints along the tendons. In black patients, the lesions may appear hyperpigmented. (By courtesy of Dr Frank C. Arnett, University of Texas at Houston Health Sciences Center.)

Patients often report exacerbation or development of new lesions after sun exposure (Callen 1987), and this has also been documented after therapeutic ultraviolet or experimental solar-simulated light (Cheong *et al*. 1994; Euwer and Sontheimer 1994). The distribution also suggests photosensitivity, but can occur without sun exposure (Franks 1988).

Heliotrope

The heliotrope rash, found in about 30 to 60 per cent of dermatomyositis, is a purplish, lilac-coloured suffusion (resembling a heliotrope flower) around the eyes, particularly the upper eyelids and surrounding area (Fig. 4), often associated with periorbital oedema. It is characteristic but not pathognomonic of dermatomyositis, since a similar appearance may occasionally be seen in allergy, trichinosis, lupus, or other conditions (Euwer and Sontheimer 1994). It may be difficult to see in black patients.

Fig. 4 A severe heliotrope rash of dermatomyositis, with the characteristic lilac colour and accompanying periorbital oedema. (By courtesy of the Department of

Dermatology, University of Oklahoma Health Sciences Center.)

Other cutaneous features

Periungual telangiectasia and/or haemorrhages are also seen in dermatomyositis (Franks 1988). Nailfold capillaries may show marked changes similar to those in scleroderma, including thrombosis and haemorrhage, giant capillary loops, and capillary loss, that may parallel disease activity (Fig. 5). The cuticles may be thickened, roughened, and irregular (Fig. 6) (Caro 1988).

Fig. 5 Nailfold vascular changes in a patient with amyopathic dermatomyositis for 4 years. Capillary dilatation, dropout, and haemorrhage are evident without microscopy. Gottron's papules were present over the proximal interphalangeal joints of the index and middle fingers, with suggestive lesions visible over the distal interphalangeal joints. (By courtesy of Dr Lela Lee, Department of Dermatology, University of Oklahoma Health Sciences Center.)

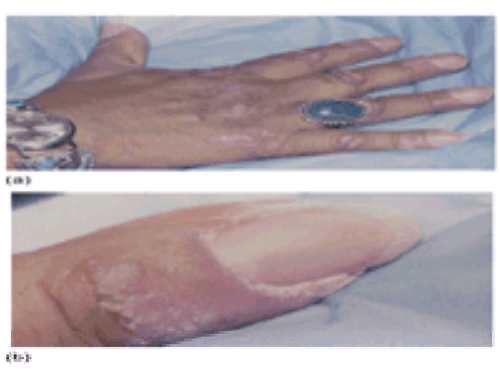


Fig. 6 (a) Hand of a patient with anti-Mi-2-positive dermatomyositis with severe weakness of 2 months duration. Typical Gottron's papules are seen over the metacarpophalangeal, and proximal and distal interphalangeal joints, with more scale than in Fig. 7. There is cuticular hyperkeratosis around many nails, and mechanic's hand change on the thumb. (By courtesy of Dr E. Taylor-Albert, University of Oklahoma Health Sciences Center). (b) Thumb of the patient in (a). Cuticular changes are evident. A Gottron's papule is over the interphalangeal joint.

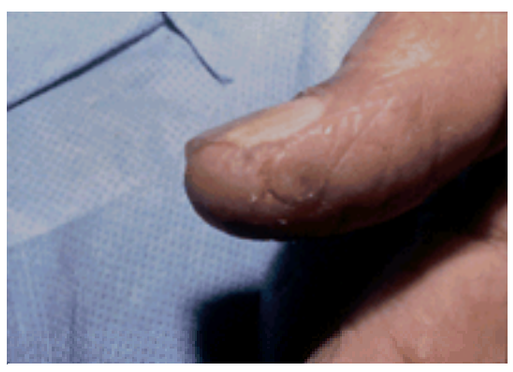


Fig. 7 Thumb of a patient with anti-Jo-1-positive polymyositis. The edge of the thumb shows fissuring and some hyperkeratosis as in mechanic's hands. (By courtesy of Dr Frank C. Arnett, University of Texas at Houston Health Sciences Center.)

Hyperkeratosis and scaling with fissuring and hyperpigmentation may appear as dirty horizontal lines along the lateral and palmar aspects of the fingers ('mechanic's hands') (Fig. 7). This lesion was associated strongly with antisynthetase autoantibodies (Love *et al*. 1991) (see below). Although often seen without other signs of dermatomyositis, the histology resembles dermatomyositis (Mitra *et al*. 1994).

Calcinosis can occur in adults but is more common in juvenile dermatomyositis. It can be extensive, and usually occurs late. Cutaneous vasculitis, especially of the fingers, is not infrequent in adult dermatomyositis (Ramirez *et al*. 1990); cutaneous ulcers associated with severe vasculopathy of juvenile dermatomyositis (Roberts and Fink 1988) may rarely be seen in adults (Fig. 8). Other rare manifestations include panniculitis (Fusade *et al*. 1993), and erythroderma (Ramirez *et al*. 1990).

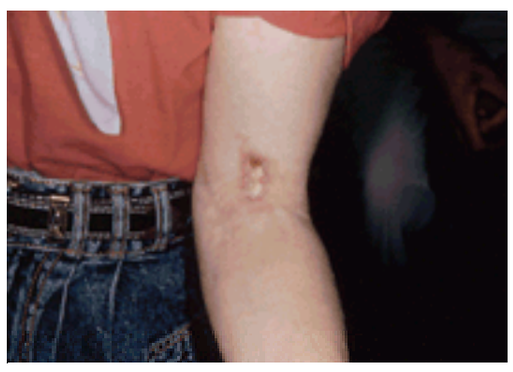


Fig. 8 Cutaneous ulcer in a 32-year-old woman with dermatomyositis. Endothelial cell swelling and change (vasculopathy) without inflammatory invasion of the vessel wall (vasculitis) was seen by biopsy. (By courtesy of the Department of Dermatology, University of Oklahoma Health Sciences Center.)

Amyopathic dermatomyositis

There is increasing recognition of patients with typical cutaneous dermatomyositis who do not develop myositis (Fig. 9). Some of the patients have subclinical myositis demonstrable by testing (see below), and some later develop overt myositis, but some have no sign of myositis for as long as 12 years (Rockerbie *et al*. 1989; Stonecipher *et al*. 1993). 'Amyopathic dermatomyositis' (or 'dermatomyositis sine myositis') is applied to those with rash alone for at least 2 years without treatment. Such patients cannot satisfy criteria for dermatomyositis, which are based on the myositis, and alternative criteria and a separate classification category have been proposed (Euwer and Sontheimer 1993). The risk of malignancy and systemic complications appears similar to usual dermatomyositis (Euwer and Sontheimer 1993; Euwer and Sontheimer 1994).

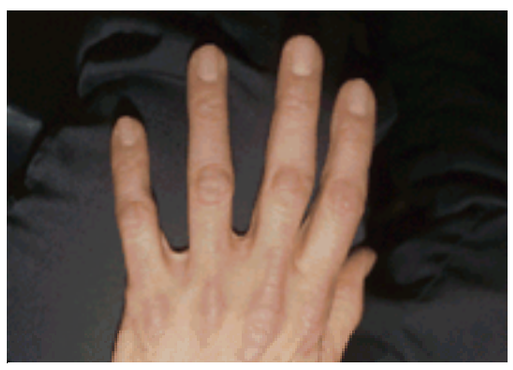


Fig. 9 Skin over the finger joints of a patient with cutaneous dermatomyositis for less than 6 months, but without evident muscle disease, showing characteristic lesions. The metacarpophalangeal region shows Gottron's papules (raised erythematous lesions with fine scale) with additional erythema extending along the tendons, most evident on the middle finger, consistent with linear extensor erythema. Classic Gottron's papules have formed over the proximal interphalangeal and, to a lesser extent, the distal interphalangeal joints. (By courtesy of Dr Lela Lee, Dept. of Dermatology, University of Oklahoma Health Sciences Center.)

Manifestations in other systems

Systemic signs

Fatigue and malaise are common and must be distinguished from muscle weakness. Weight loss may be impressive in some patients (Tymms and Webb 1985). Fevers are seen in about 40 per cent overall, but are associated strongly with antisynthetases (87 per cent, versus 23 per cent without antibody, Love *et al*. 1991).

Pulmonary disease

Pulmonary involvement resulting from muscle weakness, treatment, or the underlying disease (Table 2), occurs in 40 to 50 per cent of patients (Dickey and Myers 1984), contributing to morbidity and mortality (Arsura and Greenberg 1988).

Pulmonary	
Due to weakness:	
Ventilatory failure	
Aspiration pneumonia	
Due to treatment:	
Hypersensitivity pneumonitis	
Opportunistic infection	
Due to disease:	
Interstitial lung disease	
Pulmonary hypertension	
Pulmonary vasculitis	
Cardiac	
Heart block	
Arrhythmias	
Cardiomyopathy	
Gastrointestinal	
Oesophagus (dysphagia)	
Striated muscle dysfunction	
Oropharyngeal dysfunction	
Lower oesophageal dysfunction	
Stomach, intestines	
Decreased motility	
Arthritis	
Mild, responsive	
Deforming, minimally erosive	

Table 2 Features of polymyositis/dermatomyositis in addition to muscle and skin involvement

Respiratory muscle weakness

Clinical respiratory muscle weakness develops in 4 to 7 per cent of patients (Dickey and Myers 1984), but a measurable decrease in respiratory muscle strength may be more common (Braun *et al*. 1983). Both inspiratory and expiratory muscles may be affected, including the diaphragm. Total lung capacity and vital capacity are decreased, while residual volume may be increased. Respiratory failure requiring assisted ventilation may result, and may develop rapidly. It is usually responsive to or prevented by treatment. Most of these patients have had involvement of the pharyngeal and tongue muscles with dysphagia and impaired speech. Patients at risk should be monitored closely, commonly in hospital, using serial pulmonary function testing to judge respiratory muscle strength (expiratory pressure, peak flow, etc.). Rarely, weakness may involve respiratory muscles selectively or disproportionately, and be the presenting feature (Sano *et al*. 1994). Pharyngeal and tongue involvement also increase risk of aspiration, as does impaired cough and difficulty turning or sitting up in bed. Most patients with aspiration have had dysphagia (Lakhanpal *et al*. 1987).

Pulmonary complications of treatment

New interstitial infiltrates in patients with polymyositis/dermatomyositis on immunosuppressive agents may cause diagnostic confusion between the disease itself (see below) and effects of treatment, including opportunistic infections and drug reactions. Methotrexate hypersensitivity pneumonitis can occur in polymyositis/dermatomyositis, which can present acutely with fever, cough, dyspnoea, bilateral interstitial infiltrates, and with lymphocytic infiltrates on biopsy. It usually improves after withdrawal of the drug, but corticosteroids may be required. Pulmonary reactions may rarely occur with azathioprine.

Interstitial lung disease

Interstitial lung disease is found in 10 to 30 per cent of polymyositis and dermatomyositis overall (Hochberg *et al*. 1986; Love *et al*. 1991), and may be more easily diagnosed by pulmonary function testing than by chest radiography. It is much more frequent in patients with antisynthetases. Interstitial lung disease in polymyositis/dermatomyositis is similar to idiopathic interstitial lung disease (Targoff 1990). A small proportion have a fulminant course with fever and rapidly progressive dyspnoea that may be fatal within weeks. A second group has a more chronic course, and a third group may have asymptomatic test abnormalities. Chest radiography shows a reticulonodular pattern, often more prominent in the lower lobes (Fig. 10), and pulmonary function tests show a restrictive defect with decreased diffusing capacity and hypoxaemia with exercise, early signs that can be used to assess progress.

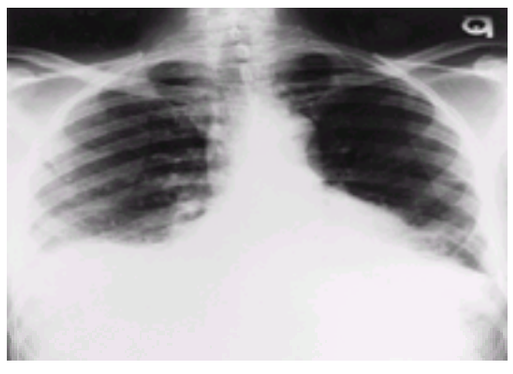


Fig. 10 Chest radiograph from a patient with anti-Jo-1-positive polymyositis that shows severe interstitial fibrosis affecting predominantly the lower lobes.

Interstitial lung disease may occur in polymyositis/dermatomyositis of any type (including cancer-associated). The severity of the interstitial lung disease is unrelated to that of the myositis, even occurring in amyopathic dermatomyositis (Euwer and Sontheimer 1994). It may precede myositis in up to 40 per cent of cases (Schwarz *et al*. 1976), and limitations from interstitial lung disease may mask muscle weakness in others. Elevated creatine kinase, myositis-specific autoantibodies, or the dermatomyositis rash may be clues to underlying polymyositis/dermatomyositis.

Lung histology shows interstitial mononuclear cell infiltrates (Fig. 11) with a variable amount of fibrosis. Often there is a loss of type I cells, proliferation of type II cells, and increased numbers of free alveolar macrophages. Tazelaar *et al*. (1990) identified four histological patterns among 15 patients: bronchiolitis obliterans with organizing pneumonia (6 patients), usual interstitial pneumonitis, diffuse alveolar damage, and cellular interstitial pneumonia. Those with bronchiolitis obliterans had the best prognosis, as in other studies (Hsue *et al*. 1993), while all three with diffuse alveolar damage died. Direct immunofluorescence has been negative for immunoglobulin or complement deposition in most studies.

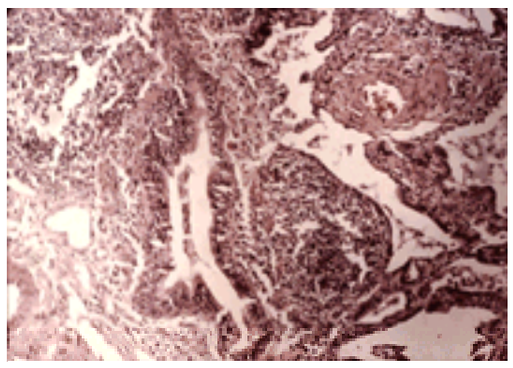


Fig. 11 Lung pathology from a patient with anti-Jo-1-positive polymyositis, showing severe mononuclear cell infiltration and thickening of the interstitium.

There may be medial and/or intimal thickening of small pulmonary arteries or arterioles as in pulmonary hypertension, which may occur in association with interstitial fibrosis (Schwarz *et al*. 1976). Pulmonary vasculitis may also occur (Lakhanpal *et al*. 1987). Hebert *et al*. (1990) found pulmonary hypertension in 7 of 11 patients using echocardiography, but the clinical significance is unclear.

There have been at least 21 case reports of spontaneous pneumomediastinum in dermatomyositis (adult and juvenile, not polymyositis), usually associated with interstitial lung disease, often with normal creatine kinase (Matsuda *et al*. 1993). The mechanism for this potentially fatal complication is unclear.

Cardiac disease

The potential for cardiac involvement in polymyositis/dermatomyositis was recognized early, but was considered unusual. More recent studies have found signs of cardiac involvement in excess of 70 per cent of cases (Taylor *et al*. 1993). It is commonly asymptomatic, but may contribute to mortality (Hochberg *et al*. 1986). The activity of the cardiac disease may be independent of the myositis (Rechavia *et al*. 1985). The major manifestations include conduction disturbances, arrhythmias, and myocarditis.

The frequency of conduction block varies, but only occasionally does advanced heart block occur requiring a pacemaker. Taylor *et al*. (1993) found ECG abnormalities in 81 per cent, but 58 per cent were non-specific ST and T-wave changes, and significant conduction block was uncommon. Fibrosis of the conducting system was correlated with conduction disturbance in some cases (Haupt and Hutchins 1982), and inflammation has also been seen.

The most common arrhythmias seen are extrasystoles and tachyarrhythmias, and are usually mild (Taylor *et al*. 1993). Love *et al*. (1991) found palpitations in 26 per cent, more in polymyositis (57 per cent) than dermatomyositis (19 per cent) or other subgroups. A Holter monitor should be considered in patients with palpitations or ECG abnormalities.

Congestive heart failure resulting from myocarditis is uncommon in polymyositis and dermatomyositis, and is found in about 3 per cent (Bohan *et al*. 1977). Myocarditis was seen in a quarter of patients in an autopsy study (Haupt and Hutchins 1982), where severe cardiac disease would be overrepresented. A diffuse interstitial and perivascular mononuclear cell infiltrate may be seen, similar to that in skeletal muscle, with replacement fibrosis and sometimes small vessel disease with medial smooth muscle hyperplasia (Denbow *et al*. 1979).

A reported increase in mild mitral valve prolapse, possibly from myocarditis, was not seen in other studies (Taylor *et al*. 1993). Pericardial effusions are seen in 5 to 25 per cent of polymyositis/dermatomyositis by echocardiogram, but are usually asymptomatic (Tami and Bhasin 1993). Significant pericarditis without systemic lupus erythematosus overlap is rare.

Gastrointestinal disease

Dysphagia may occur in up to 30 per cent (Tymms and Webb 1985), particularly with more severe disease. It predisposes to aspiration and has been associated with a poor prognosis. Dysphagia can result from weakness of the muscles of swallowing (pharyngeal muscles or striated muscles of the upper oesophagus), correlating with disease activity and responding to treatment (Dietz *et al*. 1980). It may be worse while recumbent, and can cause regurgitation of liquids into the nose with attempted swallowing. Changes in the voice (nasal speech or hoarseness) may be associated. Histology is similar to that of other striated muscle, dysphagia but may occasionally be the presenting or sole complaint.

Abnormalities of oesophageal motility have been reported to be common, involving both upper and lower oesophagus (DeMerieux *et al*. 1983). Involvement is similar to scleroderma, but can occur without other evidence of overlap. It is not associated with inflammation, and does not respond to treatment. Decreased motility and lower oesophageal sphincter pressure may lead to dysphagia, heartburn, reflux, and stricture. Gastric emptying may also be delayed (Horowitz *et al*. 1986).

Cricopharyngeal muscle dysfunction from inflammation and/or fibrosis may lead to a distinctive dysphagia marked by a sensation of food sticking in the back of the throat, or coughing with swallowing. It can be distinguished from weakness by cine-oesophagoscopy and oesophageal manometry, and is important to identify since it

may require surgical myotomy (Kagen *et al*. 1985).

Intestinal vasculitis with perforation, as well as pneumatosis cystoides intestinalis, well recognized in juvenile dermatomyositis, are extremely rare in adults, but reported (Ramirez *et al*. 1990).

Malignancy

Association

The link with malignancy was noted in 1916, and was at one point thought to be frequent. Later studies found a much lower frequency, and some questioned the reality of an association (Lakhanpal *et al*. 1986). Most recent studies, however, find a modest increase in malignancies within 1 to 2 years of onset in dermatomyositis, and possibly also polymyositis. Malignancy may be antecedent, concurrent, or subsequent to myositis onset.

The frequency of malignancy in dermatomyositis has ranged from 6 to 43 per cent (Bernard and Bonnetblanc 1993). A large population-based study in Sweden (Sigurgeirsson *et al*. 1992) found increased malignancy in dermatomyositis and polymyositis: 15 per cent of 392 dermatomyositis patients (relative risk, **RR**=2.4 for men, 3.4 for women) and 9 per cent of 396 polymyositis patients (RR=1.8 for men, 1.7 for women). Cancer deaths were increased in dermatomyositis but not polymyositis, supporting a true association in dermatomyositis, rather than intensive searching. The case–control study of Manchul *et al*. (1985) found that 71 adult polymyositis/dermatomyositis patients had significantly more total antecedent plus concurrent malignancies (21.1 per cent) than matched controls (5.6 per cent with inflammatory disease, 1.4 per cent others), with no difference beyond 6 months after myositis. A meta-analysis that included these studies found a significant overall association with cancer (odds ratio = 4.4 for dermatomyositis, 2.1 for polymyositis), with risk before and after onset for dermatomyositis (Zantos *et al*. 1994).

The activity of the myositis may appear linked to that of the malignancy ('paraneoplastic', 22 per cent of dermatomyositis cancers, Bonnetblanc *et al*. 1990), supporting the validity of an association. The myositis may resolve with treatment of the cancer, be resistant until it is resected, or flare with its recurrence. In the majority, cancer and myositis have an independent course (Callen 1993).

Tumours

A wide variety of tumours have been reported in polymyositis/dermatomyositis patients. Tumours that are frequent in the general population (lung, breast, etc.) are frequent in polymyositis/dermatomyositis. However, a number of studies have indicated an increase in ovarian cancer out of proportion to that of other tumours (Cox *et al*. 1990; Cherin *et al*. 1993b; and others). Sigurgeirsson *et al*. (1992) found a relative risk of 8.2 for ovarian cancer in women with dermatomyositis, and 16.7 during the 5 years after dermatomyositis diagnosis. Three of five patients of Whitmore *et al*. (1994) had amyopathic dermatomyositis; the skin disease was resistant until treatment of the cancer. Tumours were typically stage III or IV serous carcinoma. In Asia, an increase in nasopharyngeal carcinomas has been suggested (Koh *et al*. 1993), which may reflect an increase in such tumours in the population of the region.

Evaluation

The extent of testing that should be performed to uncover an occult malignancy in recent-onset polymyositis/dermatomyositis is controversial. Most of the associated malignancies show abnormalities detectable by a thorough initial evaluation, and all agree with the importance of this step (Callen 1994; Bernard and Bonnetblanc 1993). It should include careful history and physical examination, rectal examination and stool occult blood testing, breast examination and screening mammography and pelvic examination with Pap smear in women, complete blood counts, chemistries, urinalysis, and chest radiograph, and a prostate-specific antigen test in men. Any abnormalities should be pursued, and the evaluation should be repeated yearly, at least during the 2 to 3 year theoretical risk period.

In addition, most would give special attention to excluding ovarian cancer in women. Ovarian cancer is often missed in dermatomyositis patients, even sometimes by pelvic CT or ultrasound (Sigurgeirsson *et al*. 1992). Whitmore *et al*. (1994) recommend routine serial gynaecological examinations, transvaginal ultrasound, and CA-125 levels (even after oophorectomy).

There is disagreement as to the extent of further searching recommended in the absence of abnormalities. Schulman *et al*. (1991), citing examples of occult cancers, recommend routine upper and lower gastrointestinal barium studies and abdominal and chest CT, with repeat evaluation for myositis flares. Others limit additional testing to those with higher risk, resistant disease, or weight loss. The highest risk is in those aged over 45 with dermatomyositis rather than polymyositis, with no autoimmune rheumatic diseases overlap syndromes or myositis-specific autoantibodies (**MSAs**) (Love *et al*. 1991) (although cases with MSAs have occurred). Malignancy may be higher with cutaneous vasculitis, or with capillary damage even without a dermatomyositis rash (Casademont *et al*. 1993).

Other features, overlap syndromes, and associated conditions

Raynaud's phenomenon and arthralgia/arthritis, common features of autoimmune rheumatic diseases, may occur as part of polymyositis/dermatomyositis. They are more common in polymyositis/dermatomyositis when antisynthetases are present. Patients with polymyositis/dermatomyositis may have overlap syndromes in which diagnostic criteria for other conditions are also fulfilled. Conditions commonly overlapping with polymyositis/dermatomyositis include systemic lupus erythematosus (found in almost half of overlap patients by Love *et al*. 1991), scleroderma, Sjögren's syndrome (about 20 per cent), and rheumatoid arthritis (6 per cent). The distinction between an autoimmune rheumatic disease overlapping with polymyositis/dermatomyositis versus an autoimmune rheumatic disease with myositis as a manifestation, is not well defined. The relative severity of clinical features, and the serological picture may be helpful.

Renal disease is very rare in pure polymyositis/dermatomyositis without overlap, but focal mesangial proliferative glomerulonephritis has been seen (Frost *et al*. 1993), including cases with antisynthetases. Renal injury may occur from myoglobinuria.

A variety of other autoimmune conditions have been reported in association with polymyositis and dermatomyositis, including Graves' or Hashimoto's disease, inflammatory bowel disease, cryoglobulinaemia, primary biliary cirrhosis, dermatitis herpetiformis, coeliac disease, Behçet's disease, thrombotic thrombocytopenic purpura, myasthenia gravis, and others. Whether these conditions have a true association with polymyositis/dermatomyositis is unclear.

Pregnancy and polymyositis/dermatomyositis

Active polymyositis/dermatomyositis appears to confer increased risk to both the mother and the fetus, with potential for exacerbations, fetal loss and premature births. However, in established polymyositis/dermatomyositis, controlled at the time of pregnancy, most patients do not have flares, and many pregnancies are successful (Ishii *et al*. 1991; Oddis and Hill 1993). Those with myositis onset during pregnancy (and those with relapse) often have severe myositis requiring high-dose steroids, and have a high rate of adverse fetal outcome (Rosenzweig *et al*. 1989; Satoh *et al*. 1994). Thus, patients should optimally be in remission before becoming pregnant. If treatment is needed during pregnancy, prednisone is often used. Immunosuppressive medications should be avoided if possible, particularly methotrexate. Intensive evaluation for malignancy in pregnant patients who develop polymyositis/dermatomyositis without other indications would be inadvisable. No effect of polymyositis/dermatomyositis on fertility has been identified.

Laboratory investigations

Muscle factors

Enzyme

Creatine kinase

Serum levels of enzymes released from damaged muscle can be helpful for diagnosis and disease monitoring (Targoff 1988; Rider and Miller 1995). Creatine kinase is the most widely used due to its sensitivity, relative specifity for muscle, ready availability, and correlation with disease activity.

Elevations

Elevated creatine kinase levels are present in most patients (80 to 90 per cent) when first seen, and in more than 95 per cent at some time during their course (Bohan *et al*. 1977; Hochberg *et al*. 1986). The mean increase of creatine kinase (about 10-fold), and the potential to rise 100-fold or more, is greater than that of other enzymes measured. The creatine kinase usually rises with exacerbations, and can precede them by 5 to 6 weeks or more. A fall in the enzyme usually indicates improvement, and can precede recovery of strength by 3 to 4 weeks. There is a general correlation of creatine kinase level and disease activity for most individual patients over time (Kroll *et al*. 1986).

Normal creatine kinase

The creatine kinase may be normal in some patients despite active myositis. This is more common in dermatomyositis (Rider and Miller 1995) than polymyositis, possibly due to easier diagnosis of dermatomyositis when the creatine kinase is normal. Some dermatomyositis patients have normal creatine kinase because myositis has not yet developed. Lesser elevations may be seen in advanced or chronic disease, especially with severe atrophy, but elevations may still occur. Creatine kinase levels are lower in patients with autoimmune rheumatic diseases in general (Wei *et al*. 1981), possibly leading to lower levels in overlap patients. Creatine kinase has been correlated inversely with measures of inflammation in rheumatoid arthritis (Sanmarti *et al*. 1994). The enzymatic measurement of creatine kinase (generation of ATP from creatine phosphate and ADP) may be reduced falsely by an inhibitor found in some myositis sera (Kagen and Aram 1987). Steroids may lower the creatine kinase level even if they do not suppress disease activity. Levels that are within the normal range in patients with active disease may still be higher than the baseline for that individual.

One study of seven polymyositis/dermatomyositis patients with a normal creatine kinase found that the prognosis was worse than expected, with more interstitial lung disease and malignancy (Fudman and Schnitzer 1986); others have also found more interstitial lung disease (Koh *et al*. 1993; Matsuda *et al*. 1993), but many see no relation (Rider and Miller 1995).

Other causes of creatine kinase elevation

Creatine kinase may be elevated in a wide variety of conditions other than myositis (Table 3). It is released in conditions leading to muscle necrosis, but not usually simple atrophy (as in disuse, denervation, steroid myopathy, hyperthyroidism). Strenuous, prolonged exercise, particularly when unaccustomed, may raise the creatine kinase level in normal people for 2 days (Rider and Miller 1995). Physical injury to muscle can raise the creatine kinase, including intramuscular injections, electromyography, or muscle biopsy. Drugs can raise the creatine kinase level through a variety of mechanisms, including toxic effects, induction of myositis or myopathy, or decreasing excretion of the enzyme (see below). Various myopathies, as well as carrier states, may increase creatine kinase, including dystrophies, metabolic myopathies, and others. Some disease states other than myopathies may lead to increased creatine kinase, such as diabetic nephrotic syndrome associated with oedema (Taniyama *et al*. 1987). Others may raise total creatine kinase through effects on the creatine kinase-BB isoenzyme, such as gastrointestinal or lung tumours.

Table 3 Factors affecting creatine kinase (CK) levels

The normal range for creatine kinase may differ between patient groups. It is higher for men than for women, higher for black than for white people, and higher with increased muscle mass (Black *et al*. 1986). The composition of the group used to determine the normal range is therefore important in interpretation of the result.

Isoenzymes

Serum creatine kinase-MB isoenzyme can be elevated in polymyositis and dermatomyositis, but this has not correlated with cardiac involvement (Hochberg *et al*. 1986). High levels of creatine kinase-MB may occur without evidence of cardiac involvement, and severe cardiac involvement may occur despite a normal level of this isoenzyme (Targoff 1988). Regenerating or chronically stressed skeletal muscle fibres may contain significant creatine kinase-MB (Wolf 1991), and are the likely source in polymyositis/dermatomyositis.

Macro-creatine kinase type 1, a complex of antibody with creatine kinase, was found in 36 of 8322 patients tested for isoenzymes, and half had myositis (Lee *et al*. 1994). Modification of creatine kinase-MM after release leads to formation of subisoenzymes. A high MM3:MM1 subisoenzyme ratio suggests deteriorating disease (Annesley *et al*. 1985).

Other enzymes

Several other enzymes are released during muscle damage that may occasionally be helpful for disease monitoring. More than 98 per cent of patients will have an elevation of at least one serum enzyme at some time during their course (Hochberg *et al*. 1986). Aldolase is elevated usually in polymyositis and dermatomyositis, in some cases when creatine kinase is not, but is not as specific for muscle and does not correlate as well with disease activity (Bohlmeyer *et al*. 1994). Lactate dehydrogenase (LDH) is also elevated, predominantly LDH-5, although LDH-1 may increase without necessarily correlating with cardiac involvement (Targoff 1988). Aspartate transferase correlates very well with biopsy-proven muscle inflammation, and was found useful in combination with creatine kinase and aldolase in assessing enzyme elevations (Hood *et al*. 1991). Carbonic anhydrase III, an isoenzyme found exclusively in skeletal and not cardiac muscle, rises with skeletal muscle damage, including polymyositis/dermatomyositis (Osterman *et al*. 1985).

Myoglobin

Myoglobin is unique to skeletal and cardiac muscle. Very little is present normally in serum or urine, but it is detectable in the serum of most patients with active polymyositis/dermatomyositis (Targoff 1988). It can serve as a disease activity marker, rising with exacerbation and falling with remission, and can predict exacerbation in some cases. It may occasionally be elevated when the creatine kinase is not. It has the advantages of tissue specificity, rapid clearance, and sensitive, non-enzymatic detection. A new immunoturbidimetric assay may make testing more available, but is slightly less sensitive than radioimmunoassay (Lovece and Kagen 1993). It should be measured in samples taken at a standard time during the day because of diurnal variation.

Other muscle factors

Creatine is produced in the liver, pancreas, and kidney, and taken up by the muscle. Creatine excretion rises with muscle disease due to defects in uptake and retention. The creatine or the creatine/ [creatinine+creatine] ratio in a 24 h urine sample (normal less than 6 per cent) is elevated in most polymyositis/dermatomyositis patients, and can vary with disease activity. Its clinical utility is limited by many disadvantages. Elevation, found whenever muscle mass is reduced, is less specific than creatine kinase, and may persist in inactive polymyositis/dermatomyositis as a result of atrophy. It is cumbersome to collect and

difficult to obtain clinically.

Measures of muscle mass correlate with strength, but not necessarily inflammation; they can be used to provide a longer-term view. Tests that measure muscle mass, including 24 h urinary creatinine, 3-methyl histidine excretion (reflecting skeletal muscle protein turnover), and total body potassium (measuring total body mass), have been used in studies but are not currently practical for clinical use (Moxley 1994). Dual radiography/absorptiometry may be used in the future (Kanda *et al*. 1994).

Autoantibodies

Indirect immunofluorescence is positive in 50 to 80 per cent of patients with polymyositis or dermatomyositis (Reichlin and Arnett 1984; Love *et al*. 1991). Nuclear or speckled patterns are most common, nucleolar may be seen, and cytoplasmic patterns provide clues to an antisynthetase or anti-SRP (Fig. 12). A high antinuclear antibody (**ANA**) titre favours polymyositis or dermatomyositis over other myopathies or neuropathies, and a MSA strongly supports polymyositis/dermatomyositis when present (see below). Antisynthetases may alert the physician to an increased risk for interstitial lung disease, or identify underlying polymyositis/dermatomyositis in patients who present with prominent extra-muscle features.

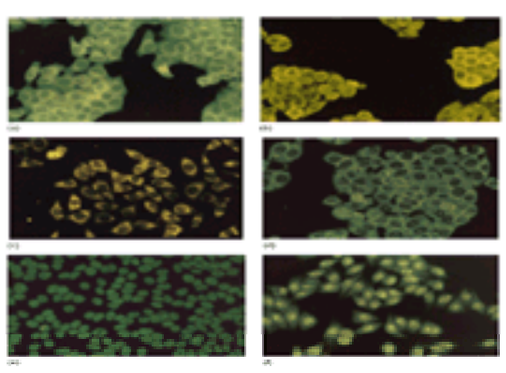


Fig. 12 Indirect immunofluorescence on HEp-2 cells using sera with myositis-associated autoantibodies. (a) Anti-Jo-1 autoantibodies (reacting with histidyl-tRNA synthetase): finely speckled cytoplasmic pattern of fluorescence. (b) Anti-PL-7 autoantibodies (antithreonyl-tRNA synthetase): cytoplasmic pattern (more homogeneous at higher concentration). (c) Anti-SRP autoantibodies (antisignal recognition particle): cytoplasmic pattern. (d) Anti-KJ autoantibodies (reacting with a translation factor): cytoplasmic pattern with slight nucleolar staining. (e) Anti-Mi-2 autoantibodies (reacting with an unidentified nuclear protein): nuclear pattern, sparing nucleoli, without cytoplasmic staining. (f) Anti-PM-Scl autoantibodies (reacting with a complex of 11 proteins): intense nucleolar staining with significant nuclear staining.

A general correlation of anti-Jo-1 titre with disease activity has been observed (Miller *et al*. 1990b), and in some cases a rise in anti-Jo-1 predicted exacerbation. Occasional disappearance of anti-Jo-1 correlates with disease remission. The usefulness of titres as an index of disease activity is not established, but they may provide support if serial measurements are available.

Other tests

The erythrocyte sedimentation rate is elevated in about half of active cases, but is correlated poorly with disease activity or response to treatment. Similarly, the C-reactive protein may be normal or slightly high despite active disease (Gabay *et al*. 1994). Rheumatoid factor is positive in about 10 to 20 per cent of patients, most commonly in the overlap group. Circulating immune complexes and cryoglobulins have been reported in some patients, but their significance is unclear. Elevated gammaglobulins may be seen; hypogammaglobulinaemia should raise the suspicion of echovirus or other infection. Patients with polymyositis with monoclonal gammopathies have been described, including some in which sarcolemmal deposition of paraprotein was found (Kiprov and Miller 1984). Complement is usually normal, but myositis has occurred in C2 deficiency (Targoff 1988). Proteinuria is usually the result of myoglobinuria, with rare exceptions (see below).

Diagnosis

The criteria of Bohan and Peter (1975) (Table 4), often with modifications such as that of Dalakas (1991), have been used widely for diagnosis and clinical studies. Muscle enzymes, electromyography and muscle biopsy remain essential in evaluation of patients and establishing the diagnosis of polymyositis and dermatomyositis. Autoantibody testing and magnetic resonance imaging (MRI) can also greatly aid in diagnosis. Recent attempts have been made to devise new criteria that take advantage of newer tests and other features (Medsger and Oddis 1995; Tanimoto *et al*. 1995). Even when the criteria are satisfied, other causes of muscle disease must be excluded.

Table 4 The diagnosis of polymyositis and dermatomyositis

Electromyography

Electromyography and nerve conduction studies cannot establish the diagnosis of polymyositis/dermatomyositis with certainty, but can demonstrate that the process is myopathic and consistent with these diseases, and can help to exclude many other neuropathies and certain myopathies. Electromyography may reveal muscle involvement in patients presenting with rash or extra-muscle features, and can identify areas of involvement to help direct biopsies (on the contralateral side) and provides some information regarding activity. Ninety per cent of patients with active polymyositis/dermatomyositis have an abnormal electomyograph (Bohan *et al*. 1977; Henriksson and Lindvall 1990). Testing of multiple muscles is important, since involvement may be limited, and to demonstrate the distribution of involvement. Paraspinal muscle involvement is common, and may be the only abnormal area.

Motor unit action potentials in polymyositis/dermatomyositis typically are myopathic (low amplitude, short duration). Polyphasic potentials (complex potentials with increased turns) are increased, and may be attributed to asynchronous firing of fibres (Bertorini 1988a). Over time, long-duration, high amplitude polyphasic potentials may be seen (attributed to reinnervation of regenerating or denervated fibres) (Uncini *et al*. 1990). Patients with myositis have early recruitment and full interference patterns (more fibres required to achieve a given force), in contrast to the decreased recruitment and interference seen in neuropathies (Greenlee 1988).

Spontaneous activity at rest is seen in three-quarters of patients (Henriksson and Lindvall 1990). Increased insertional activity generated by needle trauma to the muscle fibre is very common, as are fibrillations and positive sharp waves, sometimes more evident in the paraspinal muscles. Often associated with denervation,

fibrillations in polymyositis/dermatomyositis have been attributed to damage to intramuscular nerves, nerve endings or motor end plates, or to segmental muscle fibre necrosis that denervates the distal fibre (Greenlee 1988). Complex repetitive discharges (bizarre high-frequency discharges) may also be seen in a third to a half of patients. In contrast to myotonic discharges, these start and stop abruptly, and usually have constant amplitude. They have been attributed to inflammatory damage to the sarcolemma. Spontaneous activity has been associated with active inflammation; it is less common with chronic disease, and may subside with treatment.

Single-fibre electromyography in myositis shows increased jitter and blocking, although less prominent than in myasthenia gravis. Computerized quantitative analysis may reveal increased fibre density (more potentials per motor unit) (Bertorini 1988a). A myopathic interference pattern was the most common abnormality (83 per cent) on quantitative electromyography (Barkhaus *et al*. 1990).

Imaging

Magnetic resonance imaging (MRI)

Numerous studies in recent years have demonstrated the value of MRI in polymyositis/dermatomyositis for diagnosis and assessment of disease activity. MRI can identify sensitively areas of muscle inflammation, atrophy, or fatty replacement (Reimers *et al*. 1994). It is non-invasive and can be repeated sequentially. Use is often limited by cost or availability, but in certain situations it can provide critical information that may not be available by other methods, affecting treatment decisions (Park *et al*. 1994).

T_2-weighted images are best for showing areas of active muscle inflammation, where increased water content is seen as increased intensity; this is not seen with T_1-weighted images (Park *et al*. 1990). Increased fat is seen on both images. Some studies have used fat suppression techniques to improve image contrast (Fraser *et al*. 1991; Hernandez *et al*. 1992). Gadolinium was not helpful (Reimers *et al*. 1994). The thighs are most often studied.

Involvement seen using MRI is often focal and patchy, with differences in intensity between and within muscles, and the technique can therefore be more sensitive than biopsy for detecting clinical activity (Fraser *et al*. 1991). Park *et al*. (1990) found that the vastus lateralis was the most involved muscle in their dermatomyositis patients, with a predominance of anterior over posterior thigh muscle involvement. Fraser *et al*. (1991) did not find anterior predominance in dermatomyositis or polymyositis patients, but found a correlation of atrophy and disease duration. Fatty infiltration is more common in polymyositis than dermatomyositis, and in chronic disease.

Fraser *et al*. (1991) also noted a correlation between MR images or quantitative signal intensity scores with disease activity, although Reimers *et al*. (1994) did not. By correlating with clinical activity, MRI can help assess therapeutic response; abnormalities can return to normal within months on treatment (Fujino *et al*. 1991; Park *et al*. 1994). Persistent abnormalities can indicate activity. MRI may show high intensity in active disease, even when enzymes, electromyography, and/or biopsy are normal (Park *et al*. 1994; Stonecipher *et al*. 1994), although exceptions with negative MRI in active disease may occur (Stiglbauer *et al*. 1993).

Magnetic resonance spectroscopy (MRS)

P-31 MRS has been used in the study of various myopathies to assess energy utilization and reserve by measuring phosphate metabolites in muscle. The inorganic phosphate/phosphocreatine ratio rises in myopathies indicating a decrease in energy reserve, and is correlated with disease activity (Park *et al*. 1994). Phosphocreatine and ATP are decreased at rest, decrease further with exercise, and show delayed recovery to baseline (Park *et al*. 1990). Phosphocreatine changed more than ATP, and was thus better for assessing disease activity. However, the decline of ATP with exercise was felt to relate to the severe fatigue that may occur. The impaired energy utilization and muscle metabolism may contribute to muscle weakness in polymyositis/dermatomyositis.

While not specific, MRS may help monitor disease activity and response to therapy. It can be abnormal in some patients with normal creatine kinase (Park *et al*. 1994), but can continue to be abnormal in some patients whose inflammation has resolved (reflecting persistent muscle abnormalities). When MRI and MRS were discordant, MRS was felt to be more useful for assessing disease status. Park *et al*. (1995) also found that patients with amyopathic dermatomyositis, who have normal MRI and MRS at rest, often have subtle MRS abnormalities with exercise (reduced total oxidative capacity, V_{max}), which may explain their fatigue.

Other methods

Ultrasound of muscle was abnormal, most often with hyperechogenicity as a result of fat and atrophy, in 83 per cent of 70 patients with myositis, including some with normal creatine kinase or electromyographs (Reimers *et al*. 1993). It is less expensive and more readily available than MRI, but is less sensitive and specific (Stonecipher *et al*. 1994), and does not provide as much information. It may help direct the biopsy.

Technetium-99m and thallium uptake may be increased in muscles affected by active polymyositis. Although non-specific, this also might be useful for directing a biopsy. Gallium-67 scanning has been used to identify myositis in Lyme disease and other infections. Scanning with indium-labelled antimyosin can reveal areas of muscle necrosis in dermatomyositis (DeGeeter *et al*. 1989). With a monoclonal antibody specific for cardiac myosin, the scan can detect myocarditis in polymyositis (Le Guludec *et al*. 1993).

Biopsy

Indications

Most patients in whom polymyositis or dermatomyositis is suspected should have a muscle biopsy. It can provide the most convincing evidence supporting the diagnosis, and can definitively exclude certain relevant conditions. There are cases in which the diagnosis can be established for clinical purposes without the biopsy. With proximal weakness, elevated enzymes, and a typical electromyograph, those with a classic dermatomyositis rash, confirmed myositis-specific autoantibodies, or autoimmune rheumatic diseases overlap syndromes with specific antibodies (anti-U1RNP, anti-PM-Scl) do not usually require a biopsy, although others feel that all adults should have a biopsy (Urbano-Marquez *et al*. 1991).

Methods

Open muscle biopsy gives the best picture of the muscle architecture. A large specimen may decrease sampling error, allows proper orientation, and provides enough muscle for all studies. Complications from the procedure (bleeding, infection, nerve damage, etc.) are very low (Pamphlett 1988), but the 4 to 8 cm incision commonly used creates a significant scar, especially in obese patients. Needle biopsy, with a 0.5 cm incision, causes substantially less morbidity, is often more easily and rapidly arranged, and is adequate for diagnosis of polymyositis/dermatomyositis in most cases (O'Rourke *et al*. 1994). Multiple specimens may be taken through the same incision site for enzyme histochemistry and electron microscopy (**EM**), and to reduce sampling error (Haddad *et al*. 1994). The low morbidity allows repeat biopsy or later open biopsy if needed. However, processing of samples is more difficult, more artefact is encountered (particularly in EM), and open biopsy is required for certain functional enzyme studies.

The best information is obtained from muscle with active disease but not endstage fibrosis or atrophy. Muscles used for electromyography or intramuscular injection should not be used for biopsy. The quadriceps or deltoid are most often used due to accessibility, but focal disease may result in non-diagnostic samples. As noted, certain diagnostic procedures have been used to localize disease activity to direct the biopsy (electromyography, MRI, ultrasound). MRI is most promising since it can show focal involvement, but is not always successful (Pitt *et al*. 1993; Reimers *et al*. 1994).

EM should be performed when inclusion-body myositis or mitochondrial myopathies are considerations, or to exclude certain other conditions. Enzyme histochemistry may be helpful to exclude myophosphorylase deficiency and other metabolic myopathies which may masquerade as polymyositis. Co-ordination with the pathologist can help to optimize processing and evaluation of specimens.

Muscle pathology

Inflammation is a hallmark of myositis. The infiltrates are predominantly lymphocytes, but include macrophages, plasma cells, and sometimes eosinophils, basophils,

and neutrophils. The amount of inflammation is variable, and up to 25 per cent may show no inflammation, usually attributed to a focal process.

In polymyositis, inflammatory infiltrates more often predominate in the endomysial area around the muscle fibres, usually without perifascicular atrophy (Fig. 13). Necrosis of individual muscle fibres may be observed. The fibres appear swollen with homogeneous contents, losing the normal striations of the contractile proteins. There is invasion of mononuclear cells, phagocytosis, and regeneration (Fig. 14), the latter marked by sarcoplasmic basophilia, large vesicular, internalized nuclei, and prominent nucleoli. In later stages, there is atrophy, fibrosis, and fatty replacement. Steroid treatment may enhance type II fibre atrophy. Non-caseating granulomas were recently observed in muscle from anti-Jo-1-associated polymyositis (Moder *et al*. 1993).

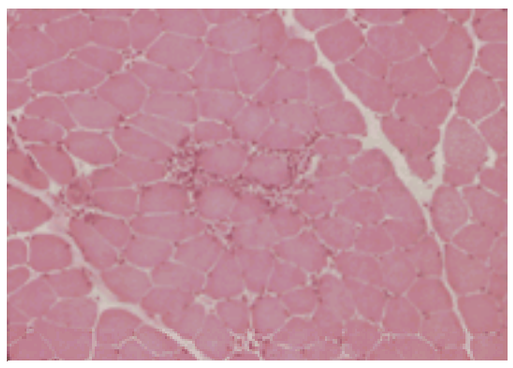


Fig. 13 Muscle biopsy (haematoxylin and eosin stain) from a patient with polymyositis. Endomysial inflammation (infiltration with mononuclear cells between fibres within the fascicle) is seen; a pattern characteristic of biopsies from patients with polymyositis.

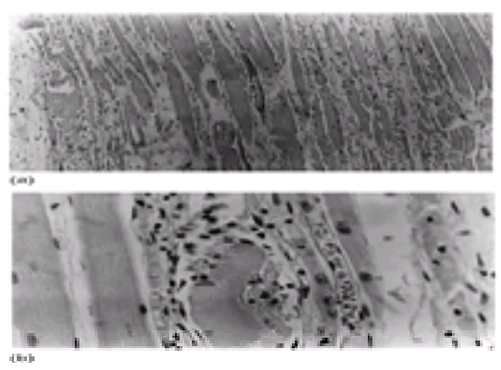


Fig. 14 Muscle biopsy from a patient with anti-Jo-1-positive polymyositis. (a) Inflammation with necrosis and degeneration of muscle fibres. (b) Necrosis with loss of characteristic striations and integrity of fibre. (By courtesy of M. Reichlin, MD.)

In typical dermatomyositis, infiltration predominates in the perimysial area (around the fascicles) and around small blood vessels, sometimes extending into the endomysial area (Fig. 15). Microvascular changes are often seen, with perifascicular atrophy (decreased fibre size at the periphery of the fascicle), a characteristic (although not specific) feature of dermatomyositis (Fig. 15 and Fig. 16). This may result from capillary loss, which is greater in the perifascicular region, or to a direct effect of perifascicular inflammation (Kalovidouris 1994). It is most common in juvenile dermatomyositis (90 per cent), and also occurs in adult dermatomyositis (50 per cent) (Bertorini 1988b). Circumscribed areas of myofibrillar loss are also typical, attributed to ischaemic damage. There is some overlap of the polymyositis and dermatomyositis patterns (Ringel *et al*. 1986). Despite prominent vascular damage, frank necrotizing vasculitis is unusual (Fig. 17).

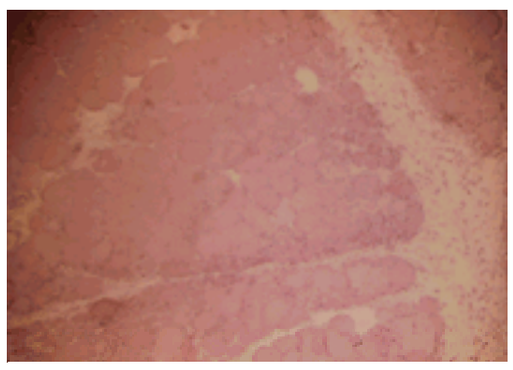


Fig. 15 Muscle biopsy (haematoxylin and eosin stain) from a patient with anti-Jo-1-positive myositis. The histological appearance is characteristic of dermatomyositis, but the patient had no cutaneous manifestations. The most intense infiltration is perimysial, with some extension into the endomysial area in a perifascicular distribution. Perifascicular atrophy (atrophy of the fibres at the periphery of the fascicle) is marked. Involvement of the vessel in the interstitial area is seen.

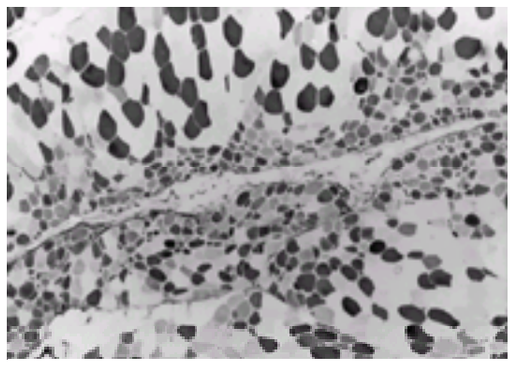


Fig. 16 Muscle biopsy (ATPase stain) from a patient with dermatomyositis. The pattern of perifascicular atrophy is evident.

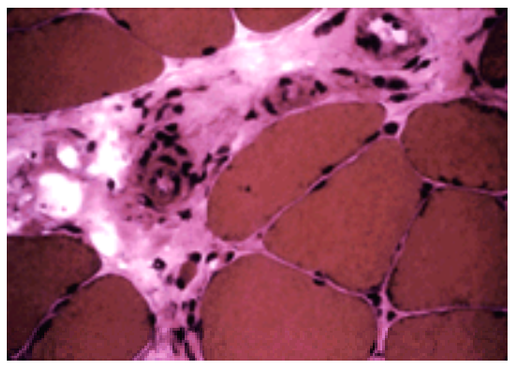


Fig. 17 Muscle biopsy from a patient with severe myositis associated with Sjögren's syndrome, showing inflammation surrounding small vessels.

EM shows endothelial cell injury, with swelling, hyperplasia, vacuolization, degeneration, and regeneration. There is endothelial cell necrosis and capillary thrombosis, and loss of capillaries resulting in decreased capillary density (Emslie-Smith and Engel 1990). The endothelial cells contain characteristic tubuloreticular inclusions ('undulating tubules'), that resemble viral structures but are believed to result from endothelial cell damage (Fidzianska and Goebel 1989). They may be seen in endothelial cells in other tissues (skin, lungs, joints, and lymphocytes). They can be induced by interferon-a, and may occur in other diseases, although usually not polymyositis.

Other EM findings are non-specific (Carpenter 1988). In dermatomyositis, Z disc streaming may be found. In polymyositis, aggregates of dense, membrane-bound material may be seen in some cases, and reduplication of the basal lamina.

Skin pathology

Skin biopsy can support the diagnosis of dermatomyositis, but generally cannot establish it. The biopsy should be taken from lesional skin (rather than as part of the muscle biopsy), usually from the chest or extremities. Liquefaction degeneration of the basal cell layer with prominent vacuolar changes is seen (Janis and Winkelman 1968). There is often a mild mononuclear cell infiltrate in the upper dermis and dermal–epidermal junction. There may be basement membrane thickening. Oedema and increased mucin can be seen in the dermis. The mucin (mostly hyaluronic acid) may be severe enough to be clinically evident (Igarashi *et al*. 1985). With poikiloderma, there is telangiectasia and epidermal atrophy. In Gottron's papules, typical dermatomyositis changes may be seen, with less atrophy, but with the addition of epidermal changes, including acanthosis and mild papillomatosis (Hanno and Callen 1985).

These findings may resemble systemic lupus erythematosus, but the dermal–epidermal infiltrate is milder in dermatomyositis, and there is more dermal mucin. Immunoglobulin deposition is much less prominent in dermatomyositis. Occasionally, mild deposition is seen at the dermal–epidermal junction in dermatomyositis lesional (but not non-lesional) skin (Kasper *et al*. 1988).

Differential diagnosis

The diagnosis of dermatomyositis is easier than that of polymyositis when the rash is florid. In polymyositis and less typical dermatomyositis, other conditions that can cause muscle weakness, myalgias, or elevated creatine kinase levels must be excluded. In all patients, tests for thyroid disease should be performed, and the role of medications should be considered. Most patients, particularly those with risk factors, should be tested for retroviral infection, and those at risk for other specific infections should be tested, such as specific exclusion of parasitic infection in those with eosinophilia or who are from endemic areas.

In patients with weakness without systemic features, other myopathies (inclusion-body myositis, metabolic or mitochondrial myopathies, drug-induced myopathies) or neuropathies (myasthenia gravis, amyotrophic lateral sclerosis) should be considered. In patients with weakness and signs of acute illness or autoimmune rheumatic disease, such as fever, arthritis, etc., the major conditions to be distinguished may be infectious or rheumatic. In a small number of cases, the patient will present with extra-muscle features only, either cutaneous lesions or autoimmune rheumatic features, resembling other autoimmune rheumatic diseases, or undifferentiated autoimmune rheumatic disease syndromes. Conditions that may be confused with polymyositis/dermatomyositis are listed in Table 5; special considerations are discussed below.

Table 5 Differential diagnosis of polymyositis/dermatomyositis

Human immunodeficiency virus (HIV)

Patients with HIV infection may develop myositis indistinguishable from polymyositis. Weakness, myalgia and elevated creatine kinase are seen, but rarely a rash. Electromyography shows myopathy, and biopsy shows inflammation with lymphocytic infiltration and necrosis (Dalakas 1993).

The pathogenesis of HIV-myositis appears similar to idiopathic polymyositis. CD8+ cytotoxic T cells and macrophages predominate, with very few CD4+ cells (Illa *et al*. 1991). As in polymyositis, there is endomysial infiltration, surrounded and invaded fibres, and expression of MHC-1 on most muscle fibres, all consistent with cell-mediated attack (Dalakas 1993). Numerous studies had failed to find HIV inside intact muscle fibres (Leon-Monzon *et al*. 1993), and it did not infect muscle cells *in vitro*, but HIV was found in cells infiltrating muscle or in degenerating fibres. However, a recent study using the more sensitive *in situ* polymerase chain reaction (**PCR**) found HIV in myocyte nuclei in four of seven cases (Seidman *et al*. 1994), plus HIV RNA indicating transcriptional activity. It remains to be shown whether myositis results from primary infection, secondary infection with a myositis-inducing virus, induction of autoimmune responses, or dysregulation.

HIV-myositis must be distinguished from the myopathy that can occur with zidovudine (Dalakas *et al*. 1990), which can cause weakness, myalgias, wasting, elevated creatine kinase, and myopathic electromyographs (Dalakas 1993). It is more frequent with longer therapy and higher doses. Zidovudine inhibits mitochondrial DNA synthesis, resulting in muscle mitochondrial toxicity *in vitro* and in experimental animals. Muscle shows ragged red fibres and abnormal mitochondrial structure and function (Mhiri *et al*. 1991). Inflammation is sometimes seen, but may relate to coexistent HIV-myositis; this combination may be required for symptomatic myopathy (Dalakas *et al*. 1990; Lane *et al*. 1993).

HIV has also been associated with a nemaline rod myopathy (Simpson and Bender 1988) and other non-inflammatory myopathy and muscle wasting. Patients with AIDS are also at risk for myositis from other infections, such as tuberculosis or microsporidia (Preston 1993), and for pyomyositis. Symptoms related to a myopathy may occur in up to 30 per cent of patients with AIDS, although a recognized inflammatory myopathy requiring treatment is much less common. Since polymyositis may

be the first manifestation of disease, testing for HIV infection is recommended in most patients with myositis.

Human T-cell lymphotropic virus type I (HTLV-1)

HTLV-1 has also been associated with myositis, with or without HTLV-1-associated myelopathy, and may be a major cause of polymyositis in endemic areas. In Jamaica, Morgan *et al*. (1989) found HTLV-1 antibodies in 11 of 13 polymyositis patients, but only in 7 of 93 others. The polymyositis was clinically similar to idiopathic polymyositis, with weakness, elevated creatine kinase, and compatible histology, and none of the patients had other HTLV-1-associated conditions. HTLV-1 was also increased in polymyositis patients in an endemic area in Japan, although not as dramatically (Higuchi *et al*. 1992). HTLV-1 is infrequent in polymyositis in non-endemic areas (Nishikai and Sato 1991; Nelson *et al*. 1994), but it may occur. HTLV-1 testing should be performed in patients at risk.

As with HIV, studies of infiltrating lymphocytes and MHC-1 expression in HTLV-1-myositis suggest T-cell mediated cytotoxicity (Leon-Monzon *et al*. 1994). One study found evidence of HTLV-1-associated proteins in muscle fibres (Wiley *et al*. 1989), but this was not confirmed (Higuchi *et al*. 1992; Leon-Monzon *et al*. 1994), and HTLV-1 did not infect muscle cells *in vitro*. The reason myositis develops in HTLV-1 infection is not known.

Bacterial pyomyositis

Pyomyositis has been well known in tropical areas for many years (Chiedozi 1979), most commonly affecting young males, with one or more spontaneous muscle abscesses, usually from *Staphylococcus aureus*. Pyomyositis in non-tropical areas is now increasingly reported, affecting older patients, often non-staphylococcal, with AIDS, diabetes, and chronic illness as predisposing factors (Rodgers *et al*. 1993; Gomez-Reino *et al*. 1994). Other muscle disease may be a factor in some cases. Although usually localized, a syndrome that could be confused with polymyositis has been seen (Wolf *et al*. 1990). Diagnosis can be made with ultrasound, CT, or MRI.

Other neurological conditions

Sensory abnormalities are against polymyositis/dermatomyositis, but diseases with exclusively diffuse motor involvement, such as amyotrophic lateral sclerosis or myasthenia gravis, can resemble polymyositis. Fasciculations and prominent atrophy may be clues in the former, and extraocular or eyelid muscle involvement in the latter.

Many myopathies are inherited and a family history of myopathy is an important point against polymyositis/dermatomyositis. Adult onset narrows the myopathies to consider, although a few associated with early onset occasionally present later (Table 5). In favour of dystrophy would be lack of inflammation on biopsy; myotonia; involvement of selective muscles (such as pectorals, biceps, triceps, etc.); and lack of response to therapy.

Mitochondrial myopathy may sometimes present in adults with pure limb myopathy without extraocular muscle involvement or encephalopathy. 'Ragged red fibres' on biopsy, due to subsarcolemmal accumulation of abnormal mitochondria, should prompt further testing with EM. Varga *et al*. (1993) described two cases of asymptomatic primary biliary cirrhosis associated with a severe, new onset mitochondrial myopathy with fatal outcome resembling polymyositis, raising the possibility of antimitochondrial antibody-induced mitochondrial myopathy.

Genetic defects in glycogen or glucose metabolism usually present with episodic fatigue, cramping, or pain related to exercise (Martin *et al*. 1994). Some, however, such as deficiencies of myophosphorylase (McArdle's disease), phosphofructokinase, myoadenylate deaminase, or acid maltase, may have an atypical presentation with late-onset progressive weakness that can be misdiagnosed as polymyositis (Higgs *et al*. 1989; Plotz 1992). The forearm ischaemic exercise test (Martin *et al*. 1994) is often used to screen for such disorders; several cause impaired rise in lactate but not ammonia, while myoadenylate deaminase causes impaired rise in ammonia but not lactate. There may also be clues on muscle biopsy (such as increased glycogen deposition). When metabolic myopathies are suspected, enzyme histochemistry and other tests for specific enzyme defects should be performed.

Drug-induced myopathies

Numerous drugs can induce myopathy (Table 6). Most have a toxic effect rather than inducing autoimmunity, and some have associated neuropathy, but the picture of weakness, myalgia, and elevated creatine kinase, often with electromyographic changes characteristic of myopathy, can look like polymyositis/dermatomyositis (Le Quintrec and Le Quintrec 1991; Zuckner 1994). D-Penicillamine is discussed later.

Table 6 Drugs causing myopathy

Colchicine may cause a non-inflammatory, vacuolar neuromyopathy (Kuncl *et al*. 1987), most often occurring with maintenance therapy at usual doses with renal insufficiency. A rise in creatine kinase is a sensitive indicator. A vacuolar myopathy may also be caused by chloroquine or less severely hydroxychloroquine, with characteristic myeloid and curvilinear bodies (Estes *et al*. 1987; Plotz 1992). Cardiomyopathy and neuropathy with loss of reflexes may be seen. Case reports of vacuolar myopathy with other drugs have appeared.

Muscle toxicity is a common side-effect with various cholesterol-lowering agents, possibly related to effects on sarcolemma (Dalakas 1992a). It can occur with HMG-CoA-reductase inhibitors (especially lovastatin) or fibrates or nicotinic acid alone, but is more severe with higher creatine kinases when combined (Pierce *et al*. 1990). Renal insufficiency may predispose to fibrate myopathy. Lovastatin myopathy is also more frequent in combination with cyclosporine, as after heart transplants.

Alcohol abuse may lead to either an acute necrotizing myopathy with prominent pain and high creatine kinase, or a chronic myopathy with proximal-muscle weakness, atrophy, and milder creatine kinase elevation (Charness *et al*. 1989). Several drugs of abuse, including cocaine, may lead to myopathy, creatine kinase elevation, or rhabdomyolysis, either due to pressure injury during unconsciousness, or without trauma (Rubin and Neugarten 1989). Abuse of ipecac in bulimia can induce a proximal, polymyositis-like myopathy, and as with other drug abuse, may be hidden (Plotz 1992).

In the eosinophilia–myalgia syndrome associated with L-tryptophan ingestion (Kaufman 1994), proximal myopathy was seen in two-thirds of cases. Aldolase was sometimes elevated, but not creatine kinase. Mononuclear and eosinophilic interstitial infiltrate (perimyositis) was seen, but fibre necrosis was rare. Similar findings occurred in the Spanish toxic oil syndrome.

Endocrine and steroid myopathies

Hypothyroidism is a common cause for elevated creatine kinase level with or without weakness, and can resemble polymyositis (Plotz 1992). Thyroid function should be tested in all patients when myositis is considered. Hypokalaemia from any cause can lead to myopathy, with creatine kinase elevation; severe hypokalaemic

myopathy has been associated with chronic licorice ingestion (Sintani *et al*. 1992). Other endocrine or metabolic problems can also cause myopathy (Table 5).

Use of corticosteroids for treatment of any condition may lead to proximal muscle weakness, associated with accentuated type II fibre atrophy (Khaleeli *et al*. 1983). Its onset is usually insidious, with lower extremity predominance. It tends to occur with higher doses for extended periods, and is more likely with multiple daily doses or longer acting, fluorinated preparations. In 10.6 per cent of patients taking dexamethasone steroid myopathy developed, with peak onset between the 9th and 12th weeks (Dropcho and Soong 1991). It usually improves if the dose is lowered. An acute form can occur associated with high-dose intravenous therapy and neuromuscular blocking agents in patients on respirators (Zuckner 1994).

In polymyositis/dermatomyositis, steroid myopathy can be confused with recurrent myositis. Elevated creatine kinase, increased spontaneous activity on electromyography, and inflammation detected by MRI or biopsy should all be absent in pure steroid myopathy. It may develop while the myositis is still active, indicating the likely need for a steroid-sparing drug. Often, if the situation is not life-threatening and remains unclear, dosage reduction is tried.

Aetiology and pathogenesis

Aetiology

Abundant evidence supports an autoimmune pathogenesis for polymyositis and dermatomyositis, but the reason they develop is unknown. It is generally felt that an inciting factor, exogenous or endogenous, acts in a genetically susceptible host, and that different inciting agents can lead to similar pictures. For example, D-penicillamine, HIV infection, various malignancies, and autoimmune rheumatic diseases, have been associated with autoimmune inflammatory myopathies that may be clinically indistinguishable from each other or from cases without recognized associations. Despite the paucity of familial cases, genetic predisposition is undoubtedly important, evident in the HLA associations, and suggested by animal models.

Potential aetiological factors

Infectious

A variety of infections can induce a syndrome resembling polymyositis/dermatomyositis in humans or animals, and unrecognized infections could be responsible for at least some idiopathic cases. Studies noting temporal variation in onset overall or for MSA-defined subgroups support this (Plotz *et al*. 1995). Mechanisms that might explain how infections such as viruses could induce autoimmune inflammatory myopathy include: persistent infection, molecular mimicry with cross-reaction of infectious agent and muscle protein, presentation or alteration of muscle antigens, production of immune complexes, effects on the immune system, or others (Targoff 1991; Behan and Behan 1993).

Picornaviruses

A variety of viruses (influenza, hepatitis, etc.) can cause myositis that resembles polymyositis, but repeated efforts to culture viruses from muscle in typical polymyositis and dermatomyositis have been unsuccessful. Picornaviruses, small RNA viruses that include enteroviruses (coxsackie, polio, echo), animal viruses such as encephalomyocarditis, and others, have been suspected as possible aetiologic factors in polymyositis/dermatomyositis (as well as myocarditis and other autoimmune conditions). One reason is their tendency to infect muscle, evident in animal models (see below) and human infection (e.g. myocarditis, rhabdomyolysis, chronic coxsackievirus A9 myopathy; Kuroda *et al*. 1986). Although self-limited in normal patients, enteroviral infection, particularly echovirus, can cause a dermatomyositis-like syndrome, usually accompanied by meningoencephalitis and sometimes by arthritis, in patients with agammaglobulinaemia (McKinney *et al*. 1987).

There is some evidence to support a viral role in polymyositis/dermatomyositis. More frequent antibodies to coxsackievirus-B, but not other viruses, were found in recent-onset juvenile dermatomyositis (83 per cent versus 25 per cent in matched controls; Christensen *et al*. 1986). Similarly, antibodies to coxsackievirus-A or B were more frequent in adult dermatomyositis, polymyositis, or myositis/malignancy (but not overlap, anti-Jo-1, or polymyositis/interstitial lung disease) compared with normal or rheumatoid arthritis controls (Nishikai 1994). Picornaviruses cannot be demonstrated by culture or immunofluorescence in typical polymyositis/dermatomyositis (Ytterberg 1994), and the picornavirus-like structures reported in early studies of diseased muscle by EM were probably not of viral origin. Data using molecular probes has been conflicting. Some studies have found evidence of coxsackievirus by hybridization (Yousef *et al*. 1990; Bowles *et al*. 1993), but others found no evidence of any enterovirus by *in situ* methods (Rosenberg *et al*. 1989; Hilton *et al*. 1994). Rosenberg *et al*. (1989) did find 60 per cent of adult dermatomyositis biopsies (but not polymyositis or controls) to be positive in muscle macrophages with a mouse picornavirus probe (Theiler's). Reactive material in the positive studies has not been characterized further, and Hilton *et al*. (1994) suggested they were detecting non-specific effects. Most studies using the PCR have been negative for enterovirus and other viruses (Leff *et al*. 1992; Leon-Monzon and Dalakas 1992; Jongen *et al*. 1993). In contrast, Behan and Behan (1993), using large samples, found polymerase chain reaction evidence of enterovirus in 56 per cent, with positives in all polymyositis/dermatomyositis groups, and in 4 per cent of controls; hybridization was seen in macrophages, endothelial cells, and occasional muscle fibres.

Thus, evidence is not consistent or convincing, and a role for ongoing picornavirus infection in polymyositis/dermatomyositis seems unlikely for most cases, although a potential role for subgroups of patients or isolated cases remains. Persistent infection is not necessarily required for an initiating role.

Retroviruses

HIV and HTLV-1 myositis discussed above suggest that polymyositis/dermatomyositis may be induced by as yet unidentified retroviruses, and understanding the mechanisms of these conditions may provide insight into the aetiology and pathogenesis of idiopathic polymyositis. No evidence of retrovirus by PCR was found (Leff *et al*. 1992; Nelson *et al*. 1994). Retroviruses are also under study in other autoimmune conditions, such as rheumatoid arthritis, Sjögren's syndrome and systemic lupus erythematosus (Kalden and Gay 1994).

Toxoplasmosis

Cases have been reported of *Toxoplasma gondii* infection causing a picture resembling polymyositis/dermatomyositis (reviewed in Ytterberg 1994), with some responding at least partially to treatment of the *Toxoplasma*. Antibodies to *Toxoplasma* were more frequent in polymyositis patients than controls, and IgM antibodies, suggesting recent or active infection, were also more frequent in polymyositis/dermatomyositis patients (24 per cent versus 3 per cent of systemic lupus erythematosus) (Magid and Kagen 1983). The significance of the tests is unclear, and a role for *Toxoplasma* in the polymyositis/dermatomyositis of these patients has not been established. Active *Toxoplasma* infection is not found in the muscle in most typical cases of these diseases by culture or biopsy. Bretagne *et al*. (1994) failed to find *Toxoplasma* DNA by a PCR method in the muscle of three patients with polymyositis/dermatomyositis with elevated titres. Reactivation of *Toxoplasma* by disease or treatment has been suggested as a cause for the positive titres (Behan *et al*. 1983). It remains possible that some cases of apparently idiopathic polymyositis/dermatomyositis result from unrecognized *Toxoplasma* or related infection.

Non-infectious causes

Drug, chemical, and environmental factors

Unlike most other drugs that cause myopathy, D-penicillamine induces a polymyositis-like inflammatory myopathy (Takahashi *et al*. 1986). The mechanism is unknown, and there is no clear relation to dose or duration. It usually responds when the drug is stopped, but steroids are often needed and deaths have occurred. Cardiac involvement can occur (Wright *et al*. 1994), as well as the dermatomyositis rash and even anti-Jo-1 (Jenkins *et al*. 1993). This emphasizes that ingested environmental agents could induce idiopathic polymyositis/dermatomyositis. A small number of other drugs, such as cimetidine and propylthiouracil, have been associated with an inflammatory myopathy.

An important lesson from D-penicillamine polymyositis/dermatomyositis is that genetic factors may affect the risk for disease among patients with equal exposure to the aetiological agent (Garlepp 1993). It was more frequent among Japanese patients (1.2 per cent) and Asian Indian patients (1.4 per cent) than Caucasian patients (0.2 to 0.4 per cent). It was associated with HLA DR4 in Caucasians, in contrast with idiopathic polymyositis/dermatomyositis. This may relate to the use of the drug in rheumatoid arthritis, but a link with DR4 was not seen in Asian Indians with rheumatoid arthritis, who did show an association with DR2 (80 versus 47 per cent)

(Taneja *et al*. 1990). Thus, immunogenetic background may affect susceptibility.

Environmental toxins are also of potential significance in the induction of autoimmune disease (Love and Miller 1993). Possible examples include cases of polymyositis developing after poisoning by the natural fish toxin ciguatera (Stommel *et al*. 1991), or with silica exposure. There has been recent focus on polymyositis/dermatomyositis after cosmetic procedures. Cukier *et al*. (1993) identified eight patients with dermatomyositis (and one with polymyositis) developing an average of 6.4 months after bovine collagen dermal implant or test exposure. However, an analysis by Rosenberg and Reichlin (1994), including review of individual cases, found the observed frequency to be lower than the estimated expected frequency. A link between polymyositis/dermatomyositis and silicone breast implants has also been postulated. Love *et al*. found that 13 women with polymyositis/dermatomyositis after implants differed in the frequency of clinical features, autoantibodies, and HLA types from 76 without implants (Love *et al*. 1992). However, its significance as an aetiological or adjuvant factor remains to be demonstrated (Houpt and Sontheimer 1994).

Malignancy

The mechanism for the association of dermatomyositis and malignancy (noted above) is unknown. The tumour might lead to polymyositis/dermatomyositis by causing immune dysregulation, producing immune complexes, or inducing specific antimuscle reactions, but a common factor might lead to both conditions (persistent viral infections, toxins, genetic predispositions, immunological abnormalities).

Other factors

A large survey to determine risk factors for developing polymyositis/dermatomyositis found an increased frequency of recent heavy muscular exertion and emotional stress in patients (Lyon *et al*. 1989), but immunization was not correlated, and recent upper respiratory infections were negatively associated.

Immunogenetics

Associations with MHC genes indicate the importance of genetic susceptibility in polymyositis/dermatomyositis (Garlepp 1993). HLA DR3 has been associated with polymyositis in several studies, such as those of Love *et al*. (1991) (DR3 in 45 per cent in polymyositis/dermatomyositis versus 23 per cent in controls), and Ehrenstein *et al*. (1992) (DR3 in 75 per cent polymyositis versus 27 per cent in controls). The association is clearer in white than in black patients. C4A null, associated with B8/DR3, showed no independent association with polymyositis/dermatomyositis (Moulds *et al*. 1990). MHC associations do not differ between clinically defined subgroups, but are stronger with myositis-specific autoantibodies than with polymyositis/dermatomyositis overall. Compared with other patients, HLA DR3 is more frequent in patients with anti-Jo-1 (Goldstein *et al*. 1990; Love *et al*. 1991), but not other antisynthetases. Both anti-Jo-1 and other antisynthetases were associated with HLA DRw52, with or without DR3. This suggested that a short sequence on the DRb1 chain may be crucial in antisynthetase production (Goldstein *et al*. 1990). DQa associations have also been reported, including an association of anti-Jo-1 with DQa4 (Gurley *et al*. 1991). Reveille *et al*. (1992) found that most black or white antisynthetase patients showed DQA1*0501 or DQA1*0401 (associated with DR52-bearing DR types), and Japanese antisynthetase patients showed DQA1*0101, *0102, or *0103. Immunoglobulin allotypes may also predispose to development of antisynthetases; patients with anti-Jo-1 had higher Gm 3;5, and the combination of Gm 3;5 and DR3 was markedly increased (92 versus 15 per cent) (Enz *et al*. 1992).

Anti-PM-Scl is most common among Caucasians, and was not found in a large study of Japanese patients (Hirakata *et al*. 1992). It is strongly associated with DR3, as first noted by Genth *et al*. (1990) (DR3 in all 12 patients with anti-PM-Scl versus 23.5 per cent of controls, $p<0.001$), and later Marguerie *et al*. (1992) (DR3 in all 22 patients with anti-PM-Scl and 50 per cent homozygous), and Oddis *et al*. (1992) (DR3 in 15 out of 20 patients with anti-PM-Scl versus 22 per cent of controls). DR7 was associated with anti-Mi-2 (75 per cent versus 16 per cent without MSA), and DR5 with anti-SRP (57 per cent versus 20 per cent) (Love *et al*. 1991).

Pathogenetic mechanisms

Autoimmunity in polymyositis/dermatomyositis is suggested by the inflammatory picture, the response to steroids, and the association with other autoimmune diseases. Studies of both the cellular and humoral immune abnormalities have supported this impression, and have provided important insight into the mechanisms involved (Kalovidouris 1994; Plotz *et al*. 1995).

Cellular immunity

Infiltrating lymphocytes

Lymphocytes in muscle inflammatory infiltrates have been characterized using monoclonal antibodies to cell surface markers (Engel *et al*. 1990). T cells, many of which are activated (expressing DR antigens), are prominent in areas of severe inflammation (Rowe *et al*. 1983). CD4+ T cells are most abundant in untreated adults with acute disease, and decrease with treatment. In moving from the perivascular to the endomysial area, the proportion of CD8+ suppressor/cytotoxic T cells increases, and that of CD4+ T cells and B cells decreases (Arahata and Engel 1984; Mantegazza *et al*. 1993). Dermatomyositis biopsies show a higher proportion of B cells and a lower proportion of CD8+ T cells than polymyositis.

In polymyositis, many non-necrotic muscle fibres are surrounded and invaded by mononuclear cells; such fibres are very rare in dermatomyositis. A majority of surrounding and most of the invading cells are CD8+ T lymphocytes (Engel and Arahata 1984), indicating direct involvement of CD8+ T cells in fibre injury. A high proportion of invading cells are activated, and most are cytotoxic rather than suppressor (Arahata and Engel 1988a). Killer and natural killer cells do not seem to contribute to muscle damage (Arahata and Engel 1988b). Necrotic fibres become infiltrated predominantly by macrophages. By EM, cytotoxic T cells and macrophages become apposed against the fibre and send 'spike-like processes' into it, which then proliferate, but the membrane remains intact (Arahata and Engel 1986). Many endomysial (not perimysial) cells may show perforin and granzyme A, granule proteins of cytotoxic T cells that may participate in fibre damage, as well as TIA-1 protein related to apoptosis (Cherin *et al*. 1993a; Orimo *et al*. 1994).

The T-cell receptors of the infiltrating lymphocytes also suggest an antigen-directed T-cell attack in polymyositis. T-cell receptor rearrangements and marked restriction in V gene usage are seen, although the V genes used differed between studies. Using PCR techniques, Mantegazza *et al*. (1993) found preferential usage of Va1, Va5, Vb1, and Vb15, while O'Hanlon *et al*. (1994a) showed predominance of Va1 and Vb6, with Jb gene conservation, among anti-Jo-1 patients with polymyositis, but not dermatomyositis. By histochemistry, V-gene usage differed between cells in the perimysial areas and those in the endomysial areas that are attacking the muscle fibres (Lindberg *et al*. 1994a). Bender *et al*. (1995), confirming this, found that sequences of T-cell receptor V genes from different invading T cells of an individual were often identical, suggesting local clonal expansion, evidently in response to a muscle antigen. One patient had a unique form of polymyositis, in which the infiltrating T cells had almost exclusively g/d-T-cell receptors (Hohlfeld *et al*. 1991b). It was hypothesized that the antigens may be heat-shock proteins, a common target of such T cells. The T cells may have been derived from a single T-cell clone (Pluschke *et al*. 1992).

Thus, the predominant mechanism for immunologically mediated muscle injury in polymyositis, but not dermatomyositis, appears to be T-cell cytotoxicity, directed at an unidentified antigen, presumably on the muscle fibre surface. The fibre is surrounded and invaded by activated, clonally expanded, cytotoxic T cells. Additional mechanisms, such as cytokines, may come into play in the final destruction of the fibre.

T-cell attack on muscle fibres does not appear to occur in dermatomyositis, but cellular immunity may still play a role, possibly contributing to the vessel injury. Saito *et al*. (1989) found that mononuclear cells (mostly CD8+ T cells) from patients with dermatomyositis, but not polymyositis, showed cytotoxicity against cultured fibroblasts in a non-MHC restricted manner.

Muscle fibres

While expression of MHC-1 antigens on muscle fibres is normally low or absent, in polymyositis and inclusion-body myositis muscle many fibres strongly express MHC-1, making them vulnerable to antigen-directed T-cell mediated cytotoxicity (Karpati *et al*. 1988). MHC-1 was seen on all fibres that were invaded by CD8+ lymphocytes (Emslie-Smith *et al*. 1989). However, it can also be seen in dermatomyositis, where surrounded and invaded fibres are rare. MHC-2 antigens have been found on muscle fibres in most active polymyositis and dermatomyositis biopsies (Kalovidouris 1994), but not in normal muscle or that of other myopathies. ICAM-1 (intercellular adhesion molecule 1) is induced on muscle fibres under attack, in the area of invasion, and may be important in adhesion of invading cells (De Bleecker and Engel 1994).

Since strength often returns with treatment of polymyositis and dermatomyositis sooner than would be expected if regeneration were required, much of the weakness may relate to muscle cell dysfunction. This may result from cytokines or other factors released by infiltrating mononuclear cells, that can interfere with calcium binding to the sarcoplasmic reticulum and suppress muscle contractility (Kalovidouris 1986). A possible additional mechanism may be relative acquired enzyme deficiencies such as myoadenylate deaminase (Sabina *et al*. 1990).

Peripheral blood lymphocytes

Early studies of peripheral blood lymphocytes in polymyositis/dermatomyositis showed increased proliferation in response to muscle extract in some studies (Targoff 1991), greater in active untreated cases, but generally of a low level, and not myositis-specific. Kalovidouris *et al*. (1989) later found significant lymphocyte stimulation in response to autologous or allogeneic muscle much more frequently in active (76 per cent) than treated myositis (12.5 per cent) or controls, consistent with sensitization of lymphocytes to muscle.

Peripheral lymphocytes from patients with polymyositis/dermatomyositis showed toxicity for muscle cells in culture in some studies (Targoff 1991), but this was not specific for polymyositis/dermatomyositis and did not require MHC compatibility, suggesting a non-antigen-specific mechanism such as natural killer cells. Recently, Hohlfeld and Engel (1991a) tested cultured muscle-infiltrating cytotoxic T cells against cultured autologous muscle, and some lines did show some evidence of cytotoxicity by chromium-51 release, but further study is needed to demonstrate clearly antigen-directed T-cell cytotoxicity in polymyositis/dermatomyositis.

In active polymyositis/dermatomyositis, a high proportion of peripheral blood T cells show activation markers (Miller *et al*. 1990a), decreasing with treatment. As expected, those with dermatomyositis show fewer activated T cells, but more B-cells than those with polymyositis. There is increased trafficking of lymphocytes to muscle. Peripheral lymphocytes show decreased mitogen responsiveness in polymyositis, and decreased autologous mixed lymphocyte responses (Plotz *et al*. 1989).

Cytokines

Cytokines could be important in the development, enhancement, or perpetuation of the autoimmune response, or in tissue injury and cytotoxicity. Lymphocytes cultured with autologous muscle elaborate cytotoxic factors (Johnson *et al*. 1972). Interferon-g, present in polymyositis/dermatomyositis muscle (Isenberg *et al*. 1986), can enhance MHC-1 and induce MHC-2 expression on cultured muscle fibres *in vitro* (Kalovidouris 1992) and enhance T-cell adhesion to muscle by increasing ICAM-1 expression (Kalovidouris *et al*. 1994). Interferon-g can inhibit proliferation and differentiation, an effect enhanced by tumour necrosis factor-a (TNF-a), thus possibly directly injuring or preventing repair of muscle fibres (Kalovidouris *et al*. 1993). Polymyositis has occurred after interferon therapy.

The serum levels of IL-2 receptor and IL-1a are elevated in active polymyositis and dermatomyositis (Wolf and Baethge 1990) and fall with treatment, possibly reflecting activated lymphocytes. IL-1 receptor antagonist was markedly elevated in some patients with active disease, and fell with treatment (Gabay *et al*. 1994), while spondylarthropathies and rheumatoid arthritis show instead higher IL-6 and C-reactive protein, possibly reflecting the relative prominence of activated T cells in polymyositis/dermatomyositis.

Humoral immunity

Microvascular injury

Direct T-cell mediated attack against muscle fibres does not seem to be important in dermatomyositis, with little endomysial infiltrate, and few surrounded and invaded non-necrotic fibres. However, the intense B-cell and CD4+ T-cell infiltrate in the perivascular area suggests a local humoral response (Arahata and Engel 1984) (Table 7).

	Dermatomyositis	Polymyositis
Skin rash	Yes	No
Microvascular injury	Yes	No
Perifascicular atrophy	Common	Uncommon
Endomysial infiltration	Less common	Common
Surrounded and invaded fibres	Rare	Frequent
Deposition of membrane attack complex	Yes	No
Anti-Mi-2	15–20%	<1%
Anti-SRP	<1%	6%

Table 7 Differences between polymyositis and dermatomyositis

Vasculopathy involving the small vessels and capillaries, with resultant ischaemic damage and perifascicular atrophy, is an important mechanism of muscle injury in dermatomyositis (Emslie-Smith and Engel 1990; Heffner 1993). Long associated with juvenile dermatomyositis, it is often found in typical adult dermatomyositis, but generally not polymyositis. Casademont *et al*. (1993) found that 87 per cent of patients with muscle capillary damage had a dermatomyositis rash.

Emslie-Smith and Engel (1990), studying muscle biopsies from adults with clinical dermatomyositis that showed little or no structural change by routine examination ('early dermatomyositis'), found definite microvascular abnormalities by EM in all cases, including endothelial cell injury, microtubular inclusions, and decreased capillary density (confirmed quantitatively), not seen in polymyositis or inclusion-body myositis. Biopsies showing overt dermatomyositis had even lower capillary density and more advanced vascular changes, including capillary necrosis.

Microvascular damage in muscle is felt to be mediated by complement. Kissel *et al*. (1986) found deposition of complement membrane attack complex (**MAC**) in the walls of muscle microvasculature, indicating local activation of complement, in most cases of juvenile dermatomyositis and some cases of adult dermatomyositis, but not polymyositis. It is greater in areas where ischaemic damage is recent (fibres with 'punched-out' central myofibrillar loss) than in those with perifascicular atrophy (a later change), and is less evident in long-standing disease (Kissel *et al*. 1991). Emslie-Smith and Engel (1990) found membrane attack complex deposition in 9 of 10 'early dermatomyositis' biopsies. Since injury to the microvasculature can occur before inflammatory infiltrates and muscle fibre damage are evident, it may be the primary target in the pathogenesis of dermatomyositis.

The factors leading to complement activation are not known. Deposition of immunoglobulin in muscle blood vessels, not seen in normal people, has been found in dermatomyositis in some studies, especially juvenile dermatomyositis and autoimmune rheumatic disease overlap, and less commonly polymyositis, but is not specific (Isenberg 1983). Its relation to the membrane attack complex deposition is unknown. Deposition has also been seen in the periphery of the fibre (sarcolemma and basement membrane), and in the fibre itself.

Autoantibodies

The clearest evidence of abnormality of humoral immunity in polymyositis/dermatomyositis is the presence of autoantibodies to nuclear and cytoplasmic antigens in up to 89 per cent of patients (Reichlin and Arnett 1984) (Fig. 12). A positive ANA is found in 40 to 80 per cent, more commonly in overlap and less in myositis with malignancy. The specificities of these antibodies have been studied in detail (Targoff 1994), and are heterogeneous, possibly reflecting heterogeneity of the disease (Table 8). About half of patients have autoantibodies of recognized specificity, some of which are found primarily in polymyositis/dermatomyositis, referred to as MSAs (Miller 1993). Others have an association with myositis, but may be found in other conditions. Most MSAs are associated with a characteristic clinical picture, and an individual generally has only a single MSA, so that MSAs can define clinical subgroups (Love *et al*. 1991).

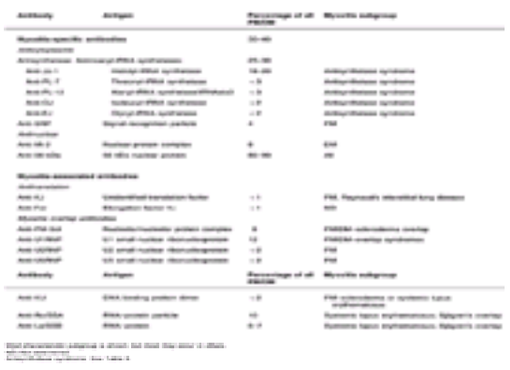


Table 8 Autoantibodies in polymyositis/dermatomyositis (PM/DM)

Myositis-specific antibodies (MSAs)

Antisynthetases

About 30 per cent of patients with polymyositis/dermatomyositis have antibodies to an aminoacyl-tRNA synthetase, an enzyme that attaches one specific amino acid to its cognate tRNAs. Five antisynthetases occur in polymyositis/dermatomyositis sera (Table 8). These enzymes are antigenically distinct, and an individual patient has antibodies to only one. By far the most common is antibody to histidyl-tRNA synthetase (anti-Jo-1) (Mathews and Bernstein 1983), found in 20 per cent of patients with polymyositis/dermatomyositis.

A group of clinical features have been associated with anti-Jo-1 in several studies (Table 9; Marguerie *et al*. 1990; Love *et al*. 1991). The limited number of patients with non-Jo-1 antisynthetases studied have had a similar clinical picture (Targoff and Arnett 1990; Targoff *et al*. 1992; Targoff *et al*. 1993). Apart from myositis, the most striking feature is interstitial lung disease, found in 50 to 90 per cent with the antibodies, but less than 10 per cent of other patients with polymyositis/dermatomyositis. The interstitial lung disease can dominate the picture, and some patients have this without overt myositis (more common with anti-PL-12). Two-thirds of patients with anti-Jo-1 have inflammatory polyarthritis (Oddis *et al*. 1990b), usually mild and responsive to treatment, but a third of these may have finger deformity, usually non-erosive, occasionally with calcinosis (rare erosive disease may represent rheumatoid arthritis overlap; O'Neill and Maddison 1993). Raynaud's phenomenon (62 per cent), fever (87 per cent), and mechanic's hands (71 per cent) are other important features (Love *et al*. 1991). Marguerie *et al*. (1990) found that sclerodactyly (72 per cent) and sicca (59 per cent) were also frequent. Their response to therapy may be less complete, with more relapses. A third to a half of the patients have dermatomyositis (some find a lower proportion among anti-Jo-1 patients). Malignancy is uncommon but has occurred. Patients with these antibodies are also immunogenetically distinctive (see above).

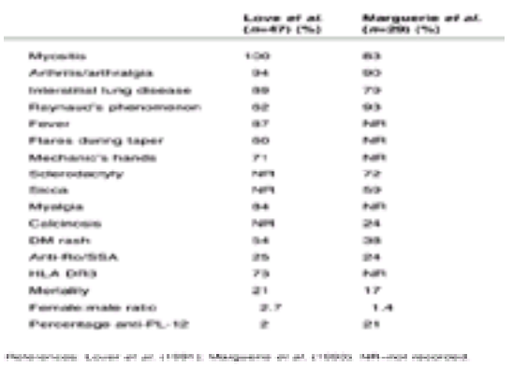


Table 9 Features of antisynthetase syndrome

The set of clinical features associated with antisynthetases has been referred to as the 'anti-synthetase syndrome' (Targoff 1994). It may resemble other autoimmune rheumatic diseases syndromes, including mixed connective tissue disease (with joint disease, scleroderma and Raynaud's phenomenon; Marguerie *et al*. 1990) or systemic lupus erythematosus (fever, pneumonitis, polyarthritis, and positive ANA). However, myositis is usually more prominent and less responsive to treatment, significant interstitial lung disease is more frequent, and features of systemic lupus erythematosus are rare in the absence of other autoantibodies.

Other anticytoplasmic myositis-specific antibodies

Antibodies to other cytoplasmic antigens are found in a small proportion of patients with polymyositis/dermatomyositis (Table 8), the most important of which is antibody to the signal recognition particle (anti-SRP), a ribonucleoprotein involved in protein translocation into the endoplasmic reticulum. Anti-SRP is found almost exclusively in adult polymyositis (Targoff *et al*. 1990), with no increase in interstitial lung disease, arthritis, or Raynaud's phenomenon. Some patients with anti-SRP have an acute, fulminant course, resistant to treatment (Love *et al*. 1991). Antibodies to translation factors have been identified rarely in patients with polymyositis/dermatomyositis, including anti-KJ, which was associated with a picture similar to the antisynthetase syndrome (Targoff *et al*. 1989). Anti-Mas, directed at a tRNA-related antigen, was identified in rare patients with inflammatory myopathy and alcoholism (Love *et al*. 1991), but is not myositis-specific. Several MSAs are directed at cytoplasmic antigens related to tRNA or protein synthesis, but more patients with polymyositis/dermatomyositis have antinuclear than anticytoplasmic antibodies.

Antinuclear myositis-specific antibody

Anti-Mi-2, which reacts with a nuclear antigen of unknown function, is strongly associated with dermatomyositis (Targoff 1994). Ninety-five per cent of patients with anti-Mi-2 have dermatomyositis, and 15 to 20 per cent with dermatomyositis have the antibody. The rash is often florid (Fig. 6), with a higher frequency of the 'V' and 'shawl' signs than others with dermatomyositis (Love *et al*. 1991). Patients with anti-Mi-2 tend to do better than those with antisynthetases. No increase in interstitial lung disease or Raynaud's phenomenon is seen.

Myositis-overlap antibodies

Several autoantibodies are seen in patients with myositis–scleroderma or other overlap syndromes. Although not MSAs, since some patients do not have myositis, these antibodies can be very helpful in evaluating a patient with suspected myositis.

Anti-PM-Scl

Anti-PM-Scl reacts with a nucleolar protein complex of unknown function containing at least 11 polypeptides. Patients with anti-PM-Scl have myositis (5 to 10 per cent of patients with polymyositis/dermatomyositis), scleroderma, or, most commonly, a myositis–scleroderma overlap syndrome, often with arthritis (Marguerie *et al*. 1992; Oddis *et al*. 1992). Cutaneous scleroderma is usually limited when present. The myositis is often mild and tends to be responsive to treatment. Raynaud's phenomenon and interstitial lung disease are common and calcinosis may be increased. Jablonska *et al*. (1993) describe 'scleromyositis', a syndrome with features of polymyositis/dermatomyositis and scleroderma with anti-PM-Scl that may resemble mixed autoimmune rheumatic disease. In Japan, anti-PM-Scl is rare, but anti-Ku is commonly found in polymyositis–scleroderma overlap (Hirakata *et al*. 1992). Anti-Ku is rare in polymyositis/dermatomyositis in the United States, but found more often

in systemic lupus erythematosus and scleroderma.

Anti-snRNPs

Anti-U1RNP is found in 10 to 15 per cent of polymyositis/dermatomyositis, often with autoimmune rheumatic disease overlap features such as Raynaud's phenomenon, arthritis or dactylitis. Myositis is often found in patients with overlap syndromes or systemic lupus erythematosus who have U1RNP. The myositis tends to be more responsive to treatment (Lundberg *et al*. 1992; Jablonska *et al*. 1993). Three per cent of patients with polymyositis/dermatomyositis have anti-Sm, usually with systemic lupus erythematosus overlap. Antibodies specific for the U2 snRNP (usually with anti-U1RNP) occur in a small number of patients with polymyositis–scleroderma (Craft *et al*. 1988), and anti-U5RNP (without anti-U1RNP/Sm) appears to be myositis-specific (Rider *et al*. 1994).

Patients with anti-U3RNP, a scleroderma-specific antibody, have a higher frequency of inflammatory myositis than others with diffuse scleroderma (Okano *et al*. 1992). Anti-Ro/SSA, often with anti-La/SSB, is found in 10 per cent of polymyositis/dermatomyositis, more commonly with antisynthetases; Sjögren's syndrome or systemic lupus erythematosus overlap may be present, but is not always evident. Antibodies associated with other autoimmune rheumatic diseases, such as anticentromere, occur in occasional patients with polymyositis/dermatomyositis, usually as part of overlap syndromes.

Other myositis-associated autoantibodies

The sera of 87 per cent of patients with polymyositis/dermatomyositis reacts with a 56-kDa protein of nuclear ribonucleoprotein particles (Arad-Dann *et al*. 1989). Anti-56 kDa is common in all subgroups, but more in dermatomyositis than polymyositis. It is not completely specific, being found in 10 per cent of other autoimmune rheumatic diseases (usually in lower titre), but was not found in serum of normal subjects or patients with other muscle diseases. The titre varied with disease activity. It may be a more general marker for myositis, and have a different significance than other MSAs.

Antiendothelial cell antibodies, of interest in view of the endothelial injury in dermatomyositis, and antibodies to heat shock proteins, occur in some patients with polymyositis and dermatomyositis, but they are common in other conditions, and their significance in polymyositis/dermatomyositis is unknown (Cervera *et al*. 1991). Antibodies to muscle proteins such as myosin occur in up to 90 per cent of polymyositis/dermatomyositis (Wada *et al*. 1983; Koga *et al*. 1987), and may vary with disease activity. Although more frequent in polymyositis/dermatomyositis than other muscle diseases, they are not specific, and could be secondary to muscle damage. The possibility that they contribute to muscle injury is of interest, but their role is unknown.

Possible relation to pathogenesis

The reason for production of these antibodies is unknown, and may relate to aetiological factors (Targoff 1994). HLA and other genetic factors seem to be important (see above). There is increasing evidence that production is driven and perpetuated by the recognized antigens (Miller *et al*. 1990c), but they may not initiate the responses. It is often speculated that myositis-inducing viruses lead to the antibodies through such mechanisms as molecular mimicry, formation of immunogenic complexes between host and viral factors, etc. (Targoff 1991). Hypotheses must explain how patients with antibodies to different members of the same enzyme family, present in every cell, have a similar clinical syndrome manifested in specific organs.

It is not known if the antibodies play a direct role in pathogenesis (Naparstek and Plotz 1993). Antisynthetases consistently inhibit antigen function, but they have not been shown to enter intact cells. They could injure muscle through immune complexes, reaction with cell surface antigens, etc., but there is as yet no evidence of this. A correlation of anti-Jo-1 titre with myositis activity (Miller *et al*. 1990b), including the development of anti-Jo-1 shortly before the onset of weakness (Miller *et al*. 1990a), suggests a pathogenetic role, but against such a role is their common occurrence in polymyositis, where cell-mediated fibre damage is most important, and low frequency in juvenile dermatomyositis, where humoral injury occurs.

The epitopes of the myositis-associated autoantibodies are under study; predominant epitopes are often found (Raben *et al*. 1994). They generally differ from those of animal antisera, and tend to be conformational (Targoff 1994). Specific immunological features of the antibodies may provide clues to their origin.

Animal models

Experimental autoimmune myositis

Certain strains of guinea pig, rat, or mouse develop myositis (but not weakness or rash) as a result of immunizations with homologous or heterologous muscle homogenates in adjuvant (Rosenberg 1993). Inflammation and necrosis occur, but no vascular changes or perifascicular atrophy. Muscle damage is mediated by cellular immune mechanisms, and lymphoid cells can transfer disease. Humoral reaction to muscle is seen, and serum could transfer disease in one (Matsubara *et al*. 1993) but not other studies. The histology resembles human polymyositis, but with more macrophages in the inflammatory infiltrate (80 to 90 per cent). Experimental autoimmune myositis demonstrates that cell-mediated immunity can mediate muscle injury, but is limited by differences from human polymyositis/dermatomyositis.

SJL/J mice are uniquely susceptible to experimental autoimmune myositis, and with aging, develop a similar and more severe myositis spontaneously. Immunological defects and neoplasms occur in these mice, which may contribute. This spontaneous model may provide insight into aetiological factors (Rosenberg 1993).

Viral models

Picornaviruses

Coxsackie and several other picornaviruses can produce a polymyositis-like myositis in mice. Neonatal Swiss mice infected with coxsackie-B1 Tuscon strain develop an acute illness followed by a chronic inflammatory myositis resembling polymyositis, with proximal muscle weakness, myopathic electromyographs, and mononuclear cell infiltration, with degenerating and regenerating muscle fibres (Strongwater *et al*. 1984). Myositis progresses beyond the period in which the virus can be cultured (6 months versus 2 weeks), but viral nucleic acid persists longer than culturable virus (Tam *et al*. 1994). Production of the myositis depends on the strain of virus and of mouse used, and requires cell-mediated immunity (Ytterberg *et al*. 1987).

Infection of adult mice with a myotropic strain of encephalomyocarditis virus produces an inflammatory myositis and myocarditis in some mouse strains (Miller *et al*. 1987). As with coxsackievirus, development of the myositis requires specific strains of virus and genetic background (susceptibility was H2 restricted). Further, viral nucleic acid could be detected at 4 weeks, when virus could no longer be cultured but muscle inflammation persisted (Cronin *et al*. 1988); thus viral persistence may be important in chronicity. The mice do not develop myositis-associated autoantibodies. Thus, picornaviruses can produce an immune-mediated myositis, supporting their potential as aetiological agents in some cases of idiopathic polymyositis/dermatomyositis.

Retroviruses

Macaque monkeys develop a syndrome resembling AIDS after infection with the D-type simian retrovirus SRV-1. Up to half of infected monkeys develop a myositis that closely resembles polymyositis as part of their disease (Dalakas *et al*. 1987), with weakness, elevated enzymes, and a biopsy showing inflammation, necrosis, phagocytosis, and fibrosis. The mechanism is unclear, although the virus could infect cultured muscle cells. The model may be useful in the study of retrovirus-associated myositis, as well as idiopathic polymyositis/dermatomyositis.

Canine dermatomyositis

A condition resembling dermatomyositis occurs spontaneously in collies and Shetland sheep dogs (Hargis and Prieur 1988), with autosomal dominant inheritance. Juvenile dogs develop muscle weakness and atrophy, difficulty eating, and an erythematous, scaly rash over the periorbital areas, face, and distal extremities. Muscle histology shows inflammation, fibre necrosis, regeneration, and atrophy, and vasculitis may occur. Serological tests are usually negative. The model has significant differences from human myositis, but is most similar to juvenile dermatomyositis. Dogs may spontaneously develop an eosinophilic myositis of the muscles of mastication. Occasionally, polymyositis is seen.

Management

Treatment should be started promptly, since significant delay has been associated with poorer outcome (see below). Treatment is most urgent, and sometimes initiated pending completion of the evaluation, in patients with rapid onset of severe weakness, dysphagia, respiratory insufficiency, myocardial involvement, or systemic signs. Most cases can be managed outside the hospital, but admission may be required for patients with respiratory insufficiency or severe dysphagia, those requiring intravenous medications, or clarification of diagnosis.

A general approach to treatment is outlined in Table 10. Guidelines for initial treatment, dosage reduction, and choice of agents should be adapted to the situation of the patient. Most patients are first given a trial of corticosteroids alone. Initial therapy with a combination of steroids and immunosuppressives also has some support (Bunch 1981). Some have recently suggested other agents for first-line treatment, such as cyclosporine (Grau *et al*. 1994) or for certain patients, intravenous gammaglobulin (Dalakas 1994a). Immunosuppressives and recently intravenous gammaglobulin are used when additional therapy is required.

Table 10 A strategy for treatment of newly diagnosed, active, idiopathic polymyositis or dermatomyositis

Corticosteroids

While not demonstrated by prospective randomized controlled trials, the effectiveness of steroids in improving muscle strength is accepted generally, and readily apparent from observation of treated patients. It is further suggested by studies demonstrating improved outcome with earlier treatment.

Initial treatment

Treatment is usually begun at high doses (prednisone 1 mg/kg per day, usually 60 to 80 mg/d, or equivalent), often in divided doses that are considered to increase effectiveness. Dalakas (1994a); Dalakas (1994b) recommends 80 to 100 mg/d as a single morning dose, feeling that this increases its benefits.

The high initial dose (prednisone more than 60 mg/d) is continued until the creatine kinase has returned to normal and the strength has substantially improved, usually 1 to 2 months (Oddis 1991). Adequate initial treatment is important, and is associated with better responses (Oddis and Medsger 1988). When begun as divided doses, the daily amount is usually consolidated to a single morning dose after the acute period or before beginning dosage taper.

Very high-dose intravenous 'pulse' methylprednisolone, in regimens such as 1 g/day for 3 days, is sometimes used initially to achieve a more immediate effect in severely ill patients, or later when disease is unresponsive to oral steroids, or for exacerbations during steroid taper (Oddis 1993). Matsubara *et al*. (1994) found that pulse therapy (0.5 g/day for 3 days per week for 1 to 9 weeks) added to oral therapy increased the rate of and shortened the time to remission. Although some have used alternate-day steroids in milder cases from the outset to reduce side-effects (Hoffman *et al*. 1983), it is less reliable, and not recommended in most cases (Dalakas 1994b).

Response

It is important to monitor patients closely during treatment for degree of improvement, new complications, or side-effects. Both muscle strength (assessed as above) and creatine kinase levels are used to judge improvement. The response to steroids in polymyositis/dermatomyositis tends to be slower than that in systemic lupus erythematosus or rheumatoid arthritis, and it may take months to achieve the full effect (Plotz *et al*. 1989), with a mean time to recovery of normal strength of more than 3 months (Tymms and Webb 1985). The creatine kinase usually normalizes weeks to months earlier than the strength, generally indicating adequate suppression of disease activity. Creatine kinase levels in the high normal range may indicate disease activity and may decrease with treatment.

Oddis and Medsger (1989) found that treatment using the following guidelines was more likely to result in prolonged suppression of disease: (i) start with 60 mg/day or more of prednisone for at least 1 month; (ii) continue initial treatment until or after the creatine kinase becomes normal; and (iii) taper steroids slowly, with average reduction from first reduction to maintenance dosage of 10 mg/month or less.

Most patients respond at least partially to treatment with steroids. If there is no improvement or if significant weakness persists, possibilities include: (i) incorrect diagnosis: the basis for the diagnosis should be reviewed, and further evaluation considered, such as repeat biopsy looking in particular for inclusion-body myositis or metabolic myopathy; (ii) malignancy: although myositis with malignancy may be quite responsive (Joffe *et al*. 1993), failure to respond or weight loss on steroids should suggest a more extensive search (Plotz 1992); (iii) steroid myopathy: when creatine kinase falls but weakness persists, not explained by the expected lag of strength behind creatine kinase, steroid myopathy should be considered; (iv) permanent loss of strength: some patients cannot recover full strength despite complete suppression of disease activity, often those with prolonged delay before treatment, atrophy, or prominent fibrosis on biopsy; (v) resistance: disease activity may be unresponsive to steroid therapy. Persistent elevation of the creatine kinase usually means that the disease has not been controlled. Occasional patients have resolution or stabilization of weakness while creatine kinase remains elevated, often at low levels. Such elevations may not reflect ongoing disease activity (Oddis 1994); MRI may help to assess activity in these situations.

Dosage reduction

When the initial goals are reached, the dosage is gradually reduced over 6 to 8 months, usually at 15 to 25 per cent per month (Table 10). Some recommend routine conversion to an alternate-day regimen after the initial high-dose daily regimen (Dalakas 1988b), usually accomplished gradually, as by reduction of the low-day dosage by about 10 mg per week. This may be advantageous in patients who have developed or are at high risk for significant steroid side-effects. Steroids are usually continued for 1 to 2 years or longer at a low, maintenance dose (Oddis 1994) that is high enough to successfully prevent recurrence but low enough to keep the risk of side-effects low (prednisone 5 to 10 mg/day or 10 to 20 mg every other day).

Less aggressive or less prolonged therapy may be adequate in some patients. In a primary care setting, Hoffman *et al*. (1983) could successfully discontinue treatment in 41 per cent of patients without recurrence. The myositis in the autoimmune rheumatic diseases overlap syndromes associated with anti-U1RNP and anti-PM-Scl may be more responsive and require less therapy than other myositis (Jablonska *et al*. 1993). However, the standard regimens should be considered more reliable for most patients with polymyositis/dermatomyositis.

Exacerbations of disease may occur during taper, and are more likely if dosage is reduced rapidly. A rise in the creatine kinase level, sometimes within the normal range, frequently precedes exacerbation (Oddis and Medsger 1989). If creatine kinase elevation occurs after initial response, causes other than disease flare should be considered, particularly when no other signs are present. Exacerbation without creatine kinase elevation may occur, even if the enzyme was elevated originally, but if weakness increases without creatine kinase elevation, steroid myopathy should be considered. Other tests, such as MRI, electromyography, or repeat needle biopsy, can help assess disease activity in these cases. Exacerbation is treated usually with an increase in steroid dosage above those which last maintained control

of the disease, although full-dose is usually not necessary, and/or addition of an immunosuppressive agent (Oddis 1991).

The side-effects of corticosteroids in polymyositis/dermatomyositis are similar to those in other situations. Steroid myopathy poses a special problem, as noted above. Steroid-induced hypokalaemia may also lead to further weakness and should be avoided. Due to the prolonged high-dose steroids used in polymyositis/dermatomyositis, osteoporosis and aseptic necrosis are significant problems, contributing to disability. Adequate calcium and vitamin D intake should be maintained. Tymms and Webb (1985) found aseptic necrosis in 8.6 per cent of patients with polymyositis/dermatomyositis receiving high-dose steroids. Gastric protection is commonly used. Diabetes, hypertension, cataracts, and opportunistic infections (including tuberculosis) may be seen.

Immunosuppressives

Immunosuppressive agents are required in about a quarter to a half of patients. Indications include: (i) steroid resistance, with failure to respond to high-dose steroids after 6 to 12 weeks, (excluding considerations mentioned above); (ii) persistent disease activity after prolonged therapy despite initial improvement; (iii) inability to taper the prednisone without recurrence; or (iv) severe steroid side-effects. Immunosuppressives would be considered earlier in those with very severe or acute disease, or factors suggesting poorer response, such as certain autoantibodies (antisynthetases or anti-SRP) (Plotz *et al*. 1995). Generally, the trend is toward earlier and more frequent use of immunosuppressive agents because of steroid side-effects and possibly improved response (Steinberg 1993).

Methotrexate and azathioprine are the immunosuppressives used most extensively. Retrospective analysis (Joffe *et al*. 1993) suggests a higher response rate and more rapid onset of effect with methotrexate, and some (Oddis 1994) use it first if an immunosuppressive is required and there are no contraindications. However, azathioprine may have less toxicity and has demonstrated efficacy in a controlled study (Bunch 1981), and some prefer it as first choice (Ramirez *et al*. 1990; Urbano-Marquez *et al*. 1991; Dalakas 1992b).

Joffe *et al*. (1993) found that patients with antisynthetases were more likely to respond to methotrexate than to azathioprine. Methotrexate should be used with caution in this setting because of the frequent interstitial lung disease in antisynthetase patients, which can lead to diagnostic confusion with hypersensitivity pneumonitis, or difficulty tolerating it. They also found that men were more likely to respond to methotrexate than azathioprine.

Methotrexate

Methotrexate has been used over a broad dosage range in polymyositis/dermatomyositis. Early studies frequently employed a high-dose regimen administered intravenously (intramuscular injections may interfere with creatine kinase monitoring in polymyositis/dermatomyositis), begun at 10 to15 mg/week, and gradually increased to 0.5 to 0.8 mg/kg per week (30 to 50 mg/week). After an adequate response, methotrexate was tapered by extending the dosing interval to 2, 3, and 4 weeks, or by decreasing the weekly dose. Metzger *et al*. (1974) found good to excellent responses in 15 of 22 patients (68 per cent), at an average of 13 weeks.

Weekly low-dose oral regimens are now more often used (Plotz *et al*. 1989; Oddis 1994), starting at 7.5 to 10 mg/week and increasing gradually as required to a maximum of 20 to 25 mg/week. Parenteral therapy should be considered for the higher dose ranges, or if gastrointestinal intolerance develops. Response times may range from 3 to 44 weeks. The prednisone dose can usually be reduced ('steroid-sparing'). Exacerbations may occur with methotrexate taper in responsive patients.

Although methotrexate toxicity in polymyositis/dermatomyositis is usually mild, severe toxicity (hepatotoxicity, fatal hypersensitivity pneumonitis) has occurred. As in rheumatoid arthritis, patient selection is important (avoid in renal insufficiency, hepatic damage, alcoholism). The monitoring guidelines for rheumatoid arthritis should be considered a minimum; the higher doses often used in polymyositis/dermatomyositis may require closer observation, and elevated transaminases from muscle injury may interefere with their use in monitoring. Zieglschmid-Adams *et al*. (1995) observed hepatic fibrosis (grade IIIA) with methotrexate in 2 of 10 patients with dermatomyositis on usual doses; diabetes was a risk factor. Euwer and Sontheimer (1994) feel that toxicity in dermatomyositis may resemble that of psoriasis more than rheumatoid arthritis. Other side-effects include stomatitis, infections, teratogenicity, bone marrow suppression, and gastrointestinal bleeding. A methotrexate-associated reversible lymphoma was reported in dermatomyositis (Kamel *et al*. 1993).

Azathioprine

Bohan *et al*. (1977) noted improvement with azathioprine in about 35 per cent of polymyositis/dermatomyositis, and a steroid-sparing effect in 50 per cent. Ramirez *et al*. (1990) noted benefit in 75 per cent of those with an adequate course. In a controlled trial (Bunch 1981), patients beginning therapy with azathioprine plus prednisone had better long-term outcome with less disability and less prednisone requirement than those receiving prednisone alone, although there was no difference in short-term outcome after 3 months. It is used orally at 1.5 to 3.0 mg/kg per day, with doses of 100 to 150 mg/d being most common. Dalakas (1992b) recommends 3 mg/kg per day, with an adequate trial requiring 3 to 6 months. Steroids may be gradually reduced as response occurs. Cytopenias or bone marrow suppression may occur, and complete blood counts including platelets must be monitored. The possibility of malignancy is a concern, but the risk appears low in polymyositis/dermatomyositis. Other side-effects are similar to those in other situations.

Alkylating agents

Cyclophosphamide and chlorambucil are generally reserved for those who have failed to respond to other agents (Steinberg 1993), or have other manifestations such as severe interstitial lung disease. They generally have a higher risk of serious side-effects, particularly increased malignancy but also bone marrow suppression, infertility, and, for cyclophosphamide, haemorrhagic cystitis.

Reports regarding the value of oral cyclophosphamide in polymyositis/dermatomyositis have been conflicting (Plotz *et al*. 1989); some found it lacked efficacy, but others had success. It is usually used at a dose of 1 to 2 mg/kg per day, with close monitoring of the white blood count. Monthly intravenous pulse cyclophosphamide has been used in an effort to limit side-effects. Cronin *et al*. (1989), used this in patients with long-standing, refractory disease, and found poor efficacy and high toxicity (fever, nausea, infections, etc.). Bombardieri *et al*. (1989), however, had a high success rate in patients without previous cytotoxic therapy using short courses (2 to 3 weeks) of 0.5 g intravenous pulses at 1 to 2 week intervals, with a repeat course for flares. De Vita and Fossaluzza (1992) had good results with a similar regimen, beginning with 3 weeks of low pulses, then 0.5 to 1 g/m^2 per month for non-responders, with individual regimens as needed.

Chlorambucil has also been successful in a small number of reported patients. Of five patients treated with chlorambucil at 4 mg/day by Sinoway and Callen (1993), all had some improvement by 4 to 6 weeks, and normal strength by 13.5 months, but with minimal effect on the skin disease. Little toxicity was seen (two with leucopenia), and the drug could be stopped in four patients.

Cyclosporin (Neoral)

Several case reports or small series have described beneficial effects of cyclosporin in polymyositis/dermatomyositis, more in children but some in adults, including in severe, resistant disease (Oddis 1994). Grau *et al*. (1994) compared their experience in 10 consecutive patients with dermatomyositis treated initially with cyclosporin (usually 5 mg/kg per day), with their previous 45 patients treated initially with prednisone. Cyclosporin-treated patients responded earlier and achieved complete remission in less than half the time (mean 8.6 weeks), with a comparable failure rate and less frequent serious toxicity. Nephrotoxicity is the greatest concern. Although many reports used doses of greater than 5 mg/kg per day, following the recent guidelines of Feutren and Mihatsch (1992) would be prudent (doses of 5 mg/kg or less per day, and creatinine kept less than 30 per cent over baseline). Tacrolimus (FK506) was also effective in a preliminary report (Oddis *et al*. 1994).

Combination therapy

Combinations of immunosuppressives, such as methotrexate and azathioprine or chlorambucil and methotrexate, often with prednisone, have been used in patients with refractory disease, or to limit toxicity by using smaller doses. The methotrexate/azathioprine combination and high-dose methotrexate with leukovorin rescue are under study (Steinberg 1993).

Intravenous gammaglobulin

There is increasing evidence of the efficacy of high-dose intravenous gammaglobulin in polymyositis and dermatomyositis, and increasing experience with its use. Dalakas *et al*. (1993) showed that intravenous gammaglobulin was significantly better than placebo in a double-blind, cross-over trial in 15 adults with

dermatomyositis. There was major improvement in 9 of 12 with intravenous gammaglobulin (0 of 11 with placebo) and none worsened (versus 5 of 11 with placebo). The rash improved in most. There are a number of reports of success in polymyositis as well (Cherin *et al*. 1991).

A major advantage of intravenous gammaglobulin is its low risk. Dalakas (1994b) recommends testing for IgA deficiency to avoid anaphylaxis, for high viscosity due to its significant rise with intravenous gammaglobulin, and for renal impairment. Side-effects are similar to those in other conditions, and usually controllable. Although transmission of hepatitis C has occurred, current risk is felt to be low.

The monthly dose of 2 g/kg per day is usually given over 2 to 5 days. The effectiveness of an individual treatment appears to be of limited duration (usually 4 to 6 weeks). Prolonged monthly courses may have continuing effect; Cherin *et al*. (1991) found significant improvement in 15 of 20 patients, with minimal side-effects, after nine treatments. Whether prolonged courses can have long-term benefit after cessation is not yet clear, although exacerbations may occur, and some have suggested tachyphylaxis (Reimold and Weinblatt 1994). Tapering and maintenance regimens are empirical. One study found lesser efficacy as a first-line therapy than as treatment for refractory disease (Cherin *et al*. 1994). The major impediment to more widespread use is the very high cost.

Dalakas *et al*. (1993) demonstrated by repeat biopsies that capillary density increased, MAC deposition was prevented, and expression of ICAM-1 on endothelial cells and MHC-1 on muscle fibres was normalized by intravenous gammaglobulin. Thus, this treatment acts on early pathogenetic mechanisms, not simply on symptoms. Basta and Dalakas (1994) showed that intravenous gammaglobulin blocked deposition of activated C3 fragments *in vitro*, suggesting that it works in dermatomyositis by binding fragments, protecting muscle capillaries from complement-mediated injury. Other mechanisms are likely to be operative in polymyositis. If persistent infection is a factor, this could be a target for intravenous gammaglobulin, but effects on the immune system are more likely (Dwyer 1992).

Experimental

Case reports and collected series have described improvement in polymyositis/dermatomyositis with plasmapheresis, including 32 of 35 patients of Dau (1981). Most reported patients have received cyclophosphamide or chlorambucil, so that the independent contribution of plasmapheresis was unclear. In a controlled, double-blind trial, Miller *et al*. (1992) found no improvement in strength or functional outcome with plasmapheresis or leucapheresis (12 treatments in 1 month). Prednisone but no immunosuppressives were used. Creatine kinase fell from direct removal, without correlation with improvement. Subgroups of patients might still benefit from these procedures, or they may enhance the benefit of other treatments.

Total body irradiation has resulted in dramatic, prolonged responses in several cases of severe, life-threatening polymyositis/dermatomyositis unresponsive to other treatments, often with minimal ill effects (Dalakas and Engel 1988). However, some have not responded (Cherin *et al*. 1992), and at least one patient died. Late malignancy is a concern. Total nodal irradiation has also been tried. Irradiation is a treatment of last resort. Successes have also be reported with thymectomy, another rarely-used last resort (Cumming 1989).

Treatment of extra-muscle manifestations

Dermatomyositis rash

The dermatomyositis rash may respond to treatment of the myositis with steroids, immunosuppressives, or intravenous gammaglobulin. If dermatomyositis lesions persist, hydroxychloroquine at a dose of 200 to 400 mg/d may be helpful (Woo *et al*. 1984). It would not be expected to help the myositis. The rash may be photosensitive, and sunscreens are recommended. Topical steroids are commonly used, but are often unsuccessful. Oddis (1994) had success with isotretinoin (0.5 to 1.0 mg/kg per day) for rash unresponsive to other measures, with caution to avoid it in pregnancy.

Treatment of amyopathic dermatomyositis is controversial. Sunscreens, topical therapy, and hydroxychloroquine are tried first. In some very severe cases refractory to these measures, steroids or immunosuppressives are justified for the cutaneous manifestations (Euwer and Sontheimer 1994). Zieglschmid-Adams *et al*. (1995), noted improvement of the skin lesions with methotrexate (average maximal dose 20 mg/week orally) in two of three patients with amyopathic dermatomyositis, and all seven with dermatomyositis; myositis improved in only four. When such treatment is withheld because of the absence of detectable myositis, the patient must be followed closely, especially in the first 2 years, to avoid delay in treatment if myositis appears.

Calcinosis

Calcinosis is more of a problem in juvenile dermatomyositis but may occasionally occur in the adults. Several treatments may be tried (Boyd and Neldner 1994), but none is of consistent or proven benefit. Surgical excision can be helpful for accessible deposits that cause problems. An inflammatory component may respond to colchicine. Treatment of the disease helps prevent calcinosis, but it does not affect established calcinosis.

Interstitial lung disease

Treatment for interstitial lung disease in polymyositis/dermatomyositis should be considered for clinically significant active and progressive disease. Interstitial lung disease may sometimes be more amenable to treatment, in myositis than in other conditions but can be unresponsive and contribute to mortality. Treatment is more effective when the biopsy shows active inflammation, or possibly bronchiolitis obliterans organizing pneumonia (Hsue *et al*. 1993). Prednisone is usually given initially, but cytotoxic agents are often needed. Cyclophosphamide has been successful in case reports (Al-Janadi *et al*. 1989). Although conventionally this is used orally for interstitial lung disease, intravenous pulse has also been used.

Gastrointestinal

Dysphagia resulting from pharyngeal weakness usually responds to treatment of the myositis, but may be resistent in some cases. In persistent dysphagia, cricopharyngeal muscle dysfunction should be considered, and cricopharyngeal myotomy may be needed (Kagen *et al*. 1985). Distal oesophageal dysmotility will generally not respond to immunosuppression, but measures similar to those used in oesophageal reflux can be helpful.

HIV myositis

Treatment of HIV-associated myositis is complicated by the increased risk of using steroids and immunosuppressives. Antiretroviral therapy may help in some patients not already receiving it, but the risk of zidovudine myopathy must be considered. For patients shown to have inflammatory myopathy by biopsy, who clinically require treatment, Dalakas (1993) recommends intravenous gammaglobulin, but steroids, daily or alternate-day, have been used despite the risk. In patients on zidovudine, especially if biopsy is unavailable for diagnosis, reduction in zidovudine dosage might be considered.

Rehabilitation

Active exercise is discouraged in the acute period when muscle inflammation is prominent. Passive range of motion exercises may be started early to help prevent joint contractures. When inflammation subsides, active exercise, cautiously introduced and slowly advanced, may help recover lost strength from disease, immobility, and steroid-related atrophy (Hicks 1988). It may help to start with isometric exercises (Oddis 1994). Active resistive exercise can be considered cautiously in stable patients even if some disease activity remains (Escalante *et al*. 1993). The patient may also require assistance with activities of daily living and assistive devices while severely disabled, and possibly vocational rehabilitation if permanent disability develops.

Prognosis

Mortality

Current survival rates for polymyositis/dermatomyositis are higher than those prior to the availability of corticosteroids, and continued improvement is apparent after their introduction. Medsger *et al*. (1971) found a 5-year survival rate of 65 per cent, and a 7-year rate of 53 per cent between 1947 and 1968, while Hochberg *et al*. (1986) found a 5-year survival rate of 80.4 per cent and an 8-year rate of 72.8 per cent between 1970 and 1981. Love *et al*. (1991) also found a low mortality rate of 14 per cent (36 per cent with cancer). Many factors contribute to this improvement, such as better use of medications and better general medical care. Also, better

recognition of polymyositis/dermatomyositis may lead to inclusion of milder cases or exclusion of other diseases.

The most common causes of mortality in polymyositis/dermatomyositis are malignancy, infection, and cardiac and lung manifestations. Factors which increase mortality include older age of onset, cardiac involvement, or interstitial lung disease (Arsura and Greenberg 1988). Dysphagia has been associated with increased mortality in some studies, possibly due to aspiration, but probably because it is found in more severely affected cases (Hochberg *et al*. 1986). Some (but not all) studies find decreased survival in myositis with malignancy, such as Bohan *et al*. (1977) (46 per cent compared with 87 to 91 per cent for polymyositis/dermatomyositis). Other possible factors suggested to increase mortality include: more severe or rapid onset, resistance to treatment, or long delay before treatment. Other measures of disease severity, such as the findings of laboratory tests or biopsy, have not been found to correlate with poor prognosis. Lilley *et al*. (1994) found that most deaths occurred within the first 6 months, and that those treated with immunosuppressives seemed to do better.

Clinical subgroups of patients with polymyositis/dermatomyositis excluding malignancy do not differ significantly in prognosis for survival. However, Love *et al*. (1991) found that those with antisynthetases and those with anti-SRP had higher mortality rates than myositis overall (21 per cent and 43 per cent versus 7 per cent).

Recovery of strength

Prognosis for recovery of full strength is worse if treatment is delayed (Joffe *et al*. 1993; Fafalak *et al*. 1994), or if the course is chronic and progressive (possibly due to delayed treatment). Patients whose treatment begins more than 6 to 12 months after onset may not recover full strength even with complete suppression of disease activity (Tymms and Webb 1985). Older patients may respond less well, and tend to require a longer duration of treatment (McKendry 1987). In general, those with adult dermatomyositis do better than those with pure polymyositis, and those with autoimmune rheumatic disease overlap syndromes do best (Lundberg *et al*. 1992; Joffe *et al*. 1993). Patients with antisynthetases usually respond initially, but patients with either antisynthetase or anti-SRP have a higher frequency of incomplete response and flare with taper (Love *et al*. 1991).

Inclusion-body myositis

Inclusion-body myositis is a distinct clinicopathological entity, comprising 15 to 28 per cent of idiopathic myositis, the most common form other than polymyositis/dermatomyositis (Dalakas 1991), although it may be overrepresented in referral centres as a result of treatment resistance. It usually begins after age 50, and is the most common muscle disease beginning in this group (Askanas *et al*. 1994b). It is two- to threefold more common in males. It can be difficult to distinguish from polymyositis, and is often misdiagnosed at first (Calabrese *et al*. 1987); this, plus insidious onset, lead to a time to diagnosis of 3 to 6 years.

Clinical picture

The onset is insidious and progression is slow. Although most patients with inclusion-body myositis have proximal muscle weakness, distal weakness is common, sometimes greater than proximal. Distal weakness occurs in 50 to 95 per cent of patients at some point (Lotz *et al*. 1989; Love *et al*. 1991). It more often involves lower extremities before upper, and spares the face. It tends to be symmetric, but Love *et al*. noted asymmetry significantly more often (61 per cent) than in polymyositis/dermatomyositis. Falling is a common problem (96 per cent, versus 1 to 17 per cent in polymyositis/dermatomyositis) possibly related to the severe quadriceps weakness or associated neuropathy. The knee reflex is reduced or absent within 5 years in most patients, probably from similar factors. Forearm thinning is notable, with resultant finger weakness (Askanas *et al*. 1994b). Distal weakness and depressed reflexes suggest a neuropathy, and a portion of cases show other neuropathic signs (Lindberg *et al*. 1994b). Dysphagia may be seen in 40 per cent of cases, occasionally as the presenting manifestation (Riminton *et al*. 1993), and myalgia in 20 per cent.

Laboratory

The creatine kinase is usually elevated but typically less than 10-fold normal. The electromyography picture usually shows myopathy similar to that of polymyositis, with myopathic potentials and fibrillations, but some show a higher than expected frequency of neuropathic (polyphasic, high-amplitude) potentials, and a portion may have abnormal nerve conduction, consistent with a neuropathic component (Chou 1993). MRI typically shows focal, predominantly anterior thigh involvement (Fraser *et al*. 1991) as observed clinically. Biopsy is required to make the diagnosis, but it is frequently missed, and repeat biopsy may be required. Criteria for the diagnosis have been proposed (Table 11) (Lotz *et al*. 1989; Calabrese and Chou 1994).

Pathological criteria
Electron microscopy:
1. Microtubular filaments
Light microscopy:
2. Lined vacuoles

Clinical criteria
1. Insidious onset of proximal muscle weakness
2. Distal muscle weakness
3. Electromyographic evidence of generalized inflammatory myopathy, with/without neuropathic features
4. Elevation of creatine kinase and/or aldolase
5. Failure of muscle weakness to substantially improve on a corticosteroid regimen of at least prednisone 40–60 mg/d

Exclusions
1. History of hereditary muscle disease[b]

Definite IBM: One pathological criterion, plus any two clinical criteria 1–5
Probable IBM: One pathological criterion, plus any one clinical criterion
[a]Reference: Calabrese and Chou (1994)
[b]Family history can support a diagnosis of hereditary inclusion-body myopathy

Table 11 Diagnostic criteria for inclusion-body myositis (IBM) [a]

Pathology

The distinctive features of inclusion-body myositis on light microscopy are vacuoles rimmed by basophilic material, and small, refractile, eosinophilic cytoplasmic and nuclear inclusions. EM shows that the vacuoles are filled with cytoplasmic degradation products including myeloid structures, membrane fragments, and debris. Cytoplasmic inclusions, usually found near vacuoles, are composed of distinctive filaments of 15 to 21 nm. Askanas *et al*. (1994a) found that they were actually paired helical filaments which resembled those of Alzheimer's, and contained hyperphosphorylated tau protein and ubiquitin. Smaller inclusions are found in the nucleus. Chou (1993) suggests that the microtubular filaments are formed in the nucleus, causing the nuclei to swell and rupture, releasing the contents. Injured nuclei become surrounded by vacuoles and multilamellar myelin bodies.

The vacuoles and inclusions are not considered specific for these conditions, and the diagnosis must be based on a combination of findings (Calabrese and Chou 1994). Similar structures are found in an autosomal-recessive familial distal myopathy (Cole *et al*. 1988) and oculopharyngeal muscular dystrophy, indicating a relationship of these conditions to inclusion-body myositis.

Inclusion-body myositis usually shows typical features of inflammatory myopathy, with endomysial inflammatory infiltrates, necrosis and regeneration. In some patients (about 10 per cent), inflammation is very sparse or not evident, despite the presence of inclusions and vacuoles (Figarella-Branger *et al*. 1990). The inflammatory features are more prominent in the early stages, and the vacuoles are more abundant later (Askanas *et al*. 1994b), possibly accounting for some initial misdiagnosis; polymyositis is not generally felt to 'evolve' into inclusion-body myositis. Grouped atrophic fibres and hypertrophied fibres are seen, reminiscent of neuropathy and usually not seen in polymyositis/dermatomyositis.

Aetiology

The aetiology is unknown, but as with polymyositis/dermatomyositis, viruses have been suspected. Although adenovirus was isolated from muscle in one case of inclusion-body myositis, viral cultures are usually negative (Chou 1993). The inclusions of inclusion-body myositis resemble viral structures, especially the nucleocapsid proteins of paramyxovirus (Chou 1986). Although Chou found specific reaction of inclusion-body myositis inclusions with mumps virus antisera, the presence of mumps virus proteins was not confirmed by immunological techniques, *in situ* hybridization (Nishino *et al*. 1989), or PCR (Leff *et al*. 1992; Calabrese and

Chou 1994).

The origin of the vacuoles and inclusions is unknown. In addition to the similar paired helical filaments noted, inclusion-body myositis shares with Alzheimer's the accumulation of several proteins, in structures of similar appearance, including Congo-red positive amyloid deposits (Mendell *et al*. 1991), and b-amyloid protein colocalized with ubiquitin (Askanas and Engel 1993; Askanas *et al*. 1993a). Accumulation of prion protein is also seen (Askanas *et al*. 1993b). These findings suggest the possibility of non-autoimmune mechanisms in inclusion-body myositis, an attractive hypothesis in view of the limited response to immunosuppression and the lack of inflammation in some cases.

Unlike polymyositis/dermatomyositis, a familial form of inclusion-body myositis may occur in a small proportion of cases, usually without inflammation, with younger average onset, referred to as hereditary inclusion-body myopathy. Cases of inclusion-body myositis with malignancy have been reported (Ytterberg *et al*. 1993), but an increased frequency of malignancy has not yet been demonstrated. A small percentage of patients have an association with autoimmune rheumatic diseases such as systemic lupus erythematosus, scleroderma, Sjögren's syndrome, or sarcoidosis, and even dermatomyositis. In the absence of such associated conditions, features of autoimmune rheumatic diseases (arthritis, etc.) are usually not found.

Garlepp *et al*. (1994) found a significant increase of DR3 (92 per cent) in 13 patients with inclusion-body myositis, including 10 with the B8/DR3/C4A*QO extended haplotype, and 2 others with the B18/C4B*QO haplotype. Love *et al*. (1991) found no significant association with DR3.

Pathogenesis

The strongest evidence of autoimmunity is the finding of non-necrotic fibres surrounded and invaded by lymphocytes (see above), and the analysis of infiltrating lymphocytes suggesting T-cell-mediated immune attack on muscle fibres. However, the possibility that this response is secondary has not been excluded. O'Hanlon *et al*. (1994b) found prominence of T-cell receptors with Vb3 and Vb6, but in most patients these were polyclonal, suggesting possible superantigen stimulation rather than the antigen-driven response seen in polymyositis. Muscle fibres in inclusion-body myositis usually show expression of MHC-1 antigen, consistent with the cytotoxic T-cell response, but it is not seen in non-inflammatory biopsies (Figarella-Branger *et al*. 1990). There is little evidence of significant humoral autoimmunity in inclusion-body myositis; autoantibodies are uncommon (23 per cent) and MSAs are not seen (Love *et al*. 1991).

The reason some patients' biopsies do not show inflammation is not known; sampling error or variation in activity over time may be factors, but inflammation may be truly absent in such patients. This is supported by the hereditary group of inclusion-body myositis, which seldom show inflammation.

Treatment

Patients with inclusion-body myositis do not respond to treatment as well as those with polymyositis. They are less likely to show any response, and those that do will not return to normal. This may result partially from the long delay between onset and diagnosis, since delayed treatment is associated with limited response. The commonly observed atrophy and fibrosis support this. Alternatively, the inflammatory process might be secondary, with a degenerative process being primary. Response is more likely when autoimmune features are seen, inflammation is more prominent by biopsy, creatine kinase is higher, or when extensive atrophy has not yet occurred (Leff *et al*. 1993; Calabrese and Chou 1994).

Recent studies indicate that treatment may slow or stop the inexorable deterioration, and a small proportion of patients may improve (Sayers *et al*. 1992; Leff *et al*. 1993). A trial of therapy is generally worthwhile, but treatment goals and expectations must be different than for polymyositis/dermatomyositis. Patients may receive extensive treatment inappropriately if this is not recognized.

Some patients benefit from prednisone (Sayers *et al*. 1992; Leff *et al*. 1993), although often temporarily and usually incompletely. However, those treated with more aggressive therapy appear to do better. The few patients of Sayers *et al*. with long-term improvement received methotrexate with prednisone. Leff *et al*. obtained stabilization or, less often, improvement with a combination of methotrexate and azathioprine or high-dose methotrexate with leukovorin rescue. Intravenous gammaglobulin produced improvement in a few reported cases (Soueidan and Dalakas 1993), but it was unsuccessful in others (Askanas *et al*. 1994b), and its role remains unclear.

Treatment usually begins with steroids, but immunosuppressives are commonly added. Improvement in creatine kinase does not necessarily predict a clinical response. Calabrese and Chou (1994) recommend starting with prednisone at 40 to 60 mg for 3 months, then adding methotrexate or azathioprine for steroid-sparing effect if there is clinical improvement, or for additional immunosuppression if warranted by clinical predictors. They would otherwise stop treatment.

Dysphagia may become a severe problem. As with polymyositis/dermatomyositis, some cases relate to cricopharyngeal dysfunction, and respond to myotomy.

Prognosis

Over time, weakness will progress in most patients (Lotz *et al*. 1989), with gradually increasing disability, but progression is very slow. Patients often need assistance with usual activities within 10 years, and may be wheelchair bound within 15, although more prolonged courses may be seen.

Chapter References

Ahlstrom, G., Gunnarsson, L.G., Leissner, P., and Sjoden, P.O. (1993). Epidemiology of neuromuscular diseases, including the postpolio sequelae, in a Swedish county. *Neuroepidemiology*, 1**2**, 262–9.

Al-Janadi, M., Smith, C.D., and Karsh, J. (1989). Cyclophosphamide treatment of interstitial pulmonary fibrosis in polymyositis/dermatomyositis. *Journal of Rheumatology*, **16**, 1592–6.

Annesley, T.M., Strongwater, S.L., and Schnitzer, T.J. (1985). MM subisoenzymes of creatine kinase as an index of disease activity in polymyositis. *Clinical Chemistry*, **31**, 402–6.

Arad-Dann, H., Isenberg, D., Ovadia, E., Shoenfeld, Y., Sperling, J., and Sperling, R. (1989). Autoantibodies against a nuclear 56-kDa protein: a marker for inflammatory muscle disease. *Journal of Autoimmunity*, **2**, 877–88.

Arahata, K. and Engel, A.G. (1984). Monoclonal antibody analysis of mononuclear cells in myopathies. I. Quantitation of subsets according to diagnosis and sites of accumulation and demonstration and counts of muscle fibers invaded by T cells. *Annals of Neurology*, **16**, 193–208.

Arahata, K. and Engel, A.G. (1986). Monoclonal antibody analysis of mononuclear cells in myopathies. III: immunoelectron microscopy aspects of cell-mediated muscle fiber injury. *Annals of Neurology*, **19**, 112–25.

Arahata, K. and Engel, A.G. (1988*a*). Monoclonal antibody analysis of mononuclear cells in myopathies V: identification and quantitation of T8+ cytotoxic and T8+ suppressor cells. *Annals of Neurology*, **23**, 493–9.

Arahata, K. and Engel, A.G. (1988*b*). Monoclonal antibody analysis of mononuclear cells in myopathies. IV: cell-mediated cytotoxicity and muscle fiber necrosis. *Annals of Neurology*, **23**,168–73.

Arsura, E.L. and Greenberg, A.S. (1988). Adverse impact of interstitial pulmonary fibrosis on prognosis in polymyositis and dermatomyositis. *Seminars in Arthritis and Rheumatism*, **18**, 29–37.

Askanas, V. and Engel, W.K. (1993). New advances in inclusion-body myositis. *Current Opinion in Rheumatology*, **5**, 732–41.

Askanas, V., Alvarez, R.B., and Engel, W.K. (1993 *a*). Beta-amyloid precursor epitopes in muscle fibers of inclusion body myositis. *Annals of Neurology*, **34**, 551–60.

Askanas, V., Bilak, M., Engel, W.K., Alvarez, R.B., Tome, F., and Leclerc, A. (1993 *b*). Prion protein is abnormally accumulated in inclusion-body myositis. *Neuroreport*, **5**, 25–8.

Askanas, V., Engel, W.K., Bilak, M., Alvarez, R.B., and Selkoe, D.J. (1994 *a*). Twisted tubulofilaments of inclusion body myositis muscle resemble paired helical filaments of Alzheimer brain and contain hyperphosphorylated tau. *American Journal of Pathology*, **144**, 177–87.

Askanas, V., Engel, W.K., and Mirabella, M. (1994*b*). Idiopathic inflammatory myopathies: inclusion-body myositis, polymyositis, and dermatomyositis. *Current Opinion in Neurology*, **7**, 448–56.

Barkhaus, P.E., Nandedkar, S.D., and Sanders, D.B. (1990). Quantitative EMG in inflammatory myopathy. *Muscle and Nerve*, **13**, 247–53.

Basta, M. and Dalakas, M.C. (1994). High-dose intravenous immunoglobulin exerts its beneficial effect in patients with dermatomyositis by blocking endomysial deposition of activated complement fragments. *Journal of Clinical Investigation*, **94**, 1729–35.

Behan, W.M. and Behan, P.O. (1993). The role of viral infection in polymyositis, dermatomyositis and chronic fatigue syndrome. *Baillière's Clinical Neurology*, **2**, 637–57.

Behan, W.M.H., Behan, P.O., Draper, I.T., and Williams, H. (1983). Does *Toxoplasma* cause polymyositis. *Acta Neuropathologica*, **61**, 246–52.

Bender, A., Ernst, N., Iglesias, A., Dornmair, K., Wekerle, H., and Hohlfeld, R. (1995). T cell receptor repertoire in polymyositis: clonal expansion of autoaggressive CD8+ T cells. *Journal of Experimental Medicine*, **181**, 1863–8.

Bernard, P. and Bonnetblanc, J.M. (1993). Dermatomyositis and malignancy. *Journal of Investigative Dermatology*, **100**, 128S–132S.

Bertorini, T.E. (1988*a*). Electromyography in polymyositis and dermatomyositis (PM/DM). In *Polymyositis and dermatomyositis* (ed. M.C. Dalakas), pp. 217–34. Butterworths, Boston.

Bertorini, T.E. (1988*b*). Histopathology of the inflammatory myopathies. In *Polymyositis and dermatomyositis* (ed. M.C. Dalakas), pp. 157–94. Butterworths, Boston.

Black, H.R., Quallich, H., and Gareleck, C.B. (1986). Racial differences in serum creatine kinase levels. *American Journal of Medicine*, **81**, 479–87.

Bohan, A. and Peter, J.B. (1975). Polymyositis and dermatomyositis. Parts 1 and 2. *New England Journal of Medicine*, **292**, 344–7, 403–7.

Bohan, A., Peter, J.B., Bowman, R.L., and Pearson, C.M. (1977). A computer-assisted analysis of 153 patients with polymyositis and dermatomyositis. *Medicine*, **56**, 255–86.

Bohlmeyer, T.J., Wu, A.H.B., and Perryman, M.B. (1994). Evaluation of laboratory tests as a guide to diagnosis and therapy of myositis. *Rheumatic Disease Clinics of North America*, **20**, 845–56.

Bombardieri, S., Hughes, G.R., Neri, R., Del Bravo, P., and Del Bono, L. (1989). Cyclophosphamide in severe polymyositis [letter]. *Lancet*, **i**, 1138–9.

Bonnetblanc, J.M., Bernard, P., and Fayol, J. (1990). Dermatomyositis and malignancy: a multicenter cooperative study. *Dermatologica*, **180**, 212–16.

Bowles, N.E. *et al*. (1993). Persistence of enterovirus RNA in muscle biopsy samples suggests that some cases of chronic fatigue syndrome result from a previous, inflammatory viral myopathy. *Journal of Medicine*, **24**, 145–60.

Boyd, A.S. and Neldner, K.H. (1994). Therapeutic options in dermatomyositis/polymyositis. *International Journal of Dermatology*, **33**, 240–50.

Braun, N.M., Arora, N.S., and Rochester, D.F. (1983). Respiratory muscle and pulmonary function in polymyositis and other proximal myopathies. *Thorax*, **38**, 616–23.

Bretagne, S., Costa, J.M., Cosnes, A., Authier, F.J., Vidaud, M., and Gherardi, R.K. (1994). Lack of *Toxoplasma gondii* DNA in muscles of patients with inflammatory myopathy and increased anti-*Toxoplasma* antibodies. *Muscle and Nerve*, **17**, 822–4.

Bunch, T.W. (1981). Prednisone and azathioprine for polymyositis: Long-term followup. *Arthritis and Rheuatism*, **24**, 45–48.

Byrne, E. and Dennett, X. (1993). Idiopathic inflammatory myopathies: clinical aspects. *Baillière's Clinical Neurology*, **2**, 499–526.

Calabrese, L.H., and Chou, S.M. (1994). Inclusion body myositis. *Rheumatic Disease Clinics of North America*, **20**, 955–72.

Calabrese, L.H., Mitsumoto, H., and Chou, S.M. (1987). Inclusion body myositis presenting as treatment-resistant polymyositis. *Arthritis and Rheumatism*, **30**, 397–403.

Callen, J.P. (1987). Dermatomyositis. *Disease-a-Month*, **33**, 237–305.

Callen, J.P. (1993). Dermatomyositis and malignancy. *Clinics in Dermatology*, **11**, 61–5.

Callen, J.P. (1994). Myositis and malignancy. *Current Opinion in Rheumatology*, **6**, 590–94.

Caro, I. (1988). A dermatologist's view of polymyositis/dermatomyositis. *Clinics in Dermatology*, **6**, 9–14.

Carpenter, S. (1988). Resin histology and electron microscopy in inflammatory myopathies. In *Polymyositis and dermatomyositis* (ed. M.C. Dalakas), pp. 195–215. Butterworths, Boston.

Casademont, J., Grau, J.M., Masanes, F., Herrero, C., and Urbano-Marquez, A. (1993). Analysis of the outcome of idiopathic inflammatory myopathies with particular emphasis on muscle capillary damage. *Scandanavian Journal of Rheumatology*, **22**, 292–8.

Cervera, R. *et al*. (1991). Antibodies to endothelial cells in dermatomyositis: Association with interstitial lung disease. *British Medical Journal*, **302**, 880–81.

Charness, M.E., Simon, R.P., and Greenberg, D.A. (1989). Ethanol and the nervous system. *New England Journal of Medicine*, **321**, 442–54.

Cheong, W.-K., Hughes, G.R.V., Norris, P.G., and Hawk, J.L.M. (1994). Cutaneous photosensitivity in dermatomyositis. *British Journal of Dermatology*, **131**, 205–8.

Cherin, P. *et al*. (1991). Efficacy of intravenous gammaglobulin therapy in chronic refractory polymyositis and dermatomyositis: an open study with 20 adult patients. *American Journal of Medicine*, **91**, 162–8.

Cherin, P., Herson, S., Coutellier, A., Bletry, O., and Piette, J.C. (1992). Failure of total body irradiation in polymyositis: report of three cases. *British Journal of Rheumatology*, **31**, 282–3.

Cherin, P., Crevon, M.C., Hauw, J.J., Galanaud, P., Herson, S., and Emilie, D. (1993 *a*). Mechanisms of lysis by cytotoxic cells in inflammatory myopathies: study of perforin and serine esterase gene expression by in situ hybridization and protein TIA-1 expression by immunohistochemistry, from muscular biopsies. *Arthritis and Rheumatism*, **36**, S84 (Abstract).

Cherin, P. *et al*. (1993*b*). Dermatomyositis and ovarian cancer: A report of 7 cases and literature review. *Journal of Rheumatology*, **20**, 1897–9.

Cherin, P. *et al*. (1994). Intravenous gamma globulin as first line therapy in polymyositis and dermatomyositis: an open study in 11 adult patients. *Journal of Rheumatology*, **21**, 1092–7.

Chiedozi, L.C. (1979). Pyomyositis: review of 205 cases in 112 patients. *American Journal of Surgery*, **137**, 255–9.

Chou, S.M. (1986). Inclusion body myositis: a chronic persistent mumps myositis? *Human Pathology*, **17**, 765–77.

Chou, S.M. (1993). Inclusion body myositis. *Baillière's Clinical Neurology*, **2**, 557–77.

Christensen, M.L., Pachman, L.M., Schneiderman, R., Patel, D.C., and Friedman, J.M. (1986). Prevalence of coxsackie B virus antibodies in patients with juvenile dermatomyositis. *Arthritis and Rheumatism*, **29**, 1365–70.

Cole, A.J., Kuzniecky, R., Karpati, G., Carpenter, S., Andermann, E., and Andermann, F. (1988). Familial myopathy with changes resembling inclusion body myositis and periventricular leucoencephalopathy. A new syndrome. *Brain*, **111**, 1025–37.

Cox, N.H., Lawrence, C.M., Langtry, J.A.A., and Ive, F.A. (1990). Dermatomyositis: disease associations and an evaluation of screening investigations for malignancy. *Archives of Dermatology*, **126**, 61–5.

Craft, J., Mimori, T., Olsen, T.L., and Hardin, J.A. (1988). The U2 small nuclear ribonucleoprotein particle as an autoantigen. Analysis with sera from patients with overlap syndromes. *Journal of Clinical Investigation*, **81**, 1716–24.

Cronin, M.E. and Plotz, P.H. (1990). Idiopathic inflammatory myopathies. *Rheumatic Disease Clinics of North America*, **16**, 655–65.

Cronin, M.E., Love, L.A., Miller, F.W., McClintock, P.R., and Plotz, P.H. (1988). The natural history of encephalomyocarditis virus-induced myositis and myocarditis in mice. Viral persistence demonstrated by in situ hybridization. *Journal of Experimental Medicine*, **168**, 1639–48.

Cronin, M.E., Miller, F.W., Hicks, J.E., Dalakas, M.C., and Plotz, P.H. (1989). The failure of intravenous cyclophosphamide therapy in refractory idiopathic inflammatory myopathy. *Journal of Rheumatology*, **16**, 1225–8.

Csuka, M.E. and McCarty, D.J. (1985). Simple method for measurement of lower extremity muscle strength. *American Journal of Medicine*, **78**, 77–81.

Cukier, J., Beauchamp, R.A., Spindler, J.S., Spindler, S., Lorenzo, C., and Trentham, D.E. (1993). Association between bovine collagen dermal implants and a dermatomyositis or a polymyositis-like syndrome. *Annals of Internal Medicine*, **118**, 920–28.

Cumming, W.J.K. (1989). Thymectomy in refractory dermatomyositis. *Muscle and Nerve*, **12**, 424.

Dalakas, M., Gravell, M., London, W.T., Cunningham, G., and Sever, J.L. (1987). Morphological changes of an inflammatory myopathy in Rhesus monkeys with simian acquired immunodeficiency syndrome (42556). *Proceedings of the Society of Experimental Biology and Medicine*, **185**, 368–76.

Dalakas, M.C. (1988*a*). A classification of polymyositis and dermatomyositis. In *Polymyositis and dermatomyositis* (ed. M.C. Dalakas), pp. 1–16. Butterworths, Boston.

Dalakas, M.C. (1988*b*). Treatment of polymyositis and dermatomyositis with corticosteroids: a first therapeutic approach. In *Polymyositis and dermatomyositis* (ed. M.C. Dalakas), pp. 235–53. Butterworths, Boston.

Dalakas, M.C. (1991). Polymyositis, dermatomyositis, and inclusion-body myositis. *New England Journal of Medicine*, **325**, 1487–98.

Dalakas, M.C. (1992*a*). Inflammatory and toxic myopathies. *Current Opinion in Neurology and Neurosurgery*, **5**, 645–54.

Dalakas, M.C. (1992*b*). Clinical, immunopathologic, and therapeutic considerations of inflammatory myopathies. *Clinics in Neuropharmacology*, **15**, 327–51.

Dalakas, M.C. (1993). Retroviruses and inflammatory myopathies in humans and primates. *Baillière's Clinical Neurology*, **2**, 659–691.

Dalakas, M.C. (1994*a*). Current treatment of the inflammatory myopathies. *Current Opinions in Rheumatology*, **6**, 595–601.

Dalakas, M.C. (1994*b*). How to diagnose and treat the inflammatory myopathies. *Seminars in Neurology*, **14**, 137–45.

Dalakas, M.C. and Engel, W.K. (1988). Total body irradiation in the treatment of intractable polymyositis and dermatomyositis. In *Polymyositis and dermatomyositis* (ed. M.C. Dalakas), pp. 281–91. Butterworths, Boston.

Dalakas, M.C., Illa, I., Pezeshkpour, G.H., Laukaitis, J.P., Cohen, B., and Griffin, J.L. (1990). Mitochondrial myopathy caused by long-term zidovudine therapy. *New England Journal of Medicine*, **322**, 1098–105.

Dalakas, M.C. *et al*. (1993). A controlled trial of high-dose intravenous immune globulin infusions as treatment for dermatomyositis. *New England Journal of Medicine*, **329**, 1993–2000.

Dau, P.C. (1981). Plasmapheresis in idiopathic inflammatory myopathy: experience with 35 patients. *Archives of Neurology*, **38**, 544–52.

De Bleecker, J.L. and Engel, A.G. (1994). Expression of cell adhesion molecules in inflammatory myopathies and Duchenne dystrophy. *Journal of Neuropathology and Experimental Neurology*, **53**, 369–76.

De Vita, S. and Fossaluzza, V. (1992). Treatment of idiopathic inflammatory myopathies with cyclophosphamide pulses: clinical experience and a review of the literature. *Acta Neurologica Belgica*, **92**, 215–27.

DeGeeter, F., Deleu, D., Debeuckelaere, S., DeConinck, A., Somers, G., and Bossuyt, A. (1989). Detection of muscle necrosis in dermatomyositis by ^{111}In-labelled antimyosin Fab fragments. *Nuclear Medicine Communications*, **10**, 603–7.

DeMerieux, P., Verity, M.A., Clements, P.J., and Paulus, H.E. (1983). Esophageal abnormalities and dysphagia in polymyositis and dermatomyositis. *Arthritis and Rheumatism*, **26**, 961–8.

Denbow, C.E., Lie, J.T., Tancredi, R.G., and Bunch, T.W. (1979). Cardiac involvement in polymyositis: a clinicopathologic study of 20 autopsied patients. *Arthritis and Rheumatism*, **22**, 1088–92.

Dickey, B.F. and Myers, A.R. (1984). Pulmonary disease in polymyositis/dermatomyositis. *Seminars in Arthritis and Rheumatism*, **14**, 60–76.

Dietz, F., Logeman, J., Sahgal, V., and Schmid, F.R. (1980). Cricopharyngeal muscle dysfunction in the differential diagnosis of dysphagia in polymyositis. *Arthritis and Rheumatism*, **23**, 491–5.

Dropcho, E.J. and Soong, S. (1991). Steroid-induced weakness in patients with primary brain tumors. *Neurology*, **41**, 1235–9.

Dwyer, J.M. (1992). Manipulating the immune system with immune globulin. *New England Journal of Medicine*, **326**, 107–16.

Ehrenstein, M.R., Snaith, M.L., and Isenberg, D.A. (1992). Idiopathic myositis: a rheumatological view. *Annals of the Rheumatic Diseases*, **51**, 41–4.

Emslie-Smith, A.M., and Engel, A.G. (1990). Microvascular changes in early and advanced dermatomyositis: a quantitative study. *Annals of Neurology*, **27**, 343–56.

Emslie-Smith, A.M., Arahata, K., and Engel, A.G. (1989). Major histocompatibility complex class I antigen expression, immunolocalization of interferon subtypes, and T cell-mediated cytotoxicity in myopathies. *Human Pathology*, **20**, 224–31.

Engel, A.G., and Arahata, K. (1984). Monoclonal antibody analysis of mononuclear cells in myopathies. II: phenotypes of autoinvasive cells in polymyositis and inclusion body myositis. *Annals of Neurology*, **16**, 209–15.

Engel, A.G., Arahata, K., and Emslie-Smith, A.M. (1990). Immune effector mechanisms in inflammatory myopathies. *Research Publications—Association for Research in Nervous and Mental Diseases*, **68**, 141–57.

Enz, L.A., Love, L.A., Targoff, I.N., Pandey, J.P., and Miller, F.W. (1992). Associations among Gm phenotypes, HLA alleles, and myositis-specific autoantibodies (MSA) in idiopathic inflammatory myopathy (IIM). *Arthritis and Rheumatism*, **35**, S52 (Abstract).

Escalante, A., Miller, L., and Beardmore, T.D. (1993). Resistive exercise in the rehabilitation of polymyositis/dermatomyositis. *Journal of Rheumatology*, **20**, 1340–44.

Estes, M.L. *et al*. (1987). Chloroquine neuromyotoxicity: clinical and pathologic perspective. *American Journal of Medicine*, **82**, 447–55.

Euwer, R.L. and Sontheimer, R.D. (1993). Amyopathic dermatomyositis: a review. *Journal of Investigative Dermatology*, **100**, 124–7S.

Euwer, R.L. and Sontheimer, R.D. (1994). Dermatologic aspects of myositis. *Current Opinion in Rheumatology*, **6**, 583–9.

Fafalak, R.G., Peterson, M.G., and Kagen, L.J. (1994). Strength in polymyositis and dermatomyositis: best outcome in patients treated early. *Journal of Rheumatology*, **21**, 643–8.

Feutren, G. and Mihatsch, M.J. (1992). Risk factors for cyclosporine-induced nephropathy in patients with autoimmune diseases. *New England Journal of Medicine*, **326**, 1654–60.

Fidzianska, A. and Goebel, H.H. (1989). Tubuloreticular structures (TRS) and cylindric confronting cisternae (CCC) in childhood dermatomyositis. *Acta Neuropathologica*, **79**, 310–16.

Figarella-Branger, D., Pellissier, J.F., Bianco, N., Devictor, B., and Toga, M. (1990). Inflammatory and non-inflammatory inclusion body myositis: Characterization of the mononuclear cells and expression of the immunoreactive class I Major histocompatibility complex product. *Acta Neuropathologica*, **79**, 528–36.

Franks, A.G., Jr. (1988). Important cutaneous markers of dermatomyositis. *Journal of Musculoskeletal Medicine*, **5**, 39–63.

Fraser, D.D., Frank, J.A., Dalakas, M.C., Miller, F.W., Hicks, J.E., and Plotz, P.H. (1991). Magnetic resonance imaging in the idiopathic inflammatory myopathies. *Journal of Rheumatology*, **18**, 1693–700.

Frost, N.A., Morand, E.F., Hall, C.L., Maddison, P.J., and Bhalla, A.K. (1993). Idiopathic polymyositis complicated by arthritis and mesangial proliferative glomerulonephritis: case report and review of the literature. *British Journal of Rheumatology*, **32**, 929–31.

Fudman, E.J. and Schnitzer, T.J. (1986). Dermatomyositis without creatine kinase elevation: a poor prognostic sign. *American Journal of Medicine*, **80**, 329–32.

Fujino, H., Kobayashi, T., Goto, I., and Onitsuka, H. (1991). Magnetic resonance imaging of the muscles in patients with polymyositis and dermatomyositis. *Muscle and Nerve*, **14**, 716–20.

Fusade, T., Belanyi, P., Joly, P., Thomine, E., Mihout, M.F., and Lauret, P. (1993). Subcutaneous changes in dermatomyositis. *British Journal of Dermatology*, **128**, 451–3.

Gabay, C. *et al*. (1994). Elevated serum levels of interleukin-1 receptor antagonist in polymyositis/dermatomyositis. A biologic marker of disease activity with a possible role in the lack of acute-phase protein response. *Arthritis and Rheumatism*, **37**, 1744–51.

Garlepp, M.J. (1993). Immunogenetics of inflammatory myopathies. *Baillière's Clinical Neurology*, **2**, 579–97.

Garlepp, M.J., Baing, B., Zilko, P.J., Ollier, W., and Mastaglia, F.L. (1994). HLA associations with inclusion body myositis. *Clinical and Experimental Immunology*, **98**, 40–45.

Genth, E. *et al*. (1990). Immunogenetic associations of scleroderma-related antinuclear antibodies. *Arthritis and Rheumatism*, **33**, 657–65.

Goldstein, R. *et al*. (1990). HLA-D region genes associated with autoantibody responses to Jo-1 (histidyl-tRNA synthetase) and other translation-related factors in myositis. *Arthritis and Rheumatism*, **33** 1240–48.

Gomez-Reino, J.J., Aznar, J.J., Pablos, J.L., Diaz-Gonzalez, F., and Laffon, A. (1994). Nontropical pyomyositis in adults. *Seminars in Arthritis and Rheumatism*, **23**, 396–405.

Grau, J.M., Herrero, C., Casademont, J., Fernandez-Sola, J., and Urbano-Marquez, A. (1994). Cyclosporine A as first choice therapy for dermatomyositis. *Journal of Rheumatology*, **21**, 381–2.

Greenlee, R. (1988). The neurologist's approach to polymyositis. *Clinics in Dermatology*, **6**, 23–35.

Gurley, R.C., Love, L.A., Targoff, I.N., Leff, R.L., Plotz, P.H., and Miller, F.W. (1991). Associations among myositis-specific autoantibodies (MSA) and HLA-DQA1 alleles. *Arthritis and Rheumatism*, **34**, S137 (Abstract).

Gurley, R.C. *et al*. (1992). Familial myositis cases emphasize the roles of genetic and environmental factors in the pathogenesis of idiopathic inflammatory myopathy (IIM). *Arthritis and Rheumatism*, **35**, S89 (Abstract).

Haddad, M.G., West, R.L., Treadwell, E.L., and Fraser, D.D. (1994). Diagnosis of inflammatory myopathy by percutaneous needle biopsy with demonstration of the focal nature of myositis. *American Journal of Clinical Pathology*, **101**, 661–4.

Hanno, R. and Callen, J.P. (1985). Histopathology of Gottron's papules. *Journal of Cutaneous Pathology*, **12**, 389–94.

Hargis, A.M. and Prieur, D.J. (1988). Animal models of polymyositis/dermatomyositis. *Clinics in Dermatology*, **6**, 120–29.

Haupt, H.M. and Hutchins, G.M. (1982). The heart and cardiac conduction system in polymyositis-dermatomyositis: a clinicopathologic study of 16 autopsied patients. *American Journal of Cardiology*, **50**, 998–1006.

Hebert, C.A., Byrnes, T.J., Baethge, B.A., Wolf, R.E., and Kinasewitz, G.T. (1990). Exercise limitation in patients with polymyositis. *Chest*, **98**, 352–7.

Heffner, R.R., Jr. (1993). Inflammatory myopathies. A review. *Journal of Neuropathology and Experimental Neurology*, **52**, 339–50.

Henriksson, K.G. and Lindvall, B. (1990). Polymyositis and dermatomyositis 1990—diagnosis, treatment and prognosis. *Progress in Neurobiology*, **35**, 181–93.

Hernandez, R.J., Keim, D.R., Chenevert, T.L., Sullivan, D.B., and Aisen, A.M. (1992). Fat-suppressed MR imaging of myositis. *Radiology*, **182**, 217–19.

Hicks, J.E. (1988). Comprehensive rehabilitative management of patients with polymyositis and dermatomyositis. In *Polymyositis and dermatomyositis* (ed. M.C. Dalakas), pp. 293–317. Butterworths, Boston.

Higgs, J.B., Blaivas, M., and Albers, J.W. (1989). McArdle's disease presenting as treatment resistant polymyositis. *Journal of Rheumatology*, **16**, 1588–91.

Higuchi, I. *et al*. (1992). Failure to detect HTLV-1 by in situ hybridization in the biopsied muscles of viral carriers with polymyositis. *Muscle and Nerve*, **15**, 43–7.

Hilton, D.A., Fletcher, A., and Pringle, J.H. (1994). Absence of coxsackie viruses in idiopathic inflammatory muscle disease by *in situ* hybridization. *Neuropathology and Applied Neurobiology*, **20**, 238–42.

Hirakata, M., Mimori, T., Akizuki, M., Craft, J., Hardin, J.A., and Homma, M. (1992). Autoantibodies to small nuclear and cytoplasmic ribonucleoproteins in Japanese patients with inflammatory muscle disease. *Arthritis and Rheumatism*, **35**, 449–56.

Hochberg, M.C., Feldman, D., and Stevens, M.B. (1986). Adult onset polymyositis/dermatomyositis: an analysis of clinical and laboratory features and survival in 76 patients with a review of the literature. *Seminars in Arthritis and Rheumatism*, **15**, 168–78.

Hoffman, G.S., Franck, W.A., Raddatz, D.A., and Stallones, L. (1983). Presentation, treatment, and prognosis of idiopathic inflammatory muscle disease in a rural hospital. *American Journal of Medicine*, **75**, 433–8.

Hohlfeld, R. and Engel, A.G. (1991*a*). Coculture with autologous myotubes of cytotoxic T cells isolated from muscle in inflammatory myopathies. *Annals of Neurology*, **29**, 498–507.

Hohlfeld, R., Engel, A.G., Ii, K., and Harper, M.C. (1991 *b*). Polymyositis mediated by T lymphocytes that express the gamma/delta receptor. *New England Journal of Medicine*, **324**, 877–81.

Hood, D., Van Lente, F., and Estes, M. (1991). Serum enzyme alterations in chronic muscle disease. *American Journal of Clinical Pathology*, **95**, 402–7.

Hopkinson, N.D., Shawe, D.J., and Gumpel, J.M. (1991). Polymyositis, not polymyalgia rheumatica. *Annals of the Rheumatic Diseases*, **50**, 321–2.

Horowitz, M., McNeil, J.D., Maddern, G.J., Collins, P.J., and Shearman, D.J.C. (1986). Abnormalities of gastric and esophageal emptying in polymyositis and dermatomyositis. *Gastroenterology*, **90**, 434–9.

Houpt, K.R. and Sontheimer, R.D. (1994). Autoimmune connective tissue disease and connective tissue disease-like illnesses after silicone gel augmentation mammoplasty. [Review]. *Journal of the American Academy of Dermatology*, **31**, 626–42.

Hsue, Y.T., Paulus, H.E., and Coulson, W.F. (1993). Bronchiolitis obliterans organizing pneumonia in polymyositis. A case report with longterm survival. *Journal of Rheumatology*, **20**, 877–9.

Igarashi, M., Aizawa, H., Tokudome, Y., and Tagami, H. (1985). Dermatomyositis with prominent mucinous skin change. Histochemical and biochemical aspects of glycosaminoglycans. *Dermatologica*, **170**, 6–11.

Illa, I., Nath, A., and Dalakas, M.C. (1991). Immunocytochemical and virological characteristics of HIV-associated inflammatory myopathies: similarities with seronegative polymyositis. *Annals of Neurology*, **29**, 474–81.

Isenberg, D.A. (1983). Immunoglobulin deposition in skeletal muscle in primary muscle disease. *Quarterly Journal of Medicine*, **207**, 297–310.

Isenberg, D.A., Rowe, D., Shearer, M., Novick, D., and Beverley, P.C.L. (1986). Localization of interferons and interleukin 2 in polymyositis and muscular dystrophy. *Clinical and Experimental Immunology*, **63**, 450–58.

Ishii, N., Ono, H., Kawaguchi, T., and Nakajima, H. (1991). Dermatomyositis and pregnancy: case report and review of the literature. *Dermatologica*, **183**, 146–9.

Jablonska, S., Chorzelski, T.P., Blaszczyk, M., Jarzabek-Chorzelska, M., Kumar, V., and Beutner, E.H. (1993). Scleroderma/polymyositis overlap syndromes and their immunologic markers. *Clinics in Dermatology*, **10**, 457–72.

Janis, J.F. and Winkelman, R.K. (1968). Histopathology of the skin in dermatomyositis: a histopathologic study of 55 cases. *Archives of Dermatology*, **97**, 640–50.

Jenkins, E.A., Hull, R.G., and Thomas, A.L. (1993). D-Penicillamine and polymyositis: the significance of the anti-Jo-1 antibody. *British Journal of Rheumatology*, **32**, 1109–10.

Joffe, M.M. *et al*. (1993). Drug therapy of the idiopathic inflammatory myopathies: predictors of response to prednisone, azathioprine, and methotrexate and a comparison of their efficacy. *American Journal of Medicine*, **94**, 379–87.

Johnson, R.L., Fink, C.W., and Ziff, M. (1972). Lymphotoxin formation by lymphocytes and muscle in polymyositis. *Journal of Clinical Investigation*, **51**, 2435–49.

Jongen, P.J., Zoll, G.J., Beaumont, M., Melchers, W.J., van de Putte, L.B., and Galama, J.M. (1993). Polymyositis and dermatomyositis: no persistence of enterovirus or encephalomyocarditis virus RNA in muscle. *Annals of the Rheumatic Diseases*, **52**, 575–8.

Kagen, L.J. and Aram, S. (1987). Creatine kinase activity inhibitor in sera from patients with muscle disease. *Arthritis and Rheumatism*, **30**, 213–17.

Kagen, L.J., Hochman, R.B., and Strong, E.W. (1985). Cricopharyngeal obstruction in inflammatory myopathy (polymyositis/dermatomyositis): report of three cases and review of the literature. *Arthritis and Rheumatism*, **28**, 630–36.

Kalden, J.R. and Gay, S. (1994). Retroviruses and autoimmune rheumatic diseases. *Clinical and Experimental Immunology*, **98**, 1–5.

Kalovidouris, A.E. (1986). Mononuclear cells from patients with polymyositis inhibit calcium binding by sarcoplasmic reticulum. *Journal of Laboratory and Clinical Medicine*, **107**, 23–8.

Kalovidouris, A.E. (1992). The role of cytokines in polymyositis: interferon-gamma induces class II and enhances class I major histocompatibility complex antigen expression on cultured human muscle cells. *Journal of Laboratory and Clinical Medicine*, **120**, 244–51.

Kalovidouris, A.E. (1994). Mechanisms of inflammation and histopathology in inflammatory myopathy. *Rheumatic Disease Clinics of North America*, **20**, 881–98.

Kalovidouris, A.E., Pourmand, R., Passo, M.H., and Plotkin, Z. (1989). Proliferative response of peripheral blood mononuclear cells to autologous and allogeneic muscle in patients with polymyositis/dermatomyositis. *Arthritis and Rheumatism*, **32**, 446–53.

Kalovidouris, A.E., Plotkin, Z., and Graesser, D. (1993). Interferon-gamma inhibits proliferation, differentiation, and creatine kinase activity of cultured human muscle cells. II. A possible role in myositis. *Journal of Rheumatology*, **20**, 1718–23.

Kalovidouris, A.E., Horn, C.A., and Plotkin, Z. (1994). The role of cytokines in polymyositis. III. Recombinant human interferon-gamma enhances T cell adhesion to cultured human muscle cells. *Arthritis and Rheumatism*, **37**, 907–14.

Kamel, O.W. *et al*. (1993). Brief report: reversible lymphomas associated with Epstein–Barr virus occurring during methotrexate therapy for rheumatoid arthritis and dermatomyositis. *New England Journal of Medicine*, **328**, 1317–21.

Kanda, F., Fujii, Y., Takahashi, K., and Fujita, T. (1994). Dual-energy X-ray absorptiometry in neuromuscular diseases. *Muscle and Nerve*, **17**, 431–5.

Karpati, G., Pouliot, Y., and Carpenter, S. (1988). Expression of immunoreactive major histocompatibility complex products in human skeletal muscles. *Annals of Neurology*, **23**, 64–72.

Kasper, C.S., White, C.L., III, and Freeman, R.G. (1988). Pathology and immunopathology of polymyositis/dermatomyositis. *Clinics in Dermatology*, **6**, 64–75.

Kasteler, J.S. and Callen, J.P. (1994). Scalp involvement in dermatomyositis: often overlooked or misdiagnosed. *Journal of the American Medical Association*, **272**, 1939–41.

Kaufman, L.D. (1994). The evolving spectrum of eosinophilia myalgia syndrome. *Rheumatic Disease Clinics of North America*, **20**, 973–94.

Khaleeli, A.A. *et al*. (1983). Corticosteroid myopathy: a clinical and pathological study. *Clinics in Endocrinology*, **18**, 155–66.

Kiprov, D.D. and Miller, R.G. (1984). Polymyositis associated with monoclonal gammopathy. *Lancet*, **ii**, 1183–6.

Kissel, J.T., Mendell, J.R., and Rammohan, K.W. (1986). Microvascular deposition of complement membrane attack complex in dermatomyositis. *New England Journal of Medicine*, **314**, 329–34.

Kissel, J.T., Halterman, R.K., Rammohan, K.W., and Mendel, J.R. (1991). The relationship of complement-mediated microvasculopathy to the histologic features and clinical duration of disease in dermatomyositis. *Archives of Neurology*, **48**, 26–30.

Koga, K., Abe, S., Hashimoto, H., and Yamaguchi, M. (1987). Western-blotting method for detecting antibodies against human muscle contractile proteins in myositis. *Journal of Immunological Methods*, **105**, 15–21.

Koh, E.T., Seow, A., Ong, B., Ratnagopal, P., Tjia, H., and Chng, H.H. (1993). Adult onset polymyositis/dermatomyositis: clinical and laboratory features and treatment response in 75 patients. *Annals of the Rheumatic Diseases*, **52**, 857–61.

Kroll, M., Otis, J., and Kagen, L.J. (1986). Serum enzyme, myoglobin and muscle strength relationships in polymyositis and dermatomyositis. *Journal of Rheumatology*, **13**, 349–55.

Kuncl, R.W., Duncan, G., Watson, D., Alderson, K., Rogawski, M.A., and Peper, M. (1987). Colchicine myopathy and neuropathy. *New England Journal of Medicine*, **316**, 1562–8.

Kuroda, Y., Neshige, R., Oda, K., and Shibasaki, H. (1986). Chronic polymyositis: presence of coxsackievirus A9 antigen in muscle. *Japanese Journal of Medicine*, **25**, 191–4.

Lakhanpal, S., Bunch, T.W., Ilstrup, D.M., and Melton, L.J. (1986). Polymyositis-dermatomyositis and malignant lesions: does an association exist? *Mayo Clinic Proceedings*, **61**, 645–53.

Lakhanpal, S., Lie, J.T., Conn, D.L., and Martin, W.J. (1987). Pulmonary disease in polymyositis/dermatomyositis: a clinicopathological analysis of 65 autopsy cases. *Annals of the Rheumatic Diseases*, **46**, 23–9.

Lane, R.J., McLean, K.A., Moss, J., and Woodrow, D.F. (1993). Myopathy in HIV infection: the role of zidovudine and the significance of tubuloreticular inclusions. *Neuropathology and Applied Neurobiology*, **19**, 406–13.

Le Guludec, D. *et al*. (1993). New application of myocardial antimyosin scintigraphy: diagnosis of myocardial disease in polymyositis. *Annals of the Rheumatic Diseases*, **52**, 235–8.

Le Quintrec, J.-S. and Le Quintrec, J.-L. (1991). Drug-induced myopathies. *Baillière's Clinical Rheumatology*, **5**, 21–38.

Lee, K.N., Csako, G., Bernhardt, P., and Elin, R.J. (1994). Relevance of macro creatine kinase type 1 and type 2 isoenzymes to laboratory and clinical data. *Clinical Chemistry*, **40**, 1278–83.

Leff, R.L. *et al*. (1991). Distinct seasonal patterns in the onset of adult idiopathic inflammatory myopathy in patients with anti-Jo-1 and anti-signal recognition particle autoantibodies. *Arthritis and Rheumatism*, **34**, 1391–6.

Leff, R.L. *et al*. (1992). Viruses in idiopathic inflammatory myopathies: absence of candidate viral genomes in muscle. *Lancet*, **339**, 1192–5.

Leff, R.L., Miller, F.W., Hicks, J.E., Fraser, D.D., and Plotz, P.H. (1993). The treatment of inclusion body myositis: a retrospective review and a randomized, prospective trial of immunosuppressive therapy. *Medicine*, **72**, 225–35.

Leon-Monzon, M. and Dalakas, M.C. (1992). Absence of persistent infection with enteroviruses in muscles of patients with inflammatory myopathies. *Annals of Neurology*, **32**, 219–22.

Leon-Monzon, M., Lamperth, L., and Dalakas, M.C. (1993). Search for HIV proviral DNA and amplified sequences in the muscle biopsies of patients with HIV polymyositis. *Muscle and Nerve*, **16**, 408–13.

Leon-Monzon, M., Illa, I., and Dalakas, M.C. (1994). Polymyositis in patients infected with human T-cell leukemia virus type I: the role of the virus in the cause of the disease. *Annals of Neurology*, **36**, 643–9.

Lilley, H., Dennett, X., and Byrne, E. (1994). Biopsy proven polymyositis in Victoria 1982–1987: analysis of prognostic factors. *Journal of the Royal Society of Medicine*, **87**, 323–6.

Lindberg, C., Oldfors, A., and Tarkowski, A. (1994*a*). Restricted use of T cell receptor V genes in endomysial infiltrates of patients with inflammatory myopathies. *European Journal of Immunology*, **24**, 2659–63.

Lindberg, C., Persson, L.I., Bjorkander, J., and Oldfors, A. (1994*b*). Inclusion body myositis: clinical, morphological, physiological and laboratory findings in 18 cases. *Acta Neurologica Scandinavia*, **89** 123–31.

Lotz, B.P., Engel, A.G., Nishino, H., Stevens, J.C., and Litchy, W.J. (1989). Inclusion body myositis: observations in 40 patients. *Brain*, **112**, 727–47.

Love, L.A. and Miller, F.W. (1993). Noninfectious environmental agents associated with myopathies. *Current Opinions in Rheumatology*, **5**, 712–18.

Love, L.A. *et al*. (1991). A new approach to the classification of idiopathic inflammatory myopathy: myositis-specific autoantibodies define useful homogeneous patient groups. *Medicine*, **70**, 360–74.

Love, L.A. *et al*. (1992). Geographical and seasonal clustering in the onset of idiopathic inflammatory myopathy (IIM) in groups defined by myositis-specific autoantibodies (MSA). *Arthritis and Rheumatism*, **35**, S40 (Abstract).

Love, L.A. *et al*. (1992). Clinical and immunogenetic features of women who develop myositis after silicone implants (MASI). *Arthritis and Rheumatism*, **35**, S46 (Abstract).

Lovece, S. and Kagen, L.J. (1993). Sensitive rapid detection of myoglobin in serum of patients with myopathy by immunoturbidimetric assay. *Journal of Rheumatology*, **20**, 1331–4.

Lundberg, I., Nennesmo, I., and Hedfors, E. (1992). A clinical, serological, and histopathological study of myositis patients with and without anti-RNP antibodies. *Seminars in Arthritis and Rheumatism*, **22**, 127–38.

Lyon, M.G., Bloch, D.A., Hollak, B., and Fries, J.F. (1989). Predisposing factors in polymyositis-dermatomyositis: results of a nationwide survey. *Journal of Rheumatology*, **16**, 1218–24.

Magid, S.K. and Kagen, L.J. (1983). Serologic evidence for acute toxoplasmosis in polymyositis-dermatomyositis. *American Journal of Medicine*, **75**, 313–20.

Manchul, L.A. *et al*. (1985). The frequency of malignant neoplasms in patients with polymyositis-dermatomyositis: a controlled study. *Archives of Internal Medicine*, **145**, 1835–9.

Manta, P., Kalfakis, N., and Vassilopoulos, D. (1989). Evidence for seasonal variation in polymyositis. *Neuroepidemiology*, **8**, 262–5.

Mantegazza, R. *et al*. (1993). Analysis of T cell receptor repertoire of muscle-infiltrating T lymphocytes in polymyositis. Restricted V alpha/beta rearrangements may indicate antigen-driven selection. *Journal of Clinical Investigation*, **91**, 2880–86.

Marguerie, C. *et al*. (1990). Polymyositis, pulmonary fibrosis and autoantibodies to aminoacyl-tRNA synthetase enzymes. *Quarterly Journal of Medicine*, **77**, 1019–38.

Marguerie, C. *et al*. (1992). The clinical and immunogenetic features of patients with autoantibodies to the nucleolar antigen polymyositis-Scl. *Medicine*, **71**, 327–36.

Martin, A., Haller, R.G., and Barohn, R. (1994). Metabolic myopathies. *Current Opinion in Rheumatology*, **6**, 552–8.

Mathews, M.B. and Bernstein, R.M. (1983). Myositis autoantibody inhibits histidyl-tRNA synthetase: a model for autoimmunity. *Nature*, **304**, 177–9.

Matsubara, S., Shima, T., and Takamori, M. (1993). Experimental allergic myositis in SJL/J mice immunized with rabbit myosin B fraction: immunohistochemical analysis and transfer. *Acta Neuropathologica*, **85**, 138–44.

Matsubara, S., Sawa, Y., Takamori, M., Yokoyama, H., and Kida, H. (1994). Pulsed intravenous methylprednisolone combined with oral steroids as the initial treatment of inflammatory myopathies. *Journal of Neurology, Neurosurgery and Psychiatry*, **57**, 1008–16.

Matsuda, Y., Tomii, M., and Kashiwazaki, S. (1993). Fatal pneumomediastinum in dermatomyositis without creatine kinase elevation. *Internal Medicine*, **32**, 643–7.

Mbauya, A.L., Plotz, P.H., Wilder, R.L., and Miller, F.W. (1993). Increased prevalence of autoimmune disease in first degree relatives of patients with idiopathic inflammatory myopathy (IIM). *Arthritis and Rheumatism*, **36**, S255 (Abstract).

McKendry, R.J.R. (1987). Influence of age at onset on the duration of treatment in idiopathic adult polymyositis and dermatomyositis. *Archives of Internal Medicine*, **147**, 1989–91.

McKinney, R.E., Katz, S.L., and Wilfert, C.M. (1987). Chronic enteroviral meningoencephalitis in agammaglobulinemic patients. *Review of Infectious Diseases*, **9**, 334–56.

Medsger, T.A., Jr. and Oddis, C.V. (1995). Classification and diagnostic criteria for polymyositis and dermatomyositis. *Journal of Rheumatology*, **22**, 581–5.

Medsger, T.A., Jr., Robinson, H., and Masi, A.T. (1971). Factors affecting survivorship in polymyositis: a life-table study of 124 patients. *Arthritis and Rheumatism*, **14**, 249–58.

Mendell, J.R., Sahenk, Z., Gales, T., and Paul, L. (1991). Amyloid filaments in inclusion body myositis: novel findings provide insight into nature of filaments. *Archives of Neurology*, **48**, 1229–34.

Metzger, A.L., Bohan, A., Goldberg, L.S., Bluestone, R., and Pearson, C.M. (1974). Polymyositis and dermatomyositis: combined methotrexate and corticosteroid therapy. *Annals of Internal Medicine*, **81**, 182–9.

Mhiri, C. *et al*. (1991). Zidovudine myopathy: a distinctive disorder associated with mitochondrial dysfunction. *Annals of Neurology*, **29**, 606–14.

Miller, F.W. (1993). Myositis-specific autoantibodies: touchstones for understanding the inflammatory myopathies. *Journal of the American Medical Association*, **270**, 1846–9.

Miller, F.W. (1994). Classification and prognosis of inflammatory muscle disease. *Rheumatic Disease Clinics of North America*, **20**, 811–26.

Miller, F.W., Love, L.A., Biswas, T., McClintock, P.R., Notkins, A.L., and Plotz, P.H. (1987). Viral and host genetic factors influence encephalomyocarditis virus-induced polymyositis in adult mice. *Arthritis and Rheumatism*, **30**, 549–56.

Miller, F.W., Love, L.A., Barbieri, S.A., Balow, J.E., and Plotz, P.H. (1990 *a*). Lymphocyte activation markers in idiopathic myositis: changes with disease activity and differences among clinical and autoantibody subgroups. *Clinical and Experimental Immunology*, **81**, 373–9.

Miller, F.W., Twitty, S.A., Biswas, T., and Plotz, P.H. (1990 *b*). Origin and regulation of a disease-specific autoantibody response: antigenic epitopes, spectrotype stability, and isotype restriction of anti-Jo-1 antibodies. *Journal of Clinical Investigation*, **85**, 468–75.

Miller, F.W., Waite, K.A., Biswas, T., and Plotz, P.H. (1990 *c*). The role of an autoantigen, histidyl-tRNA synthetase, in the induction and maintenance of autoimmunity. *Proceedings of the National Academy of Sciences USA*, **87**, 9933–7.

Miller, F.W. *et al*. (1992). Controlled trial of plasma exchange and leukapheresis in polymyositis and dermatomyositis. *New England Journal of Medicine*, **326**, 1380–84.

Mitra, D., Lovell, C.L., Macleod, T.I., Tan, R.S., and Maddison, P.J. (1994). Clinical and histological features of 'mechanic's hands' in a patient with antibodies to Jo-1—a case report. *Clinical and Experimental Dermatology*, **19**, 146–8.

Moder, K.G., Gaffey, T.A., and Matteson, E.L. (1993). Idiopathic inflammatory myopathy of the antisynthetase (Jo-1) type associated with noncaseating granulomas. *Arthritis and Rheumatism*, **36**, 1743–7.

Morgan, O. St.C., Mora, C., Rodgers-Johnson, P., and Char, G. (1989). HTLV-1 and polymyositis in Jamaica. *Lancet*, **ii**, 1184–7.

Moulds, J.M. *et al*. (1990). C4 Null genes in American whites and blacks with myositis. *Journal of Rheumatology*, **17**, 331–4.

Moxley, R.T. (1994). Evaluation of neuromuscular function in inflammatory myopathy. *Rheumatic Disease Clinics of North America*, **20**, 827–44.

Nagaraja, D., Taly, A.B., Suresh, T.G., Gourie-Devi, M., Sarala Das, and Rao, B.S.S. (1992). Epidemic of acute inflammatory myopathy in Karnataka, South India: 30 cases. *Acta Neurologica Scandinavia*, **86**, 230–36.

Naparstek, Y. and Plotz, P.H. (1993). The role of autoantibodies in autoimmune disease. *Annual Review of Immunology*, **11**, 79–104.

Nelson, P.N., Lever, A.M., Bruckner, F.E., Isenberg, D.A., Kessaris, N., and Hay, F.C. (1994). Polymerase chain reaction fails to incriminate exogenous retroviruses HTLV-I and HIV-1 in rheumatological diseases although a minority of sera cross react with retroviral antigens. *Annals of the Rheumatic Diseases*, **53**, 749–54.

Nishikai, M. (1994). Coxsackievirus infection and the development of polymyositis/dermatomyositis. *Rheumatology International*, **14**, 43–6.

Nishikai, M. and Sato, A. (1991). Human T lymphotropic virus type I and polymyositis and dermatomyositis in Japan. *Arthritis and Rheumatism*, **34**, 791–2.

Nishino, H., Engel, A.G., and Rima, B.K. (1989). Inclusion body myositis: the mumps virus hypothesis. *Annals of Neurology*, **25**, 260–64.

Oddis, C.V. (1991). Therapy for myositis. *Current Opinion in Rheumatology*, **3**, 919–24.

Oddis, C.V. (1993). Therapy for myositis. *Current Opinion in Rheumatology*, **5**, 742–8.

Oddis, C.V. (1994). Therapy of inflammatory myopathy. *Rheumatic Disease Clinics of North America*, **20**, 899–918.

Oddis, C.V. and Hill, P. (1993). Pregnancy outcome in women with inflammatory myopathy. *Arthritis and Rheumatism*, **36**, S255 (Abstract).

Oddis, C.V. and Medsger, T.A., Jr. (1988). Relationship between serum creatine kinase level and corticosteroid therapy in polymyositis-dermatomyositis. *Journal of Rheumatology*, **15**, 807–11.

Oddis, C.V. and Medsger, T.A., Jr. (1989). Current management of polymyositis and dermatomyositis. *Drugs*, **37**, 382–90.

Oddis, C.V., Conte, C.G., Steen, V.D., and Medsger, T.A., Jr. (1990 *a*). Incidence of polymyositis-dermatomyositis: a 20-year study of hospital diagnosed cases in Allegheny County, PA 1963–1982. *Journal of Rheumatology*, **17**, 1329–34.

Oddis, C.V., Medsger, T.A., Jr., and Cooperstein, L.A. (1990 *b*). A subluxing arthropathy associated with the anti-Jo-1 antibody in polymyositis/dermatomyositis. *Arthritis and Rheumatism*, **33**, 1640–45.

Oddis, C.V., Okano, Y., Rudert, W.A., Trucco, M., Duquesnoy, R.J., and Medsger, T.A., Jr. (1992). Serum autoantibody to the nucleolar antigen polymyositis-Scl: clinical and immunogenetic associations. *Arthritis and Rheumatism*, **35**, 1211–17.

Oddis, C.V., Carroll, P., Abu-Elmagd, K., McCauley, J., Fung, J.J., and Starzl, T.E. (1994). FK 506 in the treatment of polymyositis. *Arthritis and Rheumatism*, **37**, S286 (Abstract).

O'Hanlon, T.P., Dalakas, M.C., Plotz, P.H., and Miller, F.W. (1994 *a*). Predominant TCR-alpha beta variable and joining gene expression by muscle-infiltrating lymphocytes in the idiopathic inflammatory myopathies. *Journal of Immunology*, **152**, 2569–76.

O'Hanlon, T.P., Dalakas, M.C., Plotz, P.H., and Miller, F.W. (1994 *b*). The alpha beta T-cell receptor repertoire in inclusion body myositis: diverse patterns of gene expression by muscle-infiltrating lymphocytes. *Journal of Autoimmunity*, **7**, 321–33.

Okano, Y., Steen, V.D., and Medsger, T.A., Jr. (1992). Autoantibody to U3 Nucleolar ribonucleoprotein (fibrillarin) in patients with systemic sclerosis. *Arthritis and Rheumatism*, **35**, 95–100.

O'Neill, T.W. and Maddison, P.J. (1993). Rheumatoid arthritis associated with myositis and anti-Jo-1 antibody. *Journal of Rheumatology*, **20**, 141–3.

Orimo, S. *et al*. (1994). Immunohistochemical analysis of perforin and granzyme A in inflammatory myopathies. *Neuromuscular Disorders*, **4**, 219–26.

O'Rourke, K.S., Blaivas, M., and Ike, R.W. (1994). Utility of needle muscle biopsy in a university rheumatology practice. *Journal of Rheumatology*, **21**, 413–24.

Osterman, P.O., Askmark, H., and Wistrand, P.J. (1985). Serum carbonic anhydrase III in neuromuscular disorders and in healthy persons after a long-distance run. *Journal of Neurological Science*, **70**, 347–57.

Pamphlett, R. (1988). Muscle biopsy. In *Inflammatory diseases of muscle* (ed. F.L. Mastaglia), pp. 17–36. Blackwell Scientific, Oxford.

Park, J.H. *et al*. (1990). Dermatomyositis: correlative MR imaging and P-31 MR spectroscopy for quantitative characterization of inflammatory disease. *Radiology*, **177**, 473–9.

Park, J.H. *et al*. (1994). Magnetic resonance imaging and P-31 magnetic resonance spectroscopy provide unique quantitative data useful in the longitudinal management of patients with dermatomyositis. *Arthritis and Rheumatism*, **37**, 736–46.

Park, J.H. *et al*. (1995). Use of magnetic resonance imaging and P-31 magnetic resonance spectroscopy to detect and quantify muscle dysfunction in the amyopathic and myopathic variants of dermatomyositis. *Arthritis and Rheumatism*, **38**, 68–77.

Pierce, L.R., Wysowski, D.K., and Gross, T.P. (1990). Myopathy and rhabdomyolysis associated with lovastatin-gemfibrozil combination therapy. *Journal of the American Medical Association*, **264**, 71–5.

Pitt, A.M., Fleckenstein, J.L., Greenlee, R.G., Jr., Burns, D.K., Bryan, W.W., and Haller, R. (1993). MRI-guided biopsy in inflammatory myopathy: initial results. *Magnetic Resonance Imaging*, **11**, 1093–9.

Plotz, P.H. (1992). Not myositis: a series of chance encounters. *Journal of the American Medical Association*, **268**, 2074–7.

Plotz, P.H., Dalakas, M., Leff, R.L., Love, L.A., Miller, F.W., and Cronin, M.E. (1989). Current concepts in the idiopathic inflammatory myopathies: polymyositis, dermatomyositis, and related disorders. *Annals of Internal Medicine*, **111**, 143–57.

Plotz, P.H., Rider, L.G., Targoff, I.N., Raben, N., O'Hanlon, T.P., and Miller, F.W. (1995). Myositis: immunologic contributions to understanding cause, pathogenesis, and therapy. *Annals of Internal Medicine*, **122**, 715–24.

Pluschke, G., Rüegg, D., Hohlfeld, R., and Engel, A.G. (1992). Autoaggressive myocytotoxic T lymphocytes expressing an unusual gamma/delta T cell receptor. *Journal of Experimental Medicine*, 1**76**, 1785–9.

Preston, D.C. (1993). Electrophysiology of *Microsporidia* myositis in an AIDS patient. *Muscle and Nerve*, **16**, 1420–22.

Raben, N. *et al*. (1994). A motif in human histidyl-tRNA synthetase which is shared among several aminoacyl-tRNA synthetases is a coiled-coil that is essential for enzymatic activity and contains the major autoantigenic epitope. *Journal of Biological Chemistry*, **269**, 24277–83.

Ramirez, G., Asherson, R.A., Khamashta, M.A., Cervera, R., D'Cruz, D., and Hughes, G.R.V. (1990). Adult-onset polymyositis-dermatomyositis: description of 25 patients with emphasis on treatment. *Seminars in Arthritis and Rheumatism*, **20**, 114–20.

Rechavia, E., Rotenberg, Z., Fuchs, J., and Strasberg, B. (1985). Polymyositic heart disease. *Chest*, **88**, 309–11.

Reichlin, M. and Arnett, F.C. (1984). Multiplicity of antibodies in myositis sera. *Arthritis and Rheumatism*, **27**, 1150–56.

Reimers, C.D., Fleckenstein, J.L., Witt, T.N., Muller-Felber, W., and Pongratz, D.E. (1993). Muscular ultrasound in idiopathic inflammatory myopathies of adults. *Journal of Neurological Science*, **116**, 82–92.

Reimers, C.D. *et al*. (1994). Magnetic resonance imaging of skeletal muscles in idiopathic inflammatory myopathies of adults. *Journal of Neurology*, **241**, 306–14.

Reimold, A.M. and Weinblatt, M.E. (1994). Tachyphylaxis of intravenous immunoglobulin in refractory inflammatory myopathy. *Journal of Rheumatology*, **21**, 1144–6.

Reveille, J.D., Targoff, I.N., Mimori, T., Nguyen, H.C., Goldstein, R., and Arnett, F.C. (1992). MHC class II alleles associated with myositis. *Arthritis and Rheumatism*, **35**, S84 (Abstract).

Rider, L.G. and Miller, F.W. (1995). Laboratory evaluation of the inflammatory myopathies. *Clinical Diagnosis and Laboratory Immunology*, **2**, 1–9.

Rider, L.G. *et al*. (1994). Association of autoantibodies to the U5-ribonucleoprotein (U5-RNP) with idiopathic inflammatory myopathy (IIM). *Arthritis and Rheumatism*, **37**, S242 (Abstract).

Riminton, D.S., Chambers, S.T., Parkin, P.J., Pollock, M., and Donaldson, I.M. (1993). Inclusion body myositis presenting solely as dysphagia. *Neurology*, **43**, 1241–3.

Ringel, S.P., Carry, M.R., Aguilera, A.J., and Starcevich, J.M. (1986). Quantitative histopathology of the inflammatory myopathies. *Archives of Neurology*, **43**, 1004–9.

Roberts, L.J. and Fink, C.W. (1988). Childhood polymyositis/dermatomyositis. *Clinics in Dermatology*, **6**, 36–46.

Rockerbie, N.R., Woo, T.Y., Callen, J.P., and Giustina, T. (1989). Cutaneous changes of dermatomyositis precede muscle weakness. *Journal of the American Academy of Dermatology*, **20**, 629–32.

Rodgers, W.B., Yodlowski, M.L., and Mintzer, C.M. (1993). Pyomyositis in patients who have the human immunodeficiency virus. Case report and review of the literature. *Journal of Bone and Joint Surgery*, **75**, 588–92.

Rosenberg, M.J. and Reichlin, M. (1994). Is there an association between injectable collagen and polymyositis/dermatomyositis? *Arthritis and Rheumatism*, **37**, 747–53.

Rosenberg, N.L. (1993). Experimental models of inflammatory myopathies. *Baillière's Clinical Neurology*, **2**, 693–715.

Rosenberg, N.L., Rotbart, H.A., Abzug, M.J., Ringel, S.P., and Levin, M.J. (1989). Evidence for a novel picornavirus in human dermatomyositis. *Annals of Neurology*, **26**, 204–9.

Rosenzweig, B.A., Rotmensch, S., Binette, S.P., and Phillippe, M. (1989). Primary idiopathic polymyositis and dermatomyositis complicating pregnancy: diagnosis and management. *Obstetrics and Gynecology Survey*, **44**, 162–70.

Rowe, D., Isenberg, D.A., and Beverley, P.CL. (1983). Monoclonal antibodies to human leucocyte antigens in polymyositis and muscular dystrophy. *Clinical and Experimental Immunology*, **54**, 327–36.

Rubin, R.B. and Neugarten, J. (1989). Cocaine-induced rhabdomyolysis masquerading as myocardial ischemia. *American Journal of Medicine*, **86**, 551–3.

Sabina, R.L., Sulaiman, A.R., and Wortmann, R.L. (1990). Reduced transcript availability and myoadenylate deaminase (MAD) activity in polymyositis: implications for acquired MAD deficiency. *Arthritis and Rheumatism*, **33**, S70.

Saito, E., Kuroda, K., Yoshimoto, Y., Oshima, H., and Kinoshita, M. (1989). Mechanism of the damaging effect of dermatomyositis mononuclear cells on cultured human skin fibroblasts. *Journal of Rheumatology*, **16**, 1055–60.

Sanmarti, R. *et al*. (1994). Reduced activity of serum creatine kinase in rheumatoid arthritis: a phenomenon linked to the inflammatory response. *British Journal of Rheumatology*, **33**, 231–4.

Sano, M., Suzuki, M., Sato, M., Sakamoto, T., and Uchigata, M. (1994). Fatal respiratory failure due to polymyositis. *Internal Medicine*, **33**, 185–7.

Satoh, M., Ajmani, A.K., Hirakata, M., Suwa, A., Winfield, J.B., and Reeves, W.H. (1994). Onset of polymyositis with autoantibodies to threonyl-tRNA synthetase during pregnancy. *Journal of Rheumatology*, **21**, 1564–6.

Sayers, M.E., Chou, S.M., and Calabrese, L.H. (1992). Inclusion body myositis: analysis of 32 cases. *Journal of Rheumatology*, **19**, 1385–9.

Schulman, P., Kerr, L.D., and Spiera, H. (1991). A reexamination of the relationship between myositis and malignancy. *Journal of Rheumatology*, **18**, 1689–92.

Schwarz, M.L., Matthay, R.A., Sahn, S.A., Stanford, R.E., Marmorstein, B.L., and Scheinhorn, D.J. (1976). Interstitial lung disease in polymyositis and dermatomyositis: analysis of six cases and review of the literature. *Medicine*, **55**, 89–104.

Seidman, R., Peress, N.S., and Nuovo, G.J. (1994). *In situ* detection of polymerase chain reaction-amplified HIV-1 nucleic acids in skeletal muscle in patients with myopathy. *Modern Pathology*, **7**, 369–75.

Sigurgeirsson, B., Lindelof, B., Edhag, O., and Allander, E. (1992). Risk of cancer in patients with dermatomyositis or polymyositis: a population-based study. *New England Journal of Medicine*, **326**, 363–7.

Simpson, D.M. and Bender, A.N. (1988). Human immunodeficiency virus-associated myopathy: Analysis of 11 patients. *Annals of Neurology*, **24**, 79–84.

Sinoway, P.A. and Callen, J.P. (1993). Chlorambucil. An effective corticosteroid-sparing agent for patients with recalcitrant dermatomyositis. *Arthritis and Rheumatism*, **36**, 319–24.

Sintani, S., Murase, H., Tsukagoshi, H., and Shiigai, T. (1992). Glycyrrhizin (licorice)-induced hypokalemic myopathy: report of 2 cases and review of the literature. *European Neurology*, **32**, 44–51.

Soueidan, S.A. and Dalakas, M.C. (1993). Treatment of inclusion-body myositis with high-dose intravenous immunoglobulin. *Neurology*, **43**, 876–9.

Steinberg, A.D. (1993). Chlorambucil in the treatment of patients with immune-mediated rheumatic diseases. *Arthritis and Rheumatism*, **36**, 325–8.

Stiglbauer, R. *et al*. (1993). Polymyositis: MRI-appearance at 1.5 T and correlation to clinical findings. *Clinics in Radiology*, **48**, 244–8.

Stommel, E.W., Parsonnet, J., and Jenkyn, L.R. (1991). Polymyositis after ciguatera toxin exposure. *Archives of Neurology*, **48**, 874–7.

Stonecipher, M.R., Jorizzo, J.L., White, W.L., Walker, F.O., and Prichard, E. (1993). Cutaneous changes of dermatomyositis in patients with normal muscle enzymes: dermatomyositis sine myositis? *Journal of the American Academy of Dermatology*, **28**, 951–6.

Stonecipher, M.R., Jorizzo, J.L., Monu, J., Walker, F., and Sutej, P.G. (1994). Dermatomyositis with normal muscle enzyme concentrations. A single-blind study of the diagnostic value of magnetic resonance imaging and ultrasound. *Archives of Dermatology*, **130**, 1294–9.

Strongwater, S.L., Dorovini-Zis, K., Ball, R.D., and Schnitzer, T.J. (1984). A murine model of polymyositis induced by coxsackievirus Bl (Tucson strain). *Arthritis and Rheumatism*, **27**, 433–42.

Takahashi, K., Ogita, T., Okudaira, H., Yoshinoya, S., Yoshizawa, H., and Miyamoto, T. (1986). D-Penicillamine-induced polymyositis in patients with rheumatoid arthritis. *Arthritis and Rheumatism*, **29**, 560–64.

Tam, P.E., Schmidt, A.M., Ytterberg, S.R., and Messner, R.P. (1994). Duration of virus persistence and its relationship to inflammation in the chronic phase of coxsackievirus B1-induced murine polymyositis. *Journal of Laboratory and Clinical Medicine*, **123**, 346–56.

Tami, L.F. and Bhasin, S. (1993). Polymorphism of the cardiac manifestations in dermatomyositis. *Clinics in Cardiology*, **16**, 260–64.

Taneja, V., Mehra, N., Singh, Y.N., Kumar, A., Malaviya, A., and Singh, R.R. (1990). HLA-D region genes and susceptibility to D-penicillamine-induced myositis. *Arthritis and Rheumatism*, **33**, 1445–7.

Tanimoto, K. *et al*. (1995). Classification criteria for polymyositis and dermatomyositis. *Journal of Rheumatology*, **22**, 668–74.

Taniyama, M., Yoh, S., Asaba, Y., Maruyama, T., Takei, I., and Kataoka, K. (1987). Elevated serum creatine kinase level in diabetic patients with nephrotic syndrome: a role of fluid retention. *Annals of Internal Medicine*, **106**, 711–12.

Targoff, I.N. (1988). Laboratory manifestations of polymyositis/dermatomyositis. *Clinics in Dermatology*, **6**, 76–92.

Targoff, I.N. (1990). Inflammatory muscle disease. In *The lung in rheumatic diseases* (ed. G.W. Cannon and G.A. Zimmerman), pp. 303–28. Marcel Dekker, Inc., New York.

Targoff, I.N. (1991). Polymyositis. In *Systemic autoimmunity* (ed. P.E. Bigazzi and M. Reichlin), pp. 201–46. Marcel Dekker Inc., New York.

Targoff, I.N. (1994). Immune manifestations of inflammatory muscle disease. *Rheumatic Disease Clinics of North America*, **20**, 857–80.

Targoff, I.N. and Arnett, F.C. (1990). Clinical manifestations in patients with antibody to PL-12 antigen (alanyl-tRNA synthetase). *American Journal of Medicine*, **88**, 241–51.

Targoff, I.N., Arnett, F.C., Berman, L., O'Brien, C.A., and Reichlin, M. (1989). Anti-KJ: a new antibody associated with the myositis/lung syndrome that reacts with a translation-related protein. *Journal of Clinical Investigation*, **84**, 162–72.

Targoff, I.N., Johnson, A.E., and Miller, F.W. (1990). Antibody to signal recognition particle in polymyositis. *Arthritis and Rheumatism*, **33**, 1361–70.

Targoff, I.N., Trieu, E.P., Plotz, P.H., and Miller, F.W. (1992). Antibodies to glycyl-transfer RNA synthetase in patients with myositis and interstitial lung disease. *Arthritis and Rheumatism*, **35**, 821–30.

Targoff, I.N., Trieu, E.P., and Miller, F.W. (1993). Reaction of anti-OJ autoantibodies with components of the multi-enzyme complex of aminoacyl-tRNA synthetases in addition to isoleucyl-tRNA synthetase. *Journal of Clinical Investigation*, **91**, 2556–64.

Taylor, A.J., Wortham, D.C., Burge, J.R., and Rogan, K.M. (1993). The heart in polymyositis: a prospective evaluation of 26 patients. *Clinics in Cardiology*, **16**, 802–8.

Tazelaar, H.D., Viggiano, R.W., Pickersgill, J., and Colby, T.V. (1990). Interstitial lung disease in polymyositis and dermatomyositis. *American Review of Respiratory Diseases*, **141**, 727–33.

Tymms, K.E. and Webb, J. (1985). Dermatopolymyositis and other connective tissue diseases: a review of 105 cases. *Journal of Rheumatology*, **12**, 1140–48.

Uncini, A., Lange, D.J., Lovelace, R.E., Solomon, M., and Hays, A.P. (1990). Long-duration polyphasic motor unit potentials in myopathies: a quantitative study with pathological correlation. *Muscle and Nerve*, **13**, 263–7.

Urbano-Marquez, A., Casademont, J., and Grau, J.M. (1991). Polymyositis/dermatomyositis: the current position. *Annals of the Rheumatic Diseases*, **50**, 191–5.

Varga, J., Heiman-Patterson, T., Munoz, S., and Love, L.A. (1993). Myopathy with mitochondrial alterations in patients with primary biliary cirrhosis and antimitochondrial antibodies. *Arthritis and Rheumatism*, **36**, 1468–75.

Wada, K. *et al*. (1983). Radioimmunoassay for antibodies to human skeletal muscle myosin in serum from patients with polymyositis. *Clinical and Experimental Immunology*, **52**, 297–304.

Wei, N., Pavlidis, N., Tsokos, G., Elin, R.J., and Plotz, P.H. (1981). Clinical significance of low creatine phosphokinase values in patients with connective tissue diseases. *Journal of the American Medical Association*, **246**, 1921–3.

Whitmore, S.E., Rosenshein, N.B., and Provost, T.T. (1994). Ovarian cancer in patients with dermatomyositis. *Medicine*, **73** 153–60.

Wiley, C.A., Nerenberg, M., Cros, D., and Soto-Aguilar, M.C. (1989). HTLV-I polymyositis in a patient also infected with the human immunodeficiency virus. *New England Journal of Medicine*, **320**, 992–5.

Wolf, P.L. (1991). Abnormalities in serum enzymes in skeletal muscle diseases. *American Journal of Clinical Pathology*, **95**, 293–6.

Wolf, R.E. and Baethge, B.A. (1990). Interleukin-1a, interleukin-2, and soluble interleukin-2 receptors in polymyositis. *Arthritis and Rheumatism*, **33**, 1007–14.

Wolf, R.F., Sprenger, H.G., Mooyaart, E.L., Tamsma, J.T., Kengen, R.A., and Weits, J. (1990). Nontropical pyomyositis as a cause of subacute, multifocal myalgia in the acquired immunodeficiency syndrome. *Arthritis and Rheumatism*, **33**, 1728–32.

Woo, T.Y., Callen, J.P., Voorhees, J.J., Bickers, D.R., Hanno, R., and Hawkins, C. (1984). Cutaneous lesions of dermatomyositis are improved by hydroxychloroquine. *Journal of the American Academy of Dermatology*, **10**, 592–600.

Wright, G.D., Wilson, C., and Bell, A.L. (1994). D-penicillamine induced polymyositis causing complete heart block. *Clinics in Rheumatology*, **13**, 80–82.

Yousef, G.E., Isenberg, D.A., and Mowbray, J.F. (1990). Detection of enterovirus specific RNA sequences in muscle biopsy specimens from patients with adult onset myositis. *Annals of the Rheumatic Diseases*, **49**, 310–15.

Ytterberg, S.R. (1994). The relationship of infectious agents to inflammatory myositis. *Rheumatic Disease Clinics of North America*, **20**, 995–1016.

Ytterberg, S.R., Mahowald, M.L., and Messner, R.P. (1987). Coxsackievirus B1-induced polymyositis: lack of disease expression in nu/nu mice. *Journal of Clinical Investigation*, **80**, 499–506.

Ytterberg, S.R., Roelofs, R.I., and Mahowald, M.L. (1993). Inclusion body myositis and renal cell carcinoma: report of two cases and review of the literature. *Arthritis and Rheumatism*, **36**, 416–21.

Zantos, D., Zhang, Y., and Felson, D. (1994). The overall and temporal association of cancer with polymyositis and dermatomyositis. *Journal of Rheumatology*, **21**, 1855–9.

Zieglschmid-Adams, M.E., Pandya, A.G., Cohen, S.B., and Sontheimer, R.D. (1995). Treatment of dermatomyositis with methotrexate. *Journal of the American Academy of Dermatology*, **32**, 754–7.

Zuckner, J. (1994). Drug-related myopathies. *Rheumatic Disease Clinics of North America*, **20**, 1017–32.

5.9.2 Polymyositis and dermatomyositis in children

Lauren M. Pachman

Introduction

The range of inflammatory muscle disease in children includes acute viral or bacterial myositis, and chronic myositis–juvenile dermatomyositis, polymyositis, and inflammatory myopathy associated with other autoimmune rheumatic diseases or with parasitic infection. Worldwide, most of the inflammatory myopathies accompany bacterial or parasitic infections, but in North America and Europe a viral aetiology is seen more commonly.

The major advances in the study of these diseases include more definite clinical characterization of subsets and an emerging appreciation of their epidemiology as shown by their specific as well as their shared clinical and laboratory features. Recent evidence suggests that genetic and infectious factors may play a part in disease susceptibility. The primary clinical feature of both juvenile dermatomyositis and polymyositis is chronic and progressive weakness of proximal muscles. In juvenile dermatomyositis the vasculopathy and distinctive skin manifestations are associated commonly with muscle involvement; in polymyositis the skin is spared. Fulfilment of the criteria of Bohan and Peter (1975) is needed to establish the diagnosis of either type of myopathy (Table 1). The diagnosis of definite juvenile dermatomyositis is made if, in addition to the typical rash, three of the four criteria are present; of definite polymyositis, if three of the four criteria are found. Myositis is often a part of other autoimmune rheumatic diseases and, therefore, it is essential to exclude such conditions as systemic lupus erythematosus, mixed connective tissue disease, juvenile chronic arthritis (especially of systemic onset), the spondylarthropathies, and Sjögren's syndrome.

	Juvenile dermatomyositis	Polymyositis
Characteristic rash	+	–
Symmetrical proximal muscle weakness[b]	+	+
Elevated muscle-derived enzymes	+	+
Muscle histopathology[c]	+	+
Electromyographic changes; inflammatory myopathy	+	+

[a] In addition to presence (or absence) of characteristic rash and exclusion of other rheumatic diseases, three of four criteria must be met to confirm diagnosis.
[b] Exclusion of other rheumatic diseases.
[c] The histology of childhood polymyositis is different from juvenile dermatomyositis.
Definite JDMS: rash plus three or four criteria; Probable JDMS: rash plus two criteria; possible JDMS: rash plus one criteria.
Modified from Bohan and Peter (1975). *New England Journal of Medicine*, **292**, 344–7.

Table 1 Criteria for diagnosis of juvenile dermatomyositis and polymyositis in childhood [a]

Epidemiology

Demographic data

The bimodal age distribution of populations of combined polymyositis/dermatomyositis is well known with a childhood peak (at 5 to 9 years of age, 3.7 cases/million per year and at 10 to 14 years of age, 4.3 cases/million per year) and an adult peak (at 45 to 64 years or age, 10 cases/million per year) (Medsger *et al*. 1970). Children of African or Asian origin may be at increased risk for chronic myositis (Benbassat *et al*. 1980), but in the United States, Caucasian children with juvenile dermatomyositis are reported more frequently, and twice as many girls as boys are affected; in the United Kingdom and Ireland, five times as many girls were diagnosed as boys, with a incidence of 1.9/million children under the age of 16 years (Symmons *et al*. 1995). A similar trend is found in China (Wang *et al*. 1993), in contrast to Japan where a ratio of 1.3 boys to 1 girl was identified; none had associated malignancy or interstitial lung disease (Hiketa *et al*. 1992). In children, dermatomyositis occurs at least 10 to 20 times more often than polymyositis. In marked contrast to earlier reports of one-third mortality and one-third morbidity (Bitnum *et al*. 1964), both of these adverse outcomes have decreased in childhood since the advent of steroid therapy (Hochberg 1988; Ansell *et al*. 1990). In Japan, during a 10-year period (1973 to 1983) there was a 2.9 per cent mortality rate (Hidano *et al*. 1986).

Although an increased frequency of malignancy may be associated with adult dermatomyositis within 2 years of disease onset (Masi and Hochberg 1988), this does not appear to be the case in children with either juvenile dermatomyositis or polymyositis (Hidano *et al*. 1986); only sporadic cases of children with both an inflammatory myopathy and malignancy have been cited (Sherry *et al*. 1993), and a population-based survey did not find malignancy in any patient less than 16 years of age (Sigurgeirsson *et al*. 1992).

There is some data to suggest that in the United States children are more likely to have onset of disease in the early months of the year. A comparison of disease onset of juveniles with dermatomyositis and adults with polymyositis–dermatomyositis living in Memphis, Tennessee, revealed a seasonal onset (February to April) in 55 per cent of the children but not in the adults (Medsger *et al*. 1970). In contrast, a case–control study (all paediatric inflammatory myopathy, numbers of juvenile dermatomyositis unknown) conducted in a different region of the country did not document a specific season of onset (Koch *et al*. 1976). In the next decade, in north

central region, children with definite juvenile dermatomyositis (diagnosed within 4 months of onset) were more likely to have their first symptoms in the months of January to June than at other periods of time in each of 7 years (1974 to 1980) (Christensen *et al*. 1983); in Canada clustering was observed as well, suggesting an environmental influence (Rosenberg 1994). In the United Kingdom and Ireland, several clusters of disease onset were identified, the largest of which was in April and May in 1992; the timing of these clusters appeared to vary from year to year (Symmons *et al*. 1995). In Athens, Greece, of those adults (children were not included) with polymyositis–dermatomyositis who were admitted to hospital, 39 per cent were in March, April, or May with onset of symptoms in the same months (Manta *et al*. 1989). It is not yet firmly established if prevalence for season of onset is the same for both adults and children living in the same region or if the peak onset of polymyositis is in the same period as that of dermatomyositis in either age group. A national study is ongoing in the United States to determine if there is in fact seasonality or symptoms of antecedent illness in children with juvenile dermatomyositis.

Infectious agents and juvenile dermatomyositis

Several agents have been associated with the onset (and on occasion, flare) of juvenile dermatomyositis, of which the most prominent have been the RNA picornaviruses, group A b-haemolytic streptococci, and *Toxoplasma gondii*. Coxsackievirus B (CVB) 2 was isolated from the stool, and the titre of CF (complement fixing) antibody rose in a child with chronic myositis (Schiraldi and Iandolo 1978); others have measured rise in B4 antibodies (Travers *et al*. 1977). As noted above, there may be temporal, seasonal, and regional differences in agents associated with disease onset. Sera from newly diagnosed children during 1974 to 1980 from the Chicago area had an increased frequency of antibody (both neutralizing and complement fixing) to coxsackievirus B, as well as increased frequency of the histocompatibility antigens HLA B8/DR3 (Christensen *et al*. 1986) (see Genetics below), which was not reproduced by the same group of investigators in a study of 20 children with dermatomyositis from the same region with onset of disease in the years 1987 to 1992 (Pachman *et al*. 1995b). Enteroviral RNA was identified in the muscle of English patients with polymyositis/dermatomyositis (Bowles *et al*. 1987; Yousef *et al*. 1990), but other investigators have not found viral RNA in cases from the United States, either in Jo-1 positive adults (Leff *et al*. 1992), or in MRI-directed muscle biopsies of 20 newly diagnosed, untreated, children with active dermatomyositis (Pachman *et al*. 1995b). Other infectious agents have been thought to play a role in juvenile dermatomyositis, including toxoplasmosis (Schroter *et al*. 1987; Lapetina 1989) and hepatitis B (Peters *et al*. 1991), but antibody titres to these agents were not increased either in a regional study (Christensen *et al*. 1983) or in a case–control national study of new onset juvenile dermatomyositis spanning the years 1987 to 1992 (Pachman *et al*. 1992a; Pachman *et al*. 1995b). Although Theiler's murine encephalomyelitis virus (TMEV) was identified in adult polymyositis, none of the children with dermatomyositis were positive for this agent (Rosenberg *et al*. 1989). Taken together, the above data suggests that the aetiology of juvenile dermatomyositis is multifactorial, perhaps permitting a role for molecular mimicry. Evidence that may support this hypothesis is found in a report of skeletal myosin (which has sequence homology with the streptoccal type 5 M protein) stimulating lymphocytes from a child with recurrent dermatomyositis following streptococcal infections (Martini *et al*. 1992).

Clinical presentation

Cutaneous findings

In juvenile dermatomyositis, the rash may predate or follow the onset of symmetrical weakness in proximal muscles. Periorbital erythema and oedema, and/or eyelid telangiectasia, are seen in 50 to 90 per cent of affected children. The eyelid telangiectasia (as well as that elsewhere) may persist long after other signs and symptoms of disease activity have resolved. The rash has a violaceous or heliotrope hue, and is often most prominent on the eyelid, where small infarctions may be seen (Fig. 1). The rash may also cross the bridge of the nose and be precipitated by even a brief unprotected exposure to the sun. Sunlight may simply exacerbate the skin manifestations or may activate symptoms of myositis. Other areas of erythema involving the upper torso, the extensor surfaces of the arms and legs, medial malleoli of the ankles, as well as the buttocks may occur in the absence of raised serum concentrations of muscle-derived enzymes (see below). The skin over the knuckles is often either hypertrophic or pale red, evolving into colourless bands of atrophic skin (Gottron's sign). Similar patches can occur in a variety of places including over the metacarpal phalangeal joints, the extensor aspect of the elbows or knees, or the medial canthus of the eyelid. Lipoatrophy which is not tender despite a lymphohistiocytic panniculitis (Commens *et al*. 1990), as well as acanthosis nigricans, associated with insulin-resistant diabetes, have also been seen in children with juvenile dermatomyositis. Diffuse vasculopathy (nailbed telangiectasia, infarction of oral epithelium and skin folds, or digital ulceration) associated clearly with more severe disease, is correlated with the child's clinical course (Silver and Maricq 1989), and can be quantified to aid in diagnosis and in monitoring response to therapy (Pachman *et al*. 1995d). Although calcinosis without known antecedent muscle disease has been reported as a presenting sign of dermatomyositis in childhood (Martini *et al*. 1987), it may be considered to be related to disease severity and duration (Pachman *et al*. 1985). Delay in diagnosis by 6 months, which may be a consequence of difficulty in identification of the rash on pigmented skin, was associated with an increased incidence of calcinosis (Pachman *et al*. 1992b). None of 20 recently diagnosed and treated children with juvenile dermatomyositis developed calcinosis, suggesting that aggressive therapy early in the disease course may be of substantial benefit (Callen *et al*. 1994).

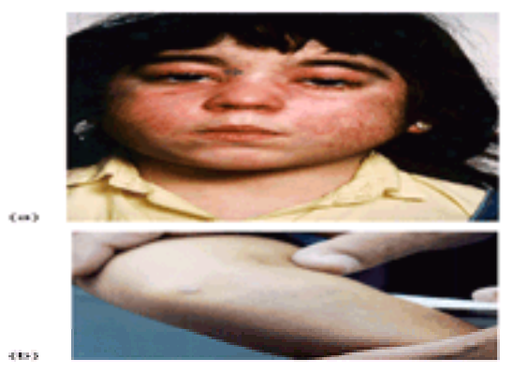


Fig. 1 (a) This young girl has the typical rash of juvenile dermatomyositis. Erythema and oedema involving the eyelids are seen, with healed microinfarcts at the medial aspect of the right upper eyelid (open arrow). The rash crosses the bridge of the nose and has areas of relative decreased vascularity as well as erythema. (b) Gottron's papules on the elbow of a 4-year-old child with recurrent juvenile dermatomyositis.

Musculoskeletal symptoms

Proximal muscle weakness as evidenced by difficulty in climbing stairs, getting up from a chair, combing hair, or using the hands to push off the body in an attempt to stand (Gower's sign) is common; weakness of the neck flexors is a particularly sensitive indicator of muscular impairment. Complaints of pain on compression of the muscles are found in over 60 per cent of cases, but are less severe than in the bacterial-related myopathies. Fatigue is common, and also expressed by children with the dystrophies (most commonly those of Duchenne and Becker). The most common symptoms are listed in Table 2. Usually the child is more comfortable when the limbs are held in the flexed position, promoting the development of flexion contractures. Inflammation can be detected by electromyography or muscle biopsy despite the appearance of clinical quiescence (Miller *et al*. 1987). The use of MRI-directed biopsies minimizes error in sampling uninvolved areas in this focal disease (Pitt *et al*. 1993), and can monitor a child's response to therapy (Yanagisawa *et al*. 1983; Keim *et al*. 1991; Hernandez *et al*. 1993). The MRI may become normal several months after the muscle enzymes return to the normal ranges (Huppertz and Kaiser 1994). P-32 spin MRI has been in use for the past few years and gives useful information about the child's muscle strength and performance (Park *et al*. 1994). Decreased bone density (associated with a depressed serum osteocalcin) is frequent in untreated juvenile dermatomyositis and places the child at risk of bony fracture (Reed *et al*. 1990), which is further augmented by steroid administration.

Symptoms	Percentage at diagnosis
Weakness	100
Rash	100
Muscle pain	72
Fever	65
Difficulty swallowing	45
Abdominal pain	37
Arthritis	36
Calcifications	22

Table 2 Juvenile dermatomyositis symptoms at diagnosis

Gastrointestinal involvement

One of the most troubling prognostic indicators is impairment of the flow of secretions associated with decreased oesophageal motility, which can be documented by radiographic contrast studies showing retained barium in a widened atonic, pyriform sinus. The swallowing of liquids may also be impaired; oesophageal reflux may result in aspiration pneumonia and appropriate precautions should be taken to prevent this (e.g. using thickened foods, raising the bed head, and bronchial drainage). Smooth muscle dysfunction can also result in decreased lower-gastrointestinal dysmotility, making constipation an annoying symptom. Involvement of the masseter may result in difficulty in chewing; chronic masseter atrophy is often apparent once the disease has become quiescent, and may endow the child with a characteristic chipmunk appearance. Vasculopathy affects any part of the gastrointestinal tract; in severe disease there is weight loss and mucosal ulceration with life-threatening perforation. In the young child, development of normal speech patterns can be disturbed; soft palate involvement is often revealed by nasal, high-pitched speech (for example by saying the alphabet) and usually resolves with a decrease in the inflammatory component of the myositis.

Cardiorespiratory abnormalities

The ECG is abnormal at disease onset in over half the children with definite juvenile dermatomyositis. Asymptomatic conduction abnormalities predominate, with an occasional complete block of the right bundle branch (Pachman and Cooke 1980), which usually resolves with decrease in disease activity. In the absence of respiratory complaints, a decrease in ventilatory capacity, with a normal diffusion of carbon dioxide was found in 78 per cent of children with juvenile dermatomyositis whose sera were subsequently found to be negative for antibody to the tRNA synthetase, Jo-1 (Pachman and Cooke 1980). The decrease in ventilatory capacity can be associated with diminished speech volume; several of our children have developed vocal cord nodules, presumably as a result of the stress of trying to be heard. Pulmonary fibrosis may occur in both adults and children with inflammatory myopathy, but it is more commonly found in individuals who carry antibody to the tRNA synthetases, of which anti-Jo-1 is the most common in adults (see Pathophysiology). Myositis-specific antibodies (MSA) are rare in children and are just starting to be identified (Rider *et al*. 1994). A young woman with Jo-1-positive myositis with onset in childhood has been recently described; her initial complaints included a reactive airway component (were called 'asthma'), but progressed to include arthritis and were accompanied by pulmonary fibrosis and a markedly decreased diffusion of carbon dioxide. The severe myopathy occurred 7 years after onset of respiratory symptoms (Bowles *et al*. 1987). In general, the few children with Jo-1 who have been identified recently are similar to adults with antisynthetase antibodies and are characterized by dyspnoea on exertion, a pulmonary perfusion deficit with evidence of pulmonary fibrosis on both histological and radiographical examination, and disease flares with the reduction of therapy. Most have severe arthritis; some have 'mechanic's hands' (Rider *et al*. 1995), which is much more common in adults, and rarely seen in children.

Genitourinary involvement

Massive breakdown of muscle elements as well as primary compromise of the renal parenchyma itself, may occur in children with an active myopathic process requiring prompt hydration and monitoring of renal function. If unchecked, renal failure can occur. On histochemical examination, cytoplasmic tubular arrays (may be a correlate of a-interferon production) are found in the renal glomerular endothelium as well as cutaneous fibroblastic cells (Pachman and Maryjowski 1984). Necrosis of the ureter has been reported, involving the middle one-third (iliac) segment, which is more vulnerable to vascular compromise as a consequence of inflammation. This vulnerability occurs because of the relatively sparse blood supply to this region compared with the upper (lumbar) or lower (pelvic) segments (Borrelli *et al*. 1988).

Eye signs

The most common finding is persistent thrombosis of dilated vessels at the margin of the upper eyelid, which may persist years after other clinical signs of inflammation have disappeared. In active disease, transient retinal exudates and 'cotton wool' spots may occur after the occlusion of small vessels, leading to intraretinal oedema with injury to retinal nerve fibres, optic atrophy, and sustained visual loss. Neovascularization of the retina with spontaneous regression has also been reported (Fong and Yeung 1990). Disease of conjunctival vessels can also lead to an avascular zone with a potential for infarction. Children treated with steroids should be monitored for both glaucoma and for the development of sublenticular cataracts, which are related to the dose of corticosterids given (Callen *et al*. 1994). If there is a family history of red–green colour blindness, the use of hydroxychloroquine should be avoided.

Other disease manifestations

Vasculopathy involving the central nervous system may be associated with depression and/or wide mood swings, which may be exacerbated by steroid therapy (see below). It is not usual for a child with juvenile dermatomyositis to present with Raynaud's phenomenon; these symptoms are found more frequently in overlap syndromes.

Differential diagnosis of juvenile dermatomyositis

General

The differential diagnosis of this inflammatory myopathy includes many of the major neuromuscular disorders of infancy and childhood as well as metabolic and infectious diseases which can be symptomatic at any age. Table 3 presents most of the potential candidates for consideration.

Table 3 Classification of the major neuromuscular disorders of infancy and childhood

Skin

Many of the other autoimmune diseases exhibit some of the cutaneous signs of juvenile dermatomyositis. As in systemic lupus, exposure to sun can exacerbate both the malar rash as well as the systemic complaints in children with dermatomyositis. Gottron's sign can be mimicked by psoriatic lesions, accompanied by healing foci of hypopigmentation found in areas usually unaffected in juvenile dermatomyositis, such as the pretibial region. These rashes may clear with sun exposure, rather than becoming more prominent. A description of a child with psoriatic arthritis and myopathy underlines the differences in clinical presentation (Thompson *et al*. 1990). Telangiectasia, a prominent feature of scleroderma, also occurs in overlap syndromes in which myositis is a component. Capillary destruction with resulting avascularity and healing can be monitored using nailfold capillary microscopy, which can help to differentiate juvenile dermatomyositis from some of the other vasculopathies (Pachman 1995). For example, a 2-year-old child had progressive unilateral focal calcinosis since the age of 5 months, as well as elevated muscle enzymes, but no muscle weakness or rash. Her capillary microscopy was abnormal, but not in the pattern seen in juvenile dermatomyositis, ultimately leading to the

correct diagnosis of progressive osseous heterotopia (Sachrison *et al*. 1995).

Muscle-derived enzymes

Skeletal muscle is rich in enzymes which can be released from the sarcoplasm to the peripheral circulation as a consequence of reduced vascular supply, trauma, or immunological cytotoxicity. Increased serum concentrations of enzymes result from either tissue necrosis or leakage through damaged cell membranes; sera to be analysed must be free of haemolytic products. Functional myopathies and those that result from denervation are accompanied by decreased muscle mass, weakness, pain, and loss of function or, in severe cases, paralysis. In these diseases, the serum creatine kinase, its muscle and brain isoenzymes, as well as the pyruvate kinase, lactate dehydrogenase, and aldolase are normal. In contrast, in some of the dystrophies, creatine kinase and aldolase may be increased by as much as 30 times the normal level; creatine kinase muscle isoenzymes are elevated after an acute myocardial infarction and after back surgery, and may also be raised in Duchenne's as well as Becker's muscular dystrophy. Increases are also seen in myocardial and skeletal muscle disease or trauma ranging from vigorous physical exercise (e.g. marathon runners) and surgical and accidental crush injuries to intramuscular injections. Drugs that enhance the permeability of muscle membranes, also result in elevated concentrations of the muscle isoenzyme form of creatine kinase; for example, aminocaproic acid, D-penicillamine, halothane (hyperpyrexia), and quinidine. Overdose with amphetamines, barbiturates, ethanol, or heroin results in a massive increase in creatine kinase (Lott and Landesman 1984). Disorders of calcium metabolism such as hypocalcaemia (either isolated or associated with rickets, hypoparathyroidism) or chronic renal failure can be accompanied by increased creatine kinase. Once the low levels of ionized calcium become normal, the rise in creatine kinase usually resolves. Aldolase has a short half-life and can be increased in viral hepatitis, metastatic liver disease, some prostate tumours, and in some of the leukaemias and anaemias (Lott and Landesman 1984). Lactate dehydrogenase has five isoenzymes: the anodal forms are increased in cardiac and renal disease, and in some diseases of skeletal muscle, while the cathodal forms are elevated in skeletal muscle and liver dysfunction. Other less specific enzymes that may also be found in higher concentrations in muscle damage are aspartate aminotransferase and alanine aminotransferase. Myoglobin is present only in skeletal and cardiac muscle. This oxygen-binding haem protein is elevated in serum from children with myopathies; persistent or massive myoglobinuria can result in renal failure, which should be prevented by infusion of adequate intravenous hydration.

Muscle complaints

Children with juvenile dermatomyositis may complain of pain on compression of the proximal muscles, or of muscle cramps, but weakness of symmetrical proximal muscles is still the predominant symptom. Other conditions associated with muscle cramps and contractures include hypothyroidism, uraemia, and electrolyte imbalance such as hypokalemia (either iatrogenic or in conjunction with familial periodic paralysis). Pretibial tenderness is seen with erythema nodosum, but is not a feature of juvenile dermatomyositis. Pain that awakens the child at night should be investigated for another cause such as malignancy, osteoid osteoma, or osteomyelitis. Muscle weakness can be see in hormonal derangements, either endogenous or iatrogenic, such as in adrenal dysfunction, or after long-term high-dose steroid administration. In addition, thyroid, pituitary, and parathyroid dysfunction may be accompanied by skeletal complaints. Metabolic muscle diseases include defects of glycolysis (e.g. phosphofructokinase deficiency) and are associated with contractures, exercise intolerance, myoglobinuria, and a positive ischaemic lactate test. There may be a defect in lipid metabolism such as a carnitine deficiency state (Breningstall 1990), which may be exacerbated by non-steroidal anti-inflammatory drugs; or a myalgia syndrome, which can be detected by a positive ischaemic ammonia test. Inclusion-body myositis, which often runs a steroid-resistant course, has also been described in children (Serratrice *et al*. 1989).

Acute infectious viral myositis in children, most frequently attributed to influenza A or B, is differentiated clinically from chronic myositis by its localization to the muscles of the calf, severe pain, and rapid resolution in 1 to 4 weeks (Mejlszenkier *et al*. 1973). The agent of this acute myositis has been isolated from cultures of muscle biopsy, accompanied by rise in complement-fixing antibody titres to influenza and by myoglobinuria, electomyographic changes, and elevated creatine kinase. As in adults with HIV or HTLV-1, children with these illnesses may also have muscle complaints.

Other autoimmune rheumatic disorders may also be accompanied by inflammatory muscle disease. For example, children with systemic-onset juvenile arthritis may have spiking fevers and an evanescent rash. Often they do not wish to be held and have increased concentrations of muscle enzymes, making the appropriate diagnosis of the specific connective tissue disease imperative. Confusion may be created by a child who has classical Gottron's papules, elevated muscle enzymes, and muscle biopsy evidence of perifascicular atrophy in the presence of antibody to RNP or PM/Scl. Such children with overlap syndrome are less likely to have complete resolution of their disease, requiring long-term therapy.

Electromyogram

Evidence of an inflammatory myopathy on electromyography is not specific to juvenile dermatomyositis and is similar in other autoimmune rheumatic disorders which have a myopathic component. Selection of a site of active involvement is facilitated by MRI identification of focally involved muscles. Once the location of the electrodes has been chosen (not the site of a future biopsy), insertional irritability, followed by spontaneous electrical activity at rest is often observed. This pattern can be also seen in the muscular dystrophies and in early acute myositis. Abnormal early full recruitment of muscle fibres with moderate effort occurs in about 45 per cent of patients with juvenile dermatomyositis, and bizarre, high-frequency discharges occur in 15 to 20 per cent of patients tested. Reduced motor unit activity is seen in Duchenne muscular dystrophy as well as in juvenile dermatomyositis. Myasthenia gravis can coexist with an inflammatory myopathy, resulting in a greater degree of instability of motor-unit potential than is found in the uncomplicated inflammatory myopathies.

Pathophysiology of juvenile dermatomyositis

Vascular findings

The vasculopathy of this disease may occur in the absence of a prominent inflammatory component, and is a characteristic of children with dermatomyositis (Banker and Victor 1966); vascular occlusion in the absence of an inflammatory infiltrate in adults with early lesions of dermatomyositis has also been well described (Emslie-Smith and Engel 1990). In juvenile dermatomyositis, damage to capillaries, venules, and small arteries causes loss of the capillary network resulting in structural change in the nailfold capillary bed as well as in muscle, with a subsequent decrease in the capillary/fibre ratio. Appropriate controls must be age, gender, and activity related, for there is a progressive increase in mean fibre diameter, capillary/fibre ratio, and number of capillaries surrounding a single fibre, together with a decrease in capillary density (capillaries/mm ^{2} of muscle fibre) which was documented when data from normal infants, children, and teenagers were compared (Carry *et al*. 1986). In adults with dermatomyositis, both large and small vessels (less than 20 µm in diameter) are involved (Kalovidouris *et al*. 1988).

Muscle pathology

The muscle pathology in juvenile dermatomyositis reflects vascular compromise and capillary dropout, with perifascicular atrophy of both type I and type II fibres (Fig. 2). Multiple satellite cells are frequently seen in atrophic fibres; focal repair takes place concomitantly with fibre atrophy (Woo *et al*. 1988). In muscle biopsies from children aged 3 to 11 years with juvenile dermatomyositis, mitochondria in muscle appeared to be increased in number, but were reduced in size and in the total content of cytochrome C oxidase as measured both histochemically and biochemically. The biochemical changes were physically associated with perifascicular atrophy, suggesting that ischaemia could lead to structural change (Woo *et al*. 1988). Low-grade ischaemia may also be related to expression of class I and class II major histocompatibility complex gene products, which also were found primarily in the perifascicular area (Karpati *et al*. 1988). In some children with juvenile dermatomyositis who despite normal serum levels of muscle-derived enzymes have elevated levels of von Willebrand-factor antigen and/or neopterin (see below), the muscle biopsy can reveal active foci of inflammation (Pachman *et al*. 1996). In contrast, in polymyositis there appears to be less primary involvement of vessels—they may be normal in number and structure. It is not known if dermatomyositis in the adult has the identical pathophysiology as in the child, but comprehensive national studies in the United States are under way which should clarify this issue. In adult dermatomyositis, the infiltrate is composed of DR+ B cells and there is an increased CD4/CD8 ratio (it should be noted that few children have been extensively studied). It appears that informed discussion of specific immunohistopathology must include classification of the patients with respect to myositis-specific antibodies, as well as the duration of inflammatory disease, the presence of previous therapy at the time of biopsy, verification that the biopsy site was appropriate (for example, positive on MRI or ultrasound), data concerning the child's HLA phenotype with respect to the presence or absence of DQA1*0501, and the specific anatomic site of the biopsy under consideration (perivascular, endomesial, perifascicular, etc.). Overall, in dermatomyositis, there appears to be a close relationship between CD4+ cells and B cells as well as macrophages, suggesting a cytotoxic mechanism, perhaps directed against immune complex-modified endothelial cells (Engel and Arahata 1986; Hohlfeld and Engel 1994). The relationship of the tissue localization of cells from the immune system to those that are circulating is just beginning to be understood (see below). The various possible mechanisms of immunological injury have been discussed in Chapter 5.9.1, but this information is poorly defined for children at different stages of disease duration and activity.

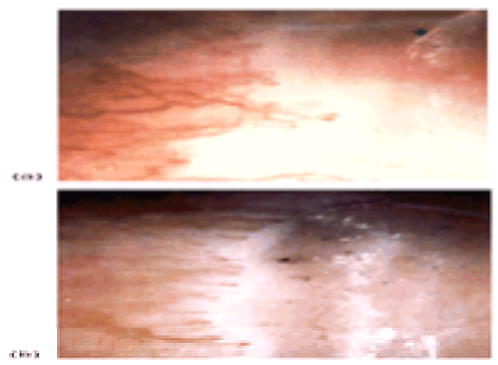


Fig. 2 (a) Nailfold capillary studies of a child with dermatomyositis of 2 years duration with severe active disease (not able to walk, flexion contractures) referred for evaluation. Her nailfolds display a prominent subcapillary venous plexus and marked avascularity with decreased numbers of end capillary loops. The few end capillaries that are present are tortuous and show terminal bushing characteristic of this disease. The black arrow indicated the edge of the nail. (b) The same nailfold 2 years and 4 months later: the child is now walking, attending school, but still requires intensive therapy. The prominent subcapillary venous plexus has been replaced by new capillary formation which has decreased the areas of avascularity. The end capillary loops are still not normal: there is increased dilation and open loop formation.

Calcification in soft tissues

The aetiology of the calcinosis must be determined and clearly differentiated from other syndromes in which calcinosis occurs, such as in heterotopic calcinosis or following trauma. In juvenile dermatomyositis, calcification in soft tissue may be a correlate of disease severity and duration (Pachman *et al*. 1985) Delay in diagnosis (Pachman *et al*. 1992b) and/or insufficient therapy (Callen *et al*. 1994) appears to be correlated with the development of calcinosis. Therefore, early recognition of the disease and institution of appropriate therapy can prevent the development of calcinosis (Callen *et al*. 1994). The therapeutic dilemma for the physician is to identify indicators of disease activity that are useful for an individual child. Those children with juvenile dermatomyositis who do develop calcifications may have one of several outcomes. The calcifications may resolve spontaneously, draining as a white, cheesy or serosanguinous exudate, and leaving dry pitted, scars. The calcifications do not decrease when chelating agents are given. In some children with persistent, active disease, the calcifications progress to become a sheath, impairing flexion and function, braking the barrier of the skin to form a site of entry of infection. Sepsis is not uncommon in this event and contributes to morbidity and mortality from this disease.

Calcinosis is accompanied by increased urinary excretion of g-carboxyglutamic acid (**GLA**), which is a component of the vitamin K-dependent coagulation pathway. This increased excretion of GLA (when compared with age-/sex-matched controls) was present both in children with juvenile dermatomyositis who had no clinical or radiological evidence of calcinosis (whose mean values for GLA excretion were twice normal) and in those who did have calcinosis (whose mean values were three time normal) (Lian *et al*. 1982). Routine measures of coagulation, including the prothrombin and activated partial thromboplastin times were normal (Lian *et al*. 1982). Calcifications are not correlated with antinuclear antibody, immune complexes, or class II HLA antigens (Pachman *et al*. 1985).

Immunological, genetic, and haematological data

Immunological data

Children with active untreated juvenile dermatomyositis are lymphopenic (O'Gorman *et al*. 1995). Despite this lymphopenia, there is a relative increase in the percentage of B cells (defined by anti-CD19 monoclonal antibody), which is correlated with a clinical disease activity score (in contrast, the percentages of CD4, CD8, and CD25 are not correlated with disease activity) (Eisenstein *et al*. 1995). This increased percentage of B cells may reflect response to medical therapy by returning to normal ranges at a time when other serological indicators of disease activity have already become normal (Pachman 1994). The CD4:CD8 cell ratio is increased, suggesting that there is a decrease in circulating CD8+ cells in the periphery (O'Gorman *et al*. 1995). Efforts to identify clonal expansion of T cells in the muscle of children with new-onset untreated juvenile dermatomyositis have not yet demonstrated a specific increase in T-cell receptor variable gene expression (V) as determined by polymerase chain reaction and sequencing of the CDR3 region in six children positive for the HLA antigen DQA1*0501 (Pachman *et al*. 1995a).

Humoral immunity may be abnormal in a minority of patients: in early disease, IgM (Pachman and Cooke 1980) or IgG (O'Gorman *et al*. 1995) may be elevated or IgA deficiency may be present (Pachman and Cooke 1980). Peripheral blood lymphocytes from children with juvenile dermatomyositis appear to have a high spontaneous rate of immunoglobulin synthesis *in vitro* (Cambridge 1990). Tests for an array of antibodies with tissue or organ specificity were negative when comparing sera from 89 patients with newly diagnosed juvenile dermatomyositis with 105 age-/sex-matched controls. Only the antinuclear antibodies (**ANA**) and antibody to the polymyositis antigen 1 (PM-1) were more frequent in patients than controls ($p<0.005$, $p< 0.001$ respectively) (Pachman *et al*. 1985; Montecucco *et al*. 1990). The ANA is speckled and cytoskeletal in pattern in 60 to 70 per cent of children (often in high titres), but negative for Jo-1 (Pachman *et al*. 1984). A similar speckled ANA was found in sera from children with a primary cardiomyopathy (who were positive on culture for CVB) and in CVB type-specific mouse monoclonal antibody, suggesting that a response to a viral antigen may play a role in juvenile dermatomyositis (Patterson *et al*. 1989). Other unconfirmed investigations identified an antibody specificity for heat shock protein 60 in the juvenile dermatomyositis sera (Patterson *et al*. 1993).

As in adults, identification of the ANA specificity in a person with an inflammatory myopathy may allow greater understanding of the aetiology and pathogenesis of theses heterogenous diseases (Love *et al*. 1991). Five children have been identified with Jo-1, the most common of the antisynthetases (Rider *et al*. 1995); in addition nine out of eleven other children had antibody to Mi-2 and their disease onset and course was the same as other ANA-positive, MSA-negative children; seven out of the nine were Hispanic, one was Asian, the other child was black. They had no obvious pattern in the season of onset of their weakness; fevers and joint contractures without associated arthritis were frequent in this group, but cardiopulmonary symptoms were uncommon (Rider *et al*. 1994). This was in contrast to the child with anti-SRP who had a sudden onset in the autumn of proximal and distal muscle weakness without the rashes characteristic of dermatomyositis, or cardiac involvement or palpitations prominent in adults with anti-SRP disease. Once her specific inflammatory myopathy was recognized, she was switched from corticosteroids (which worsened her disease) to cytotoxic agents and intravenous immunoglobulin to achieve remission (Rider *et al*. 1994).

Complement activation has been implicated in several studies which included children with juvenile dermatomyositis: the C5b–9 membrane-attack complex was localized to the intramuscular microvasculature in 10 of 12 patients (Kissel *et al*. 1986); the duration of the clinical disease was correlated with this finding (Kissel *et al*. 1991). Immune complexes appear to participate in the pathophysiology of this disease: there is data demonstrating complement activation (despite normal levels of total complement, C3 and C4) accompanied by increased levels of fibrinopeptide A and von Willebrand factor antigen (Scott and Arroyave 1987).

Another useful indicator of disease activity is neopterin, a member of the pteridine family derived from GTP via guanosinetriphosphate cyclohydrolase and released from macrophages as a consequence of T-cell dependent interactions (Barak *et al*. 1989) involving interferon-g. Early reports of increased neopterin levels in active dermatomyositis in children (Myones *et al*. 1989) were confirmed (DeBenedetti *et al*. 1993). Further studies demonstrated that neopterin levels were a correlate of a clinical disease activity score in over 65 per cent of cases (Pachman *et al*. 1995c).

Genetic studies

Records of juvenile dermatomyositis in more than one family member are sporadic, but the disease has been reported in monozygotic twins, who developed muscle-related abnormalities 2 weeks after an upper respiratory tract infection (Harati *et al*. 1986). In a large, cross-sectional study of children with juvenile dermatomyositis and their families, there was a marked increase in Caucasians with HLA B8 (relative risk 2.8, $p< 0.01$) (Friedman *et al*. 1983a) and DR3 (relative risk 3.8, $p< 0.01$); in Latin Americans, the relative risk for HLA DR3 was 18.5 ($p< 0.05$) (Friedman *et al*. 1983b). The association of the supratypes A1, Cw7, B8, and DR3 suggests that there might be a genetic component to disease susceptibility or expression. Further examination showed that there was an increased association of HLA DQA1*0501 but not C4a when 30 Caucasian children with juvenile dermatomyositis were compared with regional controls matched for HLA DR3 (Reed *et al*. 1991). This observation was sustained when other racial groups in the United States were studied (Reed and Stirling 1995), but did not appear to be true for Czech children with this disease (Vavrincova *et al*. 1993).

Haematological data

In children with juvenile dermatomyositis, the usual indicators of an acute-phase reaction are often within normal range, although children with acute severe disease or infected sites of calcinosis may have elevated values. The lymphopenia is not commonly accompanied by an abnormal platelet count, although a mild microcytic anaemia may be present. A sensitive indicator of inflammatory disease is an elevated von Willebrand factor antigen (Bowyer *et al*. 1989; Guzman *et al*. 1994) which may precede a disease flare (when muscle enzyme data is normal), or remain elevated once the enzymes have returned to normal (Scott and Arroyave 1987). The von Willebrand factor antigen is correlated with disease activity in some but not all children (Bloom *et al*. 1995). In over 50 per cent of children, levels of this antigen correlate with a clinical score of disease activity; analysis of the multimers confirms that it is endothelial in origin, reflecting endothelial cell damage (Miller *et al*. 1995).

These clues to disease activity—the MRI, neopterin, von Willebrand factor antigen, and percentage of B cells—are still imperfect guides to therapy (Pachman 1995), but may be of substantial aid in the characterization of the severity of the immunologically mediated inflammatory process (Pachman 1994).

Course and therapy

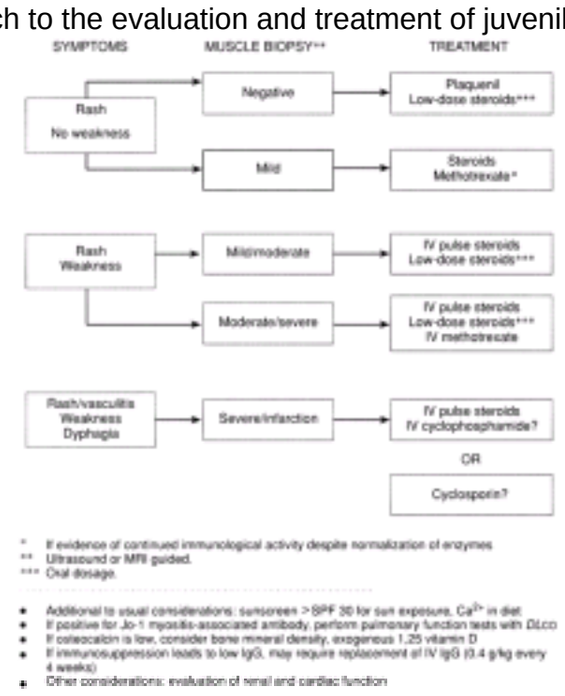
Course

The outcome of juvenile dermatomyositis has improved greatly since the 1960s when one-third of the children died, one-third were crippled, and the remainder recovered (Bitnum *et al*. 1964). Several types of disease course have been described—monocyclic, recurrent, and continuous (Spencer *et al*. 1984)—but they may be attenuated by early diagnosis and aggressive therapy. The frequency of calcinosis (which was associated with loss of mobility) has decreased from over 60 per cent of cases (Bowyer *et al*. 1983) to none (Callen *et al*. 1994). Late disease recurrence after years of apparent inactivity has been reported (Lovell and Lindsley 1986), suggesting the need for periodic monitoring with more sensitive indicators of disease activity. It is difficult to predict outcome at the onset of illness, although the magnitude of the initial creatine kinase appears to be a direct correlate of disease severity (Van Rossum *et al*. 1994). Several groups have found that prognosis is related directly to the degree of vascular involvement (Crowe *et al*. 1982; Bowyer *et al*. 1983).

Therapy

There is continuing controversy over the type, duration, and route of medication to be instituted (Malleson 1990). Recommendations are impaired because of lack of long-term outcome data; medical practices have changed in the recent past, and it is not known how this will affect the children's course in the future—both the consequences of the autoimmune disease and its therapy. Given these caveats, there is sparse data on long-term outcome. A summary of the treatment of juvenile dermatomyositis is given in Box 1.

Box 1 An approach to the evaluation and treatment of juvenile dermatomyositis



Several investigators have observed the utility of high-dose intravenous intermittent (pulse) methylprednisolone for the treatment of juvenile dermatomyositis over the past 2 decades (Miller 1980; Yanagisawa *et al*. 1983; Laxer *et al*. 1987). When children with similar clinical disease activity scores were compared, one group had been given 2 mg/kg per day of prednisone to treat their active disease, and required about 4 years until all therapy was discontinued. The other group was given intravenous intermittent high-dose methylprednisone at 30 mg/kg, supplemented with vitamin D and a calcium-sufficient diet, and low daily doses of about 0.5 mg/kg on days when they did not receive intravenous methylprednisolone. The children receiving the intravenous therapy had a shorter disease course with respect to persistence of rash (1.5 years compared with 3.9 years), weakness (1.5 years compared with 2.7 years), and did not have calcinosis or growth retardation, although the frequency of cataracts was the same in both groups (Pachman *et al*. 1994). When a subset of this group, who had a monocyclic disease course, was subjected to a cost analysis, the intravenously treated group had 2 disease-free years, but their bill was about $US10 000 higher than those given oral therapy (Klein-Gitelman *et al*. 1996). Cost differences associated with long-term outcomes are not known.

In our centre, the child with active juvenile dermatomyositis is thoroughly evaluated on admission, using the parameters appropriate to the child. Adequate hydration lessens the possibility of renal damage. If there is evidence of dysphagia or difficulty in handling secretions, then serious consideration is given to the immediate use of weekly intravenous methotrexate at a starting dose of 15 mg/m^2, to be administered immediately following the administration of high-dose intermittent pulse methylprednisolone at 30 mg/kg per day in 100 ml D5W to be given over at least 30 min with monitoring of vital signs every 15 min for 30 min after the infusion is completed. The frequency of the intravenous steroid administration is determined by the rate of response of the child, using the parameters that reflect that child's inflammatory response. Low-dose oral steroids, at 0.5 mg/kg per day are given in the morning on the days that the intravenous methylprednisolone is not infused. The protocol for each child is individual. In general, an intensive course of intravenous methylprednisolone is used until the laboratory tests become normal, with gradual reduction in therapy. The use of high-dose methylprednisolone is not without some side-effects. In an analysis of the drug usage over a 5-year period in which 213 children with various types of serious rheumatic disease were given over 2622 doses, 46 children (22 per cent) experienced an adverse reaction of which 21 had behavioural changes ranging from euphoria to emotional lability. There was one case of anaphylaxis (Klein-Gitelman and Pachman 1995).

When there is severe skin involvement, hydroxychloroquine (7 mg/kg per day) is given if there is no family history of red–green colour blindness. With milder involvement, the cutaneous symptoms often resolve within several weeks, making the use of this drug unnecessary. Topical agents to lessen dryness help the occasional pruritis as do topical steroids, which should be used sparingly (Stonecipher *et al*. 1993). For breaks in the integument, a 'skin substitute' (e.g. 'Second skin', duoderm) should be considered. Sepsis secondary to infected calcinosis must be treated aggressively.

Children who have severe onset, or who do not respond to steroids, have been treated with methotrexate for the past 2 decades (Jacobs 1977). More recently, earlier use of this drug at doses of 15 mg/m^2 per week has reduced the morbidity of the disease (Miller *et al*. 1992), and permits the use of lower doses of steroids. In active disease, intravenous administration of methotrexate assures drug absorption; as the disease becomes quiescent, oral administration is tolerated. Complaints of nausea can be circumvented by dividing the dose but giving the full amount in a 24-h period. The function of the liver and bone marrow must be monitored. If it appears that a child remains severely ill (sometimes despite the muscle enzymes becoming normal), evaluation of the immune system may help to guide therapy (e.g. the percentage of B cells remains elevated in 40 to 60 per cent) (Eisenstein *et al*. 1995). Intravenous cyclophosphamide therapy, starting at 500 mg/m^{2}every 3 weeks (following adequate hydration), with mesna for bladder protection, is instituted and the methotrexate is discontinued (Pachman 1994). As with most modes of immunosuppression, levels of IgG must be checked on a periodic basis to ensure that they are adequate; if not, replacement therapy (0.4 g/kg every 3 to 4 weeks) is needed to prevent recurrent infections.

When considering therapies other than steroids, high-dose (not replacement) intravenous gammagloblin may initially dampen the inflammatory process, especially the rash (Roifman *et al*. 1987; Lang *et al*. 1991) if given early in the disease course (Basta and Dalakas 1994), but it is unclear if prolonged control of disease activity can

be achieved with this modality alone. Plasmapheresis alone does not appear to be effective in adults (Miller *et al*. 1992) and no data is available for children. Evaluation of the efficacy of cyclosporin has been proposed (Heckmatt *et al*. 1989) but has been hampered by coexisting therapies (Pistoia *et al*. 1993). FK506 has been useful in the therapy of adults with Jo-1 myopathy (Oddis *et al*. 1994), but no published data are available for children.

At the moment there are no successful therapies for long-standing calcinosis in children with inactive disease. In those children who do have calcinosis and residual evidence of an active immunological process, an MRI may reveal more inflammation than expected, permitting more aggressive therapy, which may result in regression (and occasionally radiological resolution) of the calcinosis. Treatment with low-dose warfarin early in the disease has been suggested (Berger *et al*. 1987), but is a considerable hazard for the active child, is not useful in advanced disease (Lassoued *et al*. 1988), and has not yet been proved effective in a case–control study.

The progression of osteopenia, a consequence of disease activity as well as therapy, can be slowed with a calcium-sufficient diet, and the addition of thrice-weekly (Monday, Wednesday, Friday) 1,25-vitamin D (20 µg under 30 kg; 50 µg over 30 kg) in conjunction with disease control. Administration of vitamin D and increasing the calcium intake may aid in calcium absorption from the gastrointestinal tract in the face of steroid therapy (which inhibits calcium absorption), and diminishes the occurrence of one of the most serious consequences of steroid therapy, osteopenia, which can progress to spinal cord compression fractures (Callen *et al*. 1994).

Combined drug treatment (to suppress inflammation) and physiotherapy (gentle, passive stretching) are required in the early phase of the disease, and more intensive, graded physiotherapy is effective later in the disease, once the inflammation has abated. Prevention of sunburn, both by avoidance and barriers (clothing, UVA/UVB PABA-free sunblocks over SPF 30) helps keep the disease in remission.

In summary, juvenile dermatomyositis, characterized by genetically restricted and immunologically mediated vasculopathy, is under intense investigation. It is axiomatic that as more knowledge of the specific pathophysiology of the disease(s) is accrued, more effective therapeutic interventions will be devised.

This work was supported in part by Grant NIH/NIAMS P60-AR-30692 from the National Institutes of Health and grants from the Illinois Chapter of the Arthritis Foundation.

Chapter References

Ansell, B.A., Miller, J.J., III, Pachman, L.M., and Sullivan, D.B. (1990). Controversies in juvenile dermatomyositis. *Journal of Rheumatology*, **17**(Suppl. 22), 1–6.

Banker, B.Q. and Victor, M. (1966). Dermatomyositis (systemic angiopathy) of childhood. *Medicine*, **45**, 261–89.

Barak, M., Merzback, D., and Gruener, N. (1989). The effect of immunomodulators on PHA or IFN-gamma induced release of neopterin from purified macrophages and peripheral blood mononuclear cells. *Immunology Letters*, **21**, 317–22.

Basta, M. and Dalakas, M.C. (1994). High-dose intravenous immunoglobulin exerts its beneficial effect in patients with dermatomyositis by blocking endomysial deposition of activated complement fragments. *Journal of Clinical Investigation*, **95**, 1729–35.

Benbassat, J., Geffel, D., and Zlotnick, A. (1980). Epidemiology of polymyositis–dermatomyositis in Israel. *Israel Journal of Medical Science*, **16**, 197–200.

Berger, R.G., Featherstone, G.L., Raasch, R.H., McCartney, W.H., and Handler, N.M. (1987). Treatment of calcinosis universalis with low-dose warfarin. *American Journal of Medine*, **83**, 72–6.

Bitnum, C., Dawschnor, C.W., and Travis, L.B. (1964). Dermatomyositis. *Journal of Pediatrics*, **64**, 101–31.

Bloom, B.J., Tucker, L.B., Miller, L.C., and Schaller, J.G. (1995). Von Willebrand factor in juvenile dermatomyositis. *Journal of Rheumatology*, **22**, 320–5.

Bohan, A. and Peter, J.B. (1975). Polymyositis and dermatomyositis (parts 1 and 2). *New England Journal of Medicine*, **292**, 344–7, 403–7.

Borrelli, M.P. *et al*. (1988). Ureteral necrosis in dermatomyositis. *Journal of Urology*, **139**, 1275–7.

Bowles, N.E., Dubowitz, V., Sewry, C.A., and Archand, L.C. (1987). Dermatomyositis, polymyositis, and coxsackie-B-virus infection. *Lancet*, **ii**, 1004–7.

Bowyer, S.L., Blane, C.E., and Sullivan, D.B. (1983). Childhood dermatomyositis: factors predicting functional outcome and development of dystrophic calcification. *Journal of Pediatrics*, **103**, 882–8.

Bowyer, S.L., Ragsdale, C.G., and Sullivan, D.B. (1989). Factor VIII related antigen and childhood rheumatic diseases. *Journal of Rheumatology*, **16**, 1093–7.

Breningstall, G.N. (1990). Carnitine deficiency syndromes. *Pediatric Neurology*, **6**, 75–81.

Callen, A.M., Pachman, L.M., Hayford, J.R., Chung, A., and Ramsey-Goldman, R. (1994). Intermittent high-dose intravenous methylprednisolone (IV pulse) therapy prevents calcinosis and shortens disease course in juvenile dermatomyositis (JDMS). *Arthritis and Rheumatism*, **37**, R10.

Cambridge, G. (1990). What is the role of the immune system in juvenile dermatomyositis. In *Pediatric rheumatology update* (ed. P. Woo, P. White, and B. Ansell), p.182. Oxford University Press.

Carry, M.R., Ringel, S.P., and Starcevich, J.M. (1986). Distribution of capillaries in normal and diseased human skeletal muscle. *Muscle and Nerve*, **9**, 445–54.

Christensen, M.L., Pachman, L.M., and Maryjowski, M.L. (1983). Antibody to coxsackie-B virus: increased incidence in sera from children with recently diagnosed juvenile dermatomyositis. *Arthritis and Rheumatism*, **26**, S24.

Christensen, M.L., Pachman, L.M., Schneiderman, R., Patel, D.C., and Friedman, J.M. (1986). Prevalence of coxsackie B virus antibodies in patients with juvenile dermatomyositis. *Arthritis and Rheumatism*, **29**, 1365–70.

Commens, C., O'Neill, P., and Walker, G. (1990). Dermatomyositis associated with multifocal atrophy. *Journal of the American Academy of Dermatology*, **22**, 966–9.

Crowe, W.E., Love, K.E., Levinson, J.E., and Hilton, P.K. (1982). Clinical and pathogenetic implications of histopathology in childhood polydermatomyositis. *Arthritis and Rheumatism*, **25**, 126–39.

DeBenedetti, F., DeAmici, M., Aramini, L., Ruperto, N., and Martini, A. (1993). Correlations of serum neopterin concentrations with disease activity in juvenile dermatomyositis. *Archives of Diseases in Childhood*, **69**, 232–5.

Eisenstein, D.E., O'Gorman, M.R.G., and Pachman, L.M. (1995 *a*). Peripheral blood lymphocyte (PBLn) subsets in patients with juvenile dermatomoysitis (JDMS). *Arthritis and Rheumatism*, **38**, S361.

Eisenstein, D.M., O'Gorman, M.R.G., Donovan, M., and Pachman, L.M. (1995 *b*). Percentage of B cells in peripheral blood of patients with JDMS. *Arthritis and Rheumatism*, **38**, R15.

Emslie-Smith, A.M. and Engel, A.G. (1990). Microvascular changes in early and advanced dermatomyositis: a quantitative study. *Annals of Neurology*, **27**, 343.

Engel, A.G. and Arahata, K. (1986). Mononuclear cells in myopathies: quantitation of functionally distinct subsets, recognition of antigen-specific cell-mediated cytotoxicity in some diseases, and implications for the pathogenesis of the different inflammatory myopathies. *Human Pathology*, **17**, 704–21.

Fong, L.P. and Yeung, J. (1990). Spontaneous regression of retinal neovascularization in juvenile dermatomyositis. *Australian and New Zealand Journal of Ophthalmology*, **18**, 107–8.

Friedman, J.M. *et al*. (1983*a*). Immunogenetic studies of juvenile dermatomysitis: HLA antigens in patients and their families. *Tissue Antigens*, **21**, 45–9.

Friedman, J.M. *et al*. (1983*b*). Immunogenetic studies of juvenile dermatomyositis: HLA-DR antigen frequencies. *Arthritis and Rheumatism*, **26**, 214–16.

Guzman, J., Petty, R.E., and Malleson, P.N. (1994). Monitoring disease activity in juvenile dermatomyositis: the role of von Willebrand factor and muscle enzymes. *Journal of Rheumatology*, **21**, 739–43.

Harati, Y., Niakan, E., and Bergman, E.W. (1986). Childhood dermatomyositis in monozygotic twins. *Neurology*, **36**, 721–3.

Heckmatt, J.Z. *et al*.(1989). Effectiveness of cyclosporin for dermatomysitis. *Lancet*, **i**, 1063–6.

Hernandez, R.J., Sullivan, D.B., and Chenevert, T.L. (1993). MR imaging in children with dermatomyositis: musculoskeltal findings and correlation with clinical and laboratory findings. *American Journal of Roentgenology*, **161**, 359–66.

Hidano, A., Keneka, K., and Arai, Y. (1986). Survey of the prognosis for dermatomyositis with special reference to its associated malignancy and pulmonary fibrosis. *Journal of Dermatology*, **13**, 233–41.

Hiketa, T., Matsumoto, Y., Ohashi, M., and Sakaki, R. (1992). Juvenile dermatomyositis: a statistical study of 114 patients with dermatomyositis. *Journal of Dermatology*, **19**, 470–6.

Hochberg, M.C. (1988). Epidemiology of polymyositis/dermatomyositis. *Mount Sinai Journal of Medicine*, **55**, 447–52.

Hohlfeld, R. and Engel, A.G. (1994). The immunobiology of muscle. *Immunology Today*, **15**, 269–74.

Huppertz, H.I. and Kaiser, W.A. (1994). Serial magnetic resonance imaging in juvenile dermatomyositis—delayed normalization. *Rheumatology International*, **4**, 127–9.

Jacobs, J.C. (1977). Methotrexate and azathioprine treatment of childhood dermatomyositis. *Pediatrics*, **59**, 212–18.

Kalovidouris, A.E., Stoesz, E., Muller, J., and Kimes, T. (1988). Relationships between clinical features and distribution of mononuclear cells in muscle of patients with polymyositis. *Journal of Rheumatology*, **15**, 1401–6.

Karpati, G., Pouliot, Y., and Carpenter, S. (1988). Expression of immunoreactive major histocompatibility complex products in human skeletal muscles. *Annals of Neurology*, **23**, 64–72.

Keim, D.R., Hernandez, R.J., and Sullivan, D.B. (1991). Serial magnetic resonance imaging in juvenile dermatomyositis. *Arthritis and Rheumatism*, **34**, 1580–4.

Kissel, J.T., Mendell, J.R., and Rammohan, K.W. (1986). Microvascular deposition of complement membrane attack complex in dermatomyositis. *New England Journal of Medicine*, **314**, 329–34.

Kissel, J.T., Halterman, R.K., Rammohan, K.W., and Mendell, J.R. (1991). The relationship of complement-mediated microvasculopathy to the histologic features and clinical duration of disease in dermatomyositis. *Archives of Neurology*, **48**, 26–30.

Klein-Gitelman, M. and Pachman, L.M. (1995). IV pulse corticosteroids (CS): adverse reactions are more variable than expected in children. *Arthritis and Rheumatism*, **38**, S338.

Klein-Gitelman, M., Waters, T., and Pachman, L.M. (1996). A comparison of the cost effectiveness of IV and PO corticosteroids in the treatment of juvneile dermatomyositis (JDMS). *Arthritis and Rheumatism*, **39**, R13.

Koch, M.J., Brody, J.A., and Gillespie, M.M. (1976). Childhood polymyositis: a case-controlled study. *American Journal of Epidemiology*, **104**, 627–31.

Lang, B.A., Laxer, R.M., Murphy, G., Silverman, E.D., and Roifman, C.M. (1991). Treatment of dermatomyositis with intravenous gammaglobulin. *American Journal of Medicine*, **91**, 169–72.

Lapetina, F. (1989). Toxoplasmosis and dermatomyositis: a causal or casual relationship. *Pediatrica Medica e Chirurgica*, **11**, 197–203.

Lassoued, K., Saiag, P., Anglade, M., Roujeau, J., and Touraine, R.L. (1988). Failure of warfarin in treatment of calcinosis universalis. *American Journal of Medicine*, **84**, 795–6.

Laxer, R.M., Stein, L.D., and Petty, R.E. (1987). Intravenous pulse methylprednisolone treatment of juvenile dermatomyositis. *Arthritis and Rheumatism*, **30**, 328–34.

Leff, R.L., Miller, F.W., Greenberg, S.J., Klein, E.A., Dalakas, M.C., and Plotz, P.H. (1992). Viruses in idiopathic inflammatory myopathies: absence of candidate viral genomes in muscle. *Lancet*, **339**, 1192–5.

Lian, J.B., Pachman, L.M., Gundberg, C.M., Partridge, R.E.H., and Maryjowski, M.L. (1982). Gamma-carboxyglutamate excretion and calcinosis in juvenile dermatomyositis. *Arthritis and Rheumatism*, **25**, 1094–100.

Lott, J.A. and Landesman, P.W. (1984). The enzymology of skeletal muscle disorders. *Critical Review in Clinical Laboratory Sciences*, **20**, 153–90.

Love, L.A. *et al*. (1991). A new approach to the classification of idiopathic inflammatory myopathy: myositis-specific autoantibodies define useful homogeneous patient groups. *Medicine (Baltimore)*, **70**, 360–74.

Lovell, H.B. and Lindsley, C.B. (1986). Late recurrence of childhood dermatomyositis. *Journal of Rheumatology*, **13**, 821–2.

Malleson, P.N. (1990). Controversies in juvenile dermatomyositis. *Journal of Rheumatology*, **17**(Suppl. 22), 1.

Manta, P., Kalfakis, N., and Vassilopoulos, D. (1989). Evidence for seasonal variation in polymyositis. *Neuroepidemiology*, **8**, 262–5.

Martini, A., Ravelli, A., Viola, S., Sambugaro, R., and De Benedette, F. (1987). Calcinoisis as the presenting sign of juvenile dermatomyositis in a 14-month-old boy. *Helvetia Paediatrica*, **42**, 181–4.

Martini, A., Ravelli, A., Albani, S., Viola, S., Scotta, S., Magrini, U., and Burgio, G.R. (1992). Recurrent juvenile dermamtomyositis and cutaneous necrotizing arteritis with molecular mimicry between streptoccal type 5 protein and human skeletal myosin. *Journal of Pediatrics*, **121**, 739–42.

Masi, A.T. and Hochberg, M.C. (1988). Temporal association of polymyositis–dermatomyositis with malignancy: methodologic and clinical considerations. *Mount Sinai Journal of Medicine*, **55**, 471–8.

Medsger, T.A., Jr., Dawson, W.N., and Masi, A.T. (1970). The epidemiology of polymyositis. *American Journal of Medicine*, **48**, 715–23.

Mejlszenkier, J.D., Safran, A.E., and Healy, J.J. (1973). The myositis of influenza. *Arcives of Neurology*, **29**, 441–3.

Miller, C.H., Donovan, J.M., Maduzia, L., Chung, A., and Pachman, L.M. (1995). Relationship of von Willebrand factor antigen to disease activity in juvenile dermatomyositis. *Arthritis and Rheumatism*, **39**, R13.

Miller, F.W. *et al*.(1992). Controlled trial of plasma exchange and leukapheresis in polymyositis and dermatomyositis. *New England Journal of Medicine*, **326**, 1380–4.

Miller, J.J., III (1980). Prolonged use of large intravenous steroid pulses in the rheumatic diseases of children. *Pediatrics*, **65**, 989–94.

Miller, L.C., Michael, A.F., and Kim, Y. (1987). Childhood dermatomyositis. *Clinical Pediatrics*, **26**, 561–8.

Miller, L.C., Sisson, B.A., Tucker, L.B., DeNardo, B.A., and Schaller, J.G. (1992). Methotrexate treatment of recalcitrant childhood dermatomyositis. *Arthritis and Rheumatism*, **35**, 1143–9.

Montecucco, C., Ravelli, A., Caporali, R., Viola, S., De Gennerao, F., Albani, S., and Martini, A. (1990). Autoantibodies in juvenile dermatomyositis. *Clinical and Experimental Rheumatology*, **8**, 193–6.

Myones, B.L., Luckey, J.P., Hayford, J.R., and Pachman, L.M. (1989). Increased neopterin levels in juvenile dermatomyositis correlate with disease activity and are indicative of macrophage activation. *Arthritis and Rheumatism*, **52**, S83.

Oddis, C.V., Carroll, P., Abu-Elmagd, K., McCauley, J., Fung, J.J., and Starzl, T.E. (1994). FK506 in the treatment of polymyositis. *Arthritis and Rheumatism*, **37**, S286.

O'Gorman, M.R.G., Corrochano, V., Roleck, J., Donovan, M., and Pachman, L.M. (1995). Flow cytometric analysis of the lymphocyte subsets in pheripheral blood of children with untreated active juvenile dermatomyositis. *Clinical, Diagnostic and Laboratory Immunology*, **2**, 205–8.

Pachman, L.M. (1994). Juvenile dermatomyositis (JDMS): new clues to diagnosis and pathogenesis. *Clinical and Experimental Rheumatology*, **12**, S69–S73.

Pachman, L.M. (1995). Imperfect indications of disease activity in juvenile dermatomyositis. *Journal of Rheumatology*, **2**, 193–7.

Pachman, L.M. and Cooke, N. (1980). Juvenile dermatomyositis: a clinical and immunologic study. *Journal of Pediatrics*, **96**, 226–34.

Pachman, L.M. and Maryjowski, M.L. (1984). Juvenile dermatmyositis and polymyositis. In *Inflammatory diseases of muscle* (ed. B.A. Ansell), p.95. W.B. Saunders, Philadelphia, PA.

Pachman, L.M., Hardin, J.A., Cobb, M.A., and Arroyave, C.M. (1984). The antinuclear antibody (ANA) in juvenile dermatomyositis (JDMS) is not Jo-1, suggesting that JDMS and polymyositis (PM) are different diseases. *Arthritis and Rheumatism*, **27**, S45.

Pachman, L.M. *et al*. (1985). Immunogenetic studies of juvenile dermatomyositis. III. Study of antibody to organ-specific and nuclear antigens. *Arthritis and Rheumatism*, **28**, 151–7.

Pachman, L.M. *et al*.(1992*a*). Seasonal onset in juvenile dermatomyositis (JDMS): an epidemiological study. *Arthritis and Rheumatism*, **35**, S88.

Pachman, L.M., Hayford, J.R., Sinacore, J., Bowyer, S.L., and Hochberg, M.C. (1992 *b*). New onset juvenile dermatomyositis (JDMS): a clinical description. *Arthritis and Rheumatism*, **35**, S88.

Pachman, L.M., Callen, A.M., Hayford, J.R., Chung, A., Sinacore, J., and Ramsey-Goldman, R. (1994). Juvenile dermatomyositis (JDMS): decreased calclinoisis (Ca++) with intermittant high-dose intravenous methylprednisolone (IV pulse) therapy. *Arthritis and Rheumatism*, **37**, S429.

Pachman, L.M., Lawton, T.P., Litt, D.L., Corvera, J.S., and Pope, R.L. (1995 *a*). Lack of oligoclonality of T cell receptor (TCR) V 8 in muscle obtained by MRI directed biopsy of six children positive for HLA DQA1*0501 with active untreated juvenile dermatomyositis (JDMS). *Arthritis and Rheumatism*, **39**, R14.

Pachman, L.M. *et al*. (1995*b*). Lack of detection of enteroviral RNA or bacterial DNA in MRI directed muscle biopsies from twenty children with active untreated juvenile dermatomyositis. *Arthritis and Rheumatism*, **38**, 1513–18.

Pachman, L.M., Maduzia, L., Chung, A., Donovan, M., and Ramsey-Goldman, R. (1995 *c*). Juvenile dermatomyositis (JDMS): disease activity scores are correlated with levels of neopterin in serum. *Arthritis and Rheumatism*, **38**, R16.

Pachman, L.M., Sundberg, J., Maduzia, L., Daugherty, C., and Litt, D. (1995 *d*). Sequential studies of nailfold capillary vessels (NFC) in 10 children with juvenile dermatmyositis (JDMS): correlation with disease activity score (DAS) but not von Willebrand factor antigen (vWF:Ag). *Arthritis and Rheumatism*, **38**, S361.

Pachman, L.M., Crawford, S., Morello, F., Maduzia, L., Caliendo, J., and Heller, S. (1996). MRI directed needle biopsy for the assessment of juvenile dermatomyositis (JDMS) response to therapy: comparison of intitial and follow-up biopsies using a histological rating scate evaluating disease severity/chronicity. *Arthritis and Rheumatism*, **39**, R14.

Park, J.H. *et al*.(1994). Magnetic resonance imaging and P-31 magnetic resonance spectroscopy provide unique quantitative data useful in the longitudinal management of patients with dermatomyositis. *Arthritis and Rheumatism*, **37**, 736–46.

Patterson, B., Jacobitz, J., Pachman, L.M., Christensen, M., and Pallansch, M. (1989). A study of autoantibodies in coxsackievirus group B (COXB) infected children and in juvenile dermatomyositis (JDMS). *Clinical Research*, **30**, 964A.

Patterson, B.K. *et al*. (1993). Antinuclear antibody (ANA) positive sera from juvenile dermatomyositis (JDMS) has specificity for heat shock protein 60 (HSP-60). *Pediatric Research*, **33**, 157A.

Peters, A.M. *et al*. (1991). Renal haemodynamics of cyclosporin A nephrotoxicity in children with juvenile dermatomyositis. *Clinical Science*, **81**, 153–9.

Pistoia, V. *et al*. (1993). Cyclosporin A in the treatment of juvenile chronic arthritis and childhood polymyositis–dermatomyositis. Results of a preliminary study. *Clinical and Experimental Rheumatology*, **11**, 203–8.

Pitt, A.M., Fleckenstein, J.L., Greenlee, R.G., Jr., Burns, D.K., Bryan, W.W., and Haller, R. (1993). MRI-guided biopsy in inflammatory myopathy: initial results. *Magnetic Resonance Imaging*, **11**, 1093–9.

Reed, A.M. and Stirling, J.D. (1995). Association of the HLA-DQa1*0501 allele in multipe racial groups with juvenile dermatomysitis. *Human Immunology*, **44**, 131–5.

Reed, A.M., Haugen, M., Pachman, L.M., and Langman, C.B. (1990). Abnormalities in serum osteocalcin values in children with chronic rheumatic diseases. *Journal of Pediatrics*, **116**, 574.

Reed, A.M., Pachman, L.M., and Ober, C. (1991). Molecular genetic studies of major histocompatibility complex genes in children with juvenile dermatomyositis: increased risk associated with HLA-DQA1*0501. *Human Immunology*, **32**, 235–40.

Rider, L.G. *et al*.(1994). A broadened spectrum of juvenile myositis: myositis-specific autoantibodies in children. *Arthritis and Rheumatism*, **37**, 1534–8.

Rider, L.G. *et al*.(1995). Anti-Jo-1 autoantibodies define a clinically homogenous subset of childhood idiopathic inflammatory myopathy (IIM). *Arthritis and Rheumatism*, **38**, S362.

Roifman, C.M., Schaffer, F.M., Wachsmuth, S.E., Murphy, G., and Gelfand, E.W. (1987). Reversal of chronic polymyositis following intravenous immune serum globulin therapy. *Journal of the American Medical Association*, **258**, 513–15.

Rosenberg, A.M. (1994). Geographical clustering of childhood dermatomyositis in Saskatchewan. *Arthritis and Rheumatism*, **37**, S402.

Rosenberg, N.L., Rotbart, H.A., Abzug, M.J., Ringel, S.P., and Levin, M.J. (1989). Evidence for a novel picornavirus in human dermatomyositis. *Annals of Neurology*, **26**, 204–9.

Sachrison, B.T., Pachman, L.M., Kaplan, F.S., and Gannon, G.H. (1995). A 5 months old with progressive osseus heteroplasia (POH). *Arthritis and Rheumatism*, 38, R16.

Schiraldi, E. and Iandolo, E. (1978). Polymyositis accompanying coxsackievirus B2 infection. *Infection*, **6**, 32–4.

Schroter, H.M., Sarnet, H.B., Matheson, D.S., and Seland, T.P. (1987). Juvenile dermatomyositis induced by toxoplasmosis. *Journal of Child Neurolology*, **2**, 101–4.

Scott, J.P. and Arroyave, C. (1987). Activation of complement and coagulation in juvenile dermatomyositis. *Arthritis and Rheumatism*, **30**, 572–6.

Serratrice, G., Schiano, A., Pellissier, J.F., and Desnuelle, C. (1989). Les expressions anatomocliniques des pollymyosites chez l'enfant. *Annals Pediatrie (Paris)*, **36**, 237–43.

Sherry, D.D., Haas, J.E., and Milstein, J.M. (1993). Childhood polymyositis as a paraneoplastic phenomenon. *Pediatric Neurology*, **9**, 155–6.

Sigurgeirsson, B., Lindelöf, B., Edhag, O., and Allander, E. (1992). Risk of cancer in patients with dermatomyositis or polymyositis. A population-based study. *New England Journal of Medicine*, **326**, 363–7.

Silver, R.M. and Maricq, H.R. (1989). Childhood dermatomyositis: serial microvascular studies. *Pediatrics*, **83**, 278–83.

Spencer, C.H., Hanson, V., Singsen, B.H., Bernstein, B.H., Kornreich, H.K., and King, K.K. (1984). Course of treated juvenile dermatomyositis. *Journal of Pediatrics*, **105**, 399–408.

Stonecipher, M.R., Callen, J.P., and Jorizzo, J.L. (1993). The red face: dermatomyositis. *Clinical Dermatology*, **11**, 261–73.

Symmons, D.P.M., Sills, J.A., and Davis, S.M. (1995). The incidence of juvenile dermatomyositis: results from a nation-wide study. *British Journal of Rheumatology*, **43**, 732–5.

Thompson, G.T.D., Johnston, J.L., Barager, F.D., and Toole, J.W.P. (1990). Psoriatic arthritis and myopathy. *Journal of Rheumatology*, **17**, 395–8.

Travers, R.L., Hughes, G.R.V., Cambridge, G., and Sewell, J.R. (1977). Coxsackie B neutralization titers in polymyositis/dermatomyositis. *Lancet*, **i**, 1268.

Van Rossum, M.A.J., Hiemstra, I., Prieur, A.M., Rijkers, G.T., and Kuis, W. (1994). Juvenile dermato/polymositis: a retrospective analysis of 33 cases with special focus on initial CPK levels. *Clinical and Experimental Rheumatology*, **12**, 339–42.

Vavrincova, P., Havelka, S., Cerna, M., and Stastny, P. (1993). HLA class II alleles in juvenile dermatomyositis. *Journal of Rheumatology*, **20**(Suppl. 37), 17–18.

Wang, Y.-J., Lii, Y.-P., Lan, J.-l., Chi, C.-S., Mak, S.-C., and Scian, W.-J. (1993). Juvenile and adult dermatomyositis among the Chinese: a comparative study. *Chinese Medical Journal (Taipei)*, **52**, 285–92.

Woo, M., Chung, S.J., and Nonaka, I. (1988). Perifascicular atrophic fibres in childhood dermatomyositis with particular reference to mitochondrial changes. *Journal of Neurological Science*, **88**, 133–43.

Yanagisawa, T. *et al*.(1983). Methylprednisolone pulse therapy in dermatomyositis. *Dermatologica*, **167**, 47–51.

Yousef, G.E., Isenberg, D.A., and Mowbray, J.F. (1990). Detection of enterovirus specific RNA sequences in muscle biopsy specimens from patients with adult onset myositis. *Annals of the Rheumatic Diseases*, **49**, 310–15.

5.10 Sjögren's syndrome

Athanasios G. Tzioufas, Pierre Youinou and H. M. Moutsopoulos

Introduction

Sjögren's syndrome is a chronic autoimmune disease of unknown aetiology, characterized by lymphocyte infiltration of exocrine glands resulting in xerostomia and keratoconjunctivitis sicca. This syndrome is particularly interesting among the autoimmune diseases for two reasons. First, it has a broad clinical range extending from autoimmune exocrinopathy to extraglandular (systemic) disease affecting the lungs, kidneys, blood vessels, and muscles; it may occur alone (primary Sjögren's syndrome) or in association with other autoimmune diseases (secondary Sjögren's syndrome) (Table 1). Secondly, it is a disorder in which a benign autoimmune process can terminate in a lymphoid malignancy. Thus, Sjögren's syndrome is a 'cross-roads disease' that offers potential insight into the mechanisms whereby immunological dysregulation may predispose to malignant transformation of B cells that are already involved in an autoimmune process.

Rheumatoid arthritis
Systemic lupus erythematosus
Systemic sclerosis
Mixed connective tissue disease
Primary biliary cirrhosis
Myositis
Vasculitis
Thyroiditis
Chronic active hepatitis
Mixed cryoglobulinaemia

Table 1 Association of Sjögren's syndrome with other autoimmune diseases

Historical reviews (Moutsopoulos *et al*. 1980)

The clinical features of the disorder were first described by Hadden in 1888. Four years later, Mikulicz described the case of a German farmer who suffered from bilateral enlargement of the parotid glands. A biopsy of the parotid showed an intense, focal, lymphocytic infiltrate that is known today as the hallmark of the disease. In the 1920s Gougerot described the disease in France and in 1933, Sjögren, a Swedish ophthalmologist, wrote the classic monograph on the disease in which he emphasized that the eye manifestations are local findings of a systemic disorder. In 1953, Morgan and Castleman showed that the histopathological findings in Sjögren's syndrome and Mikulicz disease are identical. In the 1960s the diverse clinical range of the syndrome was recognized and the study of its autoantibodies was initiated. The genetic predisposition to the disease was substantiated with the study of HLA in the 1970s; in the 1980s, with progress in molecular biology, the specificity of autoantibodies to the cellular components Ro(SSA) and La(SSB), and the composition of the focal lymphocytic infiltration of the exocrine glands, were dissected.

Epidemiology

The syndrome primarily affects women (nine women to every one man), mainly in the fourth and fifth decades of life (Pavlidis *et al*. 1982). However, it can occur in people of all ages, including children and elderly persons (Drosos *et al*. 1988; Siamopoulou-Mavridou *et al*. 1989). Autopsy studies revealed that approximately 2 to 3 per cent of individuals without connective tissue disease had unexplained focal lymphocytic infiltrates of the minor labial salivary glands compatible with Sjögren's syndrome (Scott 1980). On the other hand, clinical studies from Great Britain and Greece have found that Sjögren's syndrome occurs in 3 to 5 per cent of geriatric populations (Whaley *et al*. 1972; Drosos *et al*. 1988).

In an epidemiological study of 705 Swedish adults the calculated prevalence for the disease was 2.7 per cent (Jacobsson *et al*. 1989). Symptoms of dry eyes and dry mouth were correlated with elevated levels of anti-Ro/SSA and anti-La/SSB antibodies (Jacobsson *et al*. 1992). In another study, using as a tool the recently proposed questionnaire and diagnostic criteria for Sjögren's syndrome, as suggested by the European Union concerted action (Vitali *et al*. 1992a), it was found that among 837 females (age range from 18 to 90 years), the prevalence of definite and probable Sjögren's syndrome was 0.6 per cent and 3 per cent, respectively (Tzioufas *et al*. 1994).

Aetiology and pathogenesis

Over the past two decades, research in immunopathology, autoantibodies, immunogenetics, and viruses has further refined the concepts of the pathophysiology and pathogenesis of Sjögren's syndrome. One may speculate that the disease develops in three steps: first, autoimmunity may be triggered by a given environmental factor that acts on a particular genetic background; second, the autoimmune reactivity becomes chronic through abnormal immune regulatory mechanisms; and third, the lesion occurs as a consequence of the continuing inflammatory process.

The triggering of autoimmunity

Environment

Autoimmune reactions against host tissue following viral infection have been reported in both man and experimental animals. Viruses are therefore suspected as being major contributing factors in certain autoimmune disorders. One chronic infection that is especially emphasized is with cytomegalovirus. This involves many organs; among them the salivary glands are prime targets (Hudson *et al*. 1979). Antibodies to cytomegalovirus of both IgG and IgM classes have been found in the serum of patients with primary Sjögren's syndrome, using an enzyme-linked immunosorbent assay (Shillitoe *et al*. 1982). However, no controls were used in that study

and Venables *et al*. (1985) were not able to confirm those observations.

Epstein–Barr virus (**EBV**), another herpesvirus, could well be a better candidate than cytomegalovirus. Its replication also occurs in the salivary glands during primary infection (Wolf *et al*. 1984). The virus remain latent at this site in immunologically intact adults. Expression of EBV-associated antigens in the salivary glands of some patients with Sjögren's syndrome and increased content of EBV-DNA in their saliva (Fox *et al*. 1986a) have been demonstrated. However, EBV reactivation may arise as a consequence rather than as a cause of lymphoproliferation in these patients.

Hepatitis C virus may produce a chronic lymphocytic sialadenitis, which mimics that observed in Sjögren's syndrome, since more than 50 per cent of patients infected with hepatitis C virus have been reported, in one study, to present with histological changes compatible with Sjögren's syndrome in their minor salivary glands (Haddad *et al*. 1992). On the other hand, patients with Sjögren's syndrome do not usually have antibodies to hepatitis C virus in their sera (Vitali *et al*. 1992b).

The retroviruses are another group of viruses that should be seriously considered in the pathogenesis of autoimmune diseases in general and Sjögren's syndrome in particular. Antibodies to the p24 capsid glycoprotein of human immunodeficiency virus have been detected in approximately 30 per cent of patients with Sjögren's syndrome, whilst the frequency of these antibodies in the serum of healthy, age-matched controls is 1 to 4 per cent (Talal *et al*. 1990). Transgenic mice bearing the *tax* gene of the human T-lymphotrophic virus type 1 develop an autoimmune exocrinopathy resembling that of Sjögren's syndrome: an initial increase and proliferation of the acinar epithelial cells is followed by a gradual infiltration of lymphocytes and plasma cells, leading to destruction of the acini (Green *et al*. 1989). Furthermore, a type A retroviral particle has been identified in extracts from labial gland biopsies in two out of six patients with Sjögren's syndrome after coculture with the lymphoblastoid cell line, RH9: this was distinguishable from the human immunodeficiency virus particles by several physicochemical and ultrastructural criteria (Garry *et al*. 1990).

The c-*myc* proto-oncogene is involved in the pathogenesis of B-cell malignancies and especially Burkitt's lymphoma caused by EBV. Increased expression of c- *myc* mRNA has been found in peripheral mononuclear cells of patients with Sjögren's syndrome as well as in those of normal people after stimulation (Boumpas *et al*. 1990). Skopouli *et al*. (1992), using *in situ* hybridization with specific c-*myc* probes, have demonstrated the expression of c-*myc* mRNA in minor salivary glands of patients with Sjögren's syndrome. The minor labial salivary glands of normal individuals and of patients with rheumatoid arthritis and sarcoidosis did not show this picture. Immunostaining of the hybridized tissue with monoclonal antibodies and correlation with the clinicoserological and histological findings showed that the proto-oncogene is expressed on the acinar epithelial cells and its appearance is correlated strongly with the duration of disease as well as with the intensity of the T-cell infiltration. It is not known yet whether this aberrant *c-myc* expression in the epithelial cells is a primary phenomenon resulting from a viral infection or an epiphenomenon attributable to the action of cytokines or cytokine-like molecules.

All the above data suggest viral involvement in Sjögren's syndrome. Transient or persistent infection of the epithelial cells by a putative virus may be the initiating event; accumulation of the helper/inducer memory T cells and B cells may be the second step; and monoclonal expansion of B cells under selective antigenic or T-cell-induced pressures the final step.

Genetic background (immunogenetics)

It is well known that members of the family of patients with Sjögren's syndrome have a higher prevalence of the syndrome and a higher incidence of serological autoimmune abnormalities than age- and sex-matched controls.

Numerous investigators have shown associations between primary Sjögren's syndrome and factors encoded by the major histocompatibility complex; HLA-DR3 has been reported in 50 to 80 per cent of patients with Sjögren's syndrome. The association, however, of HLA-DR3 and Sjögren's syndrome has been reported to be weaker than that of HLA-DR3 with Sjögren's syndrome and anti-Ro/SSA antibody positivity (Mann 1987).

As shown in Table 2 the ethnic origin of the patients, studied so far, influences the association with the HLA-DR phenotype. Given the linkage disequilibrium that exists between alleles of HLA loci, it is unclear whether the disease susceptibility is dependent on the associated allele or on a closely related gene. *HLA-DQ*, and to a lesser extent *HLA-DP*, alleles are tightly linked to *HLA-DR* (Navarrette *et al*. 1985). Thus, common amino acid sequences of the hypervariable region of these genes may be shared between different HLA-DR specificities, hence influencing disease susceptibility (shared epitope hypothesis) (Gregerson *et al*. 1987). The application of molecular biological techniques has made possible an understanding of the association of autoimmune diseases with different HLA haplotypes in various patient populations. In this regard, a DNA sequence-specific, oligonucleotide probe typing and a sequence analysis of Israeli Jewish and Greek non-Jewish patients with Sjögren's syndrome has been carried out (Tambur *et al*. 1993). It was found that the majority of patients in both groups presented either *DRBI* 1101* or *DRBI* 1104* alleles that were linked in a linkage disequilibrium with *DRBI*0301* and *DQA1*0501*. Molecular analysis of *DQB1* and *DQA1* alleles found in American Caucasian and American black patients with Sjögren's syndrome revealed high frequencies of *DQB1*0201* and *DQA1*0501* (Reveille *et al*. 1991). Therefore, the majority of patients with Sjögren's syndrome, independent of their racial and ethnic background, carry a common allele, the *DQA1*0501* allele. Furthermore, it has been shown that a glutamine residue at position 34 of the outermost domain of the *DQA1* and/or leucine at position 26 of the outermost domain of the *DQB1* chain have a 'gene dosage' role in the anti-Ro/SSA and anti-La/SSB antibody response. The *DQA1*0501* gene is one of the genes that possess glutamine at position 34 and is found in the majority of patients with anti-Ro/SSA and anti-La/SSB. Taken together, it appears that the DQA1*0501 molecule is probably an important determining factor for the predisposition of certain individuals to primary Sjögren's syndrome.

Ethnic group	HLA association	Reference
Caucasoid	HLA-B8	Fye *et al*. (1979)
	HLA-DW3	Chused *et al*. (1977)
	HLA-DR3	Mann (1987)
	HLA-DR3 and extragland.ular features	Vitali *et al*. (1992)
	HLA-DRw52	Mann (1987)
Greeks	HLA-DR5	Papasteriades *et al*. (1988)
Israelis	HLA-DR11 (subtype of DR5) in Israelis	Brautbar (personal communication)
Japanese	HLA-DRw53	Moriuchi *et al*. (1988)

Table 2 Ethnic variations in the association between antigens encoded by the major histocompatibility complex and Sjögren's syndrome

Development and continuation of the autoimmune process

The two major autoimmune phenomena observed in Sjögren's syndrome are the B-lymphocyte hyper-reactivity and the focal lymphoplasmacytic infiltrates in the exocrine glands. Numerous studies have sought to describe and delineate (a) the nature of these phenomena and (b) the mechanisms involved in their perpetuation.

Humoral studies

Polyclonal hyper-reactivity

The most common serological finding in Sjögren's syndrome is hypergammaglobulinaemia. The increased amount of immunoglobulins in these patients often contains a number of autoantibodies directed against non-organ-specific antigens such as other immunoglobulins (rheumatoid factor), antinuclear antibodies (which usually give a speckled pattern on immunofluorescence), cellular antigens [Ro(SSA), La(SSB), RANA], and organ-specific antigens such as salivary ductal cells, thyroid gland cells, and gastric mucosa (Harley 1987). The most common autoantibodies to cellular antigens in patients with Sjögren's syndrome are directed against two ribonucleoprotein antigens known as Ro or SSA and La or SSB. These autoantibodies are not specific for the syndrome and may be found in other autoimmune diseases, especially systemic lupus erythematosus (see Chapter 5.7.1). Anti-Ro and anti-La are detected by immunodiffusion in approximately 45 and 20 per cent, respectively, of patients with Sjögren's syndrome but in up to 95 and 85 per cent, respectively, by more sensitive techniques such as enzyme-linked immunosorbent assay. Characterization of these ribonucleoproteins has led to the observation that the fine specificity of antibodies to the polymorphic forms of Ro differs in Sjögren's

syndrome and systemic lupus (reviewed in Chapter 4.5). Thus, autoantibodies to the 52-kDa Ro protein are frequently found in serum from patients with the syndrome, while antibodies to a 60-kDa Ro protein are found more often in serum from patients with systemic lupus (Ben Chetrit *et al*. 1990).

The presence of anti-Ro(SSA)/La(SSB) autoantibodies is associated with certain clinical manifestations of primary Sjögren's syndrome: they are correlated with earlier onset and longer duration of disease, recurrent enlargement of the parotids, and with splenomegaly/lymphadenopathy and vasculitis. In addition, the incidence of these antibodies correlates with the intensity of the infiltration of minor salivary glands (Manoussakis *et al*. 1986).

Oligomonoclonal hyper-reactivity

Several patients with primary Sjögren's syndrome have been shown to have circulating monoclonal immunoglobulins. In a study of serum and urine from unselected patients with the primary syndrome, using high-resolution agarose electrophoresis combined with immunofixation, approximately, 80 per cent of those with extraglandular (systemic) disease had monoclonal light chains or immunoglobulins in the serum. Furthermore, all patients excreted monoclonal light chains in the urine. In contrast, only one-quarter of patients with disease limited to the exocrine glands had monoclonal light chains or immunoglobulins in their serum, while only 43 per cent of these excreted light chains in the urine (Moutsopoulos *et al*. 1983a; Moutsopoulos *et al*. 1985).

Subsequent analysis of cryoglobulins from patients with Sjögren's syndrome by high-resolution agarose gel electrophoresis combined with immunofixation demonstrated that these are mixed monoclonal cryoglobulins (type II), containing an IgM-k monoclonal rheumatoid factor (Tzioufas *et al*. 1986) (Fig. 1). The above data suggest that patients with Sjögren's syndrome express circulating monoclonal immunoglobulins very early in the disease, together with the polyclonal B-cell activation. Monoclonality is observed more often in those patients with systemic, extraglandular disease. This is of particular interest, as patients with extraglandular manifestations are at higher risk for developing lymphoid malignancy. In this regard, serial follow-up studies have shown that the presence of mixed monoclonal (type II) cryoglobulinaemia correlates with lymphoma in patients with primary Sjögren's syndrome (Tzioufas *et al*. 1996).

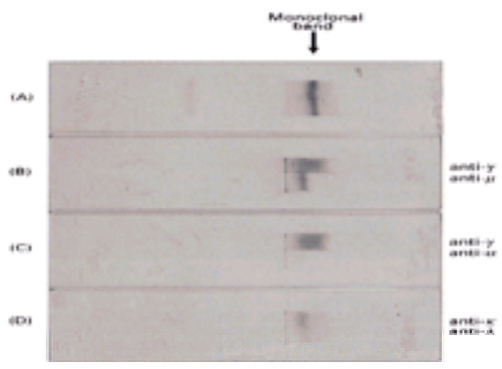


Fig. 1 High-resolution agarose gel electrophoresis of the cryoglobulins of a Sjögren's syndrome patient reveals a monoclonal band (A). Immunofixation, using anti-human k, l light and a, g, μ heavy chains, identified the monoclonal bands as an IgM-k monoclonal immunoglobulin (B, C, and D).

Schmidt *et al*. (1982) suggested that the benign lymphoepithelial lesion of salivary glands in Sjögren's syndrome, which has areas of confluent lymphoid proliferation and contains plasma cells that have monoclonal IgM-k immunoglobulins in their cytoplasm, represents an ' *in situ*' malignant lymphoma. This suggestion was further substantiated by immunogenotypic and immunophenotypic studies (Fishleder *et al*. 1987; Moutsopoulos *et al*. 1990), which showed that in the lymphoepithelial lesion there were rearrangements of oligo- or monoclonal immunoglobulin genes. In one study, three patients with benign lymphoepithelial lesion developed non-Hodgkin's lymphoma 2 to 8 years after the initial biopsy (Freimark *et al*. 1989). The presence of circulating monoclonal immunoglobulins is associated with a monoclonal B-cell expansion in the salivary glands. This has been demonstrated by a peroxidase–antiperoxidase bridge technique for the detection of intracytoplasmic immunoglobulins in the salivary lymphocytic infiltrates. Seven of 12 patients with Sjögren's syndrome with cryoprecipitable IgM-k monoclonal immunoglobulins had a predominance of k-positive plasma cells in the minor salivary glands, while patients without cryoglobulins or with polyclonal cryoglobulins had almost equal numbers of k- and l-positive cells (Moutsopoulos *et al*. 1990).

Monoclonal rheumatoid factors share cross-reactive idiotypes. Monoclonal rheumatoid factors from patients with Waldenström's macroglobulinaemia have been divided into three groups (Wa, Po, and Bla) (Kunkel et al.1973; Agnello *et al*. 1980). The light chains of the monoclonal rheumatoid factor that reacted with anti-Wa antibodies belong to the V_{IIIb} subgroup (Kipps *et al*. 1987). The 17.109 monoclonal antibody reacted with half of the monoclonal rheumatoid factors. The 17.109 idiotype is associated with the expression of the Humkv325 germ-line gene (Kipps *et al*. 1989). Cross-reactive idiotypes are shared by rheumatoid factors from autoimmune diseases. Some of them are shared by patients with Sjögren's syndrome and rheumatoid arthritis, while others (such as 17.109), have been found only in rheumatoid factors from patients with Sjögren's syndrome (Fox *et al*. 1986b). In fact, 12 out of 15 monoclonal rheumatoid factors from patients with Sjögren's syndrome reacted with the 17.109 monoclonal antibody. Furthermore, B cells containing immunoglobulin, bearing the 17.109 idiotype, were detected in the salivary gland biopsies of 11 out of 12 patients with Sjögren's. Further analysis of the B cells bearing the 17.109 idiotype in the salivary glands of patients with primary Sjögren's syndrome has shown that these are of multiclonal origin, in which somatic mutations accumulate in a non-random fashion, strongly suggesting an antigen- or T-cell-driven process in the expansion of these cells (Kipps *et al*. 1989). Two-thirds of patients with primary Sjögren's share a common idiotype, detected with a polyclonal anti-idiotype antibody directed against an IgM-k monoclonal rheumatoid factor from patients with Sjögren's syndrome. The presence of this idiotype is correlated with extraglandular manifestations, autoantibodies, and monoclonal immunoglobulins (Katsikis *et al*. 1990).

Comparative studies of three cross-reactive idiotypes in patients with primary Sjögren's syndrome have shown that these are expressed very early in the clinical spectrum of the syndrome and that their prevalence increases as the disease evolves. The presence of 17.109 and G6 (a VH1-associated cross-reactive idiotype), were found mainly in patients with Sjögren's syndrome with lymphoma, suggesting that these idiotypic determinants may serve as markers for lymphoma development in the syndrome (Tzioufas *et al*. 1996).

Lymphocyte studies

Peripheral blood lymphocytes

The absolute number of the total lymphocytes as well as T and B cells in peripheral blood does not differ substantially from that observed in normal individuals. Studies of T-lymphocyte subsets in Sjögren's syndrome were inconclusive (Fauci and Moutsopoulos 1981). Although decreased numbers of CD4+ and CD8+ T cells have been reported, this finding has not been substantiated by other investigators (Moutsopoulos and Manoussakis 1989). Peripheral blood B lymphocytes from patients with Sjögren's syndrome and normal controls, unlike lymphocytes from patients with systemic lupus, did not spontaneously secrete increased amounts of immunoglobulins. Thus, the activated B cells in patients with systemic lupus are distributed widely but in patients with Sjögren's syndrome they are probably localized to, and infiltrate, organs such as the minor labial salivary glands or the spleen.

Tissue lymphocytes

During the 1970s there some efforts were made to determine the composition of the lymphocytic infiltrates in the salivary glands but the results were controversial (Moutsopoulos and Talal 1987).

The application of molecular biological techniques, including the use of monoclonal antibodies and nucleic acid probes, heralded a more precise understanding of the immunopathological lesion in the exocrine glands of patients with Sjögren's syndrome. Several studies using monoclonal antibodies against specific lymphocyte markers have evaluated the composition of round-cell infiltrates in the labial salivary glands of primary Sjögren's syndrome. It was shown that the majority of the infiltrating lymphocytes are T cells, while B lymphocytes constitute 20 to 25 per cent of the round cells. Monocytes, macrophages, as well as natural killer cells, are less that 5 per cent (Moutsopoulos and Talal 1987). Studies of T-lymphocyte subpopulations have shown that 60 to 70 per cent of the T lymphocytes bear the CD4 phenotype and that the majority of them exhibit the memory/inducer marker (CD45 Ro). Almost all infiltrating T cells express the a T-cell receptor (**TCR**) (Skopouli *et al*. 1991). Analysis of the receptor repertoire of the infiltrating T lymphocytes from minor salivary-gland biopsies of patients with Sjögren's syndrome by a quantitative

polymerase chain reaction revealed that the repertoire of the *TCR V* gene was not restricted, although V_{b2} and V_{b13} were predominantly expressed in the inflammatory infiltrates (Sumida *et al*. 1992). Interestingly, V_{b2}- and V_{b13}-positive T cells can be stimulated by minor lymphocyte determinants of bacterial toxins, the so-called superantigens (Kappler *et al*. 1989). Thus, exposure to these molecules (e.g. after a bacterial infection) may lead to the stimulation and expansion of T cells expressing these two genes. In this regard, the junctional sequences of cDNA encoding the V_{b2} and V_{b13} genes of TCR from T lymphocytes infiltrating the minor salivary glands of patients with Sjögren's syndrome were investigated (Yonaha *et al*. 1992). Despite the fact that V_{b2}- and V_{b13}-positive T cells were polyclonal, the junctional usage was found to be restricted, supporting the notion that autoreactive T cells that contribute to the immunopathological lesion in Sjögren's syndrome are of oligoclonal origin.

The activation status of T cells was evaluated by searching for membrane expression of HLA class II molecules, interleukin-2 (**IL-2**) receptor (**r**), the lymphocyte function-associated antigen 1, and IL-2 production (Moutsopoulos and Talal 1987 ; Skopouli *et al*. 1991). Although none of the tissue lymphocytes was positive for IL-2 and IL-2r when monoclonal antibodies were used, studies using *in situ* hybridization with oligonucleotide mRNA probes demonstrated both IL-2 and its receptor in the infiltrating lymphocytes of the labial salivary glands of patients with primary Sjögren's syndrome (Boumba *et al*. 1995).

B lymphocytes infiltrating the labial salivary glands are activated, since they are able to produce increased amount of immunoglobulins with autoantibody activity (Anderson *et al*. 1972). In addition, evaluation with an immunoperoxidase technique of the isotypes of intracytoplasmic immunoglobulins of the plasma cells infiltrating the salivary glands of patients with Sjögren's syndrome showed that the IgG and IgM isotype predominates, in contrast to the plasma cells of the normal salivary glands, where the IgA isotype is dominant (Lane *et al*. 1983). This observation prompted some investigators to support the notion that quantitation of cells containing IgA and IgG intracytoplasmic immunoglobulins may serve as diagnostic criterion with high specificity and sensitivity for Sjögren's syndrome (Bodeuitsch *et al*. 1992).

B-cell activation is one of the most prominent immunoregulatory aberrations in patients with Sjögren's syndrome. This aberration can follow various stages of evolution. It begins as polyclonal activation, evolves to polyclonal–oligoclonal–monoclonal activation, and ends up as malignant monoclonal proliferation.

Immunopathology

As clearly shown in the previous sections, both the B and the T lymphocytes that contribute to the tissue lesion of primary Sjögren's syndrome are activated. This is of particular interest since the classic antigen-presenting cells, monocytes/macrophages, are poorly represented in this lesion. These findings suggest that possibly another cell may play the part of the antigen presenter in the immunopathological lesion of Sjögren's syndrome. Several recent studies suggest that the glandular or acinar epithelial cells may be play that part; these are summarized in the following list.

1. Histopathological studies in newly diagnosed cases of Sjögren's syndrome reveal that the focal lymphocytic infiltrates start around the ducts.
2. Staining of the labial salivary glands with anticlass-II HLA monoclonal antibodies showed that the ductal and acinar epithelial cells inappropriately express these molecules (Skopouli *et al*. 1991). Interferon-g and tumour necrosis factor-a have been shown to up-regulate the expression of both histocompatibility antigen classes on the surface of epithelial and other cells. These cytokines are produced locally by the activated T cell. Therefore it is not known whether the HLA-DR expression and possible antigen presentation by epithelial cells predates, or is a consequence of, the lymphocytic infiltration.
3. Studies on the expression of proto-oncogene mRNA in the minor salivary glands of patients with Sjögren's syndrome revealed that c- *myc*, in contrast to c-*fos* and c-*jun*, is selectively expressed by the epithelial glandular cells (Skopouli *et al*. 1992). Since the expression of the c-*myc* is so restricted, this phenomenon cannot be attributed to microenvironmental factors. The expression of HLA class II antigen and c- *myc* by epithelial cells may indicate a specific way of activating these cells.
4. Acinar epithelial cells coexpress accessory adhesion molecules and autoantigens, which in conjunction with the expression of class II antigen may potentially prime an autoimmune response. In fact, translocation and membrane localization of the nuclear antigen La/SSB has been observed in conjunctival epithelial cells of patients with Sjögren's syndrome (Yannopoulos *et al*. 1992). In addition, the infiltrating lymphocytes express a diverse array of cell adhesion molecules (lymphocyte function-associated antigen 1 and 3, CD2), while the intercellular adhesion molecule 1 was expressed on acinar epithelial cells adjacent to sites of intense inflammation (St Clair *et al*. 1992).
5. In recent years, cytokine production in the immunopathological lesion of Sjögren's syndrome has been studied extensively by *in situ* hybridization and reverse-transcriptase, quantitative, polymerase chain reaction (Fox *et al*. 1994; Boumba *et al*. 1995). Both techniques demonstrated the presence of proinflammatory cytokines IL-1 and IL-6 in the labial salivary glands of patients with Sjögren's syndrome. These cytokines were produced by both infiltrating and epithelial cells, which reinforces the concept that epithelial cells are active counterparts in the inflammatory response rather than targets of the immune-mediated injury.

Pathology

The common finding in all affected organs in patients with Sjögren's syndrome is a potentially progressive lymphocytic infiltration. These infiltrates cause functional disability of the affected organs, producing the various clinical manifestations.

So far, the salivary glands are the best-studied organs, because (i) they are affected in almost all patients and (ii) they are readily accessible. Microscopic examination of the enlarged major salivary glands reveals a benign lymphoepithelial lesion, characterized by lymphocytic replacement of the salivary epithelium and the presence of epimyoepithelial islands, which are composed of keratin-containing epithelial cells. Sometimes, the salivary gland biopsy does not show benign lymphoepithelial lesions, but instead contains various degrees of focal lymphocytic infiltration [defined as focal aggregates of 50 or more lymphocytes and histiocytes (Daniels *et al*. 1987)]. It is not known whether the focal lymphocytic infiltration is a precursor of the benign lymphoepithelial lesion.

The need for a practical and easy way to assess the salivary component of Sjögren's syndrome led to the introduction of labial gland biopsy. The histopathological characteristics of the biopsy of the minor salivary glands include (i) focal aggregates of at least 50 lymphocytes/plasma cells and macrophages, adjacent to and replacing the normal acini, and (ii) the consistent presence of these foci in all or most of the glands in the specimen (Fig. 2). Larger foci often show the formation of germinal centres but epimyoepithelial islands are very uncommon. These pathological lesion are, in fact, typical findings for a chronic lymphocytic sialadenitis (Batsakis and Howard 1982). However, a biopsy of minor salivary glands can be very specific for Sjögren's syndrome if it is obtained through normal-appearing mucosa, includes 5 to 10 glands, separated from the surrounding connective tissue, and shows focal lymphocytic infiltrates in all or most of the glands in the specimen, with a focus score above a chosen diagnostic threshold. Therefore, several methods of scoring the number of foci have been applied. Chisholm and Mason (1968) used a semiquantitative method to assess inflammation in salivary gland biopsies from 40 patients with several autoimmune diseases; only in patients with Sjögren's syndrome was there more than one focus of lymphocytes per 4 mm^2 of gland. Using a modification of this method, Greenspan *et al*. (1974) enumerated scores from 1 to 12 foci/4 mm^2 and found a significant positive correlation between a higher score and larger foci in biopsies of minor salivary glands of patients with Sjögren's syndrome. In another study, grading of 86 biopsies from primary and secondary Sjögren's syndrome showed, by qualitative criteria, larger lymphocytic foci in the primary syndrome (Tarpley *et al*. 1974). Biopsies from labial and sublingual salivary glands have been compared: there was a better correlation between infiltration of the ductal structure and the focus score in the sublingual glands (Pennec *et al*. 1990).

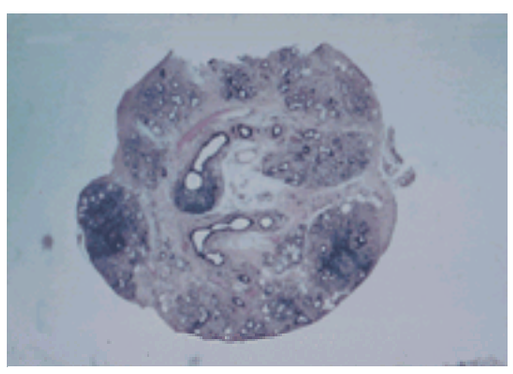


Fig. 2 Minor salivary gland biopsy of a patient with Sjögren's syndrome, showing moderate and large focal lymphocytic infiltrates around the acini and ducts.

In a recent study, comparison of patterns in the minor salivary gland biopsy taken from 618 patients with keratoconjunctivitis sicca and suspected Sjögren's syndrome showed that suspected Sjögren's syndrome was better associated with focal lymphocytic sialadenitis rather than chronic sialadenitis or even xerostomia (Daniels and Whitcher 1994).

Although there is no perfect diagnostic criterion for the salivary component of Sjögren's syndrome, the finding of the characteristic focal sialadenitis in biopsies of minor salivary glands is the best single criterion in terms of its disease specificity, convenience, availability, and low risk.

Animal models

There are several animals in which clinical manifestations resembling Sjögren's syndrome can be experimentally or spontaneously induced. Experimental models of sialoadenitis have been produced in rats, guinea-pigs, and mice using salivary gland extracts, adjuvants, and allogeneic antisera (Hoffman and Walker 1987). Infection of rats with a specific coronavirus results in swelling of the neck and ophthalmic lesions (Innes and Stanton 1961). Several days after exposure, the affected rats present with necrosis of ductal epithelial cells and infiltrates consisting of histiocytes, lymphocytes, and polymorphonuclear leucocytes. These findings are more prominent in lacrimal glands; submaxillary and parotid salivary glands are less commonly affected. Viral particles can be found by electron microscopy in ductal epithelial cells.

Spontaneously induced disease with features resembling those of human Sjögren's syndrome has been recognized in autoimmune strains of mice that develop a lupus-like syndrome. These include the NZB, NZB/NZW, and MRL/1pr, MRL/n (lacking the *lpr* gene) strains. All experimental animals have various degrees of lymphoplasmacytic infiltration in the lacrimal and salivary glands, with the milder form in the NZB mice and the more severe form in the MRL/n mice. Antibodies to Ro(SSA) and/or La(SSB) autoantigens were not detected in their sera.

Other experimental models of 'Sjögren's-like disease' include the canine keratoconjunctivitis sicca and the chicken dysgammaglobulinaemia that is associated with Coombs'-positive haemolytic anaemia, cryoglobulins, rheumatoid factor and infiltration of several parenchymal organs with mononuclear cells (Hoffman and Walker 1987).

In the last few years, the lymphoproliferation in Sjögren's syndrome has been studied in **SCID** (severe combined immunodeficient) mice. The CB17 scid/scid mice are born with severe combined immunodeficiency, lacking mature T and B lymphocytes. Injection of peripheral blood mononuclear cells from anti-La(SSB)-positive patients with Sjögren's syndrome into these mice resulted in the development of lymphoid infiltrates in several tissues consistent with disseminated lymphoid neoplasia (Whittingham *et al*. 1991).

In another study, after implantation of minor salivary gland biopsies under the kidney capsules of SCID mice, the proportion of human CD4+ T cells gradually decreased while the number of CD19+ B cells increased. The animals died after 6 to 12 weeks with human lymphoid tumours (Kang *et al*. 1991). These results suggest that, in an immunoincompetent environment, B cells of patients with Sjögren's syndrome tend to proliferate and develop malignancy.

Clinical picture

Sjögren's syndrome can occur alone (primary) or in association with other autoimmune diseases (secondary). In most patients the primary syndrome runs a rather slow and benign course. The initial manifestation can be non-specific (Table 3) and usually 8 to 10 years elapse from the initial symptoms to the full-blown development of the syndrome (Pavlidis *et al*. 1982).

	n	%
Sicca manifestations:		
Subjective xerophthalmia	63	47.0
Subjective xerostomia	56	42.5
Parotid gland enlargement	24	24.0
Dyspareunia	7	5.0
Fever/fatigue	13	10.0
Arthralgias, arthritis	37	28.0
Raynaud's phenomenon	28	21.0
Lung involvement	2	1.5
Kidney involvement	2	1.5

Table 3 Initial manifestations of primary Sjögren's syndrome (data from 132 Greek patients with Sjögren's syndrome)

Glandular manifestations

Oral component

Symptoms and signs

The principal oral symptom of Sjögren's syndrome is dryness (xerostomia). Patients describe this as difficulty in swallowing dry food, inability to speak continuously, changes in sense sensation, a burning sensation, an increase in dental caries (Fig. 3), and problems in wearing complete dentures. Examination shows a dry, erythematou **s**, sticky oral mucosa and often dental caries, while saliva from the major glands is either not expressible or is cloudy. Atrophy of the filiform papillae of the tongue is apparent (Fig. 4) and in some cases overgrowth of *Candida* is observed.

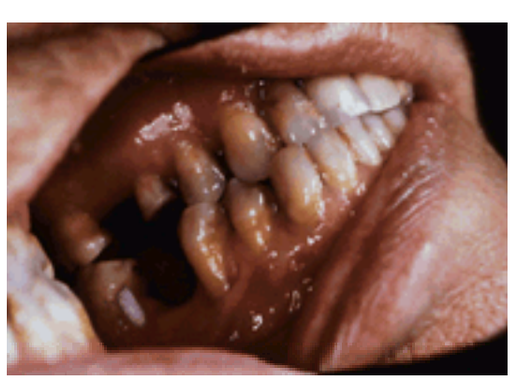


Fig. 3 Dental caries in Sjögren's syndrome; note also the remarkable periodontitis in both lower and upper teeth.

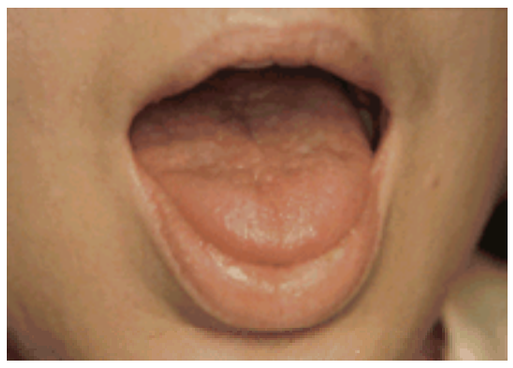


Fig. 4 Dry tongue of a patient with Sjögren's syndrome. Note the remarkable atrophy of filiform papillae.

Enlargement of the parotids or other major salivary glands occurs in two-thirds of patients with the primary syndrome, but is uncommon in those with Sjögren's syndrome and rheumatoid arthritis (see below). In many patients the salivary gland enlargement is episodic, but some have chronic enlargement. The swelling of the parotids may begin unilaterally but often becomes bilateral (Fig. 5). In Table 4, conditions other than Sjögren's syndrome causing parotid enlargement are depicted.

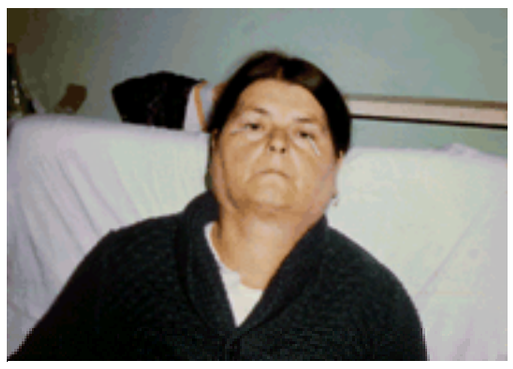


Fig. 5 Bilateral parotid-gland enlargement in a 65-year-old patient with Sjögren's syndrome.

Unilateral
Salivary gland neoplasm
Bacterial infection
Chronic sialadenitis
Bilateral
Viral infection (mumps, influenza, Epstein–Barr, coxsackie A, cytomegalovirus, HIV)
Sjögren's syndrome
Sarcoidosis
Miscellaneous (diabetes mellitus, hyperlipoproteinaemia, hepatic cirrhosis, chronic pancreatitis, acromegaly, gonadal hypofunction)
Recurrent parotitis of childhood

Table 4 Differential diagnosis of parotid gland enlargement

Diagnosis

A variety of medical conditions other than Sjögren's syndrome can cause xerostomia (Table 5). To evaluate the oral component of the syndrome, various tests are used with different specificity and sensitivity for the disease.

Drugs:
Psychotherapeutic
Parasympatholytic
Antihypertensive
Viral infections
Dehydration:
Diabetes mellitus
Trauma
Psychogenic
Irradiation
Congenital (absent or malformed glands)

Table 5 Causes of xerostomia

Sialometry

Salivary flow rates can be measured clinically for whole saliva or for separate secretions from the parotid or submandibular and sublingual glands, with or without stimulation. Patients with clinically overt Sjögren's syndrome have reduced flow. However, flow rates depend on many factors such as age, sex, medication, and time of day. Therefore, setting a cut-off point between the normal and abnormal is difficult because of the wide range of flow rates among normal individuals (Skopouli *et al*. 1989).

Sialography

This is a radiocontrast method of assessing anatomical changes in the salivary ductal system. It has been widely used in patients with Sjögren's syndrome, in whom various degrees of sialectasis have been found (Fig. 6). Sialography with oil-based contrast material shows an increased incidence of sialectasis in these patients (Daniels *et al*. 1987). However, sialography causes pain and swelling of the parotid glands and sometimes allergic reactions to the radiopaque material. In addition, some have suggested that sialography is insensitive and non-specific (Moutsopoulos and Talal 1989). However, Vitali *et al*. (1988) described sialography with water-soluble media in 84 patients with primary and secondary Sjögren's syndrome and compared it with the findings of minor salivary gland biopsy, as well as with the patients' clinical and the serological picture. They reported that sialography was as sensitive and specific as the biopsy. Furthermore, hypergammaglobulinaemia, anti-Ro(SSA) antibodies, extraglandular manifestations, and parotid swelling were all correlated with both sialographic and histological abnormalities, suggesting that

both tests are necessary for the evaluation of salivary gland involvement in Sjögren's syndrome.

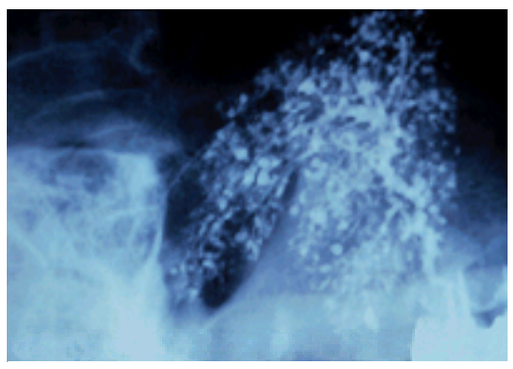


Fig. 6 Sialography of the parotid gland of a patient with Sjögren's syndrome. The retention of the sialographic medium at the end-points of the duct system reveals several degrees of sialectasis.

Scintigraphy

Isotope scanning provides a functional evaluation of all the salivary glands by observing the rate and density of uptake of the [$^{99}Tc^{m}$]pertechnetate and the time taken for it to appear in the mouth during a 60-min period after intravenous injection. In patients with Sjögren's syndrome the uptake of the label by the glands and the secretion of labelled saliva is delayed or absent. An abnormal scan was found to correlate with the presence of reduced salivary flow, the sialographic picture, and the intensity of lymphocytic infiltrates in minor salivary glands (Daniels *et al*. 1987).

In a study of 320 patients with oral dryness who had primary or secondary Sjögren's syndrome, graft-versus-host disease and other autoimmune diseases, scanning had high sensitivity but no well-established disease specificity (Parrago *et al*. 1987).

Sialochemistry

Chemical and immunological factors in saliva of patients with Sjögren's syndrome have been examined extensively in the past. So far, the results are conflicting and controversial (Baum and Fox 1987), offering very limited diagnostic value.

Ocular component

Ocular involvement is a major glandular manifestation of Sjögren's syndrome. Diminished secretion of tears leads to the destruction of the corneal and bulbar conjunctival epithelium termed keratoconjunctivitis sicca. The patients usually complain of a burning, foreign-body sensation, a sandy or scratchy sensation under the lids, itchiness, redness, and photosensitivity. Clinical signs include dilation of the bulbar conjunctival vessels, pericorneal injection, irregularity of the corneal image, and sometimes enlargement of the lacrimal glands. All tests for the evaluation of this condition are very sensitive but not specific for Sjögren's syndrome, as keratoconjunctivitis sicca may occur in a number of other conditions.

The Schirmer's test is used for the evaluation of tear secretion. The test is made with strips of filter paper 30 mm in length. The strip is slipped beneath the inferior lid, with the remainder of the paper hanging out (Fig. 8). After 5 min the length of paper wetted is measured. Wetting of less than 5 mm is a strong indication of diminished secretion. The presence of decreased tear secretion is not diagnostic of keratoconjunctivitis sicca. In contrast, it can easily be diagnosed using rose bengal staining. Rose bengal is an aniline compound that stains the devitalized or damaged epithelium of both the cornea and conjunctiva. Slit-lamp examination after rose bengal staining shows a punctate pattern of filamentary keratitis (Fig. 8).

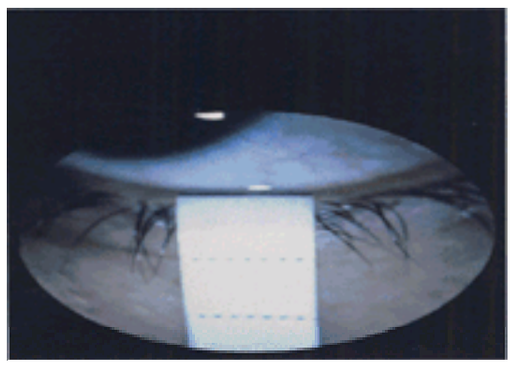


Fig. 7 The Schirmer's test in a patient with Sjögren's syndrome: the filter-paper strip is slipped under the lid, with part hanging out; the length of paper wetted, measured after 5 min, is less than 5 mm in this case.

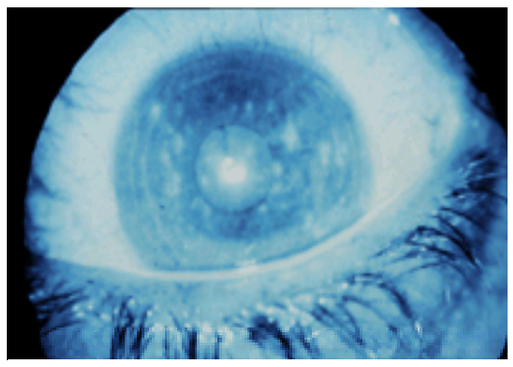


Fig. 8 Slit-lamp examination of the eye of a patient with Sjögren's syndrome after rose bengal staining. The retention of the stain in the corneal conjunctiva shows damaged epithelium and filaments, which are diagnostic of keratoconjunctivitis sicca.

The break-up time of tears is another useful measure. A drop of fluoroscein is instilled in the eye and the time between the last blink and appearance of dark, non-fluorescent areas in the tear film is measured. An overly rapid break-up of the tear film indicates an abnormality of either the mucin or the lipid layer (Kincaid 1987).

Systemic (extraglandular) manifestations

Extraglandular (systemic) manifestations are seen in one-third of patients with primary Sjögren's syndrome (Moutsopoulos *et al*. 1980) (Table 6). These patients complain most often of being easily fatigued, of low-grade fever, and of myalgias and arthralgias.

Clinical manifestation	*n*	%
Arthralgias/arthritis	79	60
Raynaud's phenomenon	49	37
Lymphadenopathy	19	14
Vasculitis	15	11
Lung involvement	18	14
Kidney involvement	12	9
Liver involvement	8	6
Splenomegaly	4	3
Peripheral neuropathy	3	2
Myositis	1	1
Lymphoma	0	0

Table 6 Incidence of extraglandular manifestations in primary Sjögren's syndrome (data from 132 Greek patients)

Arthritis

Seventy per cent of patients with primary Sjögren's syndrome experience an episode or episodes of arthritis during the course of their disease. In some cases it can precede the sicca manifestations. Joint symptoms and signs include arthralgias, myalgias, morning stiffness, intermittent synovitis, and chronic polyarthritis that sometimes lead to Jacoub's arthropathy (Maini 1987). In contrast to rheumatoid arthritis, radiographs of the hand usually do not reveal pathological changes (Castro-Poltronieri and Alarcon-Segovia 1983; Tsampoulas *et al*. 1990).

Raynaud's phenomenon

This occurs in 35 per cent of patients with primary Sjögren's and usually precedes sicca manifestations by many years. Patients with the primary syndrome and Raynaud's phenomenon present with swollen hands, but, in contrast to those with scleroderma, they do not experience digital ulcers and telangiectasias are not seen. Radiographs of the hands of these patients may show small tissue calcifications (Skopouli *et al*. 1990). Non-erosive arthritis has also been shown to be significantly more frequent in patients with Raynaud's phenomenon than in those without (Youinou *et al*. 1990a).

Skin involvement

Cutaneous lesions are seen frequently in patients with primary Sjögren's syndrome. Patients with dry skin complain of dermal stinging and itching. Nasal dryness with crusting, vaginal dryness syndrome with dyspareunia, and cheilitis have also been described. Other cutaneous manifestations include skin hyper- or hypopigmentation, patchy alopecia, and hypersensitivity vasculitis (Fye and Talal 1984). Some patients with Sjögren's syndrome may present with annular erythema affecting mainly the face and the trunk; it extends centrifugally and fades without leaving pigmentation (Teramoto *et al*. 1989).

Pulmonary involvement

Manifestations from the trachea to the pleura have been described in patients with Sjögren's syndrome. These are frequent but rarely important clinically. They can present with a range of symptoms from dry cough secondary to dryness of the tracheobronchial mucosa (xerotrachea) to dyspnoea from interstitial disease or even airway obstruction (Constantopoulos and Moutsopoulos 1987).

Interstitial lung disease in Sjögren's syndrome was thought to be a very common manifestation, since chest radiographs revealed an interstitial pattern in approximately half of patients with Sjögren's syndrome studied. High-resolution CT scanning of the lungs performed in 21 patients with the most prominent abnormalities, however, demonstrated that in 14 patients the main findings were either thickened bronchial walls at the segmental level or a mild interstitial pattern distributed around the bronchi. Transbronchial biopsy performed in 12 of these patients disclosed bronchiolar lymphoid infiltrates and follicular bronchiolitis (bronchus-associated lymphoid tissue hyperplasia). Fibrosing alveolitis was absent from all tissue specimens (Papiris *et al*. 1994).

Pseudolymphoma or frank lymphoma should always be suspected when lung nodules or hilar and/or mediastinal lymphadenopathy are found on chest radiographs (Constantopoulos and Moutsopoulos 1987).

There are differences in respiratory manifestations between the primary and secondary syndrome. In the latter, the respiratory involvement is a reflection of the primary rheumatic disorders. In fact, pleural effusions are usually found in Sjögren's syndrome associated with other rheumatic disorders and not in the primary syndrome.

Gastrointestinal and hepatobiliary features

Patients with Sjögren's syndrome often complain of dysphagia, owing to dryness of the pharynx and oesophagus, or to abnormal oesophageal motility. Nausea and epigastric pain are also common complaints. Biopsies of gastric mucosa show chronic atrophic gastritis and lymphocytic infiltrates (Buchanan *et al*. 1966), similar to those described in minor salivary glands. In addition, patients with Sjögren's syndrome have hypopepsinogenaemia, an elevated serum gastrin, low serum vitamin B_{12}, and antibodies to parietal cells (Trevino *et al*. 1987).

Acute or chronic pancreatitis has been reported rarely. In contrast, subclinical pancreatic involvement is a rather common finding, as illustrated by the fact that hyperamylasaemia is found in one-quarter of patients with the syndrome (Tsianos *et al*. 1984).

The prevalence and nature of liver involvement was studied in 300 patients with Sjögren's syndrome (Skopouli *et al*. 1994). Seven per cent had antimitochondrial antibodies and 5 per cent had elevated liver enzymes. Patients with antimitochondrial antibodies had elevated liver enzymes and no evidence of hepatitis B and C viral infection. Eleven out of 17 patients with antimitochondrial antibodies underwent liver biopsy. The histopathological picture disclosed in seven specimens was an 'autoimmune cholangiitis', that is a chronic granulomatous cholangiitis affecting small and medium-sized bile ducts and a mild periportal inflammation, without, however, evidence of 'piecemeal' necrosis (Fig. 9). These histological features are similar with those observed in stage I of primary biliary cirrhosis.

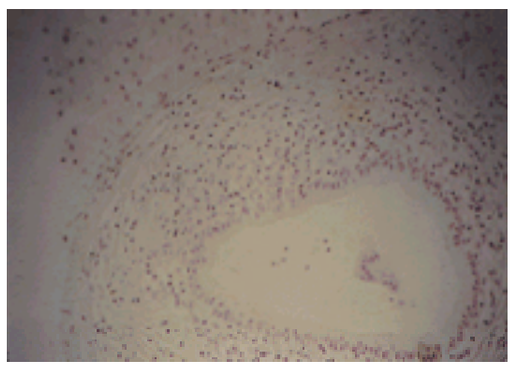


Fig. 9 Periportal inflammation in liver biopsy of a patient with Sjögren's syndrome.

There is also a high incidence of Sjögren's syndrome in patients with primary biliary cirrhosis: sicca manifestations have been described in approximately half of a group of patients with primary biliary cirrhosis; among these, 10 per cent had severe clinical features of dryness (Tsianos *et al*. 1990).

Renal involvement

Overt kidney disease is found in approximately 10 per cent of patients with Sjögren's syndrome (Kassan and Talal 1987), while approximately 35 per cent have an abnormal urine acidification test. Interstitial disease is the most common renal finding. Most of the patients present with hyposthenuria and hypokalaemic, hyperchloraemic distal tubular acidosis. Distal tubular acidosis can be silent or can present with recurrent renal colic and/or hypokalaemic muscular weakness. Untreated renal tubular acidosis leads to renal stones, nephrocalcinosis, and compromised renal function (Moutsopoulos *et al*. 1991) (Fig. 10). Less commonly, these patients have proximal tubular acidosis with Fanconi syndrome (Kassan and Talal 1987); renal biopsy reveals interstitial lymphocytic infiltration. Glomerunophritis in Sjögren's syndrome has been described in few patients (Moutsopoulos *et al*. 1978); the histological type may be membranous or membranoproliferative. In all cases a consistent serological finding was cryoglobulinaemia associated with hypocomplementaemia.

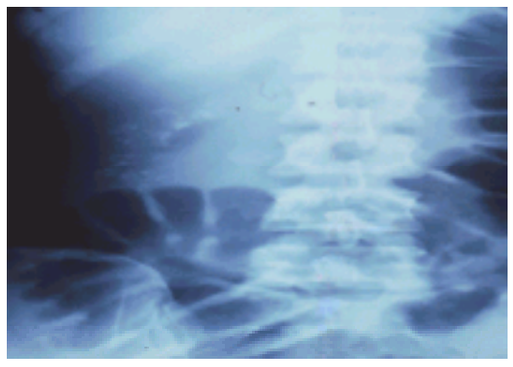


Fig. 10 Abdominal radiograph of a patient with Sjögren's syndrome reveals nephrocalcinosis of the right kidney.

Vasculitis

Vascular involvement is found in approximately 5 per cent of patients with Sjögren's syndrome. It affects small and medium-sized vessels. The most common manifestations are purpura (Fig. 11), recurrent urticaria, skin ulcerations, and mononeuritis multiplex. However, cases of systemic vasculitis with visceral involvement affecting kidney, lung, gastrointestinal tract, spleen, breast, and the reproductive tract have been described (Alexander 1987a; Tsokos *et al*. 1987). There are two histopathological types of vasculitis, according to the type of the infiltrating cell—the mononuclear and the neutrophil type. The latter is associated with hypergammaglobulinaemia, high titres of rheumatoid factor, antibodies to Ro(SSA) cellular antigen, and hypocomplementaemia (Alexander 1987a). In another classification of vascular involvement (Tsokos *et al*. 1987), the small-vessel vasculitis was of the hypersensitivity type, that is leucocytoclastic and lymphocytic, while the medium-vessel vasculitis was acute necrotizing and simulated polyarteritis nodosa but without the formation of aneurysms. Endarteritis obliterans was seen in patients with a long-standing history of vasculitis.

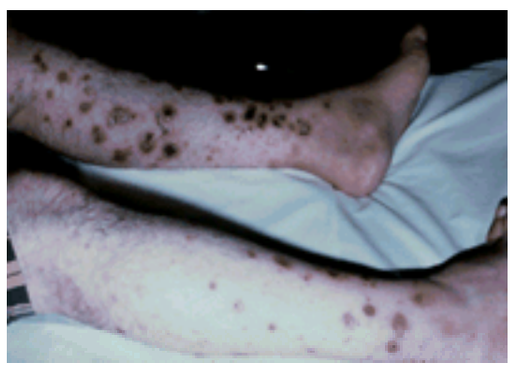


Fig. 11 The lower extremities of a patient with Sjögren's syndrome with hypergammaglobulinaemia and cryoglobulins show diffuse, palpable purpura.

Neuromuscular involvement

Neurological manifestations of Sjögren's syndrome include peripheral sensory or sensorimotor neuropathy as a consequence of vascular involvement. Cranial neuropathy, usually affecting single nerves such as the trigeminal or the optic, has been well documented. Involvement of the central nervous system in the syndrome is a matter of considerable controversy. Over the last decade some investigators have described a high proportion of such involvement, which before had been unrecognized internationally. They found that this disease was multifocal, recurrent, and progressive. The clinical signs included hemiparesis, hemisensory deficits, seizures, movement disorders, and transverse myelopathy. Some patients presented with diffuse brain injury expressed as encephalopathy, aseptic meningitis, and dementia (Alexander 1987b). On the other hand, others have failed to demonstrate severe central involvement in 55 patients with primary and 50 with secondary Sjögren's syndrome (Binder *et al*. 1988). Eighteen of these patients had mild neurological abnormalities confined to the secondary syndrome; all were characteristics of the underlying rheumatic disorders. A study of 63 consecutive patients with primary Sjögren's syndrome revealed that 17 had a mild sensory or sensorimotor neuropathy while one patient with past history of hypertension had had a mild episode (Andonopoulos *et al*. 1990a), suggesting that peripheral neuropathy is a rather common finding in Sjögren's syndrome whereas central nervous disease must be rare.

Many patients with the primary syndrome complain of myalgia, but muscle enzymes are usually normal or slightly elevated. Severe polymyositis with extensive necrosis of muscle fibres and invasion of macrophages into the affected muscle has been described in Sjögren's syndrome (Leroy *et al*. 1990). This kind of myositis seems to respond well to pulse cyclophosphamide therapy.

Other manifestations

Autoimmune thyroid disease has been described in some patients with primary Sjögren's syndrome. In a study by Karsh *et al*. (1980), half of the patients with Sjögren's syndrome presented with antithyroid antibodies and signs of altered thyroid function as evaluated by a basal thyroid-hormone stimulation test.

Sjögren's syndrome may also associated with interstitial cystitis (Van de Merwe *et al*. 1993), which is a non-bacterial disease of the bladder producing constant or intermittent, long-lasting symptoms, such frequent micturition, nocturia, and suprapubic or perineal pain. Bladder biopsy discloses intense inflammation in the mucosa and submucosa with lymphoid cells and mast cells. Lymphoid cell infiltrates contain CD4+ T cells as well as B-cell nodules with germinal centres. Detrusor fibrosis

can be seen in the later stages of the disease (Harrington *et al*. 1990).

Mild normochromic and normocytic anaemia is a common finding; leucopenia and thrombocytopenia are relatively rare features. An elevated erythrocyte sedimentation rate is found in approximately 70 per cent of patients (Moutsopoulos *et al*. 1980). In contrast, C-reactive protein is not detected in patients with primary Sjögren's syndrome but is found in those with Sjögren's syndrome and rheumatoid arthritis (Table 7) (Moutsopoulos *et al*. 1983b).

	n	%
Anaemia (Hb < 35)	27/130	21
Leucopenia (WBC < 3500 cells/mm³)	8/130	6
Thrombocytopenia (PLT < 10 000)	2/132	2
ESR (>25 mm/h)	76/125	61
C-reactive protein (< 8 mg/l)	6/130	5
Cryoglobulinaemia	34/111	28
ANA (> 1:80)[a]	119/132	92
Rheumatoid factor (> 1:40[b])	80/131	61
Antimitochondrial antibodies	8/132	6
Anti-Ro (SSA)	75/131	57
Anti-La (SSB)	48/131	36
Anti-nRNP	4/131	3
Anti-Sm	2/131	2

[a]Hep-2 cells as substrate.
[b]Latex fixation.

Table 7 Laboratory findings in Greek patients with primary Sjögren's syndrome (data tabulated from 132 patients)

Lymphoproliferative disease

Patients with Sjögren's syndrome have a 44 times higher relative risk of developing lymphoma, compared with age-, sex-, and race-matched normal controls (Kassan *et al*. 1978). Immunohistological studies in biopsies of such patients with lymphoma show that these are primarily of B-cell origin, usually expressing IgM-k in their cytoplasm (Zulman *et al*. 1978). The lymphomas are of two major types, either of highly undifferentiated B cells or well-differentiated immunocytomas.

Lymphomas may differ by location and grading. In our patients, among eight lymphomas in patients with Sjögren's syndrome, six were low-grade immunocytomas and two intermediate-grade non-Hodgkin's lymphomas. Five of the immunocytomas were diagnosed from biopsies of minor salivary or lacrimal glands. Two of the patients with immunocytomas showed spontaneous regression, while the other two developed high-grade lymphoma 3 and 5 years later (Pavlidis *et al*. 1992). Therefore, the clinical picture of Sjögren's syndrome-associated lymphoma appears to be diverse, suggesting that the therapeutic approach should be guided by the stage and the grade of the disease.

Other organs that may be affected by the lymphomas are the reticuloendothelial system, lungs, kidneys, and the gastrointestinal tract.

The diagnostic interpretation of a tissue with lymphoproliferative infiltration is sometimes very difficult. The term pseudolymphoma describes lesions that show tumour-like clusters of lymphoid cells but do not meet criteria for malignancy. Despite the use of modern molecular techniques, such as immunogenotyping (Cleary *et al*. 1987), pseudolymphoma remains in ill-defined clinicopathological entity and lymphoma should always be suspected in a patient with lymphadenopathy, organomegaly, or enlargement of major salivary glands.

Secondary Sjögren's syndrome

The association of Sjögren's syndrome with rheumatoid arthritis was first described by Henrik Sjögren in 1939. During the following years it became evident that sicca manifestations can also be found in other autoimmune rheumatic diseases, such as systemic lupus erythematosus and progressive systemic sclerosis. In addition, manifestations of Sjögren's syndrome have been described in polymyositis, polyarteritis nodosa, and primary biliary cirrhosis (see Table 1) (Moutsopoulos *et al*. 1980).

The incidence of clinically overt Sjögren's syndrome in patients with rheumatoid arthritis is around 5 per cent. Using a special questionnaire, however, 20 per cent of patients with rheumatoid arthritis registered complaints of dry eyes and/or xerostomia (Andonopoulos *et al*. 1987). The diagnosis of rheumatoid arthritis usually preceded that of Sjögren's syndrome by many years. Patients with the arthritis and the syndrome usually present with keratoconjunctivitis sicca, while enlargement of the parotids or other major salivary glands is less common than in primary Sjögren's syndrome. In addition, extraglandular features of the primary syndrome, such as lymphadenopathy, renal involvement, and Raynaud's phenomenon, are uncommon in Sjögren's syndrome associated with rheumatoid arthritis. Such clear differences in the natural history and the clinical manifestations of the syndrome in the presence and absence of rheumatoid arthritis are not usually found in other autoimmune diseases associated with sicca manifestations. In fact, patients with primary Sjögren's syndrome and those with systemic lupus may have similar disease manifestations such as arthralgias, rash, peripheral neuropathy, and glomerulonephritis. These observations prompted Heaton (1959) to conclude that Sjögren's syndrome was a benign form of systemic lupus. The diagnosis is usually obtained histologically; approximately 20 per cent of patients with systemic lupus have lymphocytic infiltrates in biopsies from their minor salivary glands (Andonopoulos *et al*. 1990b).

Dry eyes and mouth are found in approximately 20 per cent of unselected patients with scleroderma (Medsger 1987). Subjective xerostomia could be due to fibrosis of the exocrine glands. In fact, in biopsies from minor salivary glands of 44 unselected patients with scleroderma, 38 per cent had fibrosis while only 22 per cent had lymphocytic infiltration compatible with Sjögren's syndrome (Andonopoulos *et al*. 1989).

Diagnosis and differential diagnosis

Since the initial definition and the proposed criteria of Sjögren's syndrome by Bloch *et al*. (1965), several sets of criteria have been used by different groups for the diagnosis of Sjögren's syndrome. As a result, patients with Sjögren's syndrome had often been missed at diagnosis, or classified incorrectly due both to the great variability at disease presentation and to the lack of well-defined and commonly accepted diagnostic criteria.

Recently, a prospective concerted action involving 26 centres in 12 European countries led to a study with the goal of obtaining validated criteria for the diagnosis of Sjögren's syndrome. The study resulted in: (i) the validation of a simple 6-item questionnaire for determination of dry eyes and mouth, useful in the initial screening for Sjögren's syndrome, and (ii) the definition of a new set of criteria for Sjögren's syndrome. The sensitivity and specificity of both questionnaire and diagnostic criteria were determined, exhibiting good discrimination between patients and controls. Hence, using this set of criteria, a general agreement can be reached on the diagnosis of Sjögren's syndrome (Vitali *et al*. 1992a).

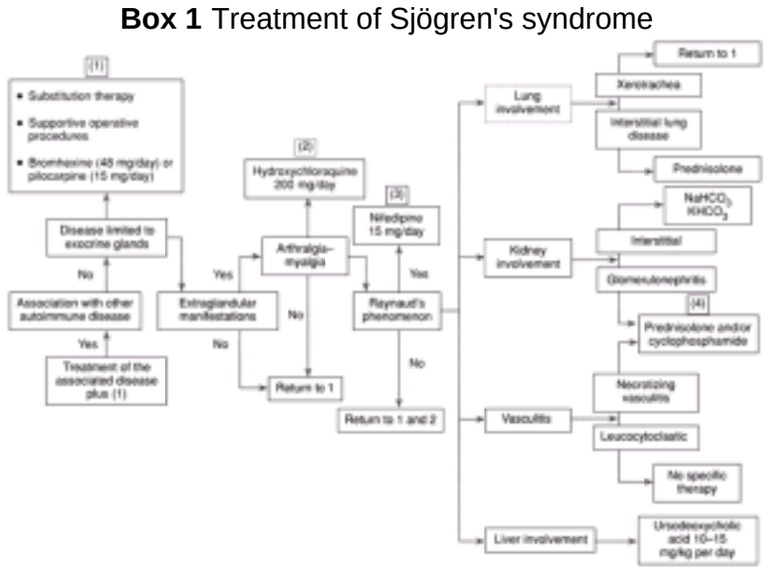
Differential diagnosis must be, of course, from other diseases responsible for dry eyes, xerostomia, and parotid enlargement. Sarcoidosis in one disease that can mimic the clinical picture of Sjögren's syndrome (Drosos *et al*. 1989). However, the biopsy of minor salivary glands reveals non-caseating granulomas in sarcoidosis, while there is a lack of autoantibodies to Ro(SSA) or La(SSB). Other medical conditions that can mimic the syndrome are lipoproteinaemias (types II, IV, and V), chronic graft-versus-host disease, amyloidosis, and, more recently, patients with human immunodeficiency virus (**HIV**) infection (Moutsopoulos and Talal 1989). In fact, sicca manifestations, with parotid enlargement, pulmonary involvement and lymphadenopathy, have been reported in 12 patients with HIV infection. These patients has an increased prevalence of HLA-DR5 alloantigen. The virus has been detected in lymphocytes from labial salivary glands in two of six patients with HIV infection, but the two diseases seem to be different, as patients with HIV infection have no autoantibodies to Ro(SSA) and La(SSB) autoantigens, and the lymphocytic infiltrates in their salivary glands are not as prominent as in Sjögren's syndrome and consist of CD8+ T cells (Youinou *et al*. 1990b).

Therapy—prognosis (Box 1)

Sjögren's syndrome is a chronic, multisystem disease. Therefore, patients with Sjögren's syndrome should be followed regularly for significant functional deterioration, signs of complications and significant changes in the course of the disease. The patient should be informed that regular outpatient visits, and close collaboration with

the outpatient clinics for rheumatology, ophthalmology, and oral medicine, give the most satisfactory results.

Box 1 Treatment of Sjögren's syndrome



Sicca manifestations are of unknown aetiology. Hence, treatment of Sjögren's syndrome is aimed at symptomatic relief and limiting the damaging local effects of chronic xerostomia and keratoconjunctivitis sicca by substitution of the missing secretions (Moutsopoulos and Vlachoyiannopoulos 1993).

Keratoconjunctivitis sicca is treated with fluid replacement supplied as often as necessary. To replace deficient tears, there are several readily available ophthalmic preparations (Tearisol; Liquifilm; 0.5 per cent methylcellulose; Hypo Tears). In severe cases, it may be necessary for patients to use these as often as every 30 min. If corneal ulceration is present, eye-patches and boric acid ointment are recommended. Certain drugs that may cause further deterioration of lacrimal and salivary function, such as diuretics, antihypertensive drugs and antidepressants, should be avoided. The low levels of humidity in air-conditioned environments, as well as windy or dry climates, must be avoided. Soft contact lenses may help to protect the cornea, especially in the presence of filaments. However, the lenses themselves require wetting and the patients must be followed very carefully due to the increased risk of infection.

Treatment of xerostomia is difficult. Stimulation of salivary flow by sugar-free, highly flavoured lozenges has been found to be rather helpful. Most patients carry water and use sugarless lemon drops or chewing-gum. These must be sugar free, because of the risk of rampant dental caries. Adequate oral hygiene after meals is essential for the prevention of dental disease. Topical oral treatment with fluoride enhances dental mineralization and retards damage to tooth surfaces. In rapidly progressive dental disease, fluoride can be directly applied to the teeth from plastic trays that are used at night. Propionic acid gels may be used to treat vaginal dryness. Bromhexine given orally at high doses (48 mg/day) has been suggested to improve sicca manifestations. However, frequent ingestion of fluids, particularly with meals, is often the best solution.

Pilocarpine hydrochloride (5 mg, three times daily) can also improve sicca manifestations, via its muscarinic, cholinergic activity (Fox 1992). Flushing and sweating are possible side-effects.

Patients with Sjögren's syndrome often complain of parotid gland swelling. If the gland becomes tender with permanent enlargement, infection should be ruled out and treatment with tetracycline orally should be recommended (500 mg, four times daily). Local moist heat and non-steroidal anti-inflammatory drugs are usually helpful in resolving this problem. If the gland remains tender, lymphoma should be ruled out by biopsy.

Preliminary studies showed that hydroxychloroquine, which is efficacious and safe in other autoimmune diseases, may be useful in treating Sjögren's patients. A dose of 200 mg/day partially corrects hypergammaglobulinaemia and decreases the titre of IgG antibodies to La/SSB antigen. Furthermore, hydroxychloroquine decreased the erythrocyte sedimentation rate and increased the haemoglobin (Fox *et al*. 1988).

Corticosteroids (prednisolone 0.5–1 mg/kg per day) or other immunosuppressive agents (i.e. cyclophosphamide) are indicated for the treatment of life-threatening extraglandular manifestations, particularly when renal or severe pulmonary involvement and systemic vasculitis are present.

In conclusion, Sjögren's syndrome remains an incurable disease without a single therapeutic approach that can change its natural course. The prognosis of the disease depends on the extension and type of the systemic features, and the appearance of a lymphoma. The treatment and prognosis of malignant lymphoma depends on the histological type, the location, and the extension. Decisions about chemotherapy and/or radiotherapy should be guided by experienced oncologists.

Chapter References

Agnello, V. *et al*. (1980). Evidence for a subset of rheumatoid factors that cross react with DNA-histone and have a distinct cross-idiotype. *Journal of Experimental Medicine*, **151**, 1514–27.

Alexander, E. L. (1987*a*). Inflammatory vascular disease in Sjögren's syndrome. In *Sjögren's syndrome: clinical and immunological aspects* (ed. N. Talal, H. M. Moutsopoulos, and S. S. Kassan), pp. 102–24. Springer, Berlin.

Alexander, E. L. (1987*b*). Neuromuscular complications of primary Sjögren's syndrome. In *Sjögren's syndrome: clinical and immunological aspects* (ed. N. Talal, H. M. Moutsopoulos, and S. S. Kassan), pp. 61–82. Springer, Berlin.

Anderson, L. G. *et al*. (1972). Salivary gland immunoglobulin and rheumatoid factor synthesis in Sjögren's syndrome: natural history and response to treatment. *American Journal of Medicine*, **49**,49–54.

Andonopoulos, A. P., Drosos, A. A., Skopouli, F. N., Acritidis, N. C., and Moutsopoulos, H. M. (1987). Secondary Sjögren's syndrome in rheumatoid arthritis. *Journal of Rheumatology*, **1**, 1098–103.

Andonopoulos, A. P., Drosos, A. A., Skopouli, F. N., and Moutsopoulos, H. M. (1989). Sjögren's syndrome in rheumatoid arthritis and progressive systemic sclerosis: a comparative study. *Clinical and Experimental Rheumatology*, **7**,203–5.

Andonopoulos, A. P., Lagos, G., Drosos, A. A., and Moutsopoulos, H. M. (1990*a*). The spectrum of neurological involvement in Sjögren's syndrome. *British Journal of Rheumatology*, **29**, 21–3.

Andonopoulos, A. P., Skopouli, F. N., Dimou, G. S., Drosos, A. A., and Moutsopoulos, H. M. (1990 *b*). Sjögren's syndrome in systemic lupus erythematosus. *Journal of Rheumatology*, **17**, 201–4.

Batsakis, J. G. and Howard, D. R. (1982). Sjögren's syndrome: an immune response associated disorder. *Clinical Laboratory Annals*, **1**, 171–88.

Baum, B. J. and Fox, P. C. (1987). Chemistry of saliva. In *Sjögren's syndrome*: *clinical and immunological aspects* (ed. N. Talal, H. M. Moutsopoulos, and S. S. Kassan), pp. 25–34. Springer, Berlin.

Ben-Chetrit, E., Fox, R. I., and Tan, E. M. (1990). Dissociation of immune response to the SSA (Ro) 52 kD and 60 kD polypeptides in systemic lupus erythematosus and Sjögren's syndrome. *Arthritis and Rheumatism*, **33**, 349–55.

Binder, A., Snaith, M. L., and Isenberg, D. (1988). Sjögren's syndrome: a study of its neurological complications. *British Journal of Rheumatology*, **27**, 275–80.

Bloch, K. J., Buchanan, W. W., Wohl, M. J., and Bunim, J. J. (1965). Sjögren's syndrome: a clinical, pathological and serological study of sixty-two cases. *Medicine (Baltimore)*, **44**, 187–231.

Bodeuitsch, C. *et al*. (1992). Quantitative immunohistologic criteria are superior to the lymphocytic focus score criterion for the diagnosis of Sjögren's syndrome. *Arthritis and Rheumatism*, **35**, 1075–87.

Boumba, D., Skopouli, F. N., and Moutsopoulos, H. M. (1995). Cytokine mRNA expression in the labial salivary gland tissues from patients with primary Sjögren's syndrome. *British Journal of Rheumatology*, **34**, 326–33.

Boumpas, D. T. *et al*. (1990). c-*myc*, proto-oncogene expression in peripheral blood mononuclear cells from patients with primary Sjögren's syndrome. *Arthritis and Rheumatism*, **33**, 49–56.

Buchanan, W. W., Cox, A. G., Harden, R. M. G., Glen, A. I. M., Anderson, J. R., and Gray, K. G. (1966). Gastric studies in Sjögren's syndrome. *Gut*, **7**, 351–9.

Castro-Poltronieri, A. and Alarcon-Segovia, D (1983). Articular manifestation of primary Sjögren's syndrome. *Journal of Rheumatology*, **10**, 485–90.

Chisholm, D. M. and Mason, D. K. (1968). Labial salivary gland biopsy in Sjögren's syndrome. *Journal of Clinical Pathology*, **21**,656–60.

Chused, T. M., Kassan, S. S., Opelz, G., Moutsopoulos, H. M., and Terasaki, P. I. (1977). Sjögren's syndrome associated with HLA DW3. *New England Journal of Medicine*, **296**,895–7.

Cleary, M. L., Calini, N., Levy, N., Talal, N., and Sklar, J. (1987). Detection of lymphoma by analysis of immunoglobulin gene rearrangements. In *Sjögren's syndrome*: *clinical and immunological aspects* (ed. N. Talal, H. M. Moutsopoulos, and S. S. Kassan), pp. 137–43. Springer, Berlin.

Constantopoulos, S. H. and Moutsopoulos, H. M. (1987). The respiratory system in Sjögren's syndrome. In *Sjögren's syndrome*: *clinical and immunological aspects* (ed. N. Talal, H. M. Moutsopoulos, and S. S. Kassan), pp. 83–9. Springer, Berlin.

Daniels, T. E. and Whitcher, J. P. (1994). Association of patterns of labial salivary gland inflammation with keratoconjunctivitis sicca. *Arthritis and Rheumatism*, **37**, 869–77.

Daniels, T. E., Aufdemorte, T. B., and Greenspan, J. S. (1987). Histopathology of Sjögren's syndrome. In *Sjögren's syndrome*: *clinical and immunological aspects* (ed. N. Talal, H. M. Moutsopoulos, and S. S. Kassan), pp. 266–86. Springer, Berlin.

Drosos, A. A., Antonopoulos, A. P., Costopoulos, J. S., Papadimitriou, C., and Moutsopoulos, H. M. (1988). Prevalence of primary Sjögren's syndrome in an elderly population. *British Journal of Rheumatology*, **27**, 123–7.

Drosos, A. A., Constantopoulos, S. H., Phsychos, D., Stefanou, D., Papadimitriou, C. S., and Moutsopoulos, H. M. (1989). The forgotten cause of sicca complex; sarcoidosis. *Journal of Rheumatology*, **16**, 1548–51.

Fauci, A. S. and Moutsopoulos, H. M. (1981). Polyclonally triggered B-cells in the peripheral blood of normal individuals and in patients with SLE and primary Sjögren's syndrome. *Arthritis and Rheumatism*, **24**, 577–84.

Fishleder, A., Tubbs, R., Hesse, B., and Levine, H. (1987). Uniform detection of immunoglobulin gene rearrangements in benign lymphoepithelial lesions. *New England Journal of Medicine*, **316**, 1118–21.

Fox, R. I. (1992). Treatment of patients with Sjögren's syndrome. *Rheumatic Disease Clinics of North America*, **18**, 699–709.

Fox, R. I., Pearson, G., and Vaughan, J. H. (1986 *a*). Detection of Epstein–Barr virus-associated antigens and DNA in salivary gland biopsies from patients with Sjögren's syndrome. *Journal of Immunology*, **137**, 3162–7.

Fox, R. I., Chen, P., Carson, D. A., and Fong, S. (1986 *b*). Expression of a cross-reactive idiotype on rheumatoid factor in patients with Sjögren's syndrome. *Journal of Immunology*, **136**, 477–83.

Fox, R. I., Chan E., Bentol, L., Fong, S., Freidlaender, M., and Howell, F. W. (1988). Treatment of primary Sjögren's syndrome with hydroxychloroquine. *American Journal of Medicine*, **85**(Suppl. 4A), 62–7.

Fox, R. I. *et al*. (1994). Cytokine mRNA expression in salivary biopsies of Sjögren's syndrome. *Journal of Immunology*, **152**,5532–9.

Freimark, B., Fantozzi, R., Bone, R., Bordin, G., and Fox, R. (1989). Detection of clonally expanded salivary gland lymphocytes in Sjögren's syndrome. *Arthritis and Rheumatism*, **32**, 859–69.

Fye, K. H. and Talal, N. (1987). Skin manifestations of Sjögren's syndrome. In *Dermatology in general medicine*(ed. T. B. Fitzpatrick, A. Z. Eisen, K. Walff, I. M. Freedberg, and K. F. Austen), pp. 1883–7. McGraw-Hill, New York.

Fye, K. H., Terasaki, P. I., Moutsopoulos, H. M., Daniels, T. E., Michalski, J. P., and Talal, N. (1976). Association of Sjögren's syndrome with HLA-B8. *Arthritis and Rheumatism*, **19**, 883–6.

Garry, R. F., Fermin, C. D., Hart, D. J., Alexander, S. S., Donehower, L. A., and Luo-Zhang, H. (1990). Detection of a human intracisternal A-type retroviral particle antigenically related to HIV. *Science*, **250**, 1127–9.

Green, J. E., Hinricks, S. H., Vogel, J., and Jay, G. (1989). Exocrinopathy resembling Sjögren's syndrome in HTLV-1 *tax* transgenic mice. *Nature*, **341**, 72–4.

Greenspan, J. S., Daniels, T. E., Talal, N., and Sylvester, R. A. (1974). The histopathology of Sjögren's syndrome in labial salivary gland biopsies. *Oral Surgery, Oral Medicine, Oral Pathology*, **37**, 217–19.

Gregerson, P. K., Silver, J., and Winchester, R. J. (1987). The shared epitope hypothesis. An approach to understanding the molecular genetics of susceptibility to rheumatoid arthritis. *Arthritis and Rheumatism*, **30**, 1205–13.

Haddad, J. *et al*. (1992). Lymphocytic sialadenitis of Sjögren's syndrome associated with chronic hepatitis C virus liver disease. *Lancet*, **339**, 321–3.

Harley, J. B. (1987). Autoantibodies in Sjögren's syndrome. In *Sjögren's syndrome*: *clinical and immunological aspects* (ed. N. Talal, H. M. Moutsopoulos, and S. S. Kassan), pp. 218–34. Springer, Berlin.

Harrington, D. S., Fall, M., and Johanson, S. E. (1990). Interstitial cystitis: bladder mucosa lymphocyte immunophenotyping and peripheral blood flow cytometry analysis. *Journal of Urology*, **144**, 868–71.

Heaton, J. M. (1959). Sjögren's syndrome and systemic lupus erythematosus. *British Medical Journal*, **1**, 466–9.

Hoffman, R. W and Walker, S. E. (1987). Animal models of Sjögren's syndrome. In *Sjögren's syndrome*: *clinical and immunological aspects* (ed. N. Talal, H. M. Moutsopoulos, and S. S. Kassan), pp. 266–86. Springer, Berlin.

Hudson, J., Chanther, J., Lok, L., Misra, V., and Muller, M. (1979). Model systems for analysis of latent CMV infections. *Canadian Journal of Microbiology*, **25**, 245–50.

Innes, J. R. M. and Stanton, M. F. (1961). Acute disease of the submaxillary and harderian glands (sialodacryoadenitis) of rats with cytomegaly and no inclusion bodies. *American Journal of Pathology*, **38**, 455–68.

Jacobsson, L. *et al*. (1989). Dry eyes or mouth—an epidemiological study in Swedish adults with special reference to primary Sjögren's syndrome. *Journal of Autoimmunity*, **2**, 521–7.

Jacobsson, L., Hansen, B. U., Manthrope, R., Hardgrave, K., Neas, B., and Harley, J. B. (1992). Association of dry eyes and dry mouth with anti Ro/SSA and anti La/SSB autoantibodies in normal adults. *Arthritis and Rheumatism*, **35**, 1492–501.

Kang, H., Pisa, P., Moore, K., Abrams, J., and Fox, R. I. (1991). Sjögren's syndrome pseudolymphoma induced in Scid/hu mice chimeras. *Clinical and Experimental Rheumatology*, **9**, 334 (Abstr.).

Kappler, J. *et al*. (1989). VB specific stimulation of human T cells by staphylococcal toxins. *Science*, **244**, 811–13.

Karsh, J., Pavlidis, N., Neintraub, B. D., and Moutsopoulos, H. M. (1980). Thyroid disease in Sjögren's syndrome. *Arthritis and Rheumatism*, **23**, 1326–9.

Kassan, S. S. and Talal, N. (1987). Renal disease with Sjögren's syndrome. In *Sjögren's syndrome*: *clinical and immunological aspects* (ed. N. Talal, H. M. Moutsopoulos, and S. S. Kassan), pp. 96–102. Springer, Berlin.

Kassan, S. S. *et al*. (1978). Increased risk of lymphoma in sicca syndrome. *Annals of Internal Medicine*, **89**, 888–92.

Katsikis, P. D., Youinou, P. Y., Galanopoulou, V., Tzioufas, A. G., and Moutsopoulos, H. M. (1990). Monoclonal process in primary Sjögren's syndrome and rheumatoid factor associated cross reactive idiotype. *Clinical and Experimental Immunology*, **82**, 509–14.

Kincaid, M. C. (1987). The eye in Sjögren's syndrome. In *Sjögren's syndrome: clinical and immunological aspects* (ed. N. Talal, H. M. Moutsopoulos, and S. S. Kassan), pp. 25–34, Springer, Berlin.

Kipps, T. J., Fong, S., Tomhave, E., Chen, P. P., Goldfien, R. D., and Carson, D. A. (1987). High-frequency expression of a conserved kappa light-chain variable region gene in chronic lymphocytic leukemia. *Proceedings of the National Academy of Sciences (USA)*, **84**, 2916–20.

Kipps, T. J., Tomhave, E., Chen, P. P., and Fox, R. I. (1989). Molecular characterization of a major autoantibody associated cross-reactive idiotype in Sjögren's syndrome. *Journal of Immunology*, **142**, 4261–8.

Kunkel, H. G., Agnello V., Joslin, F. G., Winchester, R. J., and Capra, J. D. (1973). Cross-idiotypic specificity among monoclonal IgM proteins with anti-gammaglobulin activity. *Journal of Experimental*

Medicine, **137**, 331–42.

Lane, H. C., Callahay, T. R., Jaffe, E. S., Fauci, A. S., and Moutsopoulos, H. M. (1983). The presence of intracytoplasmic infiltrates of the minor salivary glands of patients with primary Sjögren's syndrome. *Clinical and Experimental Rheumatology*, **1**, 237–9.

Leroy, J. P., Drosos, A. A., Yannopoulos, D. I., Youinou, P., and Moutsopoulos, H. M. (1990). Intravenous pulse cyclophosphamide therapy in myositis and Sjögren's syndrome. *Arthritis and Rheumatism*, **33**, 1579–81.

Maini, R. N. (1987). Relationship of Sjögren's syndrome to rheumatoid arthritis. In *Sjögren's syndrome: clinical and immunological aspects* (ed. N. Talal, H. M. Moutsopoulos, and S. S. Kassan), pp. 165–77. Springer, Berlin.

Mann, D. (1987). Immunogenetics of Sjögren's syndrome. In *Sjögren's syndrome: clinical and immunological aspects* (ed. N. Talal, H. M. Moutsopoulos, and S. S. Kassan), pp. 235–43. Springer, Berlin.

Manoussakis, M. N., Tzioufas, A. G., Pange, P. J. E., and Moutsopoulos, H. M. (1986). Serological profiles in subgroups of patients with Sjögren's syndrome. *Scandinavian Journal of Rheumatology*, **61**(Suppl.), 89–92.

Medsger, T. A., Jr (1987). Sjögren's syndrome and systemic sclerosis (scleroderma). In *Sjögren's syndrome: clinical and immunological aspects*, (ed. N. Talal, H. M. Moutsopoulos, and S. S. Kassan), pp. 182–7. Springer, Berlin.

Moriuchi, J. *et al*. (1986). Association between HLA and Sjögren's syndrome in Japanese patients. *Arthritis and Rheumatism*, **29**, 1518–21.

Moutsopoulos, H. M. and Manoussakis M. N. (1989). Immunopathogenesis of Sjögren's. Facts and fancy. *Autoimmunity*, **5**, 17–24.

Moutsopoulos, H. M. and Talal, N. (1987). Immunologic abnormalities in Sjögren's syndrome. In *Sjögren's syndrome: clinical and immunological aspects*, (ed. N. Talal, H. M. Moutsopoulos, and S. S. Kassan), pp. 258–65. Springer, Berlin.

Moutsopoulos, H. M. and Talal, N. (1989). New developments in Sjögren's syndrome. *Current Opinion in Rheumatology*, **1**,332–8.

Moutsopoulos, H. M. and Vlachoyiannopoulos, P. G. (1993). What would I do if I had Sjögren's syndrome. *Rheumatology Review*, **2**, 17–23.

Moutsopoulos, H. M., Ballow, J. E., Lawley, T. J., Stahl, N. I., Antonovych, T. T., and Chused, T. M. (1978). Immune complex glomerulonephritis in sicca syndrome. *American Journal of Medicine*, **64**, 955–60.

Moutsopoulos, H. M. *et al*. (1980). Sjögren's syndrome (sicca syndrome): current issues. *Annals of Internal Medicine*, **92**, 212–26.

Moutsopoulos, H. M., Steinberg, A. D., Fauci, A. S., Lane, H. C., and Papadopoulos, N. M. (1983 *a*). High incidence of free monoclonal l light chains in the sera of patients with Sjögren's syndrome. *Journal of Immunology*, **130**, 2263–5.

Moutsopoulos, H. M., Elkon, K. B., Mavridis, A. K., Acritidis, N. C., Hughes, G. R. V., and Pepys, M. B. (1983 *b*). Serum C-reactive protein in primary Sjögren's syndrome. *Clinical and Experimental Rheumatology*, **1**, 57–8.

Moutsopoulos, H. M., Costello, R., Drosos, A. A., Mavridis, A. K., and Papadopoulos, N. M. (1985). Demonstration and identification of monoclonal proteins in the urine of patients with Sjögren's syndrome. *Annals of the Rheumatic Diseases*, **44**, 109–12.

Moutsopoulos, H. M., Tzioufas, A. G., Bai, M. K., Papadopoulos, N. M., and Papadimitriou, C. S. (1990). Serum IgMk monoclonicity in patients with Sjögren's syndrome is associated with an increased proportion of k-positive plasma-cells infiltrating the labial minor salivary glands. *Annals of the Rheumatic Diseases*, **49**, 929–31.

Moutsopoulos, H. M., Cledes, J., Skopouli, F. N., Elisaf, M., and Youinou, P. (1991). Nephrocalcinosis in Sjögren's syndrome. *Journal of Internal Medicine*, **230**, 187–91.

Navarette, C. *et al*. (1985). Genetic and functional relationships of the HLA-DR and HLA-DQ antigens. *Immunogenetics*, **21**, 97–101.

Papasteriades, C., Skopouli, F. N., Drosos, A. A., Andonopoulos, A. P., and Moutsopoulos, H. M. (1988). HLA alloantigen association in Greek patients with Sjögren's syndrome. *Journal of Autoimmunity*, **1**, 85–90.

Papiris, S. A., Skopouli, F. N., Maniati, M. A., Constantopoulos, C. H., and Moutsopoulos, H. M. (1994). Bronchiolitis in primary Sjögren's syndrome. In *Sjögren's syndrome. State of the art* (ed. M. Homma, S. Sugai, T. Tojo, N. Miyakaka, and M. Akizuki), pp. 431–2. Kugler Publications, Amsterdam.

Parrago, G., Rain, G. D., Brochierion, C., and Rocher, F. (1987). Scintigraphy of the salivary glands in Sjögren's syndrome. *Journal of Clinical Pathology*, **40**, 1463–7.

Pavlidis, N. A., Karsh, J., and Moutsopoulos, H. M. (1982). The clinical picture of primary Sjögren's syndrome. *Journal of Rheumatology*, **40**, 1463–7.

Pavlidis, N. A. *et al*. (1992). Lymphoma in Sjögren's syndrome. *Medical Pediatric Oncology*, **20**, 279–83.

Pennec, Y. L., Leroy, J. P., Jouquan, J., Lelong, A., Katsikis, P., and Youinou, P. (1990). Comparison of labial and subclinical salivary gland biopsies in the diagnosis of Sjögren's syndrome. *Annals of the Rheumatic Diseases*, **49**, 37–9.

Reveille, J. D., Macleod, M. J., Whittington, K., and Arnett, F. C. (1991). Specific amino acid residues in the second hypervariable region of HLA-DQA1 and DQB1 chain genes promote the Ro(SSA)/La(SSB) autoantibody responses. J *ournal of Immunology*, **146**, 3871–5.

Schmidt, U., Helbron, D., and Lennert, K. (1982). Development of malignant lymphoma in myoepithelial sialadenitis (Sjögren's syndrome). *Virchow's Archives*, **395**, 11–43.

Scott, J. (1980). Quantitative observations on the histology of human salivary glands obtained post mortem. *Journal of Biologie Buccale*, **8**, 187–92.

Shillitoe, E. J., Daniels, T. E., Whitcher, J. P., Strand, D. V., Talal, N., and Greenspan, J. S. (1982). Antibody to cytomegalovirus in patients with Sjögren's syndrome as detected by enzyme-linked immunosorbent assay. *Arthritis and Rheumatism*, **25**, 260–5.

Siamopoulou-Mavridou, A., Drosos, A. A., and Andonopoulos, A. P. (1989). Sjögren's syndrome in childhood: report of two cases. *European Journal of Pediatrics*, **18**,523–4.

Skopouli, F. N., Siouna-Fatourou, H. I., Ziciadis, C., and Moutsopoulos, H. M. (1989). Evaluation of unstimulated whole saliva flow rate and stimulated parotid flow as confirmatory tests for xerostomia. *Clinical and Experimental Rheumatology*, **7**, 127–9.

Skopouli, F. N. *et al*. (1990). Raynaud's phenomenon in primary Sjögren's syndrome. *Journal of Rheumatology*, **17**, 618–20.

Skopouli, F. N., Fox, P. C., Galanopoulou, V., Atkinson, J. C., Jaffe, E. C., and Moutsopoulos, H. M. (1991). T-cell subpopulation in the labial minor salivary gland histopathologic lesion of Sjögren's syndrome. *Journal of Rheumatology*, **18**, 210–14.

Skopouli, F. N., Kousvelari, E., Mertz, P., Jaffe, E. S., Fox, P. C., and Moutsopoulos, H. M. (1992). c- *myc* mRNA expression in minor salivary glands of patients with Sjögren's syndrome. *Journal of Rheumatology*, **19**, 693–9.

Skopouli, F. N., Barbatis, C., and Moutsopoulos, H. M. (1994). Liver involvement in primary Sjögren's syndrome. *British Journal of Rheumatology*, **33**, 745–8.

St Clair, E. W., Angellilo, J. C., and Signer, K. H. (1992). Expression of cell adhesion molecules in the salivary gland microenvironment of Sjögren's syndrome. *Arthritis and Rheumatism*, **35**, 62–6.

Sumida, T. *et al*. (1992). T-cell receptor repertoire of infiltrating T-cells in lips of Sjögren's syndrome patients. *Journal of Clinical Investigation*, **89**, 681–5.

Talal, N., Dauphinee, M. J., Dang, H., Alexander, S. S., Hart, D. J., and Garry, R. F. (1990). Detection of serum antibodies to retroviral proteins n patients with primary Sjögren's syndrome (autoimmune exocrinopathy). *Arthritis and Rheumatism*, **33**, 774–81.

Tambur, A. R. *et al*. (1993). Molecular analysis of HLA class II genes in primary Sjögren's syndrome: a study of Israeli and Greek non-Jewish patients. *Human Immunology*, **36**,235–42.

Tarpley, T. M., Anderson, L. G., and White, C. L. (1974). Minor salivary gland involvement in Sjögren's syndrome. *Oral Surgery, Oral Medicine, Oral Pathology*, **37**, 64–74.

Teramoto, N. *et al*. (1989). Annular erythema: a possible association with primary Sjögren's syndrome. *Journal of the American Academy of Dermatology*, **20**, 596–601.

Trevino, H., Tsianos, E. B. and Schenkers, S. (1987). Gastrointestinal and hepatobiliary features in Sjögren's syndrome. In *Sjögren's syndrome. Clinical and immunological aspects* (ed. N. Talal, H. M.

Moutsopoulos, and S. S. Kassan), pp. 89–95. Springer, Berlin.

Tsampoulas, C. G. *et al*. (1990). Hand radiographic changes in patients with primary and secondary Sjögren's syndrome. *Scandinavian Journal of Rheumatology*, **15**, 333–9.

Tsianos, E. B., Tzioufas, A. G., Kita, M. D., Tsolas, O., and Moutsopoulos, H. M. (1984). Serum isoamylases in patients with autoimmune rheumatic diseases. *Clinical and Experimental Rheumatology*, **2**, 235–8.

Tsianos, E. B. *et al*. (1990). Sjögren's syndrome in patients with primary biliary cirrhosis. *Hepatology*, **11**, 730–4.

Tsokos, M., Lazarou, S. A., and Moutsopoulos, H. M. (1987). Vasculitis in primary Sjögren's syndrome: histologic classification and clinical presentation. *American Journal of Clinical Pathology*, **88**, 26–31.

Tzioufas, A. G., Manoussakis, M. N., Costello, R., Silis, M., Papadopoulos, N. M., and Moutsopoulos, H. M. (1986). Cryoglobulinemia in autoimmune rheumatic diseases: evidence of circulating monoclonal cryoglobulins in patients with primary Sjögren's syndrome. *Arthritis and Rheumatism*, **29**, 1098–104.

Tzioufas, A. G., Boumba, D. S., Skopouli, F. N., and Moutsopoulos, H. M. (1996). Mixed monoclonal cryoglobulinemia and monoclonal rheumatoid factor cross-reactive idiotypes as predicting factors for lymphoma development in primary Sjögren's syndrome. *Arthritis and Rheumatism*, **39**, 767–72.

Tzioufas, A. G., Staikos, P., and Moutsopoulos, H. M. (1994). Epidemiology of Sjögren's syndrome in Greece. In *Sjögren's syndrome. State of the art* (ed. M. Homma, S. Sugai, T. Tojo, N. Miyakaka, and M. Akizuki), pp. 395–6. Kugler Publications, Amsterdam.

Van de Merwe, J. P., Kamerling, R., Arendsen, H. S., Mulder, A. H., and Hooijkaas, H. (1993). Sjögren's syndrome in patients with interstitial cystitis. *Journal of Rheumatology*, **20**, 962–6.

Venables, P. J. W., Ross, M. G. R., Charles, P. J., Melson, R. D., Griffith, P. D., and Maini, R. N. (1985). A seroepidemiological study of cytomegalovirus and Epstein–Barr virus in rheumatoid arthritis and sicca syndrome. *Annals of the Rheumatic Diseases*, **44**, 742–6.

Vitali, C. *et al*. (1986). HLA antigens in Italian patients with primary Sjögren's syndrome. *Annals of the Rheumatic Diseases*, **45**, 412–16.

Vitali, C. *et al*. (1988). Parotid sialography and minor salivary gland biopsy in the diagnosis of Sjögren's syndrome: a comparative study of 84 patients. *Journal of Rheumatology*, **15**, 262–7.

Vitali, C. *et al*. (1992*a*). Preliminary criteria for the classification of Sjögren's syndrome. Results of a prospective concerted action supported by the European Community. *Arthritis and Rheumatism*, **36**, 340–8.

Vitali, C. *et al*. (1992*b*). Anti hepatitis C virus antibodies in primary Sjögren's syndrome: false positive results are related to hyper-g-globulinemia. *Clinical and Experimental Rheumatology*, **10**, 103–4.

Whaley, K. *et al*. (1972). Sjögren's syndrome and autoimmunity in a geriatric population. *Age and Ageing*, **1**, 197–200.

Whittingham, S., Nasseli, G., Hicks, J. D., O'Brien C., and Sculley, T. B. (1991). Studies on Sjögren's syndrome in Scid mice. *Clinical and Experimental Rheumatology*, **9**, 332 (Abstr.).

Wolf, H., Haus, M., and Wilmes, E. (1984). Persistence of Epstein–Barr virus in the parotid gland. *Journal of Virology*, **51**, 795–8.

Yannopoulos, D. I. *et al*. (1992). Conjunctival epithelial cells from patients with Sjögren's syndrome inappropriately express major histocompatibility complex molecules, La/SSB antigen and heat shock proteins. *Journal of Clinical Immunology*, **12**, 259–65.

Yonaha, F., Sumida, T., Maeda, T., Tomioka, H., Koike, T., and Yoshioda, S. (1992). Restrictive junctional usage of T-cell receptor Vb2 And Vb13 genes, which are overpresented on infiltrating T-cells in the lips of patients with Sjögren's syndrome. *Arthritis and Rheumatism*, **35**, 1362–6.

Youinou, P., Pennec, Y. L., Katsikis, P., Jouquan, J., Faugment, P., and LeGoll, P. (1990 *a*). Raynaud's phenomenon in primary Sjögren's syndrome. *British Journal of Rheumatology*, **29**, 205–7.

Youinou, P., Moutsopoulos, H. M., and Pennec, Y. L. (1990*b)*. Clinical features of Sjögren's syndrome. *Current Opinion in Rheumatology*, **2**, 687–93.

Zulman, J., Jaffe, R., and Talal, N. (1978). Evidence that the malignant lymphoma of Sjögren's syndrome is a monoclonal B-cell neoplasm. *New England Journal of Medicine*, **299**, 1215–20.

5.11.1 Classification of vasculitis

David G. I. Scott and Richard A. Watts

Introduction

Vasculitis means inflammation of blood vessels. Implicit in this definition is that the blood vessel is the primary site of inflammation. The blood vessel wall is thus infiltrated with inflammatory cells and perivascular cuffing does not equate with vasculitis. The consequence of such inflammation is often destruction of the vessel wall which is seen on histology as fibrinoid necrosis (Fig. 1). It is for this reason that many use the term 'necrotizing vasculitis'.

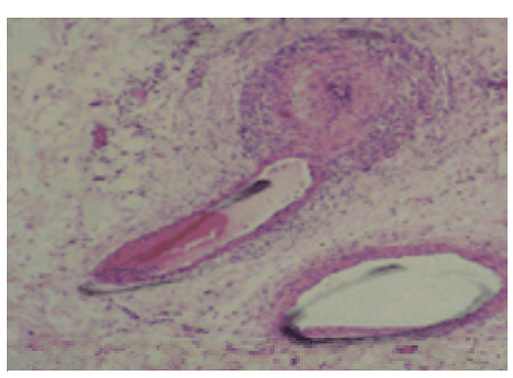


Fig. 1 Histology showing typical necrotizing vasculitis

The vasculitides are a heterogeneous group of relatively uncommon diseases which can arise *de novo* (e.g. polyarteritis nodosa, Wegener's granulomatosis) or as a secondary feature of an established clinical disease such as rheumatoid arthritis or systemic lupus erythematosus.

The consequences of such vascular inflammation depend upon the size, site, and the number of blood vessels involved. Vasculitis can occasionally be ocalized and clinically insignificant but more commonly is generalized and potentially life threatening, especially when small muscular arteries are involved. Muscular arteries may develop focal or segmental lesions. The former (affecting part of the vessel wall) may lead to aneurysm formation and possibly rupture; segmental lesions (affecting the whole circumference) are more common and lead to stenosis or occlusion with distal infarction. Haemorrhage or infarction of vital internal organs are the most serious problems of systemic vasculitis and explain the poor prognosis of untreated polyarteritis nodosa (Frohnert and Sheps 1967) or of arteritis complicating rheumatoid arthritis (Schmid *et al*. 1961; Bywaters and Scott 1963; Scott *et al*. 1981). Small vessel vasculitis, by contrast, most commonly affects the skin and rarely causes dysfunction of internal organs. Widespread small-vessel vasculitis may cause problems, especially in the kidney, when sufficient numbers of adjacent vessels are affected with significant release of inflammatory mediators or where overall perfusion is threatened.

Epidemiology

The epidemiology of the systemic vasculitides is documented poorly. Many studies have been from tertiary referral centres with the problems of referral bias and uncertainty of denominator population or have involved small populations. We have estimated the incidence of the major forms of systemic vasculitis in a stable, ethnically homogeneous population—the Norwich (United Kingdom) Health Authority (414 000 adults) over a 6-year period between 1988 and 1994 (Watts *et al*. 1995). The overall annual incidence of systemic vasculitis (excluding giant cell arteritis) was 38.6 per million (95 per cent confidence intervals 31.3 to 47.2). The commonest systemic vasculitides were systemic rheumatoid vasculitis—12.5 per million per year (95 per cent confidence intervals 8.5 to 17.7) and Wegener's granulomatosis—8.5 per million per year (95 per cent confidence intervals 5.2 to 12.9). Details of the incidence of this and the other commoner systemic vasculitides are shown in Table 1. These data suggest that the overall incidence of systemic vasculitis is significantly greater than previously thought (estimated at 10 per million per year). Whether this represents a genuine increased incidence with time or increased physician awareness, especially in association with the introduction of the antineutrophil cytoplasmic antibody (**ANCA**) test is uncertain.

	Number of patients	Annual incidence (per million; 95% confidence intervals)
Systemic rheumatoid vasculitis	31	12.5 (8.5–17.7)
Wegener's granulomatosis	21	8.5 (5.2–12.9)
Churg–Strauss syndrome	6	2.4 (0.9–5.3)
Microscopic polyangiitis	6	2.4 (0.9–5.3)
Henoch–Schönlein purpura	3	1.2 (0.3–3.5)
Systemic lupus erythematosus[a]	9	3.6 (1.7–6.9)
Miscellaneous[b]	8	
Unclassified	12	
Systemic vasculitis (overall)[c]	96	38.6 (31.3–47.2)

The adult population of the Norwich Health Authority is 414 000
[a]Arteritis diagnosed on clinical or histological evidence.
[b]Wegener, sarcoma, cryoglobulinaemia.
[c]Overall figure excludes giant cell arteritis and localized cutaneous vasculitis.

Table 1 Estimated annual incidence of the commoner vasculitides in the Norwich Health Authority, 1988 to 1994

Classification of vasculitis

Classification means distribution in classes or groups according to a method or system. Classification of vasculitis is confusing because of the considerable overlap between the different vasculitic syndromes and because the cause of the vasculitis is usually unknown.

Historical

Kussmaul and Maier (1866) are credited with the first description of 'periarteritis nodosa' in 1866 when they described a 'new disease' characterized by numerous nodules along the course of small muscular arteries. Earlier descriptions by Rokitansky (1852), Pelletan (1810), and possibly by Michaelis and Matani (1755) referenced by Lamb (1914), suggest that formal recording of the disease is at least 200 years old.

In the early part of the twentieth century there were large numbers of reports of vasculitis labelled as periarteritis nodosa, including patients with rheumatic fever (von Glahn and Pappenheimer 1926), hypersensitivity states in animals and man (Rich 1942; Rich and Gregory 1943), and most cases with histological evidence of vasculitis affecting any size of vessel, leading to considerable confusion.

In 1952, Zeek reviewed the literature relating to vasculitis and periarteritis nodosa and used the generic term 'necrotizing angiitis' to indicate the specific damage to the blood vessel wall rather than the presence of anti-inflammatory cells alone (see above); she classified these into five distinct entities: (i) hypersensitivity angiitis, (ii) allergic granulomatous angiitis, (iii) rheumatic arteritis, (iv) periarteritis nodosa, and (v) temporal arteritis. Almost all modern classifications are based on this system, which essentially combined histological changes and clinical features. Refinements included use of the term polyarteritis nodosa in the 1950s, the identification of patients with and without lung involvement (Rose and Spencer 1957), and the relationship between vasculitis and granuloma formation discussed by Alarcon-Segovia and Brown (1964) , who considered the vasculitides to represent 'a continuous spectrum of tissue changes ranging from pure necrosis and granuloma formation to pure angiitis'. Two notable omissions from Zeek's classification were Wegener's granulomatosis and Takayasu's arteritis; these were not fully described in the English literature until after 1953 (Lie 1988).

1970 to 1990

In the 1970s it was realized that there was a considerable overlap in the size of arteries involved (Gilliam and Smiley 1976). At this time Fauci advocated a classification scheme which, apart from including more conditions, also included Churg–Strauss syndrome in an overlapping group of systemic necrotizing vasculitides (Fauci *et al*. 1978).

During the 1980s, classification schemes developed based on the histopathological changes, in particular the size of the predominant vessel involved (e.g. Scott 1988; Lie 1988; Lie 1994). Lie also included a group of infectious vasculitides and vasculitis 'look-alikes'. The latter group includes atheroembolism which can present with clinical features very similar to polyarteritis nodosa. Furthermore he divided the vasculitides into primary and secondary vasculitis, that is vasculitis of unknown aetiology and that occurring secondary to either an infection or some other disease process, typically an autoimmune rheumatic disease.

Current classifications

In 1993 an international consensus conference was convened in Chapel Hill (**CHCC**) which developed definitions for the nomenclature of the systemic vasculitides based on clinical and laboratory features (Table 2) and a new classification scheme based on vessel size (Jeanette *et al*. 1994).

Table 2 Names and definitions of vasculitides adopted by the Chapel Hill Consensus Conference on the nomenclature of systemic vasculitis [a]

This and other current classification schemes have not addressed pathogenic mechanisms, in particular the relationship between ANCA and systemic vasculitis. Antibodies directed against proteinase 3 are associated strongly with Wegener's granulomatosis and those against myeloperoxidase with microscopic polyangiitis.

All previous studies required either histopathological or angiographic evidence of vasculitis. The increasing use of tests for ANCA, in particular, has resulted in some patients having a diagnosis of vasculitis made without biopsy or angiographic confirmation and this has resulted in changes in diagnostic emphasis. For example the increase in Wegener's granulomatosis seen in Leicester (United Kingdom) during the 1980s was attributed partly to increased diagnostic awareness following introduction of assays for ANCA (Andrews *et al*. 1990). If a classification system is to be useful then it should reflect aetiopathogenesis and/or approaches to treatment.

We feel that because of these developments the CHCC system is inadequate and we have developed a modification (Table 3) which reflects not only dominant vessel size but also ANCA (Scott and Watts 1994). Unlike the CHCC classification we have split Wegener's granulomatosis, Churg–Strauss syndrome, and microscopic polyangiitis from the rest because: (i) they often involve small arteries; (ii) they are diseases most commonly associated with ANCA; (iii) they are associated with a high risk of glomerulonephritis; and (iv) they are diseases which respond best to immunosuppresion with cyclophosphamide. The aetiology of these diseases is probably unrelated to immune complex formation, in contrast to pure small-vessel vasculitis such as Henoch–Schönlein purpura and essential mixed cryoglobulinaemia. Our classification also reflects broad therapeutic strategies (Table 4), with the medium and small vessel group responding best to immunosuppresion with cyclophosphamide in addition to corticosteroids, the large vessel group requiring moderate- to high-dose corticosteroids, usually alone, and the small vessel group only sometimes requiring corticosteroids at a low dose.

Dominant vessel involved	Primary	Secondary
Large arteries	Giant cell arteritis Takayasu's arteritis Isolated CNS angiitis	Arthritis associated with RA Infection (e.g. syphilis)
Medium arteries	Classical PAN Kawasaki disease	Infection (e.g. hepatitis B) Hairy cell leukaemia
Small vessels and medium arteries	Wegener's granulomatosis* Churg–Strauss syndrome* Microscopic polyangiitis*	Vasculitis 2° to RA, SLE, SS Drugs, malignancy Infection (e.g. HIV)
Small vessels (leucocytoclastic)	Henoch–Schönlein purpura Essential mixed cryoglobulinaemia Cutaneous leucocytoclastic angiitis	Drugs*, malignancy Infection (e.g. hepatitis B,C)

*Diseases most commonly associated with ANCA (antimyeloperoxidase and antiproteinase 3 antibodies), a significant risk of renal involvement, and which are most responsive to immunosuppression with cyclophosphamide.
†e.g. sulphonamides, penicillins, thiazide diuretics, and many others.
CNS, central nervous system; HIV, human immunodeficiency virus; PAN, polyarteritis nodosa; RA, rheumatoid arthritis; SLE, systemic lupus erythematosus; SS, Sjögren's syndrome.
From Scott and Watts (1994).

Table 3 Classification of systemic vasculitis

Dominant vessel involved	Corticosteroids alone	Cyclophosphamide + corticosteroids	Others
Large arteries	+++	–	+
Medium arteries	+	++	++[a]
Small vessels and medium arteries	+	+++	–
Small vessels	+	–	++

[a]Includes plasmapheresis, antiviral therapy for hepatitis B associated vasculitis, and intravenous immunoglobulin for Kawasaki disease.
Reproduced from Scott and Watts (1994) with permission.

Table 4 Relationship between vessel size and response to treatment

Despite these developments in classification and definition of the vasculitides a number of problems still remain. There is still considerable overlap between the groups (Fig. 2). For example palpable purpura due to leucocytoclastic vasculitis of small vessels affects up to 20 per cent of patients with 'polyarteritis nodosa' as defined in studies by Cohan and Adu (Cohan *et al*. 1980; Adu *et al*. 1987), but the CHCC definition of polyarteritis nodosa specifically excludes such microscopic/small vessel involvement. Necrotizing arteritis can involve the temporal arteries, mimicking giant cell arteritis (Morgan and Harris 1978; Fraha and Abu-Haider 1979). Giant cell arteritis itself can affect smaller vessels in the breast (Potter *et al*. 1981) and mimic carcinoma, and also involve the posterior ciliary artery resulting in blindness. Classification should thus be based on dominant vessel size and not be over-restrictive. The CHCC classification has not received universal acceptance because of these problems (Lie 1994).

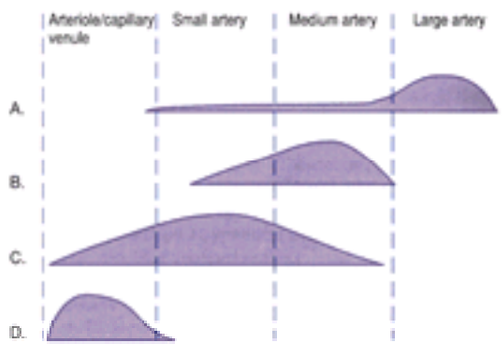


Fig. 2 Relationship between vessel size and classification.

Localized vasculitis produces particular problems in classification. It is important to exclude systemic symptoms and so confirm the benign nature of the vasculitis with long-term follow-up studies. Localized 'polyarteritis nodosa' has been described in a number of sites, including the gall bladder, uterus, and skin (Remigo and Zaino 1970; Borrie 1972; Diaz-Perez and Winkelmann 1974; Scott *et al*. 1982). Whether any of the described cases represents vasculitis truly localized to an individual vessel is debatable, but the lack of progression of cutaneous polyarteritis in long-term studies emphasizes that these cases should be classified and treated differently to their systemic counterparts.

The concept of localized giant-cell arteritis is more difficult. Most cases have arteritis restricted to the head and neck, but there are sporadic reports of giant cell arteritis at other sites, including the aorta, breast, and skin (Klein *et al*. 1975; Potter *et al*. 1981; Goldberg *et al*. 1987; Smith *et al*. 1988). The most convincing case of localized giant-cell arteritis was a mesenteric arteritis leading to bowel perforation; there was no evidence of giant cell arteritis at other sites and the patient remained well after surgery, requiring no drug therapy during an 18 month follow-up (Smith *et al*. 1988). Such cases stress the importance of the word 'systemic' in classification of arteritis, and the importance of detailed investigations, including histology, before embarking on potentially harmful drug treatment.

Rheumatoid vasculitis also causes problems with classification (see Table 3). A wide range of vessels may be involved, from digital capillaries to medium-sized arteries. Aortitis involving the aortic arch has been described in a few cases, so classification can involve all four groups. Aneurysms have been described rarely (indicating true involvement of medium-sized arteries) but we have seen arteritic changes on angiograms in four patients with severe systemic rheumatoid vasculitis.

Classification criteria for individual vasculitides

The American College of Rheumatology (**ACR**) has now published criteria for the classification of vasculitis (Fries *et al*. 1990). This series of papers describes 807 patients with seven different vasculitic diseases—polyarteritis nodosa, Wegener's granulomatosis, Churg–Strauss vasculitis, hypersensitivity vasculitis, Henoch–Schönlein purpura, giant cell arteritis and Takayasu's arteritis, seen in 48 medical centres in North America. These patients have been analysed to look for classification criteria for each specific disease; that is, those clinical findings that both identify the disease and separate it from others. The authors stress the classification criteria provide a standard way of evaluating and describing groups of patients in therapeutic, epidemiological, or other studies. A list of the proportion of patients in a group that fulfils the criteria gives considerable information about the clinical status of the patients included. The ACR criteria were presented in two forms: a traditional table and a tree format. The sensitivity and specificity rates varied considerably: 71.0 to 95.3 per cent for sensitivity and 78.7 to 99.7 per cent for specificity (Fries *et al*. 1990; Hunder *et al*. 1990). The most sensitive and specific criteria were found in Churg–Strauss syndrome, giant cell arteritis and Takayasu's arteritis; hypersensitivity vasculitis was the least well-defined condition. The criteria were not tested against the general population or against patients with other autoimmune rheumatic diseases or rheumatic conditions. These papers are important as for the first time they enable comparison of patients reported in an uniform manner—a problem which hitherto has bedevilled the study of vaculitis.

Polyarteritis nodosa and microscopic polyangiitis

Polyarteritis nodosa is a multisystem disease characterized by inflammation and necrosis of medium-sized muscular arteries leading to aneurysm formation and organ infarction (Kussmaul and Maier 1866). Davson in 1948 showed that in some patients who at autopsy had extrarenal vasculitis involving small and medium arteries, the

only detectable renal lesion was capillary inflammation (Davson *et al*. 1948). These patients have been said to have 'microscopic' polyarteritis. The main clinical feature in these patients is rapidly progressive renal failure. This is in contrast to patients in whom the dominant involvement is of medium-sized arteries, the condition now called 'classical' polyarteritis nodosa, with organ infarction as the main clinical feature (e.g. gut, nerve, or renal infarction). Microscopic polyarteritis is associated with ANCA, usually of myeloperoxidase specificity, unlike classical polyarteritis nodosa which is usually ANCA negative.

The ACR (1990) described criteria for classical polyarteritis nodosa but not for microscopic polyarteritis. The classification criteria for polyarteritis nodosa are given in Table 5 and have greater than 80 per cent specificity and sensitivity (Lightfoot *et al*. 1990). A classification tree was also constructed using the following six criteria: angiographic abnormality (Fig. 3), biopsy-proven granulocyte or mixed leucocyte infiltrate in arterial wall, neuropathy, sex, weight loss greater than 6.5 kg, and elevation of serum hepatic enzymes. This gave a sensitivity of 87.3 per cent and specificity of 89.3 per cent. These ACR criteria do not have an absolute requirement for a tissue diagnosis or arteriographic abnormality to make the diagnosis of polyarteritis nodosa.

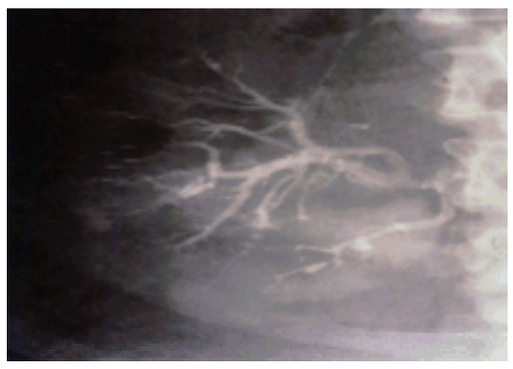


Fig. 3 Renal angiogram showing typical small aneurysms.

Criterion	Definition
1. Weight loss >4 kg	Loss of 4 kg or more of body weight since illness began, not due to dieting or other factors
2. Livedo reticularis	Mottled reticular pattern over the skin of portions of the extremities or torso
3. Testicular pain or tenderness	Pain or tenderness of the testicles, not due to infection, trauma, or other causes
4. Myalgias, weakness, or leg tenderness	Diffuse myalgias (excluding shoulder and hip girdle) or weakness of muscles or tenderness of leg muscles
5. Mononeuropathy or polyneuropathy	Development of mononeuropathy, multiple mononeuropathies, or polyneuropathy
6. Diastolic BP >90 mmHg	Development of hypertension with the diastolic BP higher than 90 mmHg
7. Elevated blood urea or creatinine	Elevation of BUN >40 mg/dl or creatinine >1.5 mg/dl, not due to dehydration or obstruction
8. Hepatitis B virus	Presence of hepatitis B surface antigen or antibody in serum
9. Arteriographic abnormality	Arteriogram showing aneurysms or occlusions of the visceral arteries, not due to atherosclerosis, fibromuscular dysplasia, or other non-inflammatory causes
10. Biopsy of small or medium-sized artery containing PMN	Histological changes showing the presence of granulocytes or granulocytes and mononuclear leucocytes in the artery wall

For classification purposes, a person shall be said to have polyarteritis nodosa if at least three of these ten criteria are present. The presence of any two or more criteria yields a sensitivity of 85.9%, and a specificity of 86.6%. BP, blood pressure; BUN, blood urea nitrogen; PMN, polymorphonuclear neutrophils.
Reproduced from Lightfoot *et al*. (1990) with permission.

Table 5 ACR 1990 criteria for the classification of polyarteritis nodosa (traditional format)

The CHCC specifically addressed the definition of both classical and microscopic polyarteritis (Table 2) (Jeanette *et al*. 1994). They defined classical polyarteritis nodosa as a disease of medium- and small-sized arterities without glomerulonephritis or vasculitis in arterioles, capillaries, or venules. This complete exclusion of small vessel disease has caused controversy because some patients show considerable overlap. A segmental necrotizing glomerulonephritis can occur in some patients with aneurysms of medium arteries. More commonly segmental necrotizing glomerulonephritis occurs without evidence of interlobar or interlobular arteritis and this explains why the concept of microscopic polyarteritis nodosa was introduced (Wainwright and Davson 1950). The glomerulonephritis in microscopic polyangiitis and Wegener's granulomatosis is similar to other vasculitic diseases, including systemic lupus erythematosus and Henoch–Schönlein purpura, except for the very low levels of immunoglobulin or complement deposition. The term 'pauci-immune' glomerulonephritis is sometimes used. The CHCC defined microscopic polyarteritis as a necrotizing vasculitis predominately affecting small vessels but small- and medium-sized arteries may be involved. They prefer, as we do, the term microscopic polyangiitis for this pattern of disease, emphasizing the difference between it and classical polyarteritis nodosa.

Polyarteritis nodosa has always been a rare disease with an estimated annual incidence of 2 to 10 per million (Sack *et al*. 1975; Scott *et al*. 1982; Kurland *et al*. 1984). Detailed analysis of the patients in these studies suggests that they included some patients with microscopic polyangiitis and Churg–Strauss syndrome, and hence that the annual incidence may be less than they reported. Using the CHCC definition of classical polyarteritis nodosa we have not seen a single case since 1988, suggesting that it has become a very rare disease. However, comparing the ACR criteria for polyarteritis nodosa with the CHCC definitions we have seen five patients with microscopic polyangiitis who fulfil the ACR criteria for polyarteritis nodosa and a further five with microscopic polyangiitis who do not. Seven of these ten patients with microscopic polyangiitis are ANCA positive.

The annual incidence of microscopic polyangiitis in Leicester (United Kingdom) increased from 0.5 to 3.3 per million following the introduction of the ANCA test in 1987 (Andrews *et al*. 1990). Data from our epidemiological studies in Norwich suggest that the annual incidence of microscopic polyangiitis is very similar at 2.4 per million (Watts *et al*. 1995) Whether these changes in incidence represent a true change or reflect changes in diagnostic patterns is uncertain.

Wegener's granulomatosis (see also Chapter 5.11.2)

Wegener first described this form of granulomatous necrotizing vasculitis in 1936 (Wegener 1936). The disease is characterized by necrotizing granulomata of the upper and lower respiratory tract, a necrotizing systemic vasculitis, and a focal glomerulonephritis (Godman and Churg 1954). A more limited form of the disease has been described with lesions restricted to the upper and lower respiratory tract (Carrington and Liebow 1966). This was thought to represent a more benign form of the disease because of the absence of overt renal disease, although 5 out of 16 patients died from non-renal disease within 1 year of diagnosis. Respiratory tract disease may precede systemic and renal vaculitis by many months or years. DeRemee classified Wegener's granulomatosis on the basis of the organs involved (ELK)—E standing for ear, nose, throat; L for lung involvement; and K for kidney involvement (DeRemee *et al*. 1976). This classification has proved useful in staging patients with the disease, but presupposes a progression from E–L–K, which does not always occur. Patients may present with pulmonary or renal involvement. More recently this progression has been viewed as a continous spectrum which may be entered at any point (Luqmani *et al*. 1994a).

The ACR 1990 classification criteria for the diagnosis of Wegener's granulomatosis are shown in traditional format in Table 6 (Leavitt *et al*. 1990). A classification tree was also constructed with five criteria: the same four criteria plus haemoptysis. This gave a sensitivity of 87.1 per cent and specificity of 93.6 per cent.

Criterion	Definition
1. Nasal or oral inflammation	Development of painful or painless oral ulcers, or purulent or bloody nasal discharge
2. Abnormal chest radiograph	Chest radiograph showing the presence of nodules, fixed infiltrates, or cavities
3. Urinary sediment	Microhaematuria (>5 red blood cells per high power field) or red cell casts in urine sediment
4. Granulomatous inflammation on biopsy	Histological changes showing granulomatous inflammation within the wall of an artery or in the perivascular or extravascular area (artery or arteriole)

For purposes of classification, a person shall be said to have Wegener's granulomatosis if at least two of these four criteria are present. The presence of any two or more criteria yields a sensitivity of 88.2% and a specificity of 92.0%.
Reproduced from Leavitt *et al*. (1990) with permission.

Table 6 ACR 1990 criteria for the classification of Wegener's granulomatosis (traditional format)

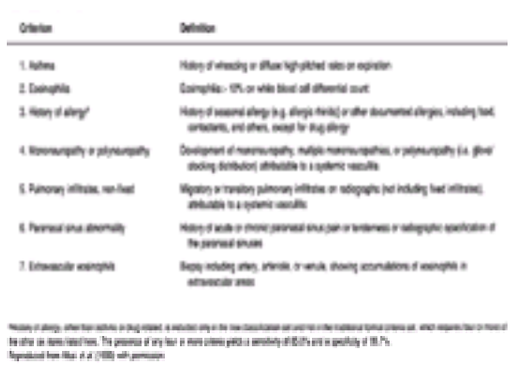
The CHCC definition restricts the term Wegener's granulomatosis to patients with necrotizing granulomatous inflammation (Table 2) (Jeanette *et al*. 1994). In the earlier stages of disease there may be overlap between Wegener's granulomatosis and microscopic polyangiitis, and subsequently the illness may evolve with development of new inflammatory lesions. Neither the CHCC nor the ACR classification scheme address the issue of limited disease.

The incidence of Wegener's granulomatosis appears to be increasing. In 1980 to 1986 in Leicester (United Kingdom) the annual incidence of Wegener's granulomatosis was estimated to be 0.7 per million, while in 1987 to 1989 the incidence was 2.8 per million (Andrews *et al*. 1990). In Norwich during the period 1988 to 1994 the estimated annual incidence was 8.5 per million (Carruthers *et al*. 1996). This does not appear to be due to changing diagnostic criteria as we have compared the clinical features of our patients to those seen at the National Institutes of Health (**NIH**, United States; 1962 to 1983) and Birmingham (United Kingdom; 1981 to 1991) and did not find any significant differences between the three groups. The Leicester group used the same diagnostic criteria as the NIH. It is, however, noteworthy that as the annual incidence of Wegener's granulomatosis has apparently increased that of classical polyarteritis nodosa has decreased, suggesting that some patients formerly labelled as polyarteritis nodosa may now be reclassified as Wegener's granulomatosis.

Churg–Strauss syndrome

In 1951, Churg and Strauss described the postmortem features of 13 patients who died following an illness characterized by asthma, eosinophilia, fever, and a systemic illness (Churg and Strauss 1951). Histologically there is a granulomatous necrotizing vasculitis. Chumbley, in 1977, reported a series of 30 patients with Churg–Strauss syndrome, and stressed the relative infrequent occurrence of renal involvement (Chumbley *et al*. 1977). Lanham, in 1984, provided a clinical definition of Churg–Strauss syndrome as a triad of asthma, eosinophilia (greater than 1×10^{9}/l), and a systemic vasculitis involving two or more extrapulmonary organs (Lanham *et al*. 1984). In their experience extravascular granulomata were not essential for the diagnosis of Churg–Strauss syndrome. They also noted a triphasic pattern of illness with allergic rhinitis, evolving into asthma; followed by peripheral blood eosinophilia and eosinophilic tissue infiltrates; and finally a systemic vasculitis phase. The granulomata may be localized and associated with a variety of systemic manifestations. Furthermore the necrotizing vasculitis may be indistinguishable from that found in classical polyarteritis nodosa and/or microscopic polyangiitis.

The ACR criteria for the diagnosis of Churg–Strauss syndrome are given in Table 7 (Masi *et al*. 1990). These criteria do not include some common clinical features of Churg–Strauss syndrome such as rash or cardiac involvement, as they gave poor discrimination. The combination of asthma and eosinophilia are both sensitive and highly specific for the diagnosis of Churg–Strauss syndrome. A classification tree was also constructed using three of six criteria that gave a sensitivity of 95 per cent and a specificity of 99.2 per cent.



Criterion	Definition
1. Asthma	History of wheezing or diffuse high-pitched rales on expiration
2. Eosinophilia	Eosinophils > 10% on white blood cell differential count
3. History of allergy[a]	History of seasonal allergy (e.g. allergic rhinitis) or other documented allergies, including food, corticosteroids, and others, except for drug allergy
4. Mononeuropathy or polyneuropathy	Development of mononeuropathy, multiple mononeuropathies, or polyneuropathy (i.e. glove/stocking distribution) attributable to a systemic vasculitis
5. Pulmonary infiltrates, non-fixed	Migratory or transitory pulmonary infiltrates on radiographs (not including fixed infiltrates) attributable to a systemic vasculitis
6. Paranasal sinus abnormality	History of acute or chronic paranasal sinus pain or tenderness or radiographic opacification of the paranasal sinuses
7. Extravascular eosinophils	Biopsy including artery, arteriole, or venule, showing accumulations of eosinophils in extravascular areas

[a]History of allergy, either from asthma, or drug-related, as instructed only on the new classification and used not in the traditional formal criteria set, which requires four or more of the criteria listed here. The presence of any four or more criteria yields a sensitivity of 85.07% and a specificity of 99.7%.
Reproduced from Hoffman *et al*. (1992) with permission

Table 7 ACR 1990 criteria for the classification of Churg–Strauss syndrome

The CHCC definition of Churg–Strauss syndrome (Table 2) includes the presence of asthma and eosinophilia (Jeanette *et al*. 1994). The presence of conspicious eosinophils in inflammatory infiltrates is not alone a discriminating feature as they can occur in other types of vasculitis including Wegener's granulomatosis and microscopic polyangiitis.

The annual incidence of Churg–Strauss syndrome has been estimated to be 2.4 per million (Watts *et al*. 1995).

Hypersensitivity vasculitis

Vasculitis occurring secondary to allergic or hypersensitivity mechanisms has been considered to be a distinct entity since the late 1940s (Zeek *et al*. 1948) and was included in Zeek's original classification (Zeek 1952). Distinguishing features were considered to be the prominent cutaneous involvement and frequent precipitation by serum or drugs. Pathologically the disease involves small vessels with leucocytoclasis. The lesions are all of the same age and can be induced experimentally. However, often there is no clear evidence of an inciting agent and a similar picture can be seen in vasculitis associated with connective tissue disease (rheumatoid arthritis, systemic lupus erythematosus, Sjögren's syndrome), malignancy, Henoch–Schönlein purpura, and cryoglobulinaemia.

The ACR 1990 criteria for the diagnosis of hypersensitivity vasculitis were relatively insensitive (71.0 per cent) and non-specific (78.5 per cent). This reflects the occurrence of a hypersensitivity-like picture in other conditions, particularly Henoch–Schönlein purpura and microscopic polyangiitis, and some authorities consider these conditions to represent different forms of hypersensitivity to foreign antigens (Heng 1985).

The CHCC did not use the term hypersensitivity vasculitis but considered that the 'categories of microscopic polyangiitis and cutaneous leucocytoclastic angiitis probably best equate with the most common usage of hypersensitivity vasculitis' (Jeanette *et al*. 1994).

Henoch–Schönlein purpura

Schönlein initially described acute purpura and arthritis occurring in children in 1837 (Schönlein 1837). Subsequently Henoch described the additional features of colicky abdominal pain and nephritis in 1874 (Henoch 1874). The classical presentation is with purpura, arthritis, haemorrhagic gastrointestinal involvement, and glomerulonephritis. It occurs most often in children, but rarely, adults of any age may be affected. Significant IgA deposition is usually seen in renal or skin biopsies. Henoch–Schönlein purpura is often considered to be a form of hypersensitivity vasculitis.

The ACR criteria readily distinguish Henoch–Schönlein purpura from other forms of vasculitis, although it was recognized that there were considerable similarities between Henoch– Schönlein purpura and hypertensitivity vasculitis, both in terms of clinical features and the classification criteria (Mills *et al*. 1990). Palpable purpura was a common criterion for both diseases. Also, although arthritis and nephritis are typical features of Henoch–Schönlein purpura, neither was sensitive for the diagnosis, and haematuria as a sign of nephritis was neither sensitive nor specific. A classification tree was also constructed with similar criteria and yielded a sensitivity of 89.4 per cent and a specificity of 88.1 per cent (Table 8).

Criterion	Definition
1. Palpable purpura	Slightly raised "palpable" haemorrhagic skin lesions, not related to thrombocytopenia
2. Age < 20 years	Patient 20 years or younger at onset of first symptom
3. Bowel angina	Diffuse abdominal pain, worse after meals, or the diagnosis of bowel ischaemia, usually including bloody diarrhoea
4. Mail granulocytes on biopsy	Histological changes showing granulocytes in the walls of arterioles or venules

For purposes of classification a patient shall be said to have Henoch–Schönlein purpura if at least two of these four criteria are present. The presence of any two or more criteria yields a sensitivity of 87.1% and a specificity of 87.7%.
Reproduced from Mills et al. (1990) with permission.

Table 8 ACR 1990 criteria for the classification of Henoch–Schonlein purpura (traditional format)

The CHCC definition of Henoch–Schönlein purpura is a small vessel vasculitis with IgA-dominant immune deposits (Table 2) (Jeanette *et al*. 1994). These deposits distinguish Henoch–Schönlein purpura from microscopic polyangiitis. The CHCC did not consider other types of immune-complex mediated disease apart from cryoglobulinaemic vasculitis.

The annual incidence of Henoch–Schönlein purpura in children is 135 per million (Stewart *et al*. 1988). Our data from Norwich suggest that the annual incidence in adults is 1.2 per million (Watts *et al*. 1995).

Takayasu's arteritis

Takayasu's arteritis is a chronic, granulomatous large-vessel arteritis which was described in 1908 by Takayasu (Takayasu 1908), but Savory had described the association between the absence of radial pulses and ocular abnormalities in 1856 (Savory 1856).

Ishikawa proposed diagnostic criteria for Takayasu's arteritis in 1988 and suggested that an age of less than 40 years should be an obligatory criterion for the diagnosis (Ishikawa 1986). In this study of 96 Japanese patients with Takayasu's arteritis two major criteria were proposed: arteriographic evidence of left or right mid-subclavian artery stenosis or occlusion. In addition minor diagnostic criteria were described: high erthrocyte sedimentation rate, carotid artery tenderness, hypertension, aortic regurgitation, or arteriographic evidence of lesions in other branches of the aorta. The presence of both major, or one major and two minor, or four or more minor criteria suggested a diagnosis of Takayasu's arteritis with a sensitivity of 84 per cent (Ishikawa 1986). These patients were, however, only compared with 12 patients having other aortic disease and not to patients with other forms of vasculitis.

The ACR classification criteria for the diagnosis of Takayasu's arteritis are given in Table 9 (Arend *et al*. 1990). Takayasu's arteritis is clearly distinguished from giant cell arteritis by age. A high erthrocyte sedimentation rate, carotid artery tenderness, and/or hypertension lack specificity and sensitivity to differentiate patients with Takayasu's arteritis from other forms of arteritis. A classification tree was also constructed with five of the same six criteria omitting claudication of a limb. This gave a sensitivity of 92.1 per cent and specificity of 97.0 per cent.

Criterion	Definition
1. Age at disease onset < 40 years	Development of symptoms or findings related to Takayasu's arteritis at age < 40 years
2. Claudication of extremities	Development and worsening of fatigue and discomfort in muscles of one or more extremity while in use, especially the upper extremities
3. Decreased brachial artery pulse	Decreased pulsation of one or both brachial arteries
4. BP difference > 10 mmHg	Difference of > 10 mmHg in systolic blood pressure between arms
5. Bruit over subclavian arteries or aorta	Bruit audible on auscultation over one or both subclavian arteries or abdominal aorta
6. Arteriogram abnormality	Arteriographic narrowing or occlusion of the entire aorta, its proximal branches, or large arteries in the proximal upper or lower extremities, not due to arteriosclerosis, fibromuscular dysplasia, or similar causes; changes usually focal or segmental

For purposes of classification a patient shall be said to have Takayasu's arteritis if at least three of these six criteria are present. The presence of any three or more criteria yields a sensitivity of 90.5% and a specificity of 97.8%. BP: blood pressure.
Reproduced from Arend et al. (1990) with permission.

Table 9 ACR 1990 criteria for the classification of Takayasu's arteritis (traditional format)

The CHCC considered Takayasu's arteritis to be a granulomatous inflammation of the aorta and its branches 'usually' occurring in patients younger than 50 years (Table 2) (Jeanette *et al*. 1994). They recognized that age is a useful discriminator between Tayakasu's arteritis and giant cell arteritis.

Takayasu's arteritis is rare in the United States with an annual incidence of up to 2.6 per million (Hall *et al*. 1985), but much more common in Japan and the Far East. In the United Kingdom we have not seen a single case from the Norwich Health District in 7 years. However, 7 cases were seen in the West Midlands from a population of 6 million during an 8-year period in the 1990s, suggesting a much lower annual incidence (0.14 per million) (D.G.I. Scott, unpublished data).

Kawasaki disease

Kawasaki disease (mucocutaneous lymph node syndrome) is an acute vasculitis of unknown aetiology that primarily affects infants and young children, and was first described in Japan in 1967 (Kawasaki 1967). Coronary vasculitis is a major cause of morbidity and mortality.

There is no laboratory test for the diagnosis of Kawasaki disease and the diagnosis is therefore based on standard clinical criteria (Table 10) (Rauch and Hurwitz 1985). There are several problems with these criteria. Some patients who do not fulfil the criteria develop coronary artery disease; also the mucocutaneous features are very variable. Burns and colleagues compared 280 patients with acute Kawasaki disease with 42 patients evaluated for Kawasaki disease in whom other diagnoses were eventually made (Burns *et al*. 1991). They found considerable overlap between the two groups with 46 per cent of the control group satisfying the criteria for Kawasaki disease. The most common alternative diagnoses were measles and group A b-haemolytic streptococcal infection.

Fever of at least 5 days duration

Presence of at least four of the five following conditions

1. Bilateral non-exudative conjunctival injection
2. One of the following changes in oropharynx:
   injected or fissured lips, injected pharynx, or 'strawberry tongue'
3. One of the following extremity changes:
   erythema of the palms or soles, oedema of the hands or feet, or periungual desquamation
4. Rash, primarily polymorphous but not vesicular
5. Acute non-suppurative cervical lymphadenopathy

Illness not explained by an other known disease process

From Rauch and Hurwitz (1985).

Table 10 Diagnostic criteria for Kawasaki disease

The ACR did not develop criteria for Kawasaki disease. The CHCC considered Kawasaki disease and felt that the presence of mucocutaneous lymph node syndrome is the defining feature that separates Kawasaki disease from juvenile polyarteritis nodosa (Jeanette *et al*. 1994). The absence of involvement of vessels smaller than arteries in their view distinguishes Kawasaki disease from microscopic polyangiitis, this is however controversial as involvement of small vessels has been described (Lie 1994).

Other vasculitides

The CHCC also considered essential cryoglobulinaemic vasculitis, cutaneous leucocytoclastic vasculitis, and giant cell arteritis. Classification criteria for giant cell arteritis are considered elsewhere.

Cryoglobulins are plasma proteins which reversibly precipitate in the cold and have been classified on the basis of the type of immunoglobulin contained within the cryoprecipitate. Types II and III contain two types of immunoglobulin (mixed cryoglobulinaemia) and are associated with a small vessel vasculitis. They may occur in the absence of underlying disease (essential mixed cryoglobulinaemia).

Cutaneous leucocytoclastic vasculitis is confined to patients in whom there is no evidence of systemic involvement. These patients have a much better prognosis. Other systemic vasculitides must be excluded rigorously. Those patients with elevated levels of immune complexes are at greater risk of immune-complex mediated disease such as Henoch–Schönlein purpura and cryoglobulinaemic vasculitis, whilst those with ANCA are at risk of developing an ANCA-associated disease (e.g. Wegener's granulomatosis).

The development of the ACR criteria and the CHCC definitions have been important steps in improving the classification and definition of the systemic vasculitides. However, despite the use of these criteria a significant number of patients cannot be classified;. in our experience approximately 10 per cent. It is possible that some of these will evolve into typical forms of systemic vasculitis, but others may represent rarer forms which are currently poorly documented and understood.

Secondary vasculitis

Vasculitis occurring as a consequence of an autoimmune rheumatic disease is relatively common and is in our experience the commonest cause of systemic vasculitis, particularly secondary to rheumatoid arthritis and systemic lupus erythematosus (Table 1). It may also occur in dermatomyositis, scleroderma, overlap syndromes, Sjögren's syndrome, rheumatic fever, and relapsing polychondritis. Classification criteria are not well established except for vasculitis occurring secondary to rheumatoid arthritis.

Rheumatoid arthritis

Vasculitis complicating rheumtoid arthritis was first described in 1898, in a patient with involvement of the vasa nervorum (Bannatyne 1898). The association was clearly established in the 1940s and 1950s (Bywaters 1949; Bywaters 1957). The classical features are peripheral gangrene and mononeuritis multiplex (Hart *et al*. 1957; Pallis and Scott 1965). However, pericarditis, scleritis, nodules, and systemic disease occur frequently, and more recent series have described a wider spectrum of disease (Scott *et al*. 1981; Luqmani *et al*. 1994b). The size of vessel involved ranges from the aorta to capillaries. Small vessel vasculitis can occur in isolation as small nail-edge or nailfold lesions. These lesions are generally considered to be benign but can herald or coexist with major arterial disease (Bywaters 1957; Bywaters and Scott 1963). Systemic rheumatoid vasculitis occurs more commonly in males with longstanding rheumatoid arthritis who are strongly seropositive for rheumatoid factor (Scott *et al*. 1981; Luqmani *et al*. 1994b).

Scott and Bacon proposed criteria for the definition of systemic rheumatoid vasculitis (Table 11) (Scott and Bacon 1984). An obligatory criterion was the presence of established rheumatoid arthritis as defined by the ACR. This should prevent confusion with other forms of either primary or secondary vasculitis as such patients are unlikely to meet the ACR criteria for the diagnosis of rheumatoid arthritis. However, these criteria have not been validated formally.

The presence in a patient with rheumatoid arthritis of one or more of:

1. Mononeuritis multiplex or acute peripheral neuropathy
2. Peripheral gangrene
3. Biopsy evidence of acute necrotizing arteritis plus systemic illness (e.g. fever, weight loss)
4. Deep cutaneous ulcers or active extra-articular disease (e.g. pleurisy, pericarditis, scleritis) if associated with typical digital infarcts or biopsy evidence of vasculitis

Other causes of such lesions such as diabetes mellitus and atherosclerosis should be excluded. Patients with nailfold or digital infarcts alone are excluded.
From Scott and Bacon (1984).

Table 11 Classification criteria for systemic rheumatoid vasculitis

In our experience systemic rheumatoid vasculitis as defined by the Scott and Bacon criteria is amongst the commonest forms of systemic vasculitis with an annual incidence of 12.5 per million (Watts *et al*. 1994).

Aetiology

The 'best' classification system should involve the aetiological agent responsible for the vascular damage. This is unknown in the majority of cases of vasculitis, but even when it appears to be known there is still considerable confusion. For example, hepatitis B surface antigen has been associated with pure cutaneous vasculitis, cryoglobulinaemic vasculitis, glomerulonephritis, arthritis, and polyarteritis nodosa (Scott 1986). There is also one report of hepatitis B associated with giant cell arteritis (Bacon *et al*. 1975). A similar picture has now emerged with the vasculitis associated with human immunodeficiency virus infection, with descriptions of leucocytoclastic vasculitis, eosinophilic vasculitis, polyarteritis nodosa, granulomatous angiitis, and lymphomatoid granulomatosis (Siefert 1989).

The most frequently postulated pathogenetic mechanism for the production of vasculitis is the deposition of circulating immune complexes. This has been documented particularly in the small vessel vasculitides, though circulating immune complexes have been detected in almost all diseases associated with vasculitis. The activity of these complexes and their relation to pathogenesis is often doubtful; for example complement activation is rare in the group associated with systemic necrotizing arteritis, except in those cases associated with hepatitis B infection. Complement activation is seen much more frequently in small vessel vasculitis (cryoglobulinaemia, Henoch–Schönlein purpura) and vasculitis complicating rheumatoid arthritis and systemic lupus erythematosus.

Other proposed mechanisms include *in situ* formation of immune complexes, direct invasion of the vessel wall by antigen (including virus), antibodies to myelocyte lysosomal enzymes (see below), cytotoxic antibodies to endothelium, and cell-mediated immune reactions to cell-wall antigens. The importance of these mechanisms to classification is as yet undefined, with the possible exception of ANCA.

ANCA

Antineutrophil cytoplasmic antibodies were first described in 1985 as sensitive and specific markers for Wegener's granulomatosis (Van der Woude *et al*. 1985). Since then ANCA have been described in other types of systemic vasculitis (e.g. microscopic polyangiitis) and some non-vasculitic diseases (Gross *et al*. 1993a), and the disease associations and antigenic targets of ANCA have been defined more clearly. Two major staining patterns are seen on indirect immunofluorescence: cytoplasmic (cANCA) and perinuclear (pANCA). Proteinase 3 is the main target antigen of cANCA and is chiefly found in patients with Wegener's granulomatosis,

whilst myeloperoxidase is the predominant target antigen of pANCA and is found in microscopic polyangiitis (Falk and Jeanette 1988).

Disease associations of cANCA

cANCA are highly sensitive (81 per cent) and specific for Wegener's granulomatosis (97 per cent) and are directed against proteinase 3 (Gross *et al*. 1993b). The specificity of cANCA for biopsy-proven Wegener's granulomatosis is around 90 per cent (Gross *et al*. 1993b). In limited or initial-phase Wegener's granulomatosis (i.e. upper and lower airways disease only) 55 per cent of patients are cANCA positive, whereas in systemic disease 88 per cent of patients are cANCA positive (Gross *et al*. 1993b). cANCA can be found in types of vasculitis with close clinical and pathological relations to Wegener's granulomatosis, such as microscopic polyangiitis (15 per cent) and Churg–Strauss syndrome (25 per cent), but rarely in classical polyarteritis nodosa (2 per cent) (Hauschild *et al*. 1993). False-positive cANCA (i.e. occurring in non-vasculitic illnesses) are very rare (Gross *et al*. 1993).

Disease association of pANCA

Unlike cANCA, pANCA are not specific for a single disease but are seen in a wide spectrum of disease including: (i) systemic vasculitis, e.g. microscopic polyangiitis, Churg–Strauss syndrome, Wegener's granulomatosis; (ii) necrotizing glomerulonephritis; (iii) rheumatic disease, e.g. rheumatoid arthritis, Still's disease, Felty's syndrome; (iv) autoimmune rheumatic disease, e.g. systemic lupus erythematosus, Sjögren's syndrome; and (v) inflammatory bowel disease. pANCA is present in less than 3 per cent of patients with biopsy-proven Wegener's granulomatosis (Hauschild *et al*. 1993).

The major target autoantigen of pANCA is myeloperoxidase (therefore sometimes called myeloperoxidase-ANCA). Myeloperoxidase-ANCA are found in 70 per cent of patients with microscopic polyangiitis and up to 50 per cent of patients with Churg–Strauss syndrome. pANCA may, however, be targeted against other antigens such as lactoferrin, cathepsin G, human neutrophil elastase, and lysozyme. In these circumstances the immunofluorescence staining pattern may be slightly different from the characteristic pANCA, such a staining pattern is known as atypical or xANCA.

An association between ANCA and extra-articular manifestations of vasculitis in patients with rheumatoid arthritis or Felty's syndrome (Juby *et al*. 1992) has been reported, especially if antibodies to lactoferrin or elastase were present (Coremans *et al*. 1992). Antibodies to cathepsin G have been described in patients with ulcerative colitis (Kallenberg *et al*. 1992). xANCA are detected in sera from patients with an even broader range of diseases including infection and carcinoma (Peter 1993).

Role of ANCA in classification

A classical, cANCA staining pattern strongly suggests a diagnosis of Wegener's granulomatosis in a patient with appropriate clinical findings. However, the presence of ANCA detected by immunofluorescence alone has limited diagnostic value. Indirect immunofluorescence should be followed by antigen-specific assays for proteinase 3 and myeloperoxidase. Proteinase 3 antibodies are associated strongly with Wegener's granulomatosis, myeloperoxidase antibodies less strongly with microscopic polyangiitis. The diagnostic value of a positive ANCA by immunofluorescence which is not directed against proteinase 3 or myeloperoxidase is uncertain. A diagnosis of systemic vasculitis, including Wegener's granulomatosis and necrotizing glomerulonephritis, should not be based on the detection of specific ANCAs alone but should be considered in the correct clinical background. A tissue diagnosis is still frequently needed to confirm the diagnosis.

For broad classification purposes we have grouped Wegener's granulomatosis, microscopic polyangiitis, and Churg–Strauss syndrome together, as they share some clinical and pathological features and are the three diseases most commonly associated with ANCA (Table 3). Using CHCC definitions, other types of vasculitis are much less frequently associated, or have no association, with ANCA.

Prognosis

Twenty years ago, a precise characterization of the different types of vasculitis might have been considered of academic interest only, because of the lack of any specific, effective treatment. Immunosuppressive therapy, particularly with cyclophosphamide, has dramatically changed the outcome of many of the more severe forms of vasculitis, especially Wegener's granulomatosis, polyarteritis nodosa, and systemic rheumatoid vasculitis (Fauci *et al*. 1979; Fauci *et al*. 1983; Scott and Bacon 1984; Hoffman *et al*. 1992; Gordon *et al*. 1993). It is important to diagnose these diseases accurately, not only to instigate early appropriate treatment, but also to avoid using cytotoxic agents (with their potentially severe side-effects) in diseases where their use is unnecessary.

Despite recent therapeutic advances, the vasculitic diseases still have significant morbidity and mortality, reflecting the importance of accurate classification. Prognosis is affected particularly by the size of vessel involved and the presence (or absence) of renal involvement. Up to 40 per cent of patients with polyarteritis nodosa, 20 to 30 per cent with systemic rheumatoid vasculitis, and 10 to 20 per cent with Wegener's granulomatosis, die within a year of diagnosis. Although the terminal event is now more commonly sepsis than uncontrolled vasculitis, the importance of renal involvement in the mortality of systemic vasculitis is shown by comparing those whose serum creatinine is more than 500 mmol/l (47 per cent mortality) with those whose serum creatinine is less than 500 mmol/l (15 per cent mortality) at the time of diagnosis (Adu *et al*. 1987), stressing the danger of any delay in diagnosis. A similar picture is seen in Wegener's granulomatosis, where, in addition, there may be significant morbidity from nasal, laryngeal, or pulmonary involvement. Signifiant renal damage is rare in patients with systemic rheumatoid vasculitis. However, these patients also have a significant morbidity and mortality; the presence of arteritis, either clinically or on biopsy, is associated with a higher mortality (44 per cent) than small vessel disease (20 per cent), where cardiac disease (restrictive pericarditis, aortic regurgitation, etc.) is a common cause of death (Scott and Bacon 1987). All these figures stress the importance of accurate classification, particularly of those patients with vasculitis involving small- or medium-sized arteries.

By contrast the prognosis for small vessel vasculitis is much more favourable. Less than 10 per cent of patients with Henoch–Schönlein purpura develop significant renal involvement and of these only a few develop chronic renal disease. Some patients with essential mixed cryoglobulinaemia develop a polyarteritis-like picture years after presentation (Gorevic *et al*. 1980). Giant cell arteritis and Takayasu's arteritis are rarely fatal but may be associated with significant complications (e.g. blindness and ischaemia).

Conclusion

The systemic vasculitides are a group of important inflammatory conditions resulting in inflammation and necrosis of blood vessel walls. They are somewhat commoner than previously believed with an annual incidence approaching 40 per million. The most common forms of systemic vasculitis are systemic rheumatoid vasculitis (12.5 per million per year) and Wegener's granulomatosis (8.5 per million per year).

Classification criteria and disease definitions are now well established, which should lead to conformity between different centres. Particular problems still exist with the definition of classical polyarterits nodosa and microscopic angiitis, and it is important when these terms are employed that the classification criteria used for diagnosis are always given.

Classification systems are still evolving. Recent changes include the recognition of the importance of the dominant blood-vessel size, the distinction between primary and secondary vasculitis, and the incorporation of pathogenetic markers such as ANCA (Table 3).

Chapter References

Adu, D., Howi, A.J., Scott, D.G.I., Bacon, P.A., McGonigle, R.J.S., and Michael, J. (1987). Polyarteritis and the kidney. *Quarterly Journal of Medicine*, **62**, 221–37.

Alarcon-Segovia, D. and Brown, A.L. (1964). Classification and aetiologic aspects of necrotising angiitides. *Mayo Clinic Proceedings*, **39**, 205–22.

Andrews, M., Edmunds, M., Campbell, A., Walls, J., and Feehally, J. (1990). Systemic vasculitis in the 1980s—is there an increasing incidence of Wegener's granulomatosis and microscopic polyarteritis? *Journal of the Royal College of Physicians*, **24**, 284–8.

Arend, W.P., Michel, B.A., Block, D.A., *et al*. (1990). The American College of Rheumatology 1990 criteria for the classification of Takayasu's arteritis. *Arthritis and Rheumatism*, **33**, 1129–34.

Bacon, P.A., Doherty, S.M., and Zuckermann, A.J. (1975). Hepatitis-B antibody in poly-myalgia rheumatica. *Lancet*, **ii**, 476–8.

Bannatyne, G.A. (1898). In *Rheumatoid arthritis*, (2nd edn), p. 73. John Wright, Bristol.

Borrie, P. (1972). Cutaneous polyarteritis nodosa. *British Journal of Dermatology*, **87**, 87–95.

Burns, J.C., Mason, W.H., Gode, M.P., *et al*. (1991). Clinical and epidemiological characteristics of possible Kawasaki disease. *Journal of Paediatrics*, **118**, 680–86.

Bywaters, E.G.L. (1949). A variant of rheumatoid arthritis characterized by digital pad nodules and palmar fasciitis, closely resembling palindromic rheumatism. *Annals of the Rheumatic Diseases*, **8**, 2–30.

Bywaters, E.G.L. (1957). Peripheral vascular obstruction in rheumatoid arthritis and its relationship to other vascular lesions. *Annals of the Rheumatic Diseases*, **16**, 84–103.

Bywaters, E.G.L. and Scott, J.T. (1963). The natural history of vascular lesions in rheumatoid arthritis. *Journal of Chronic Diseases*, **16**, 905–14.

Carrington, C.B. and Liebow, A.A. (1966). Limited forms of angiitis and granulomatosis of Wegener's type. *American Journal of Medicine*, **41**, 497–527.

Carruthers, D., Watts, R.A., Scott, D.G.I., and Symmons, D.P.M. (1996). Wegener's granulomatosis—increased incidence or increased recognition? *British Journal of Rheumatology*, **35**, 142–5.

Chumbley, L.C., Harrison, R.A., and DeRemee, R.A. (1977). Allergic granulomatous angiitis (Churg–Strauss syndrome): report and analysis of 30 cases. *Mayo Clinic Proceedings*, **52**, 477–84.

Churg, J. and Strauss, L. (1951). Allergic granulomatosis, allergic angiitis and peri-arteritis nodosa. *American Journal of Pathology*, **27**, 277–301.

Cohan, R.D., Conn, D.L., and Ilstrup, D.M. (1980). Clinical features, prognosis and response to treatment in polyarteritis. *Mayo Clinic Proceedings*, **55**, 146–55.

Coremans, I.E.M., Hagen, E.C., Daha, M.R., *et al*. (1992). Antilactoferrin antibodies in patients with rheumatoid arthritis are associated with vasculitis. *Arthritis and Rheumatism*, **35**, 1466–75.

Davson, J., Ball, J., and Platt, R. (1948). The kidney in periarteritis nodosa. *Quarterly Journal of Medicine*, **17**, 175–202

DeRemee, R.A., McDonald, T.J., Harrison, E.G., and Coles, D.T. (1976). Wegener's granulomatosis: anatomic correlates—a proposed classification. *Mayo Clinic Proceedings*, **51**, 777–81.

Diaz-Perek, J.L. and Winkelmann, R.K. (1974). Cutaneous periarteritis nodosa. *Archives of Dermatology*, **110**, 407–14.

Falk, R.J. and Jeanette, J.C. (1988). Antineutrophil cytoplasmic autoantibodies with specificity for myeloperoxidase in patients with systemic vasculitis and idiopathic necrotising and crescentic glomerulonephritis. *New England Journal of Medicine*, **318**, 1651–7.

Fauci, A.S., Haynes, B.F., and Katz, P. (1978). The spectrum of vasculitis: clinical, pathologic, immunologic, and therapeutic considerations. *Annals of Internal Medicine*, **89**, 660–76.

Fauci, A.S., Katz, P., Haynes, B.F., and Wolff, S.M. (1979). Cyclophosphamide therapy of severe systemic necrotizing vasculitis. *New England Journal of Medicine*, **301**, 235–8.

Fauci, A.S., Haynes, B.F., Katz, P., and Wolff, S.M. (1983). Wegener's granulomatosis: prospective clinical and therapeutic experience with 85 patients for 21 years. *Annals of Internal Medicine*, **98**, 76–85.

Fraha, R.A. and Abu-Haider, F. (1979). Polyarteritis nodosa masquerading as temporal arteritis. *Journal of Rheumatology*, **6**, 76–9.

Fries, J.F., Hunder, G.G., Bloch, D.A., *et al*. (1990). The American College of Rheumatology 1990 criteria for the classification of vasculitis: summary. *Arthritis and Rheumatism*, **33**, 1135–6.

Frohnert, P.P. and Sheps, S.G. (1967). Long term follow up study of periarteritis nodosa. *American Journal of Medicine*, **43**, 8–14.

Gilliam, J.N. and Smiley, J.D. (1976). Cutaneous necrotising vasculitis and related disorders. *Annals of Allergy*, **37**, 328–39.

Godman, G.C. and Churg, J. (1954). Wegener's granulomatosis: pathology and review of the literature. *Archives of Pathology*, **58**, 533–53.

Goldberg, J.W., Lee, M.L., and Sajjad, S.M. (1987). Giant cell arteritis of the skin simulating erythema nodosum. *Annals of the Rheumatic Diseases*, **46**, 706–8.

Gordon, M., Luqmani, R.A., Adu, D., *et al*. (1993). Relapses in patients with a systemic vasculitis. *Quarterly Journal of Medicine*, **86**, 779–89.

Gorevic, P.D., Kasseb, H.J., Levo, Y., *et al*. (1980). Mixed cryoglobulinaemia: clinical aspects and long-term follow up of 40 patients. *American Journal of Medicine*, **69**, 287–308.

Gross, W.L., Schmitt, W.H., and Csernok, E. (1993 *a*). ANCA and associated diseases: immunodiagnostic and pathogenic aspects. *Clinical and Experimental Immunology*, **91**, 1–12.

Gross, W.L., Hauschild, S., and Mistry, N. (1993 *b*). The clinical relevance of ANCA in vasculitis. *Clinical and Experimental Immunology*, **91** (Suppl. 1), 7–11.

Hall, S., Barr, W., Lie, J.T., *et al*. (1985). Takayasu arteritis: a study of 32 North American patients. *Medicine (Baltimore)*, **64**, 89–99.

Hart, F.D., Golding, J.R., and Mackenzie, D.H. (1957). Neuropathy in rheumatoid disease. *Annals of the Rheumatic Diseases*, **16**, 471–80.

Hauschild, S., Schmitt, W.H., Csernok, E., *et al*. (1993). ANCA in systemic vasculitides, collagen vascular diseases, rheumatic disorders and inflammatory bowel diseases. In *ANCA associated vasculitides: immunodiagnostic and pathogenetic value of anti-neutrophil cytoplasmic antibodies* (ed. W.L. Gross). Plenum Press, London.

Heng, M.C.Y. (1985) Henoch–Schönlein purpura. *British Journal of Dermatology*, **112**, 235–40.

Henoch, E.H. (1874). Über ein eigenthümliche Form von Purpura. *Berliner Klinika Wochenschrift*, **11**, 641–3.

Hoffman, G.S., Kerr, G.S., Leavitt R.Y., *et al*. (1992). Wegener's granulomatosis: an analysis of 158 patients. *Annals of Internal Medicine*, **116**, 488–98.

Hunder, G.G., Arend, W.P., Bloch, D.A., *et al*. (1990). The American College of Rheumatology 1990 criteria for the classification of vasculitis: introduction. *Arthritis and Rheumatism*, **33**, 1065–7.

Ishikawa, K. (1986). Diagnostic approach and proposed criteria for the clinical diagnosis of Takayasu's arteriopathy. *American Journal of Cardiology*, **12**, 964–72.

Jeanette, J.C., Falk, R.J., Andrassy, K., *et al*. (1994). Nomenclature of systemic vasculitides. Proposal of an international consensus conference. *Arthritis and Rheumatism*, **37**, 187–92.

Juby, C., Johnson, C., and Davies, P. (1992). Antinuclear and antineutrophil cytoplasmic antibodies (ANCA) in the serum of patients with Felty's syndrome. *British Journal of Rheumatology*, **31**, 185–8.

Kallenberg, C.G.M., Mulder, A.H., and Cohen Tervaert, J.W. (1992). Antineutrophil cytoplasmic antibodies: a still growing class of autoantibodies in inflammatory disorders. *American Journal of Medicine*, **93**, 675–82.

Kawasaki, T. (1967). Acute febrile mucocutaneous syndrome with lymphoid involvement with specific desquamation of the fingers and toes in children: clinical observations in 50 cases. *Japanese Journal of Allergology*, **16**, 178–222.

Klein, R.G., Hunger, G.G., Stenson, A.W., and Sheps, S.G. (1975). Large artery involvement in giant cell (temporal) arteritis. *Annals of Internal Medicine*, **83**, 806–9.

Kurland, L.T., Chuang, T.Y., and Hunder, G. (1984). The epidemiology of systemic arteritis. In *The epidemiology of the rheumatic diseases* (ed. R.C. Lawrence and L.E. Shulman), pp. 196–205. Gower Publishing, New York.

Kussmaul, A. and Maier, R. (1886). Über eine bisher nicht beschreibene eigenthümliche Arterienerkrankung (Periarteritis nodosa), die mit Morbus Bright und rapid fortschreitender allgemeiner Muskellhmung einhergeht. *Deutsche Archive Klinical Medizin*, **1**, 484–514.

Lamb, A.R. (1914). Periarteritis nodosa—a clinical and pathological review of the disease. *Archives of Internal Medicine*, **14**, 481–516.

Lanham, J.G., Elkon, K.B., Pusey, C.D., and Hughes, G.R.V. (1984). Systemic vasculitis in asthma and eosinophilia: a clinical approach to the Churg–Strauss syndrome. *Medicine (Baltimore)*, **63**, 65–81.

Leavitt, R.Y., Fauci, A.S., Block, D.A., *et al*. (1990). The American College of Rheumatology 1990 criteria for the classification of Wegener's granulomatosis. *Arthritis and Rheumatism*, **33**, 1101–7.

Lie, J.T. (1988). Classification and immunodiagnosis of vasculitis: a new solution or promises unfulfilled? (Editorial). *Journal of Rheumatology*, **15**, 5.

Lie, J.T. (1994). Nomenclature and classification of vasculitis: plus ça change, plus c'est la mme chose. *Arthritis and Rheumatism*, **37**, 181–6.

Lightfoot, R.W., Michel, A.B., Bloch, D.A., *et al*. (1990). The American College of Rheumatology 1990 criteria for the classification of polyarteritis nodosa. *Arthritis and Rheumatism*, **33**, 1088–93.

Luqmani, R.A., Bacon, P.A., Beaman, M., *et al*. (1994*a*). Classical versus non-renal Wegener's granulomatosis. *Quarterly Journal of Medicine*, **87**, 161–7.

Luqmani, R.A., Watts, R.A., Scott, D.G.I., and Bacon, P.A. (1994 *b*). Treatment of vasculitis in rheumatoid arthritis. *Annales de Médecine Interne*, **145**, 566–76.

Masi, A.T., Hunder, G.G., Lie, J.T., *et al*. (1990). The American College of Rheumatology 1990 criteria for the classification of Churg–Strauss syndrome (allergic granulomatosis and angiitis). *Arthritis and Rheumatism*, **33**, 1094–100.

Mills, J.A., Michel, B.A., Block, D.A., *et al*. (1990). The American College of Rheumatology 1990 criteria for the classification of Henoch–Schönlein purpura. *Arthritis and Rheumatism*, **33**, 1114–21.

Morgan, G.J. and Harris, E.D. (1978). Non-giant cell temporal arteritis. *Arthritis and Rheumatism*, **21**, 362–6.

Pallis, C.A. and Scott, J.T. (1965). Peripheral neuropathy in rheumatoid arthritis. *British Medical Journal*, **i**, 1141–7.

Pelletan, P.J. (1810). Aneurismes particuliers. In *Clinique chirugical ou memoires des observations de Chirurgien Clinique*, Vol. 2, p.1. J.G. Dentu, Paris.

Peter, H.H., Metzger, D., Rump, A., and Rother, E. (1993). ANCA in diseases other than systemic vasculitis. *Clinical and Experimental Immunology*, **91** (Suppl. 1), 12–14.

Potter, B.T., Howley, E., and Thomson, D. (1981). Giant cell arteritis mimicking carcinoma of breast. *British Medical Journal*, **282**, 1665–6.

Rauch, A. and Hurwitz, E. (1985). Centers for disease control case definition of Kawasaki disease. *Paediatric Infectious Disease Journal*, **4**, 702–3.

Remigo, P. and Zaino, E. (1970). Polyarteritis nodosa of the gall bladder. *Surgery*, **67**, 427–31.

Rich, A.R. (1942). The role of hypersensitivity in periarteritis nodosa. *Bulletin of the Johns Hopkins Hospital*, **71**, 123–40.

Rich, A.R. and Gregory, J.E. (1943). The experimental demonstration that periarteritis nodosa is a manifestion of hypersensitivity. *Bulletin of the Johns Hopkins Hospital*, **72**, 65–88.

Rokitansky, K. (1852). Ueber einige der wichtigsten Krankheiten der Arterien. Deukschriften der Kais. *Akademie der Wissenschaften Besonders Abgedrucket*, **4**, 49.

Rose, G.A. and Spencer, H. (1957). Polyarteritis nodosa. *Quarterly Journal of Medicine*, **26**, 43–81.

Sack, M., Cassidy, J.T., and Bole, G.G. (1975). Prognostic factors in polyarteritis. *Journal of Rheumatology*, **2**, 411–20.

Savory, W.S. (1856). Case of a young woman in whom the main arteries of both upper extremities and of the left side of the neck were throughout completely obliterated. *Medical and Chirurgical Transactions (London)*, **39**, 205.

Schmid, F.R., Cooper, N.S., Ziff, M., and McEwan C. (1961). Arteritis in rheumatoid arthritis. *American Journal of Medicine*, **30**, 56–83.

Schönlein, H. (1837). *Allgemeine und specielle Pathologie und Therapie*, Vol. 2 (3rd edn). Herisau, Würzburg.

Scott, D.G.I. (1986). Vasculitis. In *Copeman's textbook of the rheumatic diseases* (6th edn), (ed. J.T. Scott), pp. 1292–324. Churchill Livingstone, Edinburgh.

Scott, D.G.I. (1988) Classification and treatment of systemic vasculitis (Editorial). *British Journal of Rheumatology*, **27**, 251–3.

Scott, D.G.I. and Bacon, P.A. (1984). Intravenous cyclophosphamide plus methyl prednisolone in the treatment of systemic rheumatoid vasculitis. *American Journal of Medicine*, **76**, 377–84.

Scott, D.G.I. and Bacon, P.A. (1987). Cardiac involvement in immunological diseases. *Clinics in Immunology and Allergy*, **1**, 537–75.

Scott, D.G.I. and Watts, R.A. (1994). Classification and epidemiology of systemic vasculitis. *British Journal of Rheumatology*, **33**, 897–99.

Scott, D.G.I., Bacon, P.A., and Tribe, C.R. (1981). Systemic rheumatoid vasculitis: a clinical and laboratory study of 50 cases. *Medicine (Baltimore)*, **60**, 288–97.

Scott, D.G.I., Bacon, P.A., Elliott, P.J., Tribe, C.R., and Wallington, T. (1982). Systemic vasculitis in a district general hospital 1972–1980: clinical and laboratory features, classification and prognosis of 80 cases. *Quarterly Journal of Medicine*, **203**, 292–311.

Siefert, M. (1989). The rheumatology of HIV infection. *Topical Reviews*, **11**.

Smith, J.A.E., O'Sullivan, M., Gough, J., and Williams, B.D. (1988). Small intestine perforation secondary to localized giant cell arteritis of the mesenteric arteries. *British Journal of Rheumatology*, **27**, 236–8.

Stewart, M., Savage, J.M., Bell, B., *et al*. (1988). Long term renal prognosis of Henoch– Schönlein purpura in an unselected childhood population. *European Journal of Paediatrics*, **147**, 113–15.

Takayasu, M. (1908). Cases with unusual changes of the central vessels in the retina. *Acta Societas Ophthalmologica Japonica*, **12**, 554.

van der Woude, F.J., Rasmussen, N., Lobatto, S., *et al*. (1985). Autoantibodies against neutrophils and monocytes: a tool for diagnosis and marker of disease activity in Wegener's granulomatosis. *Lancet*, **i**, 425–9.

von Glahn, W.C. and Pappenheimer, A.M. (1926). Specific lesions of peripheral blood vessels in rheumatism. *American Journal of Pathology*, **2**, 235–50.

Wainwright, J. and Davson, J. (1950). The renal appearances in microscopic form of polyarteritis nodosa. *Journal of Pathology and Bacteriology*, **62**, 189–96.

Watts, R.A., Carruthers, D.M., Symmons, D.P.M., and Scott, D.G.I. (1994). The incidence of rheumatoid vasculitis in the Norwich Health Authority. *British Journal of Rheumatology*, **33**, 832–3.

Watts, R.A., Carruthers, D.M., and Scott, D.G.I. (1995). Epidemiology of systemic vasculitis—changing incidence or definition. *Seminars in Arthritis and Rheumatism*, **25**, 28–34.

Wegener, F. (1936) Uber generalisierte, septische Gefasserkrankungen. *Verhandlungers der Deutschen Gesellschaft für Pathologie*, **29**, 202–9.

Zeek, P.M.(1952). Periarteritis nodosa: a critical review. *American Journal of Clinical Pathology*, **22**, 777–90.

Zeek, P.M., Smith, C.C., and Weeter, J.C. (1948). Studies on periarteritis nodosa III. The differentiation between the vascular lesions of periarteritis nodosa and hyper-sensitivity. *American Journal of Pathology*, **24**, 889–917.

5.11.2 Wegener's granulomatosis

Wolfgang L. Gross

Introduction

Wegener's granulomatosis has been recognized as a distinct clinicopathological entity for 60 years. Of unknown aetiology, it is a granulomatous disorder associated with systemic necrotizing vasculitis (for review, see Wegener 1990; Wegener 1990b). It involves predominantly the upper airways, the lungs, and the kidneys: 'classical' Wegener's granulomatosis represents a triad of upper airway (= **E**), lung (= **L**) and renal disease (= **K**) according to the **ELK** classification (DeRemee *et al*. 1976). However, since the disease may be confined to E or to E and L as a purely granulomatous process without clinical evidence of systemic vasculitis, these *formes frustes* of Wegener's granulomatosis have been designated 'limited disease' (without kidney involvement), 'limited forms' (E or E and L) or 'initial phase' (locoregional restricted granulomatous process without clinically apparent vasculitis) (Carrington and Liebow 1966; Gross 1989; Nölle *et al*. 1989; DeRemee 1993).

The 'classical' syndrome represents a generalized disease process that usually begins with an initial phase of variable duration characterized by symptoms limited mostly to the upper and/or lower respiratory tract (for review, see Wegener 1990; Wegener 1990b). Since more and more of these limited forms of Wegener's granulomatosis are being detected early, due partly to increased physicians' awareness and partly to more sophisticated diagnostic procedures, it has been learned that most of these 'initial phases' follow a rather subacute or chronic protracted course of unpredictable duration before transforming into the generalized (systemic) phase associated usually with high titres of antineutrophil cytoplasmic antibodies (**ANCA**). This is the well-recognized, full-blown disease characterized by systemic necrotizing vasculitis, usually with renal and pulmonary involvement. If untreated, it can turn into the fulminant form of the disease (e.g. renal–pulmonary syndrome), which carries a devastating prognosis (for review, see Gross 1989). Rheumatic complaints, and eye or peripheral nerve involvement, are frequently an ominous herald of the onset of the generalized vasculitic phase, which is usually also associated with constitutional symptoms such as malaise, weight loss, fever, and night sweats (Noritake *et al*. 1987; Gross 1989; Alcalay *et al*. 1990; Gross *et al*. 1991). In this phase, non-specific markers of inflammation, for example the erythrocyte sedimentation rate and C-reactive protein, are usually elevated. cANCA (proteinase 3 ANCA; **PR3-ANCA**), which has a 90 per cent specificity for Wegener's granulomatosis, is detectable in 95 per cent of sera from patients with generalized Wegener's granulomatosis, but in only about half of patients in the initial phase (Nölle *et al*. 1989).

Because of the potentially life-threatening course of many of the major necrotizing vasculitides, such as generalized Wegener's granulomatosis or microscopic polyangiitis, 'standard' treatment (United States National Institutes of Health standard therapy/Fauci's scheme) consisting of daily 'low-dose' (2–4 mg/day) cyclophosphamide plus glucocorticoids is generally considered to be of most benefit to the majority of patients (for review, see Hoffman 1990). This view has been challenged in recent years by data indicating (i) the variability of clinical course and aggressiveness in the same disease entity, (ii) continued morbidity despite therapy with cyclophosphamide plus glucocorticoids, (iii) a high incidence of relapse among patients thought to be 'cured', and (iv) the alarming range of toxic effects induced by the 'standard protocol' (Anderson *et al*. 1992; Frankel *et al*. 1992; Hoffman *et al*. 1992a).

Improved diagnostic procedures, including computed tomographic (**CT**) imaging of the lung and magnetic resonance imaging (**MRI**) of the head (Greenan *et al*. 1992; Muhle *et al*. 1993), immunodiagnostic methods such as ANCA assays, and the detection of cytokine/cytokine-receptor molecules such as the soluble interleukin-2 receptor, have enhanced diagnostic precision by clarifying the anatomical distribution of involvement and the clinical activity of the disease (Nölle *et al*. 1989; Salvarani *et al*. 1992; Schmitt *et al*. 1992; Stegemann *et al*. 1993). Consequently, it is now possible to design stage-adapted and disease-activity related therapeutic protocols instead of adhering to the rather rigid standard regimen developed 20 years ago.

Definition and classification

In spite of substantial efforts by many investigators, the nomenclature of the various subsets of Wegener's granulomatosis remains enigmatic (for review, see Gross 1996). A major problem is the lack of standardized terms and definitions. Thus different names are used (e.g. Wegener's granulomatosis; Wegener's vasculitis; microscopic polyarteritis of Wegener's type, etc.) for various forms of the same disease, and different interpretations of the same term (e.g. 'limited' Wegener's granulomatosis) can be found in the literature. In a first attempt to address this problem, an International Consensus Conference held in Chapel Hill in 1992 presented a proposal for both the nomenclature and definitions of vasculitides (see Table 2 of Chapter 5.11.1).

Usually, generalized ('full-blown', classical) Wegener's granulomatosis (Fig. 1) is described as a distinct clinicopathological entity characterized by lesions induced by necrotizing granulomatous inflammation in the upper (ear, nose, and throat) and the lower respiratory tract (including the lung), and by vasculitis (including kidney: glomerulonephritis) involving various additional organs (Wegener 1939). According to the original criteria of Godman and Churg (1954), Wegener's granulomatosis characteristically involves a triad of airway, lung, and renal disease.

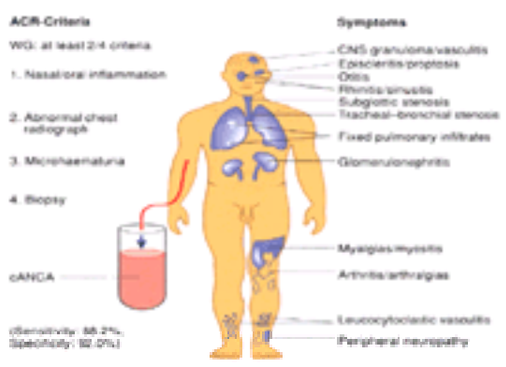


Fig. 1 Schematic representation of the characteristic clinical symptoms of Wegener's granulomatosis and the ACR 1990 classification criteria.

The 'ELK classification' proposed by DeRemee *et al*. (1976) has proved to be useful in clinical practice. Patients are classified according to the extent of the organ system involvement observed during the disease course (see Introduction; Fig. 2). The 'extended' ELK classification is now the basis for the 'Disease Extension Index' used in clinical studies (Nölle *et al*. 1989; Reinhold-Keller *et al*. 1994; Gross 1996) and, sometimes, in everyday practice.

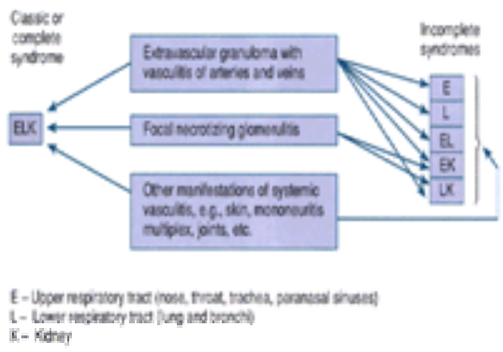


Fig. 2 Scheme for Wegener's granulomatosis in the ELK classification according to DeRemee *et al*. (1976) (with permission).

The American College of Rheumatology 1990 criteria for the classification of Wegener's granulomatosis (**ACR 1990** classification criteria) were developed by comparing 85 patients with Wegener's granulomatosis with 722 control patients with other forms of vasculitis (Leavitt *et al*. 1990). For the traditional format classification, four criteria were selected (Table 1). A classification tree was also constructed, based on the selection of five criteria. These criteria were the same as for the traditional format, plus haemoptysis. The classification tree had a sensitivity of 87.1 per cent and a specificity of 93.6 per cent.

Criteria	Definition
1. Nasal or oral inflammation	Development of painful or painless oral ulcers or purulent or bloody nasal discharge.
2. Abnormal chest radiograph	Radiograph of the chest showing the presence of nodules, fixed infiltrates, or cavities.
3. Urinary sediment	Microhaematuria (over 5 red blood cells/hpf) or red cell casts in the urine sediment.
4. Granulomatous inflammation on biopsy	Histological changes showing granulomatous inflammation within the wall of an artery or in the peri- or extravascular area (artery or arteriole)

Number of criteria present rule: For classification purposes, a patient should be said to have Wegener's granulomatosis if he/she has accumulated any two or more of these four criteria. This rule is associated with a sensitivity of 88.2 per cent and a specificity of 92.0 per cent.
From Leavitt *et al*. (1990) with permission

Table 1 ACR 1990 criteria for the classification of Wegener's granulomatosis

The Chapel Hill Consensus Conference on the Nomenclature of Systemic Vasculitis (**CHC 1992** definitions; Jennette *et al*. 1994) defined Wegener's granulomatosis as follows: granulomatous inflammation involving the respiratory tract, and necrotizing vasculitis affecting small to medium-sized vessels (e.g. capillaries, venules, arterioles, and arteries). Necrotizing glomerulonephritis is common. The disease is often associated with ANCA (see Table 2 of Chapter 5.11.1).

Neither the CHC definitions nor the ACR 1990 classification criteria addressed *formes frustes* of Wegener's granulomatosis. The criteria for definition and classification proposed for Wegener's granulomatosis (as part of the nomenclature for the vasculitides) require the presence of systemic vasculitis—but Wegener's granulomatosis may also occur as an inflammatory granulomatous process without apparent vasculitis! It must be borne in mind that neither the ACR 1990 nor the CHC 1992 constitute actual diagnostic criteria.

Epidemiology and natural history

Wegener's granulomatosis is a worldwide disease (Malaviya *et al*. 1990). New diagnostic tests (especially for ANCA) have led to an increased awareness of Wegener's granulomatosis (Andrews *et al*. 1990). Studies in the United States suggest an annual incidence of 4 per million population, while studies in England report an incidence of 8.5 per million (95 per cent confidence interval: 5.2–12.9), with an incidence for males of 10.8 and for females of 6.2 per million (Scott and Watts 1994). Wegener's granulomatosis has been observed in children and in elderly people; the peak incidence is in the fourth and fifth decades (Fauci *et al*. 1983; Hall *et al*. 1985; Weiner *et al*. 1986; McHugh *et al*. 1991; Rottem *et al*. 1993; Chakravarty *et al*. 1994).

Although, until the last decade, Wegener's granulomatosis was generally regarded as an invariably serious and fatal disease (Fauci *et al*. 1983), perceptions of it have since changed (Gross 1989; Hoffman *et al*. 1993a;). As with systemic lupus erythematosus, improved diagnostic procedures, including CT imaging of the lung and MRI of the head, as well as immunodiagnostic methods, such as ANCA assays, have enhanced diagnostic precision by identifying unclear clinical manifestations in the E or L region, and by clarifying the anatomical distribution of involvement (disease extension) and clinical disease activity (Luqmani *et al*. 1994; Reinhold-Keller *et al*. 1994). So, in addition to full-blown forms of Wegener's granulomatosis with possibly fatal outcome, indolent and/or less aggressive and life-threatening variants have been found to occur more frequently than once thought. In addition, the introduction of immunosuppressive drugs, dialysis, and renal transplantation have led to a dramatic improvement in survival rates in serious cases.

Prognosis

In their updated analysis of 158 patients with Wegener's granulomatosis, Hoffman *et al*. (1992a) emphasized the variability of the disease before diagnosis: the disease followed a confusing and indolent course in many cases (particularly in patients without renal manifestations) for up to 16 years before a definite diagnosis was established. The median and mean periods from onset of symptoms to diagnosis of Wegener's granulomatosis were 4.7 and 15 months, respectively. Diagnosis was made within 3 months after the onset of symptoms in only 42 per cent of patients; this is surprising because earlier studies had reported a median survival of only 5 months (Walton 1958). In Great Britain, 265 patients with Wegener's granulomatosis were observed between 1975 and 1985 (Anderson *et al*. 1992). The mean intervals from onset of symptoms to presentation and from presentation to diagnosis were both approximately 7 months; correct diagnosis was often missed for many years (range: less than 1–188 months). The mean survival of 4.2 years (shorter if renal disease was present) in patients receiving no drug treatment (10 per cent) indicates that these variants must have been very mild. Furthermore, it is striking that the median survival of 72 patients treated with glucocorticoids alone exceeded 12 years. Unfortunately, the investigators did not indicate the number of patients in the different treatment groups nor their disease activity and extent. Similar observations were published earlier and are reviewed elsewhere (Gross 1996).

Clinical features

Although Wegener's granulomatosis is a systemic disease with characteristic features reflected in the involvement of multiple organ systems (Table 2), it is essentially a true respiratory–renal syndrome. These two organ systems are largely responsible for the clinical course of the disease.

Organ or system	Patients			
	(number)		(%)	
	Group 1[a]	Group 2[b]	Group 1[a]	Group 2[b]
Lung	80	100	94	64
Paranasal sinuses	77	124	91	80
Kidney	72	104	85	67
Joints	57	120	67	77
Nose or nasopharynx	54	142	64	92
Ear	52	65	61	42
Eye	49	94	58	61
Skin	38	51	45	33
Nervous system	19	74	22	48
Heart	10	35	12	23

[a]From Fauci *et al*. (1983) (*n*=85). [b]From Lübeck/Bad Berleburg (*n*=155).

Table 2 Organ system involvement in Wegener's granulomatosis

Along with the well-known triad of necrotizing granulomas of the upper (E) and lower respiratory tract (L), plus glomerulonephritis (i.e. renal vasculitis; K), generalized Wegener's granulomatosis may also involve joints (A), skin (S), peripheral nerves (P), skeletal muscle (A), heart (H), brain (C) and eyes (EY), mostly via vasculitis of the small vessels (Fig. 1).

In Wegener's granulomatosis, involvement characteristically begins in the upper respiratory tract and precedes symptoms of generalized disease for a long period of time. These and pulmonary features may evolve over several months and even years, and can be followed by the overt presentation of systemic vasculitis, including glomerular disease. Because of the sometimes slow evolution, the diagnosis 'Wegener's granulomatosis' may not be made until some time after respiratory presentation.

On the other hand, Wegener's granulomatosis may sometimes start without obvious granulomatous lesions as a 'purely' small-vessel vasculitis (generalized phase). It is important to realize that there are overlapping features shared by Wegener's granulomatosis and microscopic polyangiitis, combined with the subsequent development over time of inflammatory lesions (for example, glomerulonephritis followed by skin vasculitis followed by pulmonary vasculitis—and, lastly, followed by obvious granulomatous lesions in the respiratory tract).

In any case, a strict diagnosis of Wegener's granulomatosis depends on (i) characteristic clinical symptoms, (ii) the demonstration of characteristic granulomas in biopsy material, and/or (iii) the detection of characteristic cANCA (PR3-ANCA) in the serum. It is noteworthy that each of these 'characteristics' can be observed in similar disorders, and that 'diagnosis' should not rely on any one of these variables alone.

Initial phase

['Initial stage'; E, L, EL according to the ELK classification, 'purely' granulomatous Wegener's granulomatosis; for review, see Gross (1989), DeRemee (1993), Boudes (1990).]

In most instances the ENT region (upper respiratory tract including the nose, sinuses, ears, and throat) is affected first. Characteristically, there are no clinical signs of systemic vasculitis. Patients with initial-phase Wegener's granulomatosis account for 10 per cent of those with vasculitis in a rheumatological unit.

The earliest nasal manifestations (E) usually include nasal obstruction (mucosal swelling), serosanguineous discharge, and epistaxis. Examination reveals granulation (Fig. 3), and there may be thick crusts that upon removal reveal friable mucosa or even septal perforation. In many patients the clinical diagnosis can be confirmed by the characteristic histological appearances obtained from nasal biopsy (Devaney *et al*. 1990; Del Bueno and Flint 1991). In addition to the symptoms described, facial pain, nosebleeds, nasal chondritis, saddling of the nose (Fig. 4), and involvement of paranasal sinuses (chronic sinusitis) typically occur.

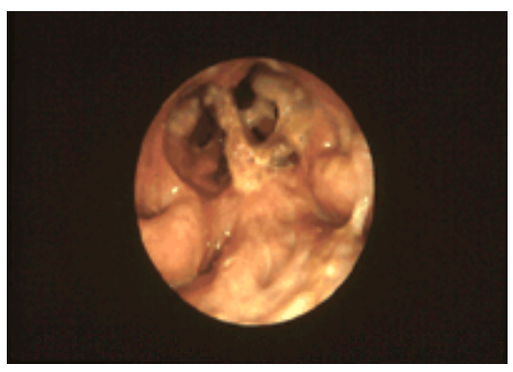


Fig. 3 Intranasal bilateral and septal mucosal disease with granulations and crusts leading to permanent serosanguinous discharge.

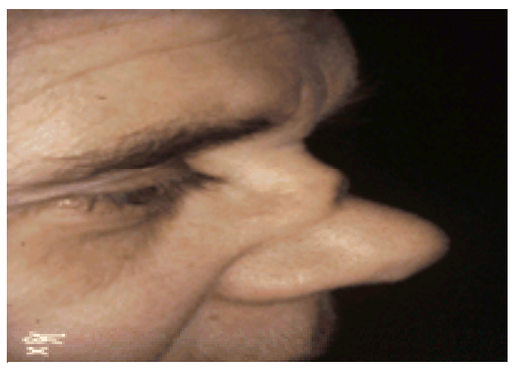


Fig. 4 Saddling of the nose due to destruction of the septal cartilage is a common finding in Wegener's granulomatosis.

In the oral cavity and oropharynx (E), gingival involvement can lead to ulcerative stomatitis, frank ulcerations (Fig. 5), or hyperplastic gingivitis (Fig. 6). Laryngeal symptoms such as hoarseness and/or increasing stridor are frequently due to reddish, friable, circumferential narrowing just below the cords and extending for 3 to 5 cm.

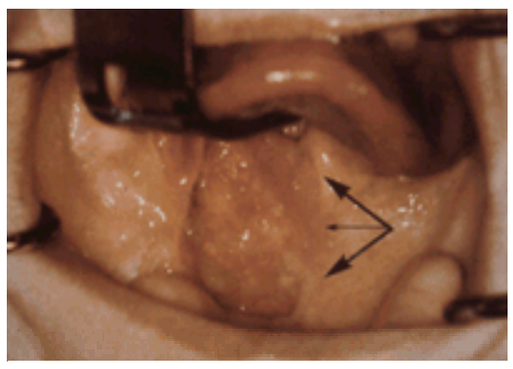


Fig. 5 Ulceration of the palate (an unusual site of manifestation for Wegener's granulomatosis) whose characteristic histological appearances led to the correct diagnosis in a patient with cavitating round nodules of the lung (see Fig. 11).

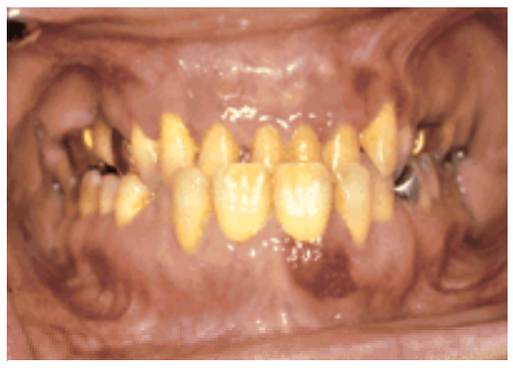


Fig. 6 Distinctive hyperplastic gingivitis originating in the interdental papilla area (biopsy-proven Wegener's granulomatosis).

Otological manifestations

The otological manifestations (E) of Wegener's granulomatosis include involvement of the external ear (chondritis, sometimes with secondary atrophy of the ear lobe), or middle ear (serous otitis media; suppurative otitis media, mastoiditis). Sometimes peripheral facial-nerve palsy occurs (D'Cruz *et al*. 1989; Murty 1990).

Eye symptoms

The eye symptoms (EY)—excluding those arising from small-vessel vasculitis seen only in generalized Wegener's granulomatosis, and which do lead to 'red eye'—derive from the granulomatous lesions (obstruction of the nasolacrimal duct) and masses developing retro-orbitally (protrusio bulbi; compression of the optic nerve; Fig. 7 and Fig. 8). Less frequently, orbital involvement may result from the spread of a purulent sinusitis, for example due to secondary bacterial infection (Charles *et al*. 1991; Satorre *et al*. 1991; Duncker *et al*. 1993).

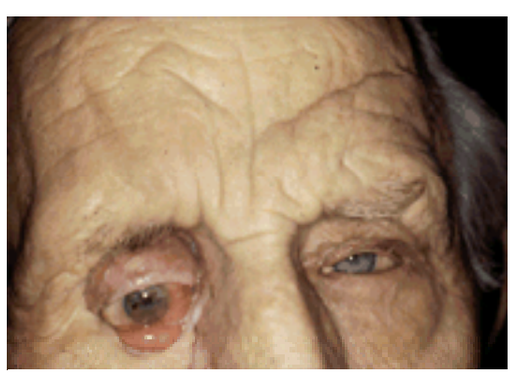


Fig. 7 Eye disease in Wegener's granulomatosis usually manifests itself in one of two ways: retro-orbital granuloma masses leading to proptosis (etc.), and small-vessel vasculitis leading to 'red eye' (see Fig. 18).

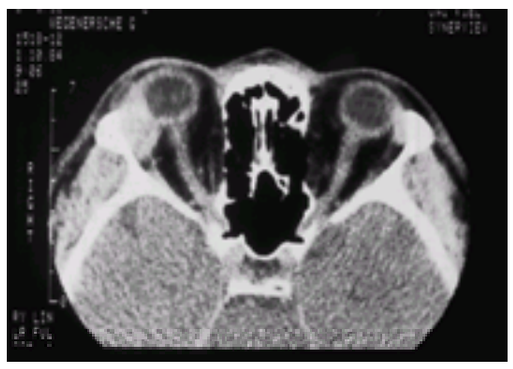


Fig. 8 Protrusio bulbi induced by retro-orbital granulomatous masses (CT scan).

Three types of involvement of the peripheral nervous system (P) have been described: disseminated vasculitis (see below: generalized Wegener's granulomatosis), contiguous granulomatous lesions, and disseminated, multicentric granulomatous lesions. The contiguous granulomatous lesion originates from the sinuses (Fig. 9) and middle ear, and spreads to the retropharyngeal area and base of the skull. This can lead to involvement of cranial nerves (I, II, III, VI, VII, VIII), and to proptosis (see above), or to diabetes insipidus, meningitis, etc. (Rosete *et al*. 1991; Greenan *et al*. 1992, Asmus *et al*. 1993; Nishino *et al*. 1993).

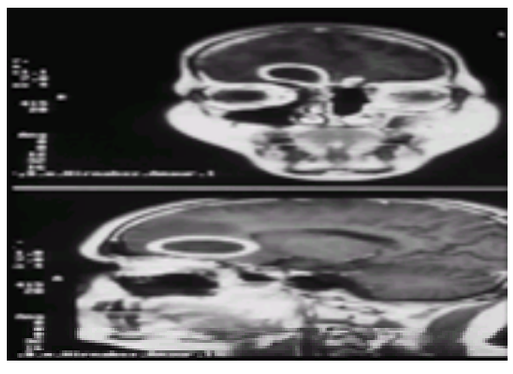


Fig. 9 Intracerebral spread of contiguous granulomatous lesion (frontal lobe) from the sinuses.

The tracheobronchial tree including the lung may also be involved in this locoregionally restricted and not yet generalized process of granulomatous inflammation: subglottic pseudotumour and/or stenosis leading to stridor and dyspnoea (Fig. 10); bronchial stenosis, which may cause atelectasis and/or obstructive pneumonia. Single or multiple nodules with or without cavitation (Fig. 11) are found incidentally in the lung of thus far asymptomatic persons (Specks and DeRemee 1990; Travis *et al*. 1991; Hoffman *et al*. 1992a).

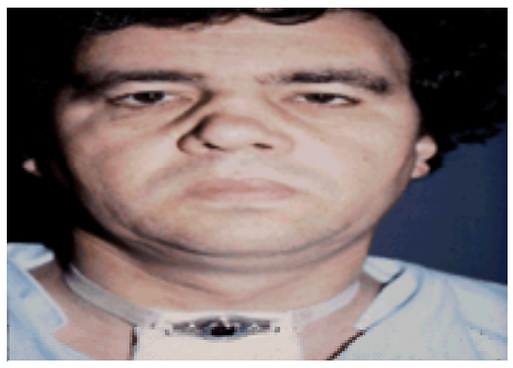


Fig. 10 Laryngeal involvement: the circumferential narrowing just below the cords led to subglottic stenoses requiring tracheostomy (see also: saddle nose deformity).

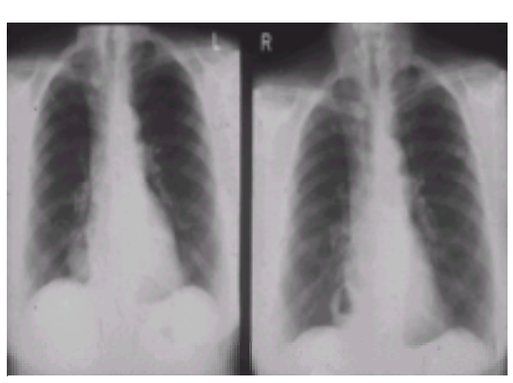


Fig. 11 Nodules in the lung usually consist of rounded lesions with well-defined margins (a); they tend to cavitate (b).

Because of the locoregionally restricted symptoms (mostly confined to the ENT region: E or lower respiratory tract including lung), the initial stage is puzzling to the diagnostician. It often takes years for Wegener's granulomatosis to be suspected and a histological diagnosis finally made. Fever can occur, perhaps caused by the underlying inflammatory disease, although it appears to be more commonly associated with secondary bacterial infection (mostly caused by *Staphylococcus aureus*) of the involved paranasal sinus (Fig. 12).

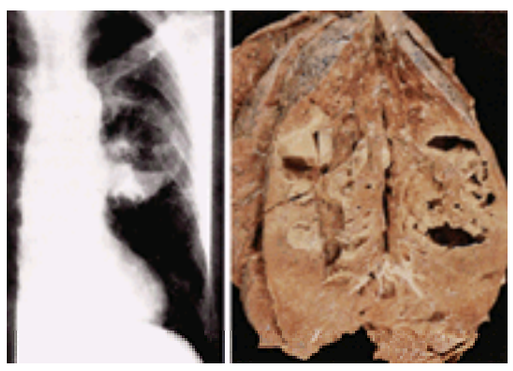


Fig. 12 Cavitary air-space consolidation initially thought to represent necrotizing bacterial pneumonia (radiographic findings and autopsy material).

Generalized Wegener's granulomatosis

P>['Systemic' Wegener's granulomatosis; E, L, K stage according to the ELK classification; for review, see Gross (1989), DeRemee (1993).]

The transition from the initial stage to the active generalized phase with its not infrequent fulminant course is heralded by the indirect symptoms of systemic vasculitis: weight loss, fever, night sweats (constitutional symptoms; similar to 'B-symptomatology' in Hodgkin's disease). Clinically, this combination of fatigue, malaise, and anorexia, together with uncharacteristic rheumatic complaints (polymyalgia, arthralgia; see below) and the striking laboratory abnormalities typical of this phase (erythrocyte sedimentation rate and C-reactive protein maximally elevated, leucocytosis, thrombocytosis), warrant close scrutiny to ensure early recognition of the imminent complications, which, if untreated, lead to the life-threatening conditions described below. Direct signs of the small-vessel vasculitis characteristically found in generalized Wegener's granulomatosis are 'red eye' (due to episcleritis, scleritis), palpable purpura of the lower extremities (due to leucocytoclastic vasculitis of the skin), peripheral neuropathy (due to the vasculitis of the vasa nervorum) and, most dangerous of all, renal involvement, which ranges from mild focal and segmental glomerulonephritis with minimal haematuria and little diminution of the glomerular filtration rate, to fulminant diffuse necrotizing and crescentic glomerulonephritis (rapidly progressive glomerulonephritis) with haematuria, pyuria, and red-cell casts, leading within several days or a few weeks to oligoanuria and dialysis. Additionally—or separately—pulmonary symptoms can arise, which may ultimately lead to pulmonary haemorrhage and severe respiratory insufficiency. In fulminant Wegener's granulomatosis, all of the symptoms described here can occur together (for example, in the form of a pulmonary–renal syndrome) or separately (alveolar

haemorrhage syndrome without glomerulonephritis; rapidly progressive glomerulonephritis without alveolar haemorrhage, etc.). In these, only rapid diagnosis and immediate immunosuppressive treatment can restrain this life-threatening condition and prevent endstage renal failure or death.

Most patients suffering from Wegener's granulomatosis, however, follow a clinical course lying between these extremes (initial-phase and fulminant generalized disease).

Otological manifestations in generalized Wegener's granulomatosis—in addition to those described for initial-phase Wegener's granulomatosis—characteristically include hearing loss due to sensorineural deafness, usually induced by the small-vessel vasculitis of the cochlear vessels and/or the vasa nervorum of the acoustic nerve. Vestibular impairment as manifested by vertigo can also occur. Similarly, peripheral facial nerve palsy can be induced by the otitis media (including mastoiditis) and the vasculitis (D'Cruz *et al*. 1989; Murty 1990).

Eye manifestations in the generalized stage of Wegener's granulomatosis are episcleritis ('red eye') (Fig. 13), vasculitis of the optic nerve, and occlusion of retinal arteries, all due to the now prominent vasculitic process, and in addition to the granulomatous lesions described above (Charles *et al*. 1991).

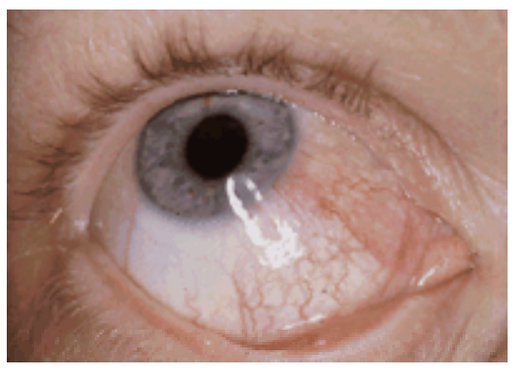


Fig. 13 Episcleritis as a characteristic clinical sign of disease activity in Wegener's granulomatosis.

The bronchial tree and lung (L) are involved usually in the systemic phase of Wegener's granulomatosis. In addition to the stenotic processes and nodules (this kind of pulmonary involvement itself is more typical of the initial phase of Wegener's granulomatosis and rather asymptomatic), localized or diffuse infiltrates should alert to the possibility of an alveolar haemorrhage resulting from the predominantly alveolar capillaritis. The three most consistent features for clinical recognition of alveolar haemorrhage are haemoptysis, infiltrates on the chest radiograph, and anaemia. The radiograph typically shows an alveolar or mixed alveolar–interstitial pattern. A distribution like that in pulmonary oedema is most common, but focal and sometimes migratory shadows are also observed (Fig. 14). However, the clinical and radiographic manifestations of immune alveolar haemorrhage are similar regardless of their aetiology. Such patients usually also have a rapidly progressive glomerulonephritis (Fig. 15) and thus typically present a pulmonary–renal vasculitic syndrome. Immunohistochemical tests on lung biopsies (e.g. via bronchoscopy) or kidney specimens reveal only a few or no immune deposits, thus excluding immune complex-mediated and antibasement membrane antibody-mediated pulmonary and pulmonary–renal vasculitic syndromes. Most patients with such 'pauci-immune' pulmonary–renal vasculitic syndromes have either cANCA (PR3-ANCA) or pANCA induced by myeloperoxidase antibodies (**myeloperoxidase-ANCA**) (Specks *et al*. 1989; Aberle *et al*. 1990; Cordier *et al*. 1990; Lombard *et al*. 1990; Dreisin 1993; Nishino *et al*. 1993).

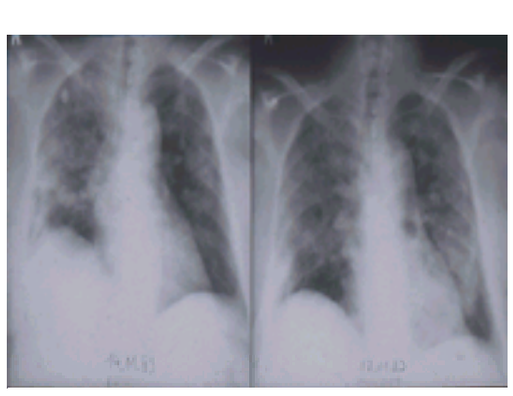


Fig. 14 Radiographic manifestation of alveolar haemorrhage in Wegener's granulomatosis.

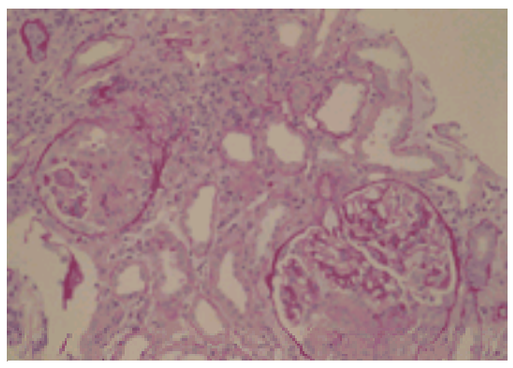


Fig. 15 Histopathology of a necrotizing and crescentic glomerulonephritis from a patient with rapidly progressive glomerulonephritis in Wegener's granulomatosis.

Abnormalities of pulmonary function include obstruction to airflow, reduced lung volumes, and abnormalities of diffusing capacity.

Kidney involvement (K) in generalized Wegener's granulomatosis shows considerable variation. At an extreme lies rapidly progressive glomerulonephritis, which is a major indicator of poor prognosis (see above). The most common manifestation of renal disease, however, is asymptomatic (micro-)haematuria with a nephritic urinary sediment and little if any renal functional impairment; this erythrocyturia is due to focal segmental glomerulonephritis. Immunohistologically, the glomeruli typically have no deposits of immunoglobulin; similarly, there is little ultrastructural evidence for immune-complex deposits. This has led to the term 'pauci-immune necrotizing and crescentic glomerulonephritis'. It is noteworthy that very similar lesions due to 'microscopic' vessel inflammation are seen in other 'ANCA-associated vasculitides' (e.g. microscopic polyangiitis) (Weiss and Crissman 1984; Wegener 1990; Falk and Jennette 1993; Jennette and Falk 1994a).

Clinical manifestations in the heart (**H** = heart involvement in the extended ELK classification) are not common in generalized Wegener's granulomatosis. However, asymptomatic pericardial effusions frequently occur and frank peri- (pan) carditis, severe granulomatous giant-cell myocarditis, and cardiomyopathy have been described (Weidhase *et al*. 1990).

Rheumatic complaints (**A** = arthralgia and/or myalgia, etc. in the extended ELK classification) ranging from mild myalgias and/or arthralgias to frank arthritis and/or myositis represent the second most frequent symptom complex (after ENT region symptoms) in generalized Wegener's granulomatosis. Musculoskeletal involvement,

particularly arthralgia and arthritis, was reported decades ago. More recently, these rheumatic manifestations were described in two-thirds of patients in a large series of cases of Wegener's granulomatosis. Twenty-eight per cent had non-erosive and non-deforming polyarthritis. Sacroiliitis was found in 3 of 50 and relapsing polychondritis in 2 of 50 patients with Wegener's granulomatosis. Rheumatoid factor was present in half of the patients with generalized Wegener's granulomatosis tested. In our own series of 186 patients with Wegener's granulomatosis (148 biopsy-proven), we found episodes of arthralgia, myalgia, or arthritis in two-thirds of cases as the presenting symptom and in three-quarters of all cases over time. Thus rheumatic complaints have been found to belong to the main symptom complex (Fig. 16). They are more frequent than symptoms of the eye, skin, or nervous system. The most common is myalgia (45 per cent), followed by frank arthritis (21 per cent), mainly in the larger joints (monoarthritis, 10 patients; oligoarthritis, 5; polyarthritis, 6). Approximately 90 per cent of patients with Wegener's granulomatosis suffering from rheumatic symptoms have a generalized form of the disease and the rheumatic symptoms usually occur together with the constitutional symptoms typically associated with active vasculitis (Noritake *et al*. 1987; Alcalay *et al*. 1990; Gross *et al*. 1991; Hoffman *et al*. 1992a) Wegener's granulomatosis can be associated with secondary relapsing polychondritis (Handrock and Gross 1993).

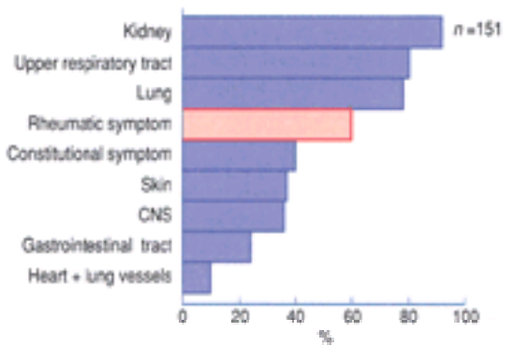


Fig. 16 Clinical symptoms in Wegener's granulomatosis.

Skin involvement (**S** = skin in the extended ELK classification) in systemic Wegener's granulomatosis usually occurs as palpable purpura due to leucocytoclastic vasculitis (Fig. 17). Less frequent are necrotic papules (Fig. 18) due to necrotizing vasculitis. Livedo reticularis and pyoderma gangrenosum are occasionally seen, sometimes as the presenting feature (Dreisin 1993; Frances *et al*. 1994).

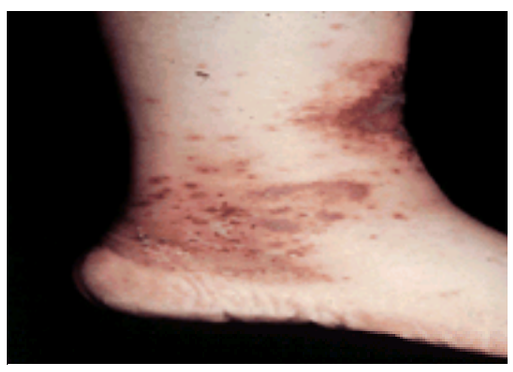


Fig. 17 Palpable purpura in a patient with fulminant Wegener's granulomatosis.

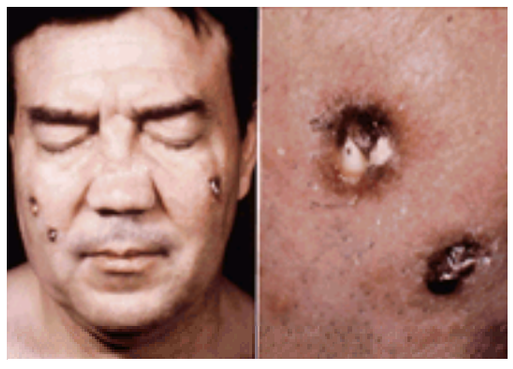


Fig. 18 Necrotic papules (and saddle-nose deformity) in Wegener's granulomatosis.

About one-third of patients with Wegener's granulomatosis have involvement of the nervous system (**C** = nervous system in the extended ELK classification). Peripheral neuropathy (excluding cranial neuropathy) occurs in about 30 per cent of patients; mononeuropathy multiplex and distal sensorimotor polyneuropathy are the leading lesions. Cerebrovascular events, seizures, and cerebritis are less frequent findings (Nishino *et al*. 1993).

In addition to this broad spectrum of symptoms, a wide variety of other manifestations, including involvement of the breast, ovaries, prostate, urethral duct, have been described, mostly in the form of case reports.

Laboratory investigations

Up to the mid-1980s, no disease-specific laboratory index for Wegener's granulomatosis existed. Disease activity was evaluated according to the clinical picture and general indices of inflammation (erythrocyte sedimentation rate, C-reactive protein) in the blood (Hoffman *et al*. 1992). In addition, the degree of normochromic normocytic anaemia, leucocytosis (usually with no or only moderate eosinophilia: less than 10 per cent), and thrombocytosis was correlated with the disease activity: all of these variables are only slightly elevated in the initial phase of Wegener's granulomatosis, and, by contrast, maximally elevated in the fulminant generalized stage, sometimes with 'leukaemoid' reactions (Gross 1996).

In contrast to the family of autoimmune rheumatic diseases (e.g. systemic lupus), generalized Wegener's granulomatosis is associated with no or only mild hypergammaglobulinaemia (minor elevations of all immunoglobulins in serum), no or only low-titre antinuclear antibody, no complement consumption (rather slight elevations), but with a rheumatic factor in up to 50 per cent of patients. In addition, cryoglobulins are characteristically absent. Indices for organ lesions, for example kidney involvement, have to be monitored by their typical tests, as for example by urine analysis (haematuria, seldom gross proteinuria), serum creatinine, etc.

Recently, cytokines, cytokine-receptor molecules, and adhesion molecules were shown to be associated with clinical disease activity when measured in serum and/or plasma (Schmitt *et al*. 1992; Stegemann *et al*. 1993; Mrowka and Sieberth 1994). Soluble interleukin-2 receptor (**sIL2-R**) is higher in patients with generalized and active disease than in those with limited and inactive disease. Surprisingly, amounts of sIL2-R are significantly elevated in patients with complete clinical remission. Further, high, amounts of sIL2-R are associated with a high probability of relapse in this group. Serum levels of soluble cell-adhesion molecules (**CAM**) (sICAM-1, sVCAM-1, sE-selectin) were found to be significantly higher in Wegener's granulomatosis than controls; in addition, sICAM-1 and sVCAM-1 decreased with clinical

remission. Surface-antigen expression of LAM-1 on granulocytes was decreased in patients with low disease activity (Riecken *et al*. 1994).

Markers of endothelial perturbation and damage (Blann *et al*. 1992; Pearson 1993; Ohdama *et al*. 1994) have been shown to be useful indicators of disease activity or progression. von Willebrand factor in plasma is secreted from endothelial cells and is widely used as one of the best markers for vasculitic disease activity. However, as with other markers (e.g. C-reactive protein), von Willebrand factor is transiently elevated during any acute-phase response to infection. Thrombomodulin, an endothelial cell-specific glycoprotein that is an important regulator of activated thrombin (converting thrombin from a procoagulant to an anticoagulant by altering its substrate specificity so that it no longer cleaves fibrinogen but activates protein C), has been found in the circulation of patients with vasculitis, including Wegener's granulomatosis. The levels of elevated soluble thrombomodulin correlate well with disease activity. In addition, autoantibodies recognizing endothelial-cell antigens are present in the serum of patients with Wegener's granulomatosis. However, they do not play a major part in the routine analysis of disease activity.

Antineutrophil cytoplasmic autoantibodies represent a class of antibodies directed not only against various constituents of neutrophil granules but even against monocytic and endothelial antigens (for review, see Jennette and Falk 1993; Gross and Csernok 1995; Kallenberg *et al*. 1994). ANCA are routinely detected by indirect immunofluorescence on ethanol-fixed neutrophils (Wiik *et al*. 1993). At least three different patterns of fluorescence can be distinguished: the (*c*lassic) cytoplasmic pattern (Fig. 19(a)) with accentuated fluorescence intensity in the area within the nuclear lobes (cANCA), the *p*erinuclear pattern (pANCA) (Fig. 19(b)), and a more diffuse cytoplasmic staining pattern (atypical ANCA). In vasculitis, cANCA is a seromarker for Wegener's granulomatosis and pANCA is a marker for microscopic polyangiitis; the cytoplasmic pattern associated with Wegener's granulomatosis is characteristically induced by PR3-ANCA, and the perinuclear pattern seen in microscopic polyangiitis is typically induced by myeloperoxidase-ANCA. However, the associations are not absolute and PR3-ANCA is sometimes seen in microscopic polyangiitis and myeloperoxidase-ANCA in Wegener's granulomatosis.

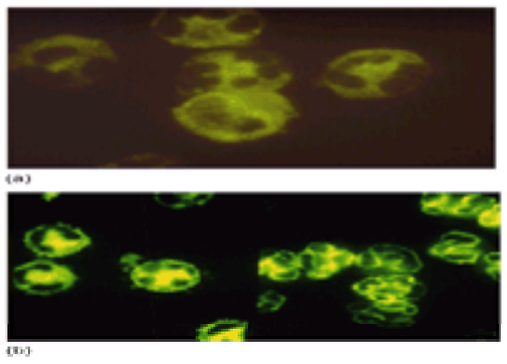


Fig. 19 Antineutrophil cytoplasmic autoantibodies: (a) cANCA and (b) pANCA fluorescence patterns.

Three major studies on cANCA comprising a total of more than 200 patients with Wegener's granulomatosis have found a 90 per cent sensitivity of the test for generalized (systemic) Wegener's granulomatosis and 75 per cent for limited Wegener's granulomatosis. It should be mentioned that these data concern patients with active disease only; during remission, cANCA were detected in a far lower percentage of the patients. However, few patients with active generalized Wegener's granulomatosis have been found to be cANCA-negative.

Various longitudinal studies have shown that titres of ANCA rise before a relapse of Wegener's granulomatosis. The rise proved to be a sensitive indicator for ensuing relapse and was detectable a mean of 49 days before the clinical manifestations. Moreover, rising titres may not only predict a relapse but can also help to differentiate relapses from the superinfections occurring not infrequently during immunosuppressive therapy. Since cANCA titres do not follow disease activity in a considerable number of patients, titres alone should not be used as a criterion for changing treatment protocols.

The specificity of cANCA for Wegener's granulomatosis has been found to be as high as 98 per cent. The application of cANCA screening to identify oligosymptomatic Wegener's granulomatosis (e.g. E, L, K, or EL, or K in the ELK classification) has led to a rise in the frequency of diagnosis of this disease. 'Initial phase' Wegener's granulomatosis can now be diagnosed using the ACR 1990 classification criteria, the definition developed by the Chapel Hill Conference in 1992, and ANCA serology. Detection of ANCA allows for the recognition (and treatment) of the underlying disease in cases presenting as Tolosa–Hunt syndrome, facial nerve paralysis, polyneuritis cranialis, peripheral neuropathy, secondary relapsing polychondritis, idiopathic necrotizing–crescentic glomerulonephritis, or renal failure of unknown origin requiring haemodialysis. It does not, however, enable recognition of acute renal failure following infectious vasculitis due to leptospirosis, as has been reported in the literature (see Gross and Csernok 1995 for review).

cANCA are found infrequently in vasculitides closely related to Wegener's granulomatosis, for example microscopic polyarteritis, Churg–Strauss syndrome, classical polyarteritis nodosa, and only exceptionally in giant-cell arteritides and Takayasu arteritis, Henoch–Schönlein purpura, cutaneous leucocytoclastic angiitis, and cryoglobulinaemic vasculitis.

A homogeneous 'cANCA' fluorescence pattern has been observed in symptomatic patients who are human immunodeficiency virus-positive and, more recently, in patients with amoebic liver abscess. In general, though, the findings of cANCA is rare in infectious disorders and one should always be aware of the possibility of a concomitant (secondary) vasculitis. Numerous recent studies, however, have shown that 'ANCA' are far from specific for Wegener's granulomatosis and have challenged the diagnostic potential of these autoantibodies. Therefore, a multinational group of experts sponsored by the European Community has attempted to work out a standardization of assays (Hagen *et al*. 1993a; Hagen *et al*. 1993b). Using their technique, one can be sure that the detection of cANCA (and/or PR3-ANCA) will help in diagnosing and evaluating disease activity in Wegener's granulomatosis.

Monitoring disease activity and disease extension

Because generalized Wegener's granulomatosis is a multisystem disease with a high rate of relapse (Hoffman *et al*. 1992a; Gordon *et al*. 1993), it is especially important that the clinical aspect of disease activity and disease extension be monitored on a regular and controlled basis by an interdisciplinary team of physicians, ear, nose, and throat surgeons, ophthalmologists, neurologists, and radiologists (Gross 1996).

At the biochemical level, the value of the erythrocyte sedimentation rate and C-reactive protein as—admittedly non-specific—indices of disease activity is accepted generally. In the initial phase of disease, both can be normal, although secondary and opportunistic infections can lead to elevated levels. cANCA titers seem to correlate with disease activity in most patients (Nölle *et al*. 1989). Consequently, changes in cANCA titre remain one of the most important laboratory markers for changes in disease activity, and the only one specific for Wegener's granulomatosis. In contrast to the erythrocyte sedimentation rate and C-reactive protein, the cANCA titre does not rise in a secondary infection. Greatly increased amounts of sIL-2R are found in patients with increased disease activity and may be an indicator of imminent relapse in those in complete clinical remission (Schmitt *et al*. 1992).

The value of MRI of the head in determining disease extension is now widely recognized (Greenan *et al*. 1992; Asmus *et al*. 1993; Muhle *et al*. 1993; Heller *et al*. 1995). MRI can demonstrate not only extensions of the granulomatous lesions within the upper respiratory tract not detectable by nasal endoscopy or conventional radiographs, but it can also disclose intracerebral granulomas (Fig. 9) and/or infarcts of grey and white matter following vasculitis.

In patients with obvious lung involvement or in those under immunosuppression who have suspected opportunistic infection, bronchoalveolar lavage can help in obtaining a histological diagnosis from transbronchial biopsy specimens (Lombard *et al*. 1990) and in demonstrating Wegener's granulomatosis-specific disease activity (marked increase in neutrophils) or *Pneumocystis carinii* (or cytomegalovirus etc.) infection (Hoffman *et al*. 1991; Sen *et al*. 1991; Barth *et al*. 1991; Jarrousse *et al*. 1993).

For patients in whom cytotoxic agents have been employed previously, continued surveillance must include the evaluation of signs and symptoms that suggest toxicity (haemorrhagic cystitis, bladder cancer, myelodysplasia, lymphoma, etc.). Because of the occurrence of 'silent' relapses, especially in the kidneys and lungs, urinalysis (sediment) and chest radiographs should be performed at close intervals if there are even minor clinical symptoms.

Pathology and pathogenesis

Wegener's granulomatosis is a clinicopathological entity closely associated with cANCA, particularly those induced by the PR3 antibodies, whose diagnosis requires the fulfilment of specific clinical and pathological criteria and/or the presence of ANCA (PR3-ANCA) in the serum.

Until fairly recently, autopsy pathology had provided most of the tissue for morphological investigations that led to the model of the classic pathological triad of granulomatous inflammation, necrosis, and vasculitis (Godman and Churg 1954). As a matter of fact, however, only tissue from open lung biopsies exhibits this triad. More commonly, only one or two elements of the triad are seen in biopsies not of the airway. Therefore, a pathological study was performed on 126 head-and-neck biopsies from 70 patients with Wegener's granulomatosis to work out criteria for the diagnosis of Wegener's granulomatosis based on such biopsies (Devaney *et al*. 1990). The following criteria were proposed:

1. If all three major pathological criteria are present, it may be considered diagnostic if there is clinical involvement of E, L, K or either L or K.
2. If two major pathological criteria are met, it may be considered diagnostic if there is typical clinical involvement of the lung and kidney as well (E, L, K); however, if only one of these sites is involved (E, L, or K), the diagnosis may be considered as probable and further biopsies should be performed.
3. If only one of the major pathological criteria is present, it may be considered as suggestive of Wegener's granulomatosis if the patient has typical clinical evidence of Wegener's granulomatosis (E, L, K) or as suspicious for Wegener's granulomatosis if the patient has only one additional site of disease (E and L or E and K).

The diagnostic value and limitations of orbital biopsy in Wegener's granulomatosis were recently delineated (Kalina *et al*. 1992).

Wegener's granulomatosis begins with granulomatous changes, as Fienberg's group was able to establish after decades of research (for review, see Wegener 1990). More recently, open lung biopsies have been studied to determine the histogenesis of pulmonary lesions and to identify early lesions (Mark *et al*. 1988). The earliest microscopic lesion is a small focus of necrosis of collagen. This is followed by an accumulation of histiocytes and their aggregation around the necrosis to form a palisading granuloma. Progression continues from micronecrosis (usually with neutrophils: 'microabscesses') to macronecrosis (wide-spread necrosis), and later to fibrosis.

Collagen in many structures of the lung was seen to have become necrotic. This included walls of blood vessels and conducting airways, alveolar walls and pleura. The primary process, therefore, is not restricted to blood vessels, although Wegener's granulomatosis is often described as a primary vasculitis.

In this study the major criteria for the histological diagnosis of Wegener's granulomatosis based on lung biopsies were outlined. The major discriminating features are palisading granulomas or palisading histiocytes in vascular walls and in extravascular tissue, microabscesses or fibrinoid necrosis in vascular walls, leucocytoclastic capillaritis, diffuse granulomatous tissue, and granulomatous bronchiolitis. Granulomatous inflammation must be found in any case.

Palisading granuloma should be distinguished from a compact, circumscribed, rounded granuloma of tuberculoid or sarcoid type. Palisading granuloma is virtually pathognomonic of Wegener's granulomatosis; by contrast, the observation of a granuloma of the tuberculoid or sarcoid type is strong evidence against Wegener's granulomatosis. Analogous to chronic granulomatous inflammation in response to a known cause (e.g. tuberculosis), palisading histiocytes and diffuse granulomatous tissue may constitute signs of a good host response in Wegener's granulomatosis. Fibrinoid necrosis of collagen within scars suggests active disease in the face of repair and preceding a host response.

From the morphological point of view, the second major feature of Wegener's granulomatosis, vasculitis, is even more polymorphic. Vasculitis generally originates in granulomas of the respiratory tract. In many cases, the anatomical picture is dominated by a necrotizing vasculitis similar to the microscopic form of polyarteritis. Vessels of many types can be affected in Wegener's granulomatosis, resulting in various clinical expressions. Less frequently, the histological picture resembles that seen in classic polyarteritis nodosa, giant-cell or Horton's arteritis, and other vasculitic disorders.

Characteristically, the acute vasculitis exhibits segmental necrotizing lesions in the walls of arteries, arterioles, capillaries, and venules. Commonly involved vessels include arteries and arterioles in skeletal and cardiac muscle as well as in liver and kidneys, postcapillary venules and arterioles in the skin, and capillaries in the pulmonary alveoli and renal glomeruli.

The third feature of the classical triad, involvement of the kidneys (for review, see Ritz *et al*. 1991; Jennette and Falk 1994b; Gross *et al*. 1995), is also characterized by a variety of clinical and anatomical pictures. As in other affected organs, the kidneys exhibit scattered granulomas in addition to variably disseminated vasculitic processes and often marked interstitial inflammatory infiltration. Glomerular involvement is frequent and is characterized histologically by segmental fibrinoid necrosis and crescent formation. No ultrastructural evidence of immune-complex localization can be found ('pauci-immune necrotizing and crescentic glomerulonephritis). Glomerular basement membranes and Bowman's capsule are often disrupted in areas of necrosis. Injury to the glomerular capillary wall is an important initiating event in the formation of glomerular crescents. Thus, the extent of renal capillaritis determines the number of crescents that ultimate lead to rapidly progressive glomerulonephritis.

Immunopathology

The aetiology of Wegener's granulomatosis is still completely unknown. In recent years, microbial infections, drugs, and tumours have been described as the 'triggering events'; in case reports viral infections (coxsackie B3, parvovirus B19) have appeared to precipitate Wegener's granulomatosis. Furthermore, an association between chronic nasal carriage of *S. aureus* and higher relapse rates in Wegener's granulomatosis has been reported (Stegemann *et al*. 1994). It has also been speculated that *S. aureus* may produce proteases with antigenic cross-reactivity with PR3 ('Wegener's autoantigen'). Sequencing of the genome of *S. aureus* will certainly be helpful in the search for possible candidates displaying molecular mimicry of PR3 and *S. aureus*.

Hydralazine, propylthiouracil, and a few other medications can induce side-effects similar to vasculitis. Both drugs have been reported to precipitate Wegener's granulomatosis (or microscopic polyarteritis). Transformation of these drugs into cytotoxic products by activated neutrophils has been described. Cytotoxicity may induce autoimmunity by exposing autoreactive lymphocytes to abnormal forms of self-material released during premature cell death (Jiang *et al*. 1994).

Recently, an association between Wegener's granulomatosis and malignancy was observed: 6.1 per cent (29/477) in a large series of patients with malignant disease. In 23 patients the tumour was diagnosed before or simultaneously with the Wegener's granulomatosis. Nearly one-third of these patients suffered from renal-cell carcinoma. Since PR3 has been found in a renal cancer cell line, autoimmunity could be induced via this pathway (for review, see Gross and Csernok 1995).

Predisposing factors are far less well accepted in Wegener's granulomatosis than in the group of collagen vascular diseases. Infrequently, Wegener's granulomatosis occurs in several members of one family (Muniain *et al*. 1986; Knudsen *et al*. 1988; Hay *et al*. 1991; Stoney *et al*. 1991; Rottem *et al*. 1993). The association between Wegener's granulomatosis and certain HLA antigens (B8, DR2) is weak. Interestingly, the persistence of ANCA in treated Wegener's granulomatosis is associated with HLA-DQw7 (Elkon *et al*. 1983; Murty *et al*. 1991). An association of severe and moderate deficiencies in protease inhibitor phenotypes with PR3-ANCA-associated vasculitis has been observed (Esnault *et al*. 1993; Elzouki *et al*. 1994). In addition, Fcg-receptor alleles may represent inheritable disease risk factors influencing the magnitude of this process, which is probably induced by neutrophil activation via ANCA (Porges *et al*. 1994).

Autoantibodies against PR3-ANCA are highly specific for Wegener's granulomatosis and since their titre often follows clinical disease activity it is believed that PR3-ANCA are of immunopathogenic relevance in this disorder. This has been extensively reviewed in recent years: for example Jennette and Falk (1992); Gross *et al*. (1993a); Gross *et al*. (1993b); Hagen *et al*. (1993); Wiik (1993); Jennette and Falk (1994a). It has been shown that the autoantibody interferes with the biological functions (e.g. by inhibition of elastinolytic activity), and with enzyme inhibitors (e.g. a $_1$-antitrypsin) of PR3. It can also activate neutrophils when PR3 is accessible on the cytoplasmic membrane by engaging the FcgRIIa receptors to produce reactive oxygen species and to degranulate. The latter event has led to the ANCA–cytokine sequence theory: As a unifying construct, it was proposed that priming doses of proinflammatory cytokines (such as those produced during infection) induce surface expression of ANCA target antigens. Binding of ANCA to these antigens leads to neutrophil degranulation and endothelial cell injury with subsequent vascular damage.

ANCA can activate primed [tumour necrosis factor-a-(**TNF-a**) pretreated] polymorphonuclear neutrophils to produce reactive oxygen species to release lysosomal enzymes (Falk *et al*. 1990) and to cause endothelial cell injury (Ewert *et al*. 1990; Savage *et al*. 1993). Proinflammatory cytokines and/or their receptors are found in serum (plasma) in active systemic vasculitis (Kekow *et al*. 1993; Roux-Lombard *et al*. 1994).

The activation pathway was once unrecognized, since PR3 (and myeloperoxidase) were thought to be located only within the azurophilic granules (Calafat *et al*. 1990). Then PR3 was detected on the plasma membrane of neutrophils using the monoclonal antibody WG M1 and electron microscopy (Csernok *et al*. 1990). Later, *in vitro* and *ex vivo* studies using flow cytometric analysis revealed that TNF-a and interleukin (**IL**) 8 act synergistically and induce a translocation of PR3 from the intragranular loci to the cell surface of neutrophils (Csernok *et al*. 1994). Apart from studies revealing upregulation of PR3 and the ability of ANCA-positive $F(ab)_2$ preparations to bind to the surface of neutrophils, to stimulate a respiratory burst, and to modulate neutrophil migration (Keogan *et al*. 1993), more recent investigations on the pathway of full neutrophil activation have shown that murine monoclonal ANCA IgG (but not IgM) binds to PR3 and stimulates the FcgRIIa ligand to activate human neutrophils via the receptor-mediated signal transduction system (Porges *et al*. 1994). Thus, ANCA-mediated neutrophil activation may occur mostly as a consequence of engagement of FcgRIIa by the Fc region of ANCA and, at least in part, via $F(ab)_2$ binding. The sum of *in vitro* and *in vivo* experimental findings supports the pathophysiological model of ANCA-mediated vasculitis depicted in Fig. 20.

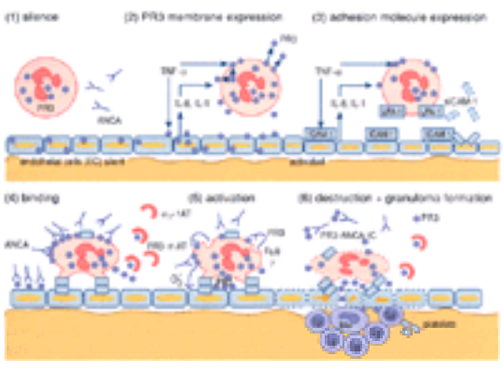


Fig. 20 Pathophysiological model of ANCA-mediated vasculitis (ANCA–cytokine sequence theory).

Mayet *et al*. (1993a) have added exciting observations to the sparse data on ANCA–endothelial cell interactions. They were able to show that TNF-a, IL-1a/b, and interferon-g lead to increased PR3 expression in the cytoplasm of endothelial cells, with translocation of PR3 to the cell membrane. Thus, PR3 located directly on the inner surface of the vessel wall becomes accessible for ANCA binding. Adhesion of polymorphonuclear neutrophils occurs via the induction of endothelial leucocyte adhesion molecule-1 (**ELAM-1**) expression on the endothelial cell surface (Mayet and Meyer zum Büschenfelde 1993). On the other hand, ANCA can bind to endothelial cells incubated with myeloperoxidase or PR3 via charge differences (cationic lysosomal proteins bind to anionic structures such as the surface of endothelial cells) and can, *in vitro*, induce complement-dependent endothelial cell lysis (Savage *et al*. 1993). Although these *in vitro* studies have demonstrated that ANCA can interact with endothelial cells, direct immunocytochemistry on lesional tissue has failed to demonstrate the binding of PR-ANCA to (Brouwer *et al*. 1994; Mrowka *et al*. 1995), or the deposition of IgG on, endothelial cells (Noronha *et al*. 1993), thus casting doubt on the immunopathological relevance of these *in vitro* findings.

PR3/myeloperoxidase deposits in tissue from ANCA-associated vasculitis have been studied by several groups (Brouwer *et al*. 1994; Mrowka *et al*. 1994; Gross *et al*. 1995). Cytokine-induced expression of adhesion molecules (e.g. LFA-1, ICAM-1, ELAM-1) facilitate close contact between neutrophils and endothelial cells, with subsequent shielding of these aggressive enzymes from their natural inhibitors (a_1 -antitrypsin, elafin, etc.; Fig. 20). Immunohistological studies of vasculitic tissue have demonstrated that both kinds of molecules are up-regulated: among the active (ANCA-associated) resident cells in glomerulonephritis, endothelial and infiltrating mononuclear cells express a variety of cytokines (IL-1, -2 and -3, TNF-a, interferon-g, platelet-derived growth factor, and transforming growth factor-b), cytokine and growth-factor receptors (**R**) (TNF-R, IL-1R type II, IL-2R, interferon-gR, and platelet-derived growth factorb-R) as well as the adhesion molecules described above (Waldherr *et al*. 1993). ICAM-1 was most abundantly present in renal lesions of Wegener's granulomatosis and the intensity of its expression correlated with the presence of glomerular crescents and the number of LFA-1 (CD11a) leucocytes. Serum levels of soluble adhesion molecules (e.g. sICAM-1) correlated with disease activity and renal functional impairment (Hauschild *et al*. 1992; Kekow *et al*. 1993; Brouwer *et al*. 1994; Mrowka and Sieberth 1994).

Activated neutrophils and the extracellular localization of lysosomal enzymes were recently observed in renal biopsies from Wegener's granulomatosis (Brouwer *et al*. 1994; Gross *et al*. 1995). Activation of neutrophils was assessed by measuring hydrogen peroxide production *in situ*; the number of activated cells in the biopsy correlated with the extent of renal functional impairment. Accordingly, PR3, myeloperoxidase, and human leucocyte elastase were localized extracellularly in renal tissue. Using a plastic embedding method, the binding of these enzymes to negatively charged structures such as endothelial cells and the glomerular basement membrane was well demonstrated, but the expression of PR3 was not found within endothelial cells. This observation was recently confirmed by Mrowka *et al*. (1995), who studied the distribution of PR3 and human leucocyte elastase in 120 renal biopsies from patients with glomerulonephritis and found these ANCA antigens expressed in ANCA-positive and -negative glomerulonephritis.

A growing body of evidence indicates that autoantibody activity in autoimmune diseases might be regulated by idiotype–anti-idiotype reactions. Observations on the clinical and *in vitro* effects of pooled human immunoglobulin and ANCA interactions (Jayne *et al*. 1993; Richter *et al*. 1993; Pall *et al*. 1994) suggest that a defect in the regulation of the idiotypic network could be involved in the production of these autoantibodies. Furthermore, network interactions can be used to develop important experimental and therapeutic agents. Our group recently generated a murine monoclonal antibody directed against a human monoclonal anti-PR3 antibody (Csernok and Gross 1993). This antibody (type Ab2b, designated 5/7) inhibits the anti-PR3 activity of cANCA in serum from patients with Wegener's granulomatosis. The anti-idiotype 5/7 is a powerful tool for studying the organization of the idiotypic network in Wegener's granulomatosis and can be used for therapeutic immunoabsorption. Furthermore, the binding of PR3-ANCA to cytokine-treated endothelial cells was blocked by the 5/7 idiotype (Mayet *et al*. 1993b). These data indicate that regulation of the idiotypic network may play an important apart in the interaction between ANCA and vascular endothelium.

Granuloma formation usually indicates a state of T-cell hyperactivity and immunohistological studies have revealed a predominance of CD4+ cells in renal biopsies of patients with Wegener's granulomatosis (Brouwer *et al*. 1991). In cellular crescents of rapidly progressive glomerulonephritis and in the peripheral blood of patients with Wegener's granulomatosis, the number of activated (CD25+) T cells is also increased. An increase in concentrations of serum cytokines (e.g. TNF-a, IL-6) during the acute phase of the disease is an indirect indicator of T-cell activation (for review, see Kekow *et al*. 1993). Elevated concentrations of sIL-2R have been detected in serum from patients with Wegener's granulomatosis, and were shown to correlate with disease activity. It has been demonstrated that even in complete remission, sIL-2R tends to be increased in Wegener's granulomatosis (Schmitt et al, 1992). However, T-cell responses to neutrophil extract did not differ between patients with vasculitis and controls: both showed only low levels of antigen-specific proliferation, and these could not be amplified by *in vitro* selection (Mathieson *et al*. 1992). In contrast to these findings, T cells from patients with Wegener's granulomatosis (PR3-ANCA positive) have been found to proliferate after exposure to highly purified PR3; however, 33 per cent of samples from healthy controls also reacted *in vitro* to PR3. In any case, the above noted extracellular localization of lysosomal proteins in renal biopsies (or elsewhere in tissue lesions in Wegener's granulomatosis) could induce the accumulation and activation of monocytes, and transform them to epithelioid and/or multinucleated giant cells characteristic of granulomas, as delineated in Fig. 20.

The IgG subclass distribution of cANCA shows a high prevalence of IgG4 antibody (Brouwer *et al*. 1991), together with increased total IgG4. This suggests repeated antigen stimulation in a T-cell response and contrasts with the IgG subclass distribution of, for example, antinuclear antibodies, which are mainly restricted to IgG1 and IgG3 (Kallenberg *et al*. 1991). Recently, a reduced CD4/CD8 T-cell ratio was reported, but this has not yet been confirmed by others (Ikeda *et al*. 1992).

Necrotizing vasculitis occurs in a number of experimental conditions involving animal models, but none of these closely resembles human systemic vasculitis (for review, see Mathieson *et al*. 1993). Brown Norway rats treated with HgCl develop a number of autoantibodies, including myeloperoxidase-ANCA, resulting in tissue injury in a number of organs, with necrotizing vasculitis especially prevalent in the gut (Mathieson *et al*. 1993). Despite similarities with human systemic vasculitis, the lack of nephritis in particular reveals the weakness of this model. The typical renal lesion of the ANCA-associated vasculitides is a pauci-immune glomerulonephritis that lacks any prominent deposition of immunoglobulin or complement, but often has extracapillary proliferation forming 'crescents' (Falk *et al*. 1990).

Brouwer *et al*. (1993) immunized Brown Norway rats with human myeloperoxidase. Five weeks after immunization the left kidney was perfused with various injurious agents, including hydrogen peroxidase plus myeloperoxidase or other products of activated neutrophils. The rats developed necrotizing crescentic glomerulonephritis and vasculitis. Since myeloperoxidase, IgG, and C3 could be detected along the glomerular basement membrane 24 h after perfusion, the 'pauci-immune' character of this vasculitis model is questionable, even when immune deposits were generally absent 4 and 10 days after perfusion. In addition, others have consistently found immune deposits along the glomerular basement membrane in both rat strains (Brown Norway and spontaneously hypertensive) used by Brouwer (Yang *et al*. 1993).

Inflammation of the vessel wall has been elicited by direct gene transfer of a foreign class I major histocompatibility complex gene, HLA-B7, to specific sites in porcine arteries. Transfer and expression of this recombinant gene was confirmed by polymerase chain reaction and immunohistochemistry, and cytolytic T cells specific for HLA-B7 were detected. This model demonstrates that vessel damage can be induced by cell-mediated immune injury (Nabel *et al*. 1992).

Treatment

Because of the possibly life-threatening course of generalized Wegener's granulomatosis there has been general agreement that the 'standard' treatment (Fauci *et al*. 1983) consisting of daily 'low-dose' (2–4 mg/day) cyclophosphamide plus glucocorticoids—undoubtedly an extremely effective therapy—is of most benefit to the majority of patients (Cupps 1990; Hoffman 1993b). This view has been challenged recently by data indicating (i) the variability of clinical course in Wegener's granulomatosis, (ii) continued morbidity despite treatment with cyclophosphamide plus glucocorticoids, (iii) a high incidence of relapse, and (iv) an alarming range of toxic effects induced by this protocol (for review, see Gross 1994).

Although the introduction of therapy with cyclophosphamide plus glucocorticoids 30 years ago improved the prognosis in Wegener's granulomatosis from a mean survival of 5 months (Walton 1958)) to a 93 per cent chance of remission (Fauci *et al*. 1983), an updated analysis (Hoffman *et al*. 1992a) now presents less favourable data on both the efficacy of the protocol (remission rate, 75 per cent) and its side-effects (serious infectious complications, 46 per cent; cyclophosphamide-induced haemorrhagic cystitis, 43 per cent; bladder cancer, 2.8 per cent; myelodysplasia, 2 per cent). These side-effects were noted earlier and are alarming because of their high rate of induced morbidity and mortality (Conn *et al*. 1988; Stillwell *et al*. 1988; Bradley *et al*. 1989; Ettinger *et al*. 1989; Anderson *et al*. 1992; Jarrousse *et al*. 1993). In addition, immediate hypersensitivity reactions to cyclophosphamide and mesna have also been observed (Knysak *et al*. 1994).

Glucocorticoids alone have failed to produce satisfactory results (Hoffman *et al*. 1992, Briedigkeit *et al*. 1993). On the other hand, high doses of glucocorticoids (in combination with daily cyclophosphamide; standard protocol in Wegener's granulomatosis) given over prolonged periods (several months) are accompanied by serious infectious complications: 50 per cent of all infections due to bacteria, pneumocystis, and fungi occurred during daily glucocorticoid therapy, 21 per cent during alternate-day therapy, 16 per cent under single-agent cytotoxic therapy, and 12 per cent during periods of no therapy (Hoffman *et al*. 1992). Such infections can contribute to vascular occlusion (Conn *et al*. 1988). Therefore glucocorticoids should be used only in conjunction with cyclophosphamide (or see below: methotrexate, azothioprine) and high initial doses of glucocorticoids should be tapered more vigorously (see Table 3) than recommended in the original report (Fauci *et al*. 1983).

The knowledge that Wegener's granulomatosis follows a two-phase course, together with the increasing identification of cases during the initial phase due to the diagnostic use of ANCA, provide a rationale for stage-adapted treatment (for review, see Hoffman 1993b; Gross and Rasmussen 1994; Gross 1996). The recognition of the toxicity and the newly discovered aspects of the disease course have led investigators to study other types of therapy for this disorder.

Trimethoprim–sulphamethoxazole

While trimethoprim–sulphamethoxazole may have the capacity to induce remission in the 'initial phase' of Wegener's granulomatosis (DeRemee *et al*. 1985; DeRemee *et al*. 1993; Reinhold-Keller *et al*. 1996), the possibility that it may actually prolong the initial phase is currently being investigated in double-blind, controlled studies. Recent experience does not indicate that trimethoprim–sulphamethoxazole can prevent relapse of generalized Wegener's granulomatosis in remission and thus it does not appear suited for treatment of classical Wegener's granulomatosis (Reinhold-Keller *et al*. 1996). Despite the lack of prospective, controlled studies, the available data on trimethoprim–sulphamethoxazole have led to its more widespread use (2 × 1 tablet, double strength) in patients in the initial stage of Wegener's granulomatosis, as described above.

Pulse cyclophosphamide

Inconsistent results were obtained in 12 patients with generalized Wegener's granulomatosis treated with pulse cyclophosphamide (Steppat and Gross 1989). Similarly, the National Institutes of Health reported only temporary benefit in 13 of 14 patients treated with the same pulse cyclophosphamide regimen; lack of response and failure to sustain improvement or tolerate continued treatment were noted in 79 per cent of patients (Hoffman *et al*. 1990). By contrast, two nephrological groups reported that the outcome after pulse cyclophosphamide therapy was similar to that following the 'standard protocol' (Falk *et al*. 1990; Haubitz *et al*. 1991). These differences may be ascribed to the fact that all of the nephrological patients had renal involvement, whereas only 50 per cent and 32 per cent of the patients investigated in the former studies showed signs of renal disease. In addition, the group studied in Chapel Hill (Falk *et al*. 1990) was mixed: only 19 of 37 patients with Wegener's granulomatosis had cANCA and a majority of the other patients had pANCA, suggesting that cases of microscopic polyangiitis had been included. Therefore, in order to determine for ourselves the clinical and/or immunological markers affecting response to pulse cyclophosphamide, we studied its efficacy in 43 patients with cANCA-positive and biopsy-proven Wegener's granulomatosis. Fifty-eight per cent of our patients did not respond to pulse cyclophosphamide treatment. Collectively, a poor response was seen in patients with generalized disease involving more than four organ systems (mainly the heart, nervous system, eyes, skin) and with constitutional symptoms and a high cANCA titre (Reinhold-Keller *et al*. 1994). We conclude, therefore, that pulse cyclophosphamide should be used in order to reduce the total amount of cyclophosphamide. Pulse cyclophosphamide can also be applied in generalized Wegener's granulomatosis with less extensive disease and a lower ANCA titre (according to the disease extension index) or in generalized Wegener's granulomatosis with formerly extensive and/or fulminant disease in partial remission (Table 3).

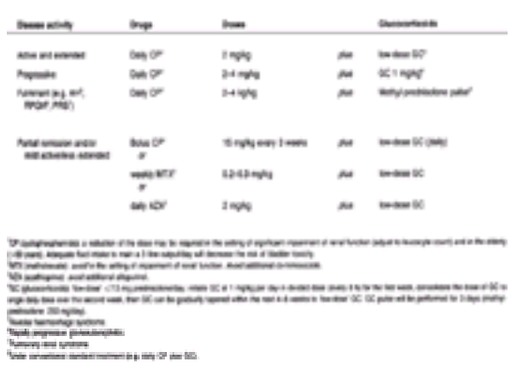


Table 3 Recommendations for drug usage in generalized Wegener's granulomatosis with respect to disease activity and extension

Methotrexate

Hoffman *et al*. (1992b) recently used 'low-dose' methotrexate (20–30 mg weekly) plus glucocorticoids in 29 patients with Wegener's granulomatosis without 'immediately life-threatening disease', although 12 (41 per cent) did have active glomerulonephritis. Treatment produced a marked improvement in 76 per cent and remission in 69 per cent. However, 3 of the 29 patients (10 per cent) suffered from *Pneumocystis carinii* pneumonia during the first month of treatment (including prednisone 60 mg/day). These data have been confirmed by Handrock *et al*. (1994). In conclusion, weekly low-dose methotrexate can be used to maintain (partial) remission in cases of Wegener's granulomatosis without renal function impairment.

Cyclosporin

Until now cyclosporin has been used only as a back-up treatment in Wegener's granulomatosis. In four patients with poor response to (oral) cyclophosphamide or with

intolerable side-effects, cyclosporin was able to control disease activity at dosages between 5 and 10 mg/kg per day (Schollmeyer *et al*. 1990). Recently, Allen *et al*. (1993) presented data on patients with active Wegener's granulomatosis undergoing combined therapy with cyclosporin and low-dose glucocorticoids. Cyclosporin was effective at initial doses of up to 5 mg/kg per day, but mild flare-ups occurred when this was lowered to 2 mg/kg per day.

Schmitt *et al*. (1993) reported on 20 patients who received renal transplants between 1982 and 1993. Treatment before transplant consisted of oral cyclophosphamide and glucocorticoids in 18 patients, and of pulse cyclophosphamide and glucocorticoids or azathioprine plus glucocorticoids in one patient each. At the time of transplant, 6 patients had symptoms of active disease. Nevertheless, 18 patients are still alive with functioning grafts (mean creatinine, 1.7 mg/per cent) and 16 are in complete remission. These results demonstrate that the rate of survival and graft function in patients with Wegener's granulomatosis is similar to that of other transplant recipients, so patients with active Wegener's granulomatosis need not be excluded from transplantation.

These data indicate that immunosuppressive therapy after transplantation must not necessarily include cyclophosphamide and that cyclosporin is effective under such circumstances.

High-dose intravenous immunoglobulin

The ability of intravenous immunoglobulin to diminish vasculitic features was described quite recently (Jayne *et al*. 1991; Tuso *et al*. 1992). Intravenous immunoglobulin was applied to 26 patients (14 with Wegener's granulomatosis, 11 with polyarteritis nodosa, 1 with rheumatoid vasculitis) requiring an intensification of therapy but mostly refractory or intolerant to glucocorticoids and cytotoxic drugs. They received a total dosage of 2 g/kg Sandoglobulin over 5 days. All patients appeared to improve after intravenous immunoglobulin, with 13/26 achieving full remission; the benefit was sustained in 18/26 at 12 months. Allowing for changes in other medications, 19/26 were in full remission after 1 year, six in partial remission, while one had died of septic complications.

In another study, intravenous immunoglobulin was given to nine patients with Wegener's granulomatosis and one with systemic pANCA-associated vasculitis, none of whom had attained complete remission under 'standard therapy' (Richter *et al*. 1995). The response was measured by blind interdisciplinary clinical assessment, cranial MRI, and immunodiagnostic analyses. All 15 patients were treated with Venimmune® (0.4 g/kg per day) for 5 days. Sixty per cent of the patients responded to therapy with improvement of single disease manifestations, but none experienced complete remission. In conclusion, intravenous immunoglobulin may play a part in Wegener's granulomatosis as an adjuvant to conventional therapy in patients with progressive disease.

Monoclonal-antibody therapy

This was introduced in 1990 for treatment of vasculitis (Mathieson *et al*. 1990). A case of severe vasculitis treated several times with an anti-CDw52 antibody (Campath 1H) responded with significant but only transient improvement. Remission lasted for more than 2 years, but this was achieved by following anti-CD52 with injection of an anti-CD4 antibody. The same group has since reported on four patients (two each with microscopic polyangiitis, Sjögren's syndrome, and Behçet's disease) who were unresponsive to immunosuppressive drugs but who received 'substantial and sustained benefit' from Campath 1H and hIgG1 anti-CD4 antibody (Lockwood *et al*. 1993). Similar results have been obtained in a few patients with systemic Wegener's granulomatosis (C. M. Lockwood, personal communication). Two other groups have reported on the successful treatment of two severe and intractable cases of relapsing polychondritis with vasculitis using a mouse and chimerical anti-CD4 monoclonal antibody (Choy *et al*. 1991; van der Lubbe *et al*. 1991).

In conclusion, preliminary studies have achieved encouraging results in patients with severe Wegener's granulomatosis refractory or intolerant to standard immunosuppressive therapy. Because such patients have a poor prognosis, these protocols appear suited as rescue regimens and should only be undertaken at experienced centres familiar with the side-effects or in trials already underway.

Azathioprine

Azathioprine has been used in the treatment of Wegener's granulomatosis and microscopic polyangiitis for many years. It has been found to be clearly less effective than cyclophosphamide for the induction of remission in generalized disease, but it seems to be valuable for maintenance of remission (Bouroncle *et al*. 1967; Fauci *et al*. 1983; Gaskin and Pusey 1992; Gordon *et al*. 1994).

Therapeutic recommendations

Recent long-term studies have revealed considerable variation in the clinical severity of Wegener's granulomatosis, while 'standard' therapy (e.g. daily cyclophosphamide plus glucocorticoids), as outlined above, is associated with considerable treatment-related morbidity, mortality, and a high incidence of relapses. On the other hand, less toxic therapeutic strategies (e.g. methotrexate) are being pursued with remarkable success. For the future, therefore, stage- and activity-dependent therapy protocols have to be developed and strategies for inducing remission must be differentiated from strategies for desperate situations (compassionate regimens). At present, many centres follow the procedures described below (see Table 3).

In initial-phase Wegener's granulomatosis, treatment can start with daily co-trimoxazole (2 × 1 double-strength tablet) if the patient can be followed easily.

In active and extended (generalized phase) Wegener's granulomatosis, most centres still begin treatment with the National Institutes of Health standard protocol (daily 2 mg/kg cyclophosphamide plus glucocorticoids). However, in contrast to the originally described concept (treatment should be continued for 1 year after the induction of remission), the therapeutic strategy is usually changed when partial remission is achieved (e.g. within the first year of treatment) to a less toxic protocol, such as pulse cyclophosphamide or weekly low-dose methotrexate (if there is no renal insufficiency) or to azathioprine (see Table 3). In addition, glucocorticoids should be tapered more vigorously (Table 3).

In the setting of (slowly) progressive Wegener's granulomatosis, by contrast, higher daily doses of cyclophosphamide (2–4 mg/kg) and glucocorticoids have to be used. The dose of cyclophosphamide will then be adjusted to maintain a total leucocyte count between 3000 and 4000/µl ('intensified' National Institutes of Health protocol) for several weeks in order to stop disease progression. In addition to an adequate fluid intake (3–4 l/day), mesna can be used to decrease the risk of cyclophosphamide-induced bladder toxicity.

In more rapidly progressive Wegener's granulomatosis (including fulminant disease), methylprednisolone pulses (250 mg/day for 3 days) are used in addition to the 'intensified' National Institutes of Health protocol. Plasmapheresis is recommended only if rapidly progressive glomerulonephritis could to lead to oligoanuria and dialysis.

In cases of generalized Wegener's granulomatosis without life- or organ-threatening disease activity, especially in the 'non-renal' variants, therapy can begin with milder protocols, for example pulse cyclophosphamide treatment every 3 weeks or weekly low-dose methotrexate, if the patient can be closely followed-up.

The adverse effects of cyclophosphamide therapy are significant (for review, see Cupps 1990). During the initial phase the leucocyte and platelet counts should be monitored at least 2 to 3 times per week. The dose of cyclophosphamide necessary to maintain a specific leucocyte count decreases as the total dose of glucocorticoids is reduced. The dose of both agents should be adapted to the age of the patient (complications arise in elderly people). The use of cyclic oestrogens may decrease the cyclophosphamide-associated toxicity in the ovary. Patients should be encouraged to maintain a fluid intake of about 3 l/day. The toxicity of glucocorticoids can be minimized by vigorous tapering (see above) and, in the event of pre-existing osteopenia, by supplementary vitamin D. A flow chart of recommended management is shown in Fig. 21.

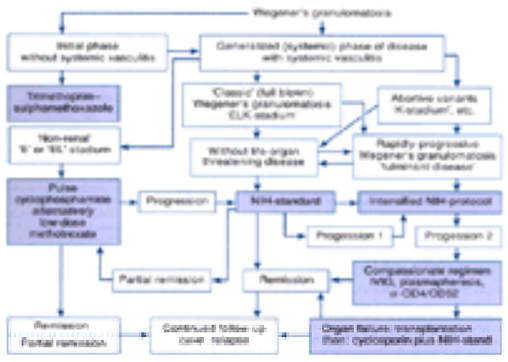


Fig. 21 Flow chart of recommended management.

It would be a grave error to assume that the available pharmacological options are enough to provide sufficient treatment. Wegener's granulomatosis obviously involves many organs and requires multidisciplinary care. Patients frequently need, for example, tympanostomies, drainage tubes for otitis media, subglottic and thoracic surgery for stenosis in the lower respiratory tract, ophthalmic surgery for obstruction of the nasolacrimal ducts, cytoscopy for cyclophosphamide-induced cystitis or bladder cancer, and renal transplantation. In addition, because of its relapsing nature, the physician treating Wegener's granulomatosis is obliged to maintain constant surveillance. This is especially true for those organs in which relapse may be asymptomatic (kidneys, lungs).

Chapter References

Aberle, D. R., Gamsu, G., and Lynch, D. (1990). Thoracic manifestations of Wegener's granulomatosis: Diagnosis and course. *Radiology*, **174**, 703–9.

Alcalay, M. *et al*. (1990). Les manifestations articulaires de la maladie de Wegener. *Revue du Rhumatism*, **57**, 845–53.

Allen, N. B., Caldwell, D. S., Rice, J. R., and McCullum, R. M. (1993). Cyclosporin A therapy for Wegener's granulomatosis. In *ANCA-associated vasculitides* (ed. W. L. Gross), pp. 473–6). Plenum, New York.

Anderson, G. *et al*. (1992). Wegener's granuloma. A series of 265 British cases seen between 1975 and 1985. A report by a sub-committee of the British Thoracic Society Research Committee. *Quarterly Journal of Medicine*, **302**, 427–38.

Andrews, M., Edmunds, M., Campbell, A., Walls, J., and Feehally, J. (1990). Systemic vasculitis in the 1980s—is there an increasing incidence of Wegener's granulomatosis and microscopic polyarteritis? *Journal of the Royal College of Physicians of London*, **24**, 284–8.

Asmus, R. *et al*. (1993). MRI of the head in Wegener's granulomatosis. In *ANCA-associated vasculitides* (ed. W. L. Gross), pp. 319–22. Plenum, New York.

Barth, J., Petermann, W., Zemke, F., Kreipe, H., and Gross, W. L. (1991). Bronchoalveolar Lavage bei Wegener'scher Granulomatose. *Pneumologie*, **45**, 570–4.

Blann, A. D., Hopkins, J., Winkles, J., and Wainwright, A. C. (1992). Plasma and serum von Willebrand factor antigen concentrations in connective tissue disorders. *Annals of Clinical Biochemistry*, **29**, 67–71.

Boudes, P. (1990). Purely granulomatous Wegener's granulomatosis: a new concept for an old disease. *Seminars in Arthritis and Rheumatism*, **19**, 365–70.

Bouroncle, B. A., Smith, E. J., and Cuppage, F. E. (1967). Treatment of Wegener's granulomatosis with Imuran. *American Journal of Medicine*, **42**, 314–18.

Bradley, J. D., Brandt, K. D., and Katz, B. P. (1989). Infectious complications of cyclophosphamide treatment for vasculitis. *Arthritis and Rheumatism*, **32**, 45–53.

Briedigkeit, L., Ulmer, M., Reinhold-Keller, E., Göbel, U., Natusch, R., and Gross, W. L. (1993). Die Therapie der Wegener'scher Granulomatose: Erfahrungen mit der konventionellen und der stadienadaptierten Behandlung bei 111 Patienten während 24 Jahren. *Innere Medizin*, **48**, 183–9.

Brouwer, E. *et al*. (1991). Immunohistology of renal biopsies in Wegener's granulomatosis (WG): clues to its pathogenesis. *Kidney International*, **39**, 1055.

Brouwer, E., Cohen Tervaert, J. W. C., van Goor, H., Huitema, M. G., Weening, J. J., and Kallenberg, C. G. M. (1993). Activated neutrophils in renal biopsies from patients with Wegener's granulomatosis. *Clinical and Experimental Immunology*, **93** (Suppl. 1), 20 (Abstr.).

Brouwer, E. *et al*. (1994). Neutrophil activation *in vitro* and *in vivo* in Wegener's granulomatosis. *Kidney International*, **45**, 1120–31.

Calafat, J., Goldschmeding, R., Ringeleing, P. L., Janssen, H., and van der Schoot, C. E. (1990). *In situ* localization by double-labeling immunoelectron microscopy of anti-neutrophil cytoplasmic autoantibodies in neutrophils and monocytes. *Blood*, **75**, 242–50.

Carrington, C. B. and Liebow, A. A. (1966). Limited forms of angiitis and granulomatosis of Wegener's type. *American Journal of Medicine*, **41**, 497–527.

Chakravarty, K., Scott, D. G. I., Blyth, J., and Courteney-Harris, R. G. (1994). Wegener's granulomatosis in the elderly—unusual presentations and misdiagnosis. *Journal of Rheumatology*, **21**, 1157–9.

Charles, S. J., Meyer, P. A. R., and Watson, P. G. (1991). Diagnosis and management of systemic Wegener's granulomatosis presenting with anterior ocular inflammatory disease. *British Journal of Ophthalmology*, **75**, 201–7.

Choy, E. H. S., Chikaza, I. C., Kingsley, G. H., and Panayi, G. S. (1991). Chimaeric anti-CD4 monoclonal antibody for relapsing polychondritis. *Lancet*, **338**, 450.

Conn, D. L., Tompkins, R. B., and Nichols, W. L. (1988). Glucocorticoids in the management of vasculitis—a double edged sword?. *Journal of Rheumatology*, **15**, 1181–2.

Cordier, J-F., Valeyre, D., Guillevin, L., Loire, R., and Brechot, J-M. (1990). Pulmonary Wegener's granulomatosis. A clinical imaging study of 77 cases. *Chest*, **97**, 906–12.

Csernok, E. and Gross, W. L. (1993). Production and characterization of an antiidiotypic monoclonal antibody that inhibits the anti-proteinase 3 activity of cANCA. *Clinical and Experimental Immunology*, **93** (Suppl. 1), 19.

Csernok, E., Lüdemann, E., Gross, W. L., and Bainton, D. F. (1990). Ultrastructural localization of proteinase 3, the target antigen of anti-cytoplasmic antibodies circulating in Wegener's granulomatosis. *American Journal of Pathology*, **137**, 1113–20.

Csernok, E., Ernst, M., Schmitt, W. H., Bainton, D. F., and Gross, W. L. (1994). Activated neutrophils express proteinase 3 on their plasma membrane *in vitro* and *in vivo*. *Clinical and Experimental Immunology*, **95**, 244–50.

Cupps, T. R. (1990). Cyclophosphamide: to pulse or not to pulse. *American Journal of Medicine*, **89**, 399–402.

D'Cruz, D. P., Baguley, E., Asherson, R. A., and Hughes, G. R. V. (1989). Ear, nose, and throat symptoms in subacute Wegener's granulomatosis. *British Medical Journal*, **299**, 419–22.

Del Buono, E. A. and Flint, A. (1991). Diagnostic usefulness of nasal biopsy in Wegener's granulomatosis. *Human Pathology*, **22**, 107–10.

DeRemee, R. A. (1993). The nosology of Wegener's granulomatosis utilizing the ELK format augmented by c-ANCA. In *ANCA-associated vasculitides* (ed. W. L. Gross), pp. 209–16. Plenum, New York.

DeRemee, R. A., McDonald, T. J., Harrison, E. G. J. R., and Coles, D. T. (1976). Wegener's granulomatosis. Anatomic correlates, a proposed classification. *Mayo Clinic Proceedings*, **51**, 777–81.

DeRemee, R. A., McDonald, T. J., and Weiland, L. H. (1985). Wegener's granulomatosis: observations on treatment with antimicrobial agents. *Mayo Clinic Proceedings*, **60**, 27–32.

Devaney, K. O., Travis, W. D., Hoffman, G., Leavitt, R., Lebovics, R., and Fauci, A. S. (1990). Interpretation of head and neck biopsies in Wegener's granulomatosis. *American Journal of Surgical Pathology*, **14**, 555–64.

Duncker, G., Nölle, B., Asmus, R., Koltze, H., and Rochels, R. (1993). Orbital involvement in Wegener's granulomatosis. In *ANCA-associated vasculitides* (ed. W. L. Gross), pp. 315–18. Plenum, New York.

Elkon, K. B., Sutherland, D. C., Rees, A. J., Hughes, G. R. V., and Batchelor, J. R. (1983). HLA antigen frequencies in systemic vasculitis: increase in HLA-DR2 in Wegener's granulomatosis. *Arthritis and Rheumatism*, **26**, 102–5.

Elzouki, A-N.Y., Segelmark, M., Wieslander, J., and Eriksson, S. (1994). Strong link between the alpha1-antitrypsin PiZ allele and Wegener's granulomatosis. *Journal of Internal Medicine*, **236**, 543–8.

Esnault, V. L. M. *et al*. (1993). Alpha1-antitrypsin genetic polymorphism in ANCA-positive systemic vasculitis. *Kidney International*, **43**, 1329–32.

Ettinger, J., Feiden, W., Hübner, G., and Schreiner, M. (1989). Progressive multifokale Leukoenzephalopathie bei Wegener'scher Granulomatose unter Therapie mit Cyclosporin A. *Klinische Wochenschrift*, **67**, 260–4.

Ewert, B. H. *et al*. (1990). Antimyeloperoxidase antibodies stimulate neutrophils to adhere to human umbilical vein endothelial cells. Abstract book of 3rd International Workshop of ANCA, Washington DC, p. 28.

Falk, R. J., Hogan, S. L., Carey, T. S., and Jennette, C. J. (1990). Clinical course of antineutrophil cytoplasmic autoantibody-associated glomerulonephritis and systemic vasculitis. *Annals of Internal Medicine*, **113**, 656–63.

Fauci, A. S., Haynes, B. F., Katz, P., and Wolff, S. M. (1983). Wegener's granulomatosis: prospective clinical and therapeutic experience with 85 patients for 21 years. *Annals of Internal Medicine*, **98**, 76–85.

Frances, C. *et al*. (1994). Wegener's granulomatosis: dermatological manifestations in 75 cases with clinicopathologic correlation. *Archives of Dermatology*, **130**, 861.

Frankel, A. H., Singer, D. R. J., Winearls, C. G., Evans, D. J., Rees, A. J., and Pusey, C. D. (1992). Type II essential mixed cryoglobulinaemia: Presentation, treatment and outcome in 13 patients. *Quarterly Journal of Medicine*, **82**, 101–24.

Gaskin, G. and Pusey, C. D. (1992). Systemic vasculitis. In *Oxford textbook of clinical nephrology* (1st edn) (ed. J. S. Cameron, A. M. Davison, J. Grunfeld, D. N. S. Kerr, and E. Ritz), pp. 612–36. Oxford University Press.

Godman, G. C. and Churg, J. (1954). Wegener's granulomatosis. *Archives of Pathology*, **58**, 533–53.

Gordon, M. *et al*. (1993). Relapses in patients with a systemic vasculitis. *Quarterly Journal of Medicine*, **86**, 779–89.

Gordon, M. *et al*. (1994). Necrotizing vasculitis—relapse despite cytotoxic therapy. *Advances in Experimental Medicine and Biology*, **336**, 477–82.

Greenan, T. J., Grossman, R. I., and Goldberg, H. I. (1992). Cerebral vasculitis: MR imaging and angiographic correlation. *Radiology*, **182**, 65–72.

Gross, W. L. (1989). Wegener's granulomatosis. New aspects of the disease course, immunodiagnostic procedures and stage-adapted treatment. *Sarcoidosis*, **6**, 15–29.

Gross, W. L. (1994). New developments in the treatment of systemic vasculitis. *Current Opinion in Rheumatology*, **6**, 11–19.

Gross, W. L. (1996). New concepts in Wegener's granulomatosis. In *The vasculitides* (ed. B. M. Ansell, P. A. Bacon, J. T. Lie, and Y. Shoenfeld), pp. 145–70. Chapman and Hall, London.

Gross, W. L. and Csernok, E. (1995). Antineutrophil cytoplasmic autoantibodies (ANCA). Immunodiagnostic and pathophysiological aspects. *Current Opinion in Rheumatology*.

Gross, W. L. and Rasmussen, N. (1994). Treatment of Wegener's granulomatosis: the view from two non-nephrologists. *Nephrology Dialysis Transplantation*, **9**, 1219–25.

Gross, W. L., Schmitt, W. H., and Csernok, E. (1991). Antineutrophil cytoplasmic autoantibody-associated diseases: a rheumatologist's view. *American Journal of Kidney Diseases*, **18**, 175–9.

Gross, W. L., Csernok, E., and Flesch, B. (1993*a*). 'Classic' anti-neutrophil cytoplasmic autoantibodies (cANCA), 'Wegener's autoantigen' and their immunopathogenic role in Wegener's granulomatosis. *Journal of Autoimmunity*, **6**, 171–84.

Gross, W. L., Schmitt, W. H., and Csernok, E. (1993*b*). ANCA and associated diseases: immunodiagnostic and pathogenetic aspects. *Clinical and Experimental Immunology*, **91**, 1–12.

Gross, W. L., Csernok, E., and Helmchen, U. (1995). Antineutrophil cytoplasmic antibodies, autoantigens and systemic vasculitis. *Acta Pathologica Microbiologica et Immunologica Scandinavica (Copenhagen)*, **103**, 81–97.

Hagen, E. C. *et al*. (1993*a*). The value of indirect immunofluorescence and solid phase techniques for ANCA detection. A report on the first phase of an international cooperative study on the standardization of ANCA assays. *Journal of Immunological Methods*, **159**, 1–16.

Hagen, C., Baillieux, B. E. P. B., van Es, L., Daha, M. R., and van der Woude, F. J. (1993 *b*). Antineutrophil cytoplasmic autoantibodies: a review of the antigens involved, the assays, and the clinical and possible pathogenetic consequences. *Blood*, **81**, 1996–2002.

Hall, S. L., Miller, L. C., Duggan, E., Mauer, S. M., Beatty, E. C., and Hellerstein, S. (1985). Wegener granulomatosis in pediatric patients. *Journal of Pediatrics*, **106**, 739–44.

Handrock, K. and Gross, W. L. (1993). Relapsing polychondritis as a secondary phenomenon of primary systemic vasculitis. *Annals of Rheumatic Diseases*, **52**, 895–7.

Handrock, K., Reinhold-Keller, E., Heller, M., Duncker, G., Rudert, H., and Gross, W. L. (1994). Beneficial effects of low-dose methotrexate (MTX) in Wegener's granulomatosis (WG). *Arthritis and Rheumatism*, **37** (Suppl.), 1152 (Abstr.).

Haubitz, M., Frei, U., Rother, U., Brunkhorst, R., and Koch, K. M. (1991). Cyclophosphamide pulse therapy in Wegener's granulomatosis. *Nephrology Dialysis Transplantation*, **6**, 531–5.

Hauschild, S., Schmitt, W. H., and Kekow, J. (1992). Hohe Serumspiegel von ICAM-1 bei der aktiven generalisierten Wegener'schen Granulomatose. *Immunität und Infektion*, **20**, 84–5.

Hay, E. M., Beaman, M., Ralston, A. J., Ackrill, P., Bernstein, R. M., and Holt, P. J. L. (1991). Wegener's granulomatosis occurring in siblings. *British Journal of Rheumatology*, **30**, 144–5.

Heller, M., Muhle, C., Reuter, M., and Schubert, F. (1995). MRI and CT in vasculitis. In *The vasculitides* (ed. B. M. Ansell, P. A. Bacon, J. T. Lie, and Y. Shoenfeld). Chapman and Hall, London.

Hoffman, G. S. (1993*a*). Wegener's granulomatosis. *Current Opinion in Rheumatology*, **5**, 11–17.

Hoffman, G. S. (1993*b*). Treatment of chronic idiopathic systemic vasculitides. In *ANCA-associated vasculitides* (ed. W. L. Gross), pp. 227–34. Plenum, New York.

Hoffman, G. S., Leavitt, R. Y., Fleisher, T. A., Minor, J. R., and Fauci, A. S. (1990). Treatment of Wegener's granulomatosis with intermittent high-dose intravenous cyclophosphamide. *American Journal of Medicine*, **89**, 403–10.

Hoffman, G. S. *et al*. (1991). Bronchoalveolar lavage analysis in Wegener's granulomatosis. *American Review of Respiratory Disease*, **143**, 401–7.

Hoffman, G. S. *et al*. (1992*a*). Wegener granulomatosis: an analysis of 158 patients. *Annals of Internal Medicine*, **116**, 488–98.

Hoffman, G. S., Leavitt, R. Y., Kerr, G. S., and Fauci, A. S. (1992 *b*). The treatment of Wegener's granulomatosis with glucocorticoids and methotrexate. *Arthritis and Rheumatism*, **35**, 1322–9.

Ikeda, M., Tsuru, S., Watanabe, Y., Kitahara, S., and Inouye, T. (1992). Reduced CD4–CD8 T cell ratios in patients with Wegener's granulomatosis. *Journal of Clinical and Laboratory Immunology*, **38**, 103–9.

Jarrousse, B. *et al*. (1993). Increased risk of Pneumocystis carinii pneumonia in patients with Wegener's granulomatosis. *Clinical and Experimental Rheumatology*, **11**, 615–21.

Jayne, D. R. W., Davies, M. J., Fox, C. J., Black, C. M., and Lockwood, C. M. (1991). Treatment of systemic vasculitis with pooled intravenous immunoglobulin. *Lancet*, **337**, 1137–9.

Jayne, D. R. W., Esnault, V. L. M., and Lockwood, C. M. (1993). ANCA anti-idiotype antibodies and the treatment of systemic vasculitis with intravenous immunoglobulin. *Journal of Autoimmunity*, **6**, 207–19.

Jennette, J. C. and Falk, R. J. (1992). Disease associations and pathogenic role of antineutrophil cytoplasmic autoantibodies in vasculitis. *Current Opinion in Rheumatology*, **4**, 9–15.

Jennette, J. C. and Falk, R. J. (1993). Pathogenic potential of anti-neutrophil cytoplasmic autoantibodies. In *ANCA-associated vasculitides* (ed. W. L. Gross), pp. 7–15. Plenum, New York.

Jennette, J. C. and Falk, R. J. (1994a). Pathogenic potential of anti-neutrophil cytoplasmic autoantibodies. *Laboratory Investigation*, **70**, 135–7.

Jennette, J. C. and Falk, R. J. (1994b). The pathology of vasculitis involving the kidney. *American Journal of Kidney Disease*, **24**, 130–41.

Jennette, J. C. *et al*. (1994). Nomenclature of systemic vasculitides. Proposal of an International Consensus Conference. *Arthritis and Rheumatism*, **37**, 187–92.

Jiang, X., Khursiga, G., and Rubin, R. L. (1994). Transformation of lupus-inducing drugs to cytotoxic products by activated neutrophils. *Science*, **266**, 810–13.

Kalina, P. H., Lie, J. T., Campbell, R. J., and Garrity, J. A. (1992). Diagnostic value and limitations of orbital biopsy in Wegener's granulomatosis. *Ophthalmology*, **99**, 120–4.

Kallenberg, C. G. M. *et al*. (1991). Autoimmunity to lysosomal enzymes: new clues to vasculitis and glomerulonephritis. *Immunology Today*, **12**, 61–4.

Kallenberg, C. G. M., Brouwer, E., Weening, J. J., and Cohen Tervaert, J. W. (1994). Antineutrophil cytoplasmic antibodies: current diagnostic and pathophysiological potential. *Kidney International*, **46**, 1–15.

Kekow, J., Szimkowiak, C. H., Sticherling, M., Schröder, J. M., Christophers, E., and Gross, W. L. (1993). Pro- and anti-inflammatory cytokines in primary systemic vasculitis. In *ANCA-associated vasculitide* (ed. W. L. Gross), pp. 341–4. Plenum, New York.

Keogan, M. T., Rifkin, I., Ronda, N., Lockwood, C. M., and Brown, D. L. (1993). Anti-neutrophil cytoplasmic antibodies (ANCA) increase neutrophil adhesion to cultured human endothelium. In *ANCA-associated vasculitides* (ed. W. L. Gross), pp. 115–19. Plenum, New York.

Knudsen, B. B., Juergensen, T., and Munch-Jensen, B. (1988). Wegener's granulomatosis in a family. *Scandinavian Journal of Rheumatology*, **17**, 225–7.

Knysak, D. J., McLean, J. A., Solomon, W. R., Fox, D. A., and McCune, W. J. (1994). Immediate hypersensitivity reaction to cyclophosphamide. *Arthritis and Rheumatism*, **37**, 1101–4.

Leavitt, R. Y. *et al*. (1990). The American College of Rheumatology 1990 criteria for the classification of Wegener's granulomatosis. *Arthritis and Rheumatism*, **33**, 1101–7.

Lockwood, C. M., Thiru, S., Isaacs, J. D., Hale, G., and Waldmann, H. (1993). Long-term remission of intractable systemic vasculitis with monoclonal antibody therapy. *Lancet*, **341**, 1620–2.

Lombard, C. M., Duncan, S. R., Rizk, N. W., and Colby, T. V. (1990). The diagnosis of Wegener's granulomatosis from transbronchial biopsy specimens. *Human Pathology*, **21**, 838–42.

Luqmani, R. A. *et al*. (1994). Birmingham Vasculitis Activity Score (BVAS) in systemic necrotizing vasculitis. *Quarterly Journal of Medicine*, **87**, 671–8.

McHugh, K., Manson, D., Eberhard, B. A., Shore, A., and Laxer, R. M. (1991). Wegener's granulomatosis in childhood. *Pediatric Radiology*, **21**, 552–5.

Malaviya, A. N. *et al*. (1990). Wegener's granulomatosis in India: not so rare.. *British Journal of Rheumatology*, **29**, 499–500.

Mark, E. J., Matsubara, O., Tan-Liu, N. S., and Fienberg, R. (1988). The pulmonary biopsy in the early diagnosis of Wegener's (pathergic) granulomatosis: a study based on 35 open lung biopsies. *Human Pathology*, **19**, 1065–71.

Mathieson, P. W. *et al*. (1990). Monoclonal-antibody therapy in systemic vasculitis. *New England Journal of Medicine*, **323**, 250–4.

Mathieson, P. W., Lockwood, C. M., and Oliveira, D. B. G. (1992). T and B cell responses to neutrophil cytoplasmic antigens in systemic vasculitis. *Clinical Immunology and Immunopathology*, **63**, 135–41.

Mathieson, P. W., Quasim, F. J., Esnault, V. L. M., and Oliveira, D. B. G. (1993). Animal models of systemic vasculitis. *Journal of Autoimmunity*, **6**, 251–64.

Mayet, W. J. and Meyer zum Büschenfelde, K. H. (1993). Antibodies to proteinase 3 increase adhesion of neutrophils to human endothelial cells. *Clinical and Experimental Immunology*, **94**, 440–6.

Mayet, W. J., Csernok, E., Szymkowiak, C., Gross, W. L., and Meyer zum Büschenfelde, K. H. (1993 *a*). Human endothelial cells express proteinase 3, the target antigen of anticytoplasmic antibodies in Wegener's granulomatosis. *Blood*, **82**, 1221–9.

Mayet, W. J., Csernok, E., Gross, W. L., and Meyer zum Büschenfelde, K. H. (1993 *b*). Inhibition of cANCA binding to human endothelial cells by an anti-idiotypic anti-PR3 antibody. *Clinical and Experimental Immunology*, **93** (Suppl. 1), 19.

Mrowka, C. and Sieberth, H. G. (1994). Circulating adhesion molecules ICAM-1, VCAM-1 and E-selectin in systemic vasculitis: marked differences between Wegener's granulomatosis and systemic lupus erythematosus. *Clinical Investigator*, **72**, 762–8.

Mrowka, C. H., Csernok, E., Gross, W. L., Feucht, H. E., Bechtel, U., and Thoenes, G. H. (1995). Distribution of the granulocyte serine proteinases proteinase 3 and elastase in human glomerulonephritis. *American Journal of Kidney Diseases*.

Muhle, C. *et al*. (1993). MRI of the head in Wegener's granulomatosis—result of a prospective study. *Clinical and Experimental Immunology*, **93** (Suppl. 1), 36 (Abstr.).

Muniain, M. A., Moreno, J. C., and Gonzalez Campora, R. (1986). Wegener's granulomatosis in two sisters. *Annals of Rheumatic Diseases*, **45**, 417–21.

Murty, G. E. (1990). Wegener's granulomatosis: otorhinolaryngological manifestations. *Clinical Otolaryngology*, **15**, 385–93.

Murty, G. E., Mains, B. T., Middleton, D., Maxwell, A. P., and Savage, D. A. (1991). HLA antigen frequencies and Wegener's granulomatosis. *Clinical Otolaryngology*, **16**, 448–51.

Nabel, E. G., Plautz, G., and Nabel, G. S. (1992). Transduction of a foreign histocompatibility gene into the arterial wall induces vasculitis. *Proceedings of National Academy of Sciences (USA)*, **89**, 5157–61.

Nishino, H., Rubino, F. A., DeRemee, A., Swanson, J. W., and Parisi, J. E. (1993). Neurological involvement in Wegener's granulomatosis: an analysis of 324 consecutive patients at the Mayo Clinic. *Annals of Neurology*, **33**, 4–9.

Nölle, B., Specks, U., Lüdemann, J., Rohrbach, M. S., DeRemee, R. A., and Gross, W. L. (1989). Anticytoplasmic autoantibodies: their immunodiagnostic value in Wegener's granulomatosis. *Annals of Internal Medicine*, **111**, 28–40.

Noritake, D. T., Weiner, S. R., Bassett, L. W., Paulus, H. E., and Weisbart, R. (1987). Rheumatic manifestations of Wegener's granulomatosis. *Journal of Rheumatology*, **14**, 949–51.

Noronha, I. L., Kruger, C., Andrassy, K., Ritz, E., and Waldherr, R. (1993). *In situ* production of TNF-alpha, IL-1 and IL-2R in ANCA-positive glomerulonephritis. *Kidney International*, **43**, 682–92.

Ohdama, S., Takano, S., Miyake, S., Kubota, T., Sato, K., and Aoki, N. (1994). Plasma thrombomodulin as a marker of vascular injuries in collagen vascular diseases. *American Journal of Clinical Pathology*, **101**, 109–13.

Pall, A. A., Varagunam, M., Adu, D., Smith, N., Richards, N. T., Taylor, C. M., and Michael, J. (1994). Anti-idiotypic activity against anti-myeloperoxidase antibodies in pooled human immunoglobulin. *Clinical and Experimental Immunology*, **93** (Suppl. 1), 12–14.

Pearson, J. D. (1993). Markers of endothelial perturbation and damage. *British Journal of Rheumatology*, **32**, 651–2.

Porges, A. J., Redecha, P. B., Kimberley, W. T., Csernok, E., Gross, W. L., and Kimberley, R. P. (1994). Anti-neutrophil cytoplasmic autoantibodies engage and activate human neutrophils via FcγRIIa.. *Journal of Immunology*, **153**, 1271–80.

Reinhold-Keller, E. *et al*. (1994). Influence of disease manifestation and antineutrophil cytoplasmic antibody titer in the response to pulse cyclophosphamide therapy in patients with Wegener's granulomatosis. *Arthritis and Rheumatism*, **37**, 919–24.

Reinhold-Keller, E., de Groot, E., Rudert, H., Nölle, B., Heller, M., and Gross, W. L. (1996). Response to trimethoprim/sulfamethoxazole in Wegener's granulomatosis depends on the phase of the disease. *Quarterly Journal of Medicine*, **89**, 15–23.

Richter, C., Schnabel, A., Csernok, E., Reinhold-Keller, E., and Gross, W. L. (1995). Treatment of Wegener's granulomatosis with intravenous immunoglobulin. In *ANCA-associated vasculitides* (ed. W. L. Gross), pp. 487–90. Plenum, New York.

Riecken, B., Gutfleisch, J., Schlesier, M., and Peter, H. H. (1994). Impaired granulocyte oxidative burst and decreased expression of leucocyte adhesion molecule-1 (LAM-1) in patients with

Wegener's granulomatosis. *Clinical and Experimental Immunology*, **96**, 43–7.

Ritz, E., Andrassy, K., Küster, S., and Waldherr, R. (1991). Wegener's granulomatosis, microscopic polyarteritis and pauciimmune crescentic necrotizing glomerulonephritis. *Contributions to Nephrology*, **94**, 1–12.

Rosete, A., Cabral, A. R., Kraus, A., and Alarcon-Segovia, D. (1991). Diabetes insipidus secondary to Wegener's granulomatosis: report and review of the literature. *Journal of Rheumatology*, **18**, 761–5.

Rottem, M., Fauci, A. S., Hallahan, C. W., Kerr, G. S., Lebovics, R., Leavitt, R. Y., and Hoffman, G. S. (1993). Wegener granulomatosis in children and adolescents: Clinical presentation and outcome. *Journal of Pediatrics*, **122**, 26–31.

Roux-Lombard, P., Lin, J. M., Peter, J. B., and Dayer, J. M. (1994). Elevated serum levels of TNF soluble receptors in patients with positive anti-neutrophil cytoplasmic antibodies. *British Journal of Rheumatology*, **33**, 428–31.

Salvarani, C. *et al*. (1992). Soluble interleukin-2 receptors in polymyalgia rheumatica/giant cell arteritis. Clinical and laboratory correlations. *Journal of Rheumatology*, **19**, 1100–6.

Satorre, J., Antle, M., O'Sullivan, R., White, V. A., Nugent, R. A., and Rootman, J. (1991). Orbital lesions with granulomatous inflammation. *Canadian Journal of Ophthalmology*, **26**, 174–95.

Savage, C., Gaskin, G., Pusey, C. D., and Pearson, J. D. (1993). Myeloperoxidase binds to vascular endothelial cells, is recognized by ANCA and can enhance complement dependent cytotoxicity. In *ANCA-associated vasculitides* (ed. W. L. Gross), pp. 121–3. Plenum, New York.

Schmitt, W. H., Heesen, C., Csernok, E., Rautmann, A., and Gross, W. L. (1992). Elevated serum levels of soluble interleukin-2 receptor in patients with Wegener's granulomatosis. *Arthritis and Rheumatism*, **35**, 1088–96.

Schmitt, W. H., Haubitz, M., Mistry, N., Brunkhorst, R., Erbslöh-Möller, B., and Gross, W. L. (1993). Renal transplantation in Wegener's granulomatosis. *Lancet*, **342**, 860.

Schollmeyer, P. and Grotz, W. (1990). Cyclosporin in the treatment of Wegener's granulomatosis (WG) and related diseases. *Acta Pathologica Microbiologica et Immunologica Scandinavica (Copenhagen)*, **19** (Suppl.), 54–5.

Scott, D. G. I. and Watts, R. A. (1994). Classification and epidemiology of systemic vasculitis. *British Journal of Rheumatology*, **33**, 897–900.

Sen, R. P., Walsh, T. P., Fisher, W., and Brock, N. (1991). Pulmonary complications of combination therapy with cyclophosphamide and prednisone. *Chest*, **99**, 143–6.

Specks, U. and DeRemee, R. A. (1990). Granulomatous vasculitis: Wegener's granulomatosis and Churg–Strauss syndrome. *Rheumatic Diseases Clinics of North America*, **16**, 377–97.

Specks, U., DeRemee, R. A., and Gross, W. L. (1989). Die Wegener'scher Granulomatose: Systemerkrankung mit bevorzugtem Befall des Respirationstrakts. *Pneumologie*, **43**, 648–59.

Stegemann, C., Tervaert, J. W. C., Huitema, M. G., and Kallenberg, C. G. M. (1993). Serum markers of T-cell activation in relapses of Wegener's granulomatosis. In *ANCA-associated vasculitides* (ed. W. L. Gross), pp. 389–92. Plenum, New York.

Stegemann, C., Cohen Tervaert, J. W., Sluiter, W. J., Manson, W. L., de Jong, P. E., and Kallenberg, C. G. M. (1994). Association of chronic nasal carriage of *Staphylococcus aureus* and higher relapse rates in Wegener's granulomatosis. *Annals of Internal Medicine*, **120**, 12–17.

Steppat, D. and Gross, W. L. (1989). Stage-adapted treatment of Wegener's granulomatosis. *Klinische Wochenschrift*, **67**, 666–71.

Stillwell, T. J., Benson, R. C., DeRemee, R. A., McDonald, T. J., and Weiland, L. H. (1988). Cyclophosphamide-induced bladder toxicity in Wegener's granulomatosis. *Arthritis and Rheumatism*, **31**, 465–70.

Stoney, P. J., Davies, W., Ho, S. F., Paterson, I. C., and Griffith, I. P. (1991). Wegener's granulomatosis in two siblings: a family study. *Journal of Laryngology and Otology*, **105**, 123–4.

Travis, W. D., Hoffman, G. S., Leavitt, R. Y., Pass, H. I., and Fauci, A. S. (1991). Surgical pathology of the lung in Wegener's granulomatosis. *American Journal of Surgical Pathology*, **15**, 315–33.

Tuso, P., Moudgil, A., Goodman, D., Kamil, E., Koyyana, R., and Jordan, S. C. (1992). Treatment of antineutrophil cytoplasmic autoantibody-positive systemic vasculitis and glomerulonephritis with pooled intravenous gammaglobulin. *American Journal of Kidney Diseases*, **5**, 504–8.

van der Lubbe, P. A., Miltenburg, A. M., and Breedveld, F. C. (1991). Anti-CD4 monoclonal antibody for relapsing polychondritis. *Lancet*, **337**, 1349.

Waldherr, R., Noronha, I. L., Niemir, Z., Krüger, C., Stein, H., and Stumm, G. (1993). Expression of cytokines and growth factors in human glomerulonephritides. *Pediatric Nephrology*, **7**, 471–8.

Walton, E. W. (1958). Giant-cell granuloma of the respiratory tract (Wegener's granulomatosis). *British Medical Journal*, 265–9.

Wegener, F. (1939). Über eine eigenartige rhinogene Granulomatose mit besonderer Beteiligung des Arteriensystems und der Nieren. *Beiträge zur Pathologie*, **102**, 36–68.

Wegener, F. (1990). Wegener's granulomatosis. Thoughts and observations of a pathologist. *European Archives of Otorhinolaryngology*, **247**, 133–42.

Weidhase, A., Gröne, H-J., Unterberg, C., Schuff-Werner, P., and Wiegand, V. (1990). Severe granulomatous giant cell myocarditis in Wegener's granulomatosis. *Klinische Wochenschrift*, **68**, 880–5.

Weiner, S. R., Paulus, H. E., and Weisbart, R. H. (1986). Wegener's granulomatosis in the elderly. *Arthritis and Rheumatism*, **29**, 1157–9.

Weiss, M. A. and Crissman, J. D. (1984). Renal biopsy findings in Wegener's granulomatosis: Segmental necrotizing glomerulonephritis with glomerular thrombosis. *Human Pathology*, **15**, 943–56.

Wiik, A. (1993). Antineutrophil cytoplasmic antibodies in Wegener's granulomatosis. *Clinical and Experimental Rheumatology*, **11**, 191–201.

Wiik, A., Rasmussen, N., and Wieslander, J. (1993). Methods to detect autoantibodies to neutrophilic granulocytes. *Manual of Biological Markers of Disease*, **A9**, 1–14.

Yang, J., Tuttle, R., Jenette, C., and Falk, R. J. (1993). Glomerulonephritis in rats immunized with myeloperoxidase (MPO). *Clinical and Experimental Immunology*, **93** (Suppl. 1), 22.

5.11.3 Classical polyarteritis nodosa, microscopic polyarteritis, and Churg–Strauss syndrome

D. Adu and Paul A. Bacon

Acknowledgement: We are grateful to Dr A.J. Howie for the photomicrographs used in Chapter 5.11.3.

Introduction and historical background

The aetiology and pathogenesis of most forms of systemic necrotizing vasculitis remain obscure. Few areas are more contentious than the classification of the necrotizing vasculitides. The difficulties are in part due to the clinical overlap within these disorders and in part because we know so little about their aetiology. Our levels of description are based on histopathology, clinical presentation, immunology, and behaviour either spontaneous or in response to treatment, and these are inexact tools for classification (see Chapter 5.11.1).

The simplest classification is into three groups of diseases based on vessel size—giant-cell arteritis of large arteries; necrotizing vasculitis of medium and small muscular arteries; and small vessel disease involving arterioles, capillaries, and venules. The necrotizing arteritis group contains both vasculitis as a complication of connective tissue disorders, such as rheumatoid arthritis, and idiopathic disease. The latter ranges from pure arteritis (classical polyarteritis nodosa) to granulomatous vasculitis (Wegener's).

In the 19th century, following Kussmaul and Maier's (1866) classical description onwards, periarteritis nodosa was considered as a rare distinct disease entity. It was characterized by visible nodular lesions at autopsy involving the muscular arteries. However, in the 20th century cases were described in which the diagnosis was made microscopically, with consequent changes in emphasis of disease descriptions. In her classic review Zeek (1952) divided the necrotizing vasculitides into five main groups. The first of these was hypersensitivity angiitis which typically involved capillaries, arterioles, and the small arteries as well as the venules and which she differentiated from periarteritis nodosa, allergic granulomatous angiitis, rheumatoid arteritis, and temporal arteritis. The organs involved were the kidney in particular, often with necrotizing glomerulonephritis, but the angiitis was usually widespread and pulmonary vessels were frequently involved. The term hypersensitivity angiitis is now used for clinical syndromes in which cutaneous manifestations predominate and in which visceral involvement is uncommon.

Pulmonary disease is a prominent feature of systemic vasculitis and studies in the 1930s separated off firstly Wegener's granulomatosis with predominant respiratory tract granuloma formation but also complicated by a vasculitis as an entity that was distinct from polyarteritis (Wegener 1936). In 1951, Churg and Strauss described the clinical syndrome of asthma, eosinophilia, and systemic vasculitis with extravascular granulomata that bears their name.

Renal lesions have been associated with polyarteritis since the comprehensive description by Kussmaul and Maier (1866). Most classic reviews have found an incidence of renal involvement in approximately 80 per cent. The morphological changes of polyarteritis nodosa—a term first used by Ferrari (1903)— were described in detail by Arkin (1930). He pointed out that in cases without aneurysms the vasculitis may only be seen microscopically, and referred to such cases as microscopic polyarteritis nodosa. This theme was taken up and expanded in a detailed clinicopathological study by Davson *et al*. (1948). They linked the presence of the primary systemic vasculitis with the occurrence of segmental necrotizing glomerulonephritis in a subgroup of patients. They again suggested that such patients had a microscopic form of polyarteritis, which they thought was distinct from the periarteritis (or polyarteritis) nodosa described by Kussmaul and Maier. They divided their patients into two groups, those with severe and widespread glomerular disease (9 out of 14) and those with minimal or no glomerular disease. Half of the first group had arterial (arcuate artery) involvement in addition to glomerular disease. These presented with a rather uniform clinical picture with pyrexia, leucocytosis, and uraemia in the presence of a normal blood pressure. The second group had variable clinical symptoms including elevated blood pressure, locomotor symptoms, abdominal pain, heart and lung involvement. Pathology here showed lesions typical of classical periarteritis and of malignant hypertension. Davson *et al*. concluded that the microscopic form of polyarteritis was different from the classical aneurysmal type. Importantly, they stated clearly that it was possible to separate microscopic polyarteritis from other forms of diffuse nephritis on both clinical and pathological grounds, establishing that the nephritis was not a coincidental or unrelated feature.

Definition of microscopic polyangiitis (polyarteritis)

It is now widely accepted on both sides of the Atlantic that the microscopic form of necrotizing angiitis with predominant involvement of the renal and pulmonary capillaries is a separate disease entity from the original descriptions of periarteritis or polyarteritis nodosa. The main question is that of nomenclature, and we favour that proposed by the Chapel Hill consensus conference (Jennette *et al*. 1994) (Table 1) but many would disagree (Lie 1994). The preferred name is now microscopic polyangiitis (microscopic polyarteritis) which has to be distinguished from classical polyarteritis nodosa. The most important distinguishing feature is the presence of vasculitis in small vessels (arterioles, venules, and capillaries) in microscopic polyangiitis and its absence in the latter. It has to be emphasized that some overlap of the size of vessel involved is seen. By definition microscopic polyangiitis must have involvement of microscopic vessels, but it is recognized that it can also involve small or medium-sized arteries. In contrast classical polyarteritis nodosa must have no involvement of microscopic vessels and therefore no glomerulonephritis. Microscopic polyangiitis and classical polyarteritis nodosa are thus differentiated by the presence or absence of small vessel involvement, rather than by the involvement of medium-sized arteries which can occur in either. The precise definitions and the question of whether involvement at all sites carry the same diagnostic weight, are still matters of discussion. For example the occurrence of small vessel involvement of the skin in classical polyarteritis nodosa with otherwise typical arterial involvement at all other sites would not change the diagnosis in our view. The definition of microscopic polyangiitis also requires few or no immune deposits in the blood vessels, in order to allow differentiation from those variants of small vessel vasculitis that do have well-defined immune complexes. This is particularly important to allow distinction from Henoch–Schönlein purpura and cryoglobulinaemic vasculitis, both with predominant cutaneous involvement. It also differentiates other forms of immune complex-mediated small vessel vasculitis such as that seen in systemic lupus erythematosus. The most frequent, clinically important aspect of microscopic polyangiitis is the glomerular lesion; similar small vessel lesions may occur in the lung. It is recognized that identical glomerular lesions may be seen in Wegener's granulomatosis but there are other important distinctions from that condition examined below.

Size of vessel	Syndrome	Granuloma	ANCA
Large	Takayasu's arteritis	+	+
	Temporal arteritis	+	–
Medium	Polyarteritis nodosa	–	+C
	Churg–Strauss syndrome	+	+P
	Kawasaki disease	–	–
Small	Wegener's granulomatosis	+	+C
	Microscopic polyangiitis	–	+P
	Leucocytoclastic vasculitis	–	–

From Jennette et al. 1994.

Table 1 Nomenclature of vasculitis proposed by the Chapel Hill consensus conference

Comparison with classical polyarteritis nodosa

The clinical distinction between the microscopic and the classical aneurysmal form of polyarteritis is not always easy. The overlapping vessel size which occurs in microscopic polyangiitis confuses interpretation of older series and makes the interrelationship of the two syndromes difficult to interpret. For example 25 per cent of renal patients in our series (Adu *et al*. 1987), all of whom had glomerulonephritis, also had renal arteritis. There is a similarity in the clinical symptoms. Even in the renal series, the majority of patients first present with non-renal symptoms. However, previous experience in a District General Hospital had emphasized the wide spectrum of severe organ involvement in classical polyarteritis nodosa (Scott *et al*. 1982). A comparison of several series of patients from renal units with those from general or overlapping departments confirmed that there are real differences (Adu *et al*. 1987). The general series, which contained many cases of classical polyarteritis nodosa, showed a higher incidence of fever, heart and lung, gastrointestinal, and peripheral nerve involvement. The difference in vessel size involved appears to dictate not just the clinical pattern of organ damage but also the prognosis. Thus visceral infarction or haemorrhage in classical polyarteritis nodosa contributes prominently to the overall outcome. The overt clinical symptoms often also contribute to an earlier diagnosis and thus institution of therapy in classical polyarteritis nodosa.

In microscopic polyangiitis there is an increase in both renal and lung disease. Pulmonary involvement in this disease has been noted in previous series and contributes to a poor prognosis (Savage *et al*. 1985; Adu *et al*. 1987). Patients may present with overt or even massive pulmonary haemorrhage in association with renal failure—so-called pulmonary renal syndrome—and these patients tend to fair badly. Less dramatic pulmonary involvement is even more frequent. The pathological lesion, a necrotizing vasculitis in small alveolar vessels, is essentially the same as in the kidney. The most striking difference between the two conditions in our series is the frequency of presentation. Eight microscopic polyangiitis cases have been seen for every classical polyarteritis nodosa case. This is not due to a biased collection from a renal centre, since the Birmingham Vasculitis Group collects cases from rheumatological as well as renal clinics and from general physician referrals. It reflects the widespread experience in Europe now that microscopic polyangiitis is common (and perhaps increasing in incidence) while classical polyarteritis nodosa is rare (and perhaps decreasing). There is also a difference in outcome. Relapse rates are high in vasculitis and this was seen in two-fifths of our classical polyarteritis nodosa cases, but only in a quarter of the microscopic polyangiitis group. However the overall prognosis is much worse in microscopic polyangiitis, which in our series showed a 40 per cent overall mortality over the decade, largely related to renal failure often associated with delayed referrals. In contrast classical polyarteritis nodosa responded well to therapy, with a 100 per cent survival in our hands despite the relapses (Gordon *et al*. 1993).

The advent of testing for antineutrophil cytoplasmic antibodies (**ANCA**) has provided another way to compare the two forms of angiitis (reviewed by Kallenberg *et al*. 1994). In microscopic polyangiitis the majority of patients (80 per cent) are ANCA positive. Specificity is directed predominantly towards myeloperoxidase and to a lesser extent against proteinase 3. In contrast, in our patients with classical polyarteritis nodosa a positive ANCA was found in 14 per cent. Others have also concluded that classical polyarteritis nodosa, without evidence of small vessel involvement, is usually ANCA negative (Cohen Tervaert *et al*. 1991; Guillevin *et al*. 1993).

Comparison with Wegener's granulomatosis

Although patients with Wegener's granulomatosis have, by definition, disease of the upper and lower respiratory tract and histological evidence of necrotizing granulomata often with an accompanying vasculitis, biopsies of lesions often show only a non-specific inflammation. A particular feature of respiratory disease in Wegener's granulomatosis is that even with treatment it is persistent and often relapsing and this helps to differentiate this disease from microscopic polyangiitis where the respiratory disease, for example pulmonary haemorrhage or epistaxis, is usually acute and transient. Wegener's granulomatosis is the classical disease which is ANCA positive, with an antibody specificity directed primarily to proteinase 3. An identical, pauci immune, necrotizing glomerulonephritis also occurs in this condition, but there is little to support the concept that microscopic polyangiitis represents a renal-limited form of Wegener's granulomatosis (Adu *et al*. 1987; Gordon *et al*. 1993).

Treatment

Therapy in these previously incurable disorders with a very high mortality has been revolutionized in the past two decades by the introduction of cytotoxic/immunosuppressive regimes. The dramatic effect of cyclophosphamide together with steroids in Wegener's granulomatosis was first established in 1971 (Novack and Pearson 1971) and these observations were later confirmed (Fauci and Wolff 1973) and extended to polyarteritis (Fauci *et al*. 1979). Recent controlled data support this beneficial effect.

Specific diseases

Classical polyarteritis

Classical polyarteritis nodosa is a systemic illness characterized by necrotizing inflammation of medium-sized arteries leading to aneurysm formation. It affects the viscera and other organs. The clinical presentation is dominated by organ infarction and haemorrhage, a neuropathy, and myalgia in the context of a systemic illness.(Frohnert and Sheps 1967; Sack *et al*. 1975; Leib *et al*. 1979; Cohen *et al*. 1980; Scott *et al*. 1982)

Age, sex, and race

In most studies there is a consistent predominance of males (male:female ratio, 1.9–2:1). The majority of patients have been Caucasian although the disease has been described in black and Asian groups. The age range is from 14 to 80 years with a peak frequency of onset in the 40s and 50s.

Clinical presentation

The clinical presentation in these patients is summarized in Fig. 1 and compared with those of patients with microscopic polyangiitis. Patients with classical polarteritis nodosa often present with systemic features comprising malaise, fever, and weight loss. About 60 per cent of patients have an arthralgia (or less commonly an arthritis) and 50 per cent of patients have a rash that is erythematous, purpuric, or vasculitic. Punched-out ulcers may occur, often around the ankles. A prominent feature in these patients is a peripheral neuropathy in 40 per cent of patients, often presenting as a mononeuritis multiplex. Gastrointestinal involvement is frequent, found in 45 per cent of our series, and carries a high morbidity and mortality. Presenting features include gut infarction, bleeding, pancreatitis, and infarction of the gallbladder, as well as non-specific abdominal pain. Renal disease in the form of an abnormal urinary sediment occurs in 55 per cent of patients, although only 24 per cent of patients have renal impairment and it is often mild. Myalgia is a prominent symptom and biopsy of affected muscles often shows a vasculitis. Rarely, testicular pain is a presenting feature of polyarteritis.

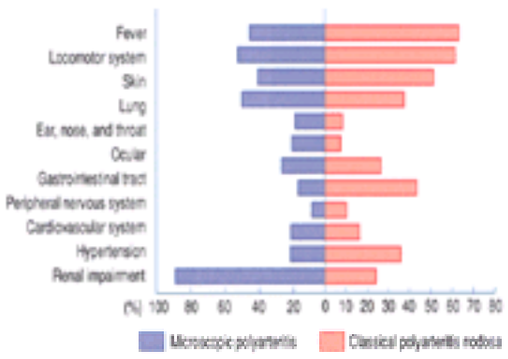


Fig. 1 Clinical features at presentation in classical polyarteritis nodosa (Frohnert and Sheps 1967; Sack *et al*. 1975; Leib *et al*. 1979; Cohen *et al*. 1980; Scott *et al*. 1982) and microscopic polyarteritis (Serra *et al*. 1984; Savage *et al*. 1985; Adu *et al*. 1987).

Hepatitis B and classical polyarteritis nodosa

Positive hepatitis B serology has been noted in between 5 and 40 per cent of different series of patients with classical polyarteritis nodosa (Frohnert and Sheps 1967; Sack *et al*. 1975; Leib *et al*. 1979; Cohen *et al*. 1980). Its prevalence appears to reflect the incidence of hepatitis B carriage in the population. It also suggests that polyarteritis nodosa is not a disease with a single aetiological agent or mechanism. In England less than 10 per cent of patients have positive hepatitis B serology (Scott *et al*. 1982) but this is considerably higher in studies from some parts of Europe (Guillevin *et al*. 1988). Clinically these patients appear more likely to have a nephrotic syndrome, which is otherwise rare in polyarteritis (Guillevin *et al*. 1988).

Localized polyarteritis

Isolated organ involvement in polyarteritis is rare. Involvement of the skin, testes, epididymis, breasts, uterus, appendix, and gallbladder have all been reported and we have seen arteritis localized to one kidney. It is not yet clear how these patients should be managed, especially if there are no systemic signs of vasculitis. The risk of progression to systemic disease in such cases is not known, although one study suggested that on long-term follow-up, the majority of patients progressed to systemic disease (Minkowitz *et al*. 1991). Close long-term follow-up is clearly important.

Microscopic polyangiitis

These patients have predominant involvement of glomerular capillaries but without granuloma formation, such as is seen in Wegener's granulomatosis. Unlike classical polyarteritis nodosa most patients with microscopic polyangiitis present with or develop severe renal disease, particularly if the diagnosis is delayed, and most studies of this disorder emanate from renal units (Serra *et al*. 1984; Savage *et al*. 1985; Coward *et al*. 1986; Adu *et al*. 1987). There may be little evidence of multisystem disease but diagnosis is aided by the detection of ANCA. In microscopic polyangiitis a perinuclear pattern of staining (**p-ANCA**) is usually seen, in contrast with the diffuse granular cytoplasmic staining (**c-ANCA**) seen in Wegener's granulomatosis (Falk and Jennette 1988; Lee *et al*. 1990). P-ANCA detects a much wider range of antigens than c-ANCA but in microscopic polyangiitis the frequent antigenic specificity is to myeloperoxidase.

Age, sex, and race

Most studies show a male predominance (male:female ratio, 1.2–1.8:1) and the overwhelming majority of patients reported are Caucasian, with only a few Negroid patients and Asians from the Indian subcontinent (Serra *et al*. 1984; Savage *et al*. 1985; Adu *et al*. 1987). The mean age at presentation is around 50 years.

Clinical features of microscopic polyangiitis

Much of the initial clinical presentation of microscopic polyangiitis is non-specific. The symptoms of microscopic polyangiitis are compared with those of classical polyarteritis nodosa in Fig. 1 (Serra *et al*. 1984; Savage *et al*. 1985; Adu *et al*. 1987). Systemic symptoms such as malaise, anorexia, fever, and weight loss predominate initially and these are accompanied by often asynchronous episodes of rash, arthralgia, myalgia, and conjunctivitis. These symptoms have often been present for months before there is a clinical suspicion of a multisystem disorder. A rash is present in approximately 40 per cent of cases, usually purpuric or vasculitic, and occurs most commonly on the limbs, in particular the hands and feet (Fig. 2). Skin ulceration is uncommon, being found in less than 5 per cent of patients, most commonly around the ankles (Fig. 2). Arthralgia or myalgia is present in about a half of patients and a quarter of these develop an asymmetric large joint arthritis.

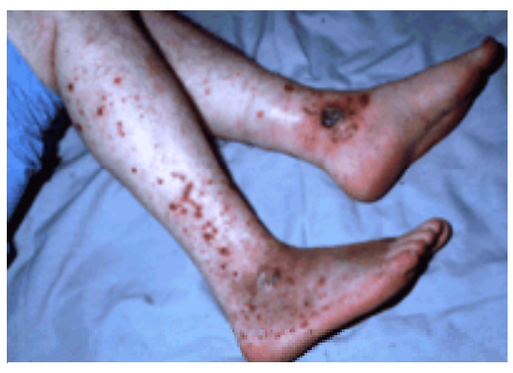


Fig. 2 Vasculitic rash and 'punched-out' ulcer on the leg of a patient with microscopic polyangiitis.

Renal diseasex

The clinical presentation of renal disease is with microscopic haematuria (80 per cent of patients), frank haematuria (6 per cent of patients) and proteinuria (80 per cent of patients). Over 90 per cent of patients have renal impairment, and in most series no fewer than 30 per cent of patients were oliguric by the time the diagnosis was made. Hypertension is infrequent, being found in 21 per cent of patients, and usually mild.

Pulmonary disease

Symptoms of lung disease are found in about 50 per cent of patients and usually take the form of haemoptysis, pleurisy, and asthma. Frank pulmonary haemorrhage develops in 5 per cent of patients (Fig. 3) and this can be severe enough to require mechanical ventilation. The importance of diffuse pulmonary haemorrhage in both microscopic polyangiitis and Wegener's granulomatosis has been emphasized (Haworth *et al*. 1985).

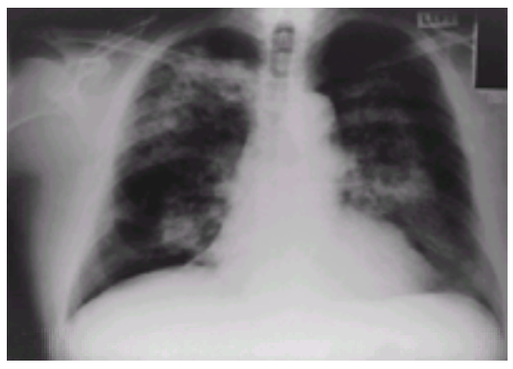


Fig. 3 Diffuse reticulonodular pulmonary shadowing in the chest radiograph of a patient with microscopic polyangiitis and pulmonary haemorrhage.

Churg–Strauss syndrome

Clinical features

In 1951, Churg and Strauss described autopsy features in 13 patients who died following an illness characterized by asthma, eosinophilia, fever, and a systemic illness. All 13 patients had had asthma for up to 10 years prior to the onset of other symptoms. Fever was a universal finding, as was eosinophilia, although the levels varied considerably within individuals over time. Anaemia, weight loss, heart failure, recurrent pneumonia, and bloody diarrhoea were common. Hypertension was seen in patients with more long-standing disease. Skin lesions were usually present, varying from purpura to nodules (biopsy of which revealed granulomata). Central and peripheral nerve involvement were common, especially peripheral neuropathy (in 8 out of 13 cases). Significant renal involvement was unusual in their patients, although mild urinary sediment abnormalities were often seen.

The spectrum of clinical features of this disorder are summarized in Fig. 4. Lanham *et al*. (1984) provided a clinical definition of Churg–Strauss syndrome as a triad of asthma, eosinophilia (more than 1.5×10^9), and a systemic vasculitis involving two or more extra-pulmonary organs. From their experience and a review of the literature they concluded that extravascular granulomas were not essential for the diagnosis of Churg–Strauss syndrome. In their study, asthma or an allergic rhinitis often preceded by years the development of a peripheral blood eosinophilia and pulmonary shadowing resembling Löffler syndrome, and this in turn was followed by a systemic vasculitis. However, these manifestations may occur simultaneously (Chumbley *et al*. 1977). Although infrequently reported, renal involvement does occur and usually presents with microscopic haematuria plus proteinuria and less commonly with renal failure or a nephrotic syndrome (Clutterbuck *et al*. 1990).

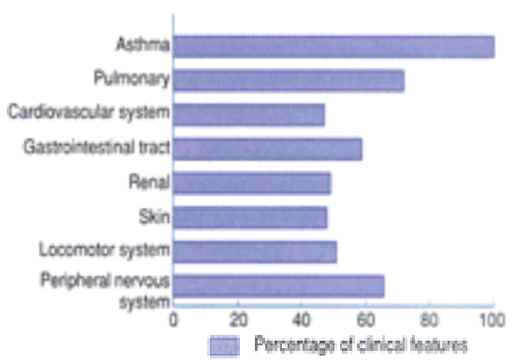


Fig. 4 Clinical features of Churg–Strauss syndrome (Lanham *et al*. 1984).

Differential diagnosis of Churg–Strauss syndrome

The prominence of asthma, an allergic rhinitis, and persistent elevation of the eosinophil count as well as the predominance of eosinophils in areas of inflammation help to differentiate Churg–Strauss syndrome from Wegener's granulomatosis, classical polyarteritis nodosa, and microscopic polyangiitis. Clinically this differentiation is not always easy (Guillevin *et al*. 1988) and the American College of Rheumatology had difficulty in classifying some of their patients with Wegener's granulomatosis, polyarteritis nodosa, and Churg–Strauss syndrome according to well-defined criteria (Masi *et al*. 1990).

Approach to the diagnosis of vasculitis

Difficulties in the classification of vasculitis often lead to delays in diagnosis. This delay is particularly marked in the elderly when it is often asociated with the rapid development of renal failure, which in turn carries a high mortality.

Clinical

The initial diagnosis is often clinical, and a high index of suspicion is required in patients with a multisystem disease, especially with a fever, skin, or renal involvement. A particular feature of systemic vasculitis is that the symptoms may be separated in time and also often fluctuate in severity. Thus for example an individual may have had an episode of arthritis followed by a uveitis and then renal disease before the possibility of a vasculitis is considered. The clinical suspicion should be pursued aggressively and with urgency in order to establish a diagnosis quickly. The overwhelming need at this stage to establish a diagnosis of a vasculitis which then requires therapy is greater than the need to establish a precise syndrome diagnosis. This may not be apparent at first presentation but subsequent progress often clarifies that. Usually there is an otherwise unexplained acute phase response with a high white blood-cell count, C-reactive protein, and erythrocyte sedimentation rate. A biopsy of clinically involved tissue is often helpful, for example skin, muscle, sural nerve, or kidney. Where there is no specific organ involvement, renal and coeliac axis angiograms should be considered.

Immunological investigations in microscopic polyangiitis, classical polyarteritis nodosa, and Churg–Strauss syndrome

Antineutrophil cytoplasmic antibodies (reviewed by Kallenberg *et al*. 1994; see also Chapter 5.11.2)

ANCA were first described by Davies *et al*. (1982) and Hall *et al*. (1984) in patients with a segmental necrotizing glomerulonephritis, some of whom had features consistent with a systemic vasculitis. The report of Van der Woude *et al*. (1985) showed that antineutrophil cytoplasmic antibodies were found commonly in the sera from patients with Wegener's granulomatosis. Subsequently these observations were extended to patients with microscopic polyangiitis (Savage *et al*. 1987). Two broad patterns of cytoplasmic staining have been recognized. In the first there is granular staining of the cytoplasm of ethanol-fixed neutrophils (c-ANCA). This pattern of staining is commonly seen in the sera of patients with Wegener's granulomatosis (Van der Woude *et al*. 1985). The target antigen for this antibody is almost entirely restricted to a 29 kDa serine proteinase, neutrophil serine proteinase 3 (Goldschmeding *et al*. 1989). The second pattern is of perinuclear staining (p-ANCA) which is seen in the sera of patients with microscopic polyangiitis and also idiopathic segmental necrotizing glomerulonephritis. A wider range of specificities is seen with p-ANCA, although the majority of these antibodies are directed against myeloperoxidase (Falk and Jennette 1988; Lee *et al*. 1990) and in a few cases against elastase and lactoferrin.

Approximately 80 per cent of patients with active, systemic Wegener's granulomatosis have c-ANCA antibodies in their sera (reviewed by Van der Woude *et al*. 1989; Specks and De Remee 1990). P-ANCA are found in the sera of approximately 80 per cent of patients with microscopic polyangiitis and about 75 per cent of patients

with Churg–Strauss syndrome (Cohen Tevaert *et al*. 1991; Guillevin *et al*. 1993). P-ANCA is less disease specific as well as less sensitive. Elevated titres are reported in Kawasaki's syndrome, rheumatoid vasculitis, and idiopathic segmental necrotizing glomerulonephritis. Between 14 and 20 per cent of patients with classical polyarteritis nodosa have a positive ANCA (Adu *et al*. 1993; Guillevin *et al*. 1993) and in patients with polyarteritis nodosa that is positive for hepatitis B, 11 per cent are reported to have a positive ANCA (Guillevin *et al*. 1993). It is important to recognize that ANCA have also been described in systemic lupus erythematosus (rarely), rheumatoid arthritis, inflammatory bowel disease, and in some infectious diseases (reviewed by Kallenberg *et al*. 1994).

ANCA in pathogenesis

The pathogenesis of vascular inflammation in systemic vasculitis is unknown. ANCA have recently been postulated to have a major role in this (reviewed by Kallenberg *et al*. 1994). Proinflammatory cytokines induce the surface expression of PR3 and MPO on neutrophils where they are accesible to react with ANCA (Falk *et al*. 1990a). Recent studies have shown that neutrophils primed with tumour necrosis factor-a are activated by ANCA to produce reactive oxygen products as well as to degranulate by the cross-linking of surface-expressed ANCA antigen with FcgR11 (Porges *et al*. 1994) and this can lead to bystander injury to endothelial cells (Ewert *et al*. 1992; Savage *et al*. 1992). MPO and PR3 bind to endothelial cells of cultured human umbilical vein by a charge mechanism and can then react with ANCA, with the potential for subsequent endothelial cell damage through complement activation or neutrophil adhesion and activation (Savage *et al*. 1992; Varagunam *et al*. 1992a). Further, it has been reported that ANCA can increase neutrophil adhesion to cultured endothelial cells, although the mechanism of this is unclear.

ANCA-negative vasculitis

The majority of patients with classical polyarteritis nodosa (more than 80 per cent) and a minority of patients with microscopic polyangiitis (20 per cent) have a negative ANCA. This means that whilst ANCA are a useful non-invasive test n the diagnosis for vasculitis they are not essential for diagnosis. Serial studies suggest that while rising ANCA titres may predict a later flare, this is not always the case (Jayne *et al*. 1995). Strictly ANCA appear to be neither necessary nor sufficient for either diagnosis or flares of vasculitis, emphasizing the complex interactions which lead to disease expression.

Antiendothelial cell antibodies

Several studies have reported the presence of antibodies to endothelial cells (**AECA**) in the sera of patients with both Wegener's granulomatosis and polyarteritis (Frampton *et al*. 1990; Varagunam *et al*. 1992b). It has been suggested that AECA are important in the pathogenesis of vasculitis and inferred that they might cause endothelial injury. Alternatively they may be a reaction to endothelial damage and thus useful in assesment. It is as yet unclear whether these antibodies are going to be useful in diagnosis or management of these disorders.

Renal and coeliac angiograms

The value of arteriography in the diagnosis of classical polyarteritis nodosa has been reviewed by Travers *et al*. (1979). The radiological findings include aneurysms and also segmental narrowing and variation in the calibre of arteries, together with pruning of the peripheral vascular tree. Interpretation of the significance of such findings requires a radiologist with special experience. Aneurysms are most found commonly in the hepatic and renal arterial tree (Fig. 5) but can also occur in cerebral and pulmonary arteries. In the kidneys, wedge-shaped areas of infarction may be seen. Angiography is helpful both for confirming the presence and for documenting the extent of a necrotizing vasculitis. It is less useful in establishing the precise disease label. Aneurysms may also occur infrequently in other vasculitic disorders such as Wegener's granulomatosis and systemic lupus erythematosus. Where there is a clinical suspicion of classical polyarteritis nodosa it is sensible to do renal and coeliac angiograms and only to proceed to a renal biopsy if no aneurysms are found.

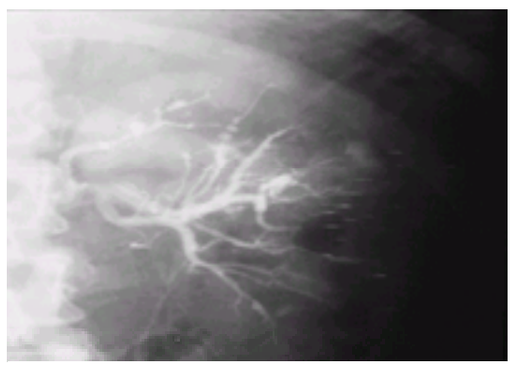


Fig. 5 Renal arteriogram in classical polyarteritis nodosa showing multiple aneurysms on renal vessels.

Pathology

Classical polyarteritis nodosa

Histology shows an arteritis with endothelial damage, fibrinoid necrosis affecting both intima and media (Fig. 6) and an inflammatory infiltrate of intima and media consisting predominantly of neutrophils but also of mononuclear cells. Often there is destruction of the internal elastic lamina. In some lesions the media and adventitia are surrounded by an infiltrate of mononuclear leukocytes and neutrophils. Immunofluorescent studies are in general negative for immunoglobulins and complement (Spargo *et al*. 1980; Cupps and Fauci 1981; Ronco *et al*. 1983).

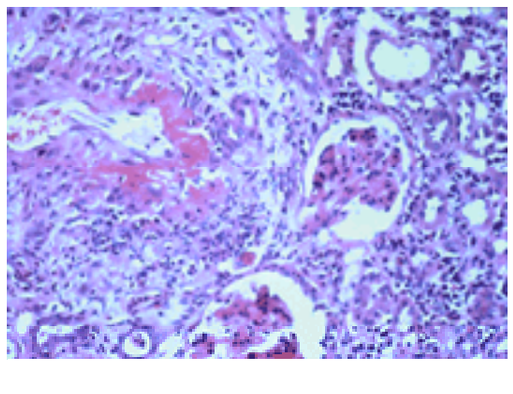


Fig. 6 Photomicrograph of a renal biopsy showing fibrinoid necrosis of an interlobular artery. Haematoxylin and eosin, original magnification × 25.

Microscopic polyangiitis—renal pathology

The characteristic renal lesion is a focal, segmental necrotizing glomerulonephritis with fibrinoid necrosis and thrombosis of segments of the glomerular tufts surrounded by neutrophils (Fig. 7) (Davson *et al*. 1948; Serra and Cameron 1985; D'Agati *et al*. 1986; Adu *et al*. 1987). Often there is rupture of the glomerular basement membrane with adjacent extracapillary proliferation. Between 15 and 20 per cent of biopsies show extensive extracapillary proliferation (crescent formation). Healed glomerular lesions are characterized by sharply defined segmental scars in which disrupted glomerular capillary loops can be seen. Glomerular immune deposits of immunoglobulin and complement are sparse and this is of value in differentiating the glomerular lesions of microscopic polyangiitis from those of

other vasculitic disorders such as systemic lupus erythematosus and Henoch–Schönlein purpura. In about 20 per cent of the patients there is an acute vasculitis affecting arterioles as well as arteries. This often takes the form of endothelial damage, with fibrinoid necrosis disrupting the internal elastic lamina, and an infiltrate of neutrophils plus mononuclear cells in the media, the adventitia, and surrounding the vessel.

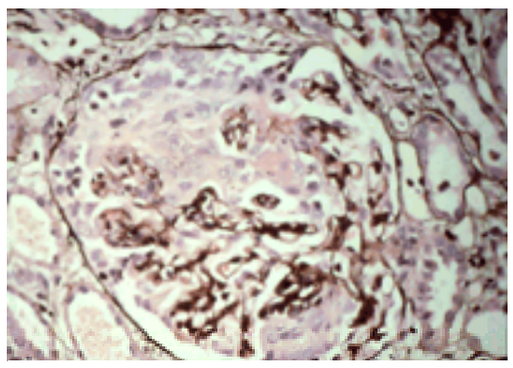


Fig. 7 Photomicrograph of a renal biopsy showing a segmental necrotizing glomerulonephritis with thrombosis and disruption of glomerular capillary loops and overlying extracapillary proliferation (crescents). Periodic acid–methenamine silver, original magnification × 100.

Most biopsies show a tubulointerstitial infiltrate with lymphocytes and eosinophils often around affected glomeruli. The severity of renal impairment correlates more closely with the severity of tubular damage than with the extent of glomerular disease.

Electron microscopy (D'Agati et al. 1986)

Electron microscopy of glomeruli and also of arteries and arterioles shows that the earliest changes are of endothelial swelling and degeneration with focal detachment of the endothelium and subendothelial deposits of fibrin. With more advanced lesions there is more widespread denudation of endothelium with intraluminal fibrin deposits and occasional thrombi. At this stage polymorphonuclear leucocytes and monocytes are found both in the vessel lumen and within and around the walls.

Leucocyte infiltrate

Studies of renal histology in patients with Wegener's granulomatosis and microscopic polyangiitis have shown infiltrating monocytes/macrophages and neutrophils within glomeruli and in crescents. CD3 and IL-2R positive T cells were also identified in crescents, in the periglomerular area, and in the interstitium (Noronha *et al*. 1993; Brouwer *et al*. 1994)

Cell adhesion molecules

Cell adhesion molecules (**CAMs**) play a critical role in the migration of leucocytes to sites of inflammation, in selecting the types of leucocytes that accumulate, and in modifying leucocyte function, thus providing a mechanism by which leucocytes accumulate in renal tissues leading to glomerular and interstitial inflammation. CAMs expressed by the human kidney include the glycoproteins, intercellular adhesion molecule-1 (ICAM-1, CD54), and vascular cell adhesion molecule-1 (**VCAM-1**), which belong to the immunoglobulin supergene family, and the selectin, E-Selectin. The glomerular endothelium in biopsies from patients with Wegener's granulomatosis and microscopic polyangiitis expresses VCAM-1, which was not seen in any of the normal biopsies studied (Bruijn and Dinklo 1993; Pall *et al*. 1996). VCAM-1 expression by the glomerular endothelium may be important in selectively recruiting monocytes/macrophages and T lymphocytes that express VLA-4 (CD49d/CD29). It is possible that this adhesion molecule is involved in the genesis of this type of glomerular inflammation.

Churg–Strauss syndrome

The histological lesions described by Churg and Strauss included extravascular necrotic lesions which were widespread but most commonly seen in the epicardium. Acute lesions showed a predominant eosinophilic infiltrate whilst more chronic lesions comprised of giant-cell granulomata. In addition there was often a necrotizing arteritis with an infiltrate comprising eosinophils and giant-cell granulomata. Subsequent studies by Chumbley *et al*. (1977) and Lanham *et al*. (1984) suggest that a granulomatous vasculitis and extravascular granulomata were not essential for the diagnosis of Churg–Strauss syndrome. As with microscopic polyangiitis, the predominant renal lesion found in 85 per cent of cases is a focal, segmental necrotizing glomerulonephritis, often with extracapillary proliferation. Eosinophil infiltration and granulomata in the renal interstitium and a necrotizing arteritis are uncommon and found in less than 10 per cent of cases (Clutterbuck *et al*. 1990).

Other investigations in patients with a vasculitis

These provide evidence of an acute phase response with anaemia, high white blood-cell count, erthrocyte sedimentation rate, C-reactive protein, platelets, and serum complement levels, and a low serum albumin (reviewed by Serra and Cameron 1985). Serum levels of factor VIII-related antigen are raised in patients with vasculitis (Woolf *et al*. 1987). An eosinophilia (more than 1.5×10^{9}/l) is a characteristic feature of Churg–Strauss syndrome, although lesser degrees of eosinophilia may be seen in patients with classical polyarteritis nodosa who usually also have pulmonary disease (Frohnert and Sheps 1967). Microscopic haematuria and proteinuria are common presenting abnormalities in patients with renal disease. Abnormal liver function tests with a raised serum alkaline phosphatase are found in approximately 50 per cent of patients with microscopic polyangiitis (Savage *et al*. 1985; Adu *et al*. 1987). Low titres of rheumatoid factor are found in both Wegener's granulomatosis (Fauci *et al*. 1983) and classical polyarteritis nodosa (Leib *et al*. 1979; Scott *et al*. 1982). Low titres of antinuclear antibodies are found in a minority of patients with classical polyarteritis nodosa (Cohen *et al*. 1980; Scott *et al*. 1982). Immune complexes, usually in low titre, have been reported in the sera of some patients with a vasculitis (Scott *et al*. 1982; Ronco *et al*. 1983; Serra *et al*. 1984) but their significance is uncertain.

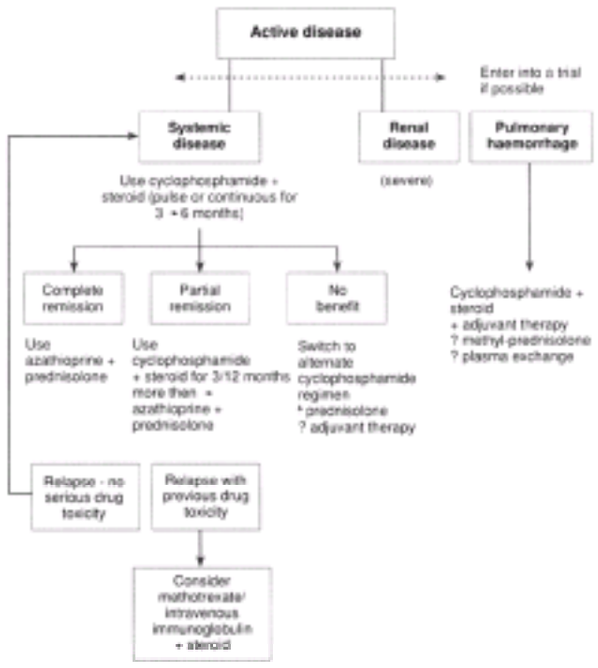
Therapy and prognosis of vasculitis

Overview of treatment

There is now a general consensus of opinion that the prognosis of idiopathic systemic necrotizing vasculitis (Wegener's granulomatosis, classical polyarteritis nodosa, and microscopic polyangiitis) has been greatly improved by the addition of cyclophosphamide to steroid treatment as originally reported by Fauci *et al*. 1979; (Fauci *et al*. 1983). The only controlled study in patients with classical polyarteritis nodosa confirmed this (Guillevin *et al*. 1991). Before the use of steroids and immunosuppressive therapy the 1-year survival rate in Wegener's granulomatosis was 20 per cent and the 2-year survival 7 per cent (Walton 1958). In classical polyarteritis nodosa the overall 5-year survival was less than 15 per cent (Frohnert and Sheps 1967; Leib *et al*. 1979). There are no historical data on the treatment of microscopic polyangiitis. Most published clinical series are relatively recent and include a large proportion of patients with renal failure. Recent studies show that over 80 per cent of patients with Wegener's granulomatosis and classical polyarteritis nodosa and over 70 per cent of patients with microscopic polyangiitis now survive for more than 5 years (reviewed by D'Amico and Sinico 1990).

Treatment of these vasculitides is essentially the same; the sole exception to this is antiviral treatment in polyarteritis nodosa induced by hepatitis B. The aims of treatment of systemic vasculitis are first to induce remission of active disease, thereby limiting the consequences of vascular injury so as to improve survival and limit disease-related morbidity. Second, to maintain remission thereby preventing organ damage from relapses. Third, to limit the consequences of the toxicity of treatment with cyclophosphamide and steroids. Now that most patients with a systemic vasculitis survive their acute illness this is an important additional aim. This makes choice of remission maintenance regimes a key issue (see Box 1).

Box 1 Algorithm for the treatment of systemic vasculitis



Treatment of classical polyarteritis nodosa and microscopic polyangiitis

There is still uncertainty as to how much cyclophosphamide and prednisolone should be given, whether they should be given continuously and orally, or as intermittent intravenous pulses, and for how long treatment should be given.

Continuous oral prednisolone and cyclophosphamide

This has been used in most studies (Fauci *et al*. 1979; Fauci *et al*. 1983; Adu *et al*. 1987). Prednisolone is used at an initial dose of 0.6 to 1 mg/kg per day and the dose tapered to around 0.15 mg/kg per day at 4 to 6 months. Cyclophosphamide is started at a dose of 1.5 to 2 mg/kg per day and the dose adjusted to avoid a leucocyte count of less than 4×10^{9}/l.

Intermittent bolus prednisolone and cyclophosphamide

Our current intermittent regime, initially intravenous and then switched to oral pulses, has been described in detail elsewhere (Bacon 1987; Bacon *et al*. 1992) and is summarized in the appendix. Following evidence of the benefits of this regime in rheumatoid vasculitis (Scott and Bacon 1984) and lupus nephritis (Austin *et al*. 1986), we have successfully used a similar regime in patients with Wegener's granulomatosis, classical polyarteritis nodosa, and microscopic polyangiitis (Bacon *et al*. 1992; Adu *et al*. 1993). Previous studies in lupus nephritis showed that pulse cyclophosphamide was as effective in improving and stabilizing renal function and was associated with a lower incidence of cyclophosphamide toxicity than the continuous oral regime (Austin *et al*. 1986). Haemorrhagic cystitis or malignancy was not seen in patients treated with pulse cyclophosphamide and amenorrhoea appeared to be less common in patients treated with pulse cyclophosphamide compared with those receiving continuous oral cyclophosphamide. Uncontrolled studies of pulse cyclophosphamide in systemic vasculitis have yielded contradictory findings. Falk *et al*. (1990b) showed that monthly intravenous cyclophosphamide plus continuous oral prednisolone for 6 months was as effective as continuous oral cyclophosphamide plus prednisolone in patients with ANCA-positive glomerulonephritis and systemic vasculitis in terms of renal improvement and patient survival. Likewise, Haubitz *et al*. (1991) found that all eight patients with Wegener's granulomatosis treated with monthly intravenous cyclophosphamide and oral prednisolone went into remission and that drug toxicity was less than in 15 patients treated with continuous oral cyclophosphamide and prednisolone. By contrast Hoffman *et al*. (1990) found that only 2 out of 11 (18 per cent) patients with Wegener's granulomatosis treated with monthly pulse cyclophosphamide and oral prednisolone achieved sustained remission. We use a pulse regime with only a 2-week gap between the initial three doses. In our ongoing controlled study, we found that pulse cyclophosphamide and steroids were equally as effective in inducing remission (Adu *et al*. 1993) and improving renal function as continuous cyclophosphamide and steroids. With both treatment regimes survival was comparable, with 90 per cent of patients with a systemic vasculitis surviving their first year of treatment. It is unlikely that this can be improved significantly. Relapses in the short term were comparable in the two treatment regimens.

Dose adjustments for renal failure and age

The metabolites of cyclophosphamide are excreted by the kidneys and accumulate in renal failure (Juma *et al*. 1981). In our experience this can lead to marrow suppression (Adu *et al*. 1987). The dose of cyclophosphamide must, therefore, be reduced in patients with renal impairment. A similar dose reduction is also advised in patients over the age of 70 years as they appear to be more sensitive than younger patients to marrow suppression with cyclophosphamide.

Duration of treatment

There is no agreed view on how long treatment should be continued. The duration of treatment is important as cyclophosphamide and steroids have substantial side-effects that have become more important as more patients survive their acute disease. On the one hand an inadequate duration of treatment might be accompanied by relapses, organ damage from vasculitis, and the consequences of this; and on the other hand overtreatment leads to a high drug toxicity. With the improved survival with current treatment, morbidity from haemorrhagic cystitis, bladder carcinoma, and lymphoma has become a major clinical problem. In Wegener's granulomatosis Fauci *et al*. (1983) recommended that treatment with cyclophosphamide should be continued for a year after remission is induced, after which the dose of cyclophosphamide and prednisolone should be tapered over several months. In patients on continuous oral prednisolone and cyclophosphamide we currently change from cyclophosphamide to azathioprine when remission is induced, usually after 4 to 6 months of treatment. Azathioprine and low-dose prednisolone is then continued with a low-dose alternative-day regime. Fuiano *et al*. (1988) also recommended keeping the patient on long-term prednisolone and azathioprine once cyclophosphamide has been discontinued. With the intermittent prednisolone and cyclophosphamide regime, cyclophosphamide is currently discontinued at 18 months and the patient changed to azathioprine and low-dose alternate-day prednisolone. The total dose of cyclophosphamide with the intermittent regime is considerably lower than with the continuous regime but we are currently studying shorter pulse regimes.

Intravenous methylprednisolone

This has been used in patients with active disease, especially when there is a crescentic glomerulonephritis. There is no compelling evidence that this confers any additional benefits when given in conjunction with continuous oral prednisolone and cyclophosphamide.

Plasmapheresis

Data from the Hammersmith Hospital suggest that plasma exchange may be of benefit in patients with a vasculitis who are also dialysis dependent (Hind *et al*. 1983; Pusey *et al*. 1991). We reserve the use of plasma exchange for patients with severe vasculitis, unresponsive to cytotoxics and steroids, or for patients with diffuse pulmonary haemorrhage or a diffuse crescentic glomerulonephritis.

Pooled, human intravenous immunoglobulin

Jayne *et al*. (1990) have suggested a beneficial effect of intravenous, pooled, normal human immunoglobulin in patients with ANCA-positive vasculitis. *In vitro* studies have shown that intravenous immunoglobulin contains anti-idiotypic antibodies to ANCA and AECA, capable of inhibiting the binding of these autoantibodies to their autoantigens. The possible mechanisms by which intravenous immunoglobulin works are still unclear. The treatment must be regarded as experimental until data from controlled studies are available.

Anti-T-lymphocyte antibodies

In a recent case report, Mathieson *et al*. (1990) successfully used a genetically engineered monoclonal antibody, Campath-1H (anti-CDw52) directed against

lymphoid cells and monocytes specifically, together with an anti-CD4 antibody against T-helper cells, to treat one patient with intractable systemic vasculitis that was ANCA negative. A subsequent study confirmed and extended these observations to four patients (Lockwood *et al*. 1993). This treatment may be of use in systemic vasculitis that is resistant to conventional treatment.

Other treatments

There are anecdotal reports that cyclosporin may be of benefit in patients with Wegener's granulomatosis resistant to cyclophosphamide (Borleffs *et al*. 1987; Gremmel *et al*. 1988) but these have not been confirmed. There appears to be no good reason at present for using this drug to treat patients with a vasculitis in preference to cyclophosphamide.

Treatment of Churg–Strauss syndrome

Lanham *et al*. (1984) found that steroids alone were beneficial, although 4 of their 16 cases were given cytotoxic agents, and 3 had additional plasma exchange. Chumbley *et al*. (1977) also found that steroids alone were useful, although the 5-year survival was only 62 per cent, suggesting that the long-term prognosis might have been improved by the addition of cytotoxic agents. In patients with renal involvement, additional cyclophosphamide has been reported to be of benefit (Clutterbuck *et al*. 1990). We and others (Guillevin *et al*. 1991) routinely add cyclophosphamide to steroids if renal or neurological disease occurs.

Side-effects of treatment

A prominent complication of treatment with steroids and cyclophosphamide is sepsis and in our earlier study this was a major cause of death (Adu *et al*. 1987). We routinely use prophylactic oral amphotericin to reduce the risk of oral candidiasis and an H_2-receptor antagonist to reduce peptic ulceration. Cystitis is a well-recognized complication of cyclophosphamide therapy with an incidence varying from 15 to 34 per cent (Stillwell and Benson 1988; Fauci *et al*. 1983). We have not seen this complication in patients on intermittent intravenous cyclophosphamide. We routinely give Mesna. By binding the acrolein metabolites of cyclophosphamide, Mesna reduces the urothelial toxicity of cyclophosphamide (Bryant *et al*. 1980). With prolonged treatment this toxicity can lead to the development of bladder tumours (Stillwell *et al*. 1988). Other malignancies associated with long-term use of cyclophosphamide include leukaemia and lymphoma (Green *et al*. 1986). Cyclophosphamide may also lead to the development of azoospermia (Fairley *et al*. 1972) and ovarian dysfunction (Miller *et al*. 1971) and, when indicated, storage of gametes may be advised. Rarely cyclophosphamide can lead to the development of an interstitial pulmonary fibrosis (Cooper *et al*. 1986).

Consequences of drug toxicity

In their follow-up report of patients with Wegener's granulomatosis treated with long-term cyclophosphamide, Hoffman *et al*. (1992) found that drug-related toxicity had become a major cause of morbidity, being found in 42 per cent of patients. On long-term follow-up, 43 per cent of their patients developed cyclophosphamide-induced haemorrhagic cystitis, 2.8 per cent a bladder cancer, 2 per cent lymphomas, 2 per cent myelodysplasia, and 57 per cent of the women ovarian failure. Overall there was a 2.4-fold increase in malignancies, a 33-fold excess risk of bladder cancer and an 11-fold excess risk of lymphoma compared with age-matched controls. Despite this very significant treatment toxicity from cyclophosphamide, the patients had active disease for 54 per cent of the total time of follow-up. A clear conclusion of these data and our own is that whilst cyclophosphamide is effective in inducing remission it appears to be relatively ineffective in maintaining remission. As more patients survive their acute illness this burden of serious drug-related toxicity becomes unacceptably high, prompting a search for safer and at least equally effective ways of maintaining remission. There are now real anxieties about the appropriateness of using long-term oral cyclophosphamide in patients with a systemic vasculitis.

Monitoring disease activity in systemic vasculitis

Clinical disease activity

Accurate assesment of disease activity makes it possible to differentiate disease activity from the effects of chronic damage with dysfunction and infection. This is important as a guide to increasing or reducing the doses of potentially toxic drugs. Several disease activity scores have been developed. These include the Kallenberg score (Kallenberg *et al*. 1990), the vasculitis activity index (Olsen *et al*. 1992), and the Birmingham vasculitis activity score (**BVAS**) (Luqmani *et al*. 1994). In our study there was reasonable correlation between these three scores. In clinical practice the BVAS score appeared to provide a precise assessment of organ involvement in vasculitis compared with a physician's global assessment (Luqmani *et al*. 1994). For longer-term follow-up, assessment of damage (accumulating, non-healing scars) and the patient's functional status are also important. In the European Community collaborative studies, a protocol combining all these aspects has been devised (Bacon *et al*. 1995).

ANCA and disease activity

There is now good evidence that ANCA titres correlate with disease activity, tending to fall in remission and rise with relapse (Cohen Tervaert *et al*. 1989; Nolle *et al*. 1989). Sequential titres of antibodies to proteinase 3 and myeloperoxidase are useful in monitoring disease activity and predicting relapses but do not always provide correct predictions (Jayne *et al*. 1995) and thus should be used with other markers of activity. The correlation with disease activity is not sufficiently close (Kerr *et al*. 1993) to justify an increase in immunosuppressant therapy with rises in ANCA titre as has been suggested (Cohen Tervaert *et al*. 1989). Some of our patients with a positive ANCA or a rise in ANCA titre have remained free of active disease for several years.

Non-specific markers of disease activity

C-reactive protein

C-reactive protein is an acute phase protein and several studies have reported that the serum levels correlate closely with disease activity in patients with Wegener's granulomatosis and microscopic polyangiitis (Hind *et al*. 1984). The serum levels of C-reactive protein rise with sepsis and both this and factor VIII-related antigen may be persistently elevated in infarction. This can limit its usefulness in monitoring disease activity in vasculitis. In addition they do not predict relapse.

Factor VIII-related antigen

Factor VIII-related antigen is a glycoprotein that is present in endothelial cells and megakaryocytes and synthesized by both cell types. It plays a role in platelet adhesion to the subendothelium, and serum levels are raised in endothelial injury. Recent studies have reported that serum levels of factor VIII-related antigen are raised in patients with vasculitis (Woolf *et al*. 1987). Our own observations are that the serum levels of this antigen rise after the onset of vascular injury and that this rise persists for several months after clinical improvement. This limits its usefulness as a marker of disease activity.

Circulating soluble adhesion molecules

Cell adhesion molecules circulate in a soluble form from proteolytic cleavage of the membrane-bound CAMs. Significantly raised serum levels of sICAM-1 and sE-selectin are found in patients with active Wegener's granulomatosis and microscopic polyangiitis and these probably reflect the extent of endothelial activation and injury (Pall *et al*. 1994). It is as yet unclear whether this is going to be useful in predicting disease activity.

Outcome

Short-term outcome

Survival has improved in all of the vasculitides and in excess of 80 per cent of patients now survive their acute illness. Although our current treatment regimes are effective in inducing remission, current strategies for maintaining remission are inadequate. We have not achieved disease- and treatment-free remission for our patients.

Relapses

With an increasing proportion of patients surviving their acute illness, it has become important to define the risks and consequences of relapses, and to balance this against the potential toxicity of continued treatment with cyclophosphamide and prednisolone. The study of Gordon *et al*. (1993) reported a high rate of relapses of up to 70 per cent over a 10-year period in patients with Wegener's granulomatosis, classical polyarteritis, and microscopic polyangiitis. In that study the tendency of the survival curves to flatten out with time (Fig. 8) contrasted markedly with the cumulative relapse rate with time (Fig. 9). Relapses occurred for up to 68 months and there was no suggestion from the actuarial curves that the rate at which relapses are occurring decreased with time. There have been suggestions that classical polyarteritis nodosa may be a self-limiting disease which tended not to recur once remission was induced. In the study of Fauci *et al*. (1979) no relapses occurred whilst patients were on cyclophosphamide and steroids. Leib *et al*. (1979), however, found that 12 per cent of patients ran a chronic course and Gordon *et al*. (1993) found that 42 per cent of patients with classical polyarteritis nodosa relapsed. Few series report on relapses in patients with microscopic polyangiitis. Savage *et al*. (1985) reported that 12 out of 33 patients with microscopic polyangiitis had 19 relapses. Serra *et al*. (1984) described 18 patients with microscopic polyangiitis who had a chronically active course which they described as a 'smouldering' vasculitis. These patients had a worse prognosis than long-term survivors with 'stable' inactive vasculitis. In the study of Gordon *et al*. (1993) relapses in microscopic polyangiitis occurred least frequently of all the vasculitides at a rate of 25.3 per cent and this compares with the rate of 36.4 per cent reported by Savage *et al*. (1985) 10 years earlier. The median time to relapse was shorter in microscopic polyangiitis (24 months) than in classical polyarteritis nodosa (33 months). A significant proportion of patients who relapsed, microscopic polyangiitis (33 per cent) and classical polyarteritis nodosa (40 per cent), did so once all treatment had been discontinued, although patients with microscopic polyangiitis had more relapses whilst on treatment albeit in a reducing dose.

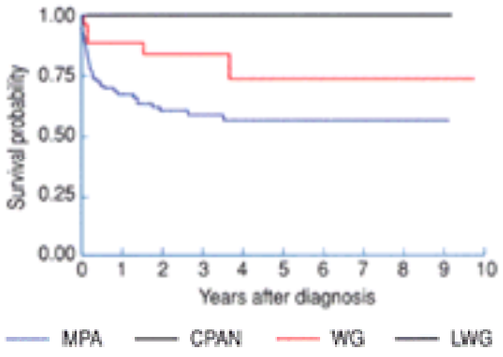


Fig. 8 Kaplan-Meier curve of the probability of survival of patients with microscopic polyangiitis (MPA), classical polyarteritis (CPAN), Wegener's granulomatosis (WG), and limited Wegener's granulomatosis (LWG). (From Gordon *et al*. 1993, with permission.)

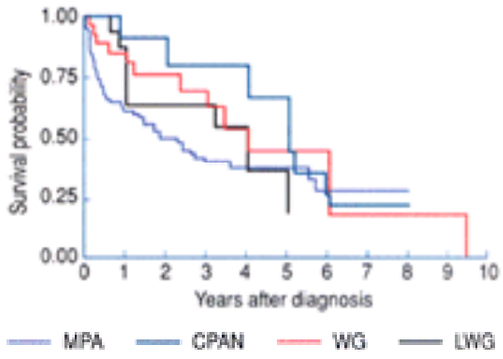


Fig. 9 Kaplan-Meier curve of the probability of surviving in terms of remaining free of relapse or of dying in patients with microscopic polyangiitis (MPA), classical polyarteritis (CPAN), Wegener's granulomatosis (WG), and limited Wegener's granulomatosis (LWG). (From Gordon *et al*. 1993, with permission.)

The clinical pattern of relapse does not necessarily mimic the original presentation, in that entirely new organs could be involved at relapse. The clinical features at relapse were in general less severe than at initial presentation in patients with microscopic polyangiitis who had significantly less systemic, renal, and pulmonary symptoms probably because they were under close supervision. Most patients at relapse had a rash and arthralgia. Patients with classical polyarteritis nodosa had more systemic symptoms at relapse. While relapse was generally a mild, non-systemic phenomenon, it may carry serious consequences including the development of renal failure and death. The vasculitides are not curable with our current treatment strategies. The implications of this are that these patients require long-term follow-up and possibly treatment.

One group of patients who had a low relapse rate were patients with microscopic polyangiitis who were on chronic dialysis and it seems either that patients with endstage renal failure have 'burned out' their disease or that renal failure is in itself immunosuppressive.

Whilst current drug treatments for inducing remission seem fairly successful it is clear that our regimes for maintaining remission are inadequate. The likelihood that relapse is a consequence of the way in which treatment modifies these diseases means that any future trial considering the efficacy of different treatments should measure not only survival as an endpoint, but also the rate of relapse and toxicity as a means of assessing the impact of treatment on patients with these diseases. There is clearly a need for prolonged follow-up in patients with a vasculitis as relapses can occur unpredictably after many years of remission.

Long-term prognosis (Fig. 8)

Overall the prognosis in patients with a vasculitis has improved over the past few years. Nevertheless, these disorders still carry a substantial mortality and morbidity.

Classical polyarteritis nodosa

The 5-year survival in classical polyarteritis nodosa ranges from 63 per cent (Guillevin *et al*. 1988) to 80 per cent (Leib *et al*. 1979) and in our own studies was 100 per cent (Gordon *et al*. 1993). Adverse prognostic factors include increasing age (Sack *et al*. 1975; Leib *et al*. 1979; Guillevin *et al*. 1988), renal failure, and gastrointestinal disease (Guillevin *et al*. 1988).

Microscopic polyangiitis

The overall prognosis in microscopic polyangiitis is much worse and in our series showed a 40 per cent mortality over a decade, largely related to renal failure often associated with delayed referrals. In other studies the 5-year survival in patients with microscopic polyangiitis ranges from 38 to 80 per cent (reviewed by D'Amico and Sinico 1990). Adverse prognostic factors include age (Adu *et al*. 1987; Fuiano *et al*. 1988) and the severity of renal failure before institution of therapy (Serra *et al*. 1984; Adu *et al*. 1987).

Churg–Strauss syndrome

The overall 5-year survival in Churg–Strauss syndrome is reported to be 62 per cent (Chumbley *et al*. 1977).

Conclusion

Progress in therapy will be linked to an increased understanding of the aetiology of disease and pathogenetic mechanisms, which probably vary from one syndrome to another despite almost identical histopathology of necrotizing vasculitis. Only when these have been defined will more specific, and hopefully less toxic, therapy be available. Nevertheless, the prognosis in vasculitis has improved enormously, so that aggressive therapy is rewarding. At present the main aim must be earlier diagnosis, since therapy is both more successful and less toxic before advanced renal impairment develops; together with the development of less-toxic long-term maintenance regimes.

Appendix

Protocol for the treatment of vasculitis with intermittent pulses of cyclophosphamide and prednisolone

Table 2 sets out the protocol for the treatment of vasculitis with intermittent pulses of cyclophosphamide and prednisolone, and should be used in conjunction with the following notes.

The induction regimen is used for up to 6 months. It can be shorter if clinical remission is achieved early but is continued for at least 3 months of treatment.

The remission regimen commences immediately after the end of the induction regimen, and is continued for up to 1 year from the start of induction.

The consolidation regimen is normally for a further 6 months, but may be continued for longer if disease activity is still apparent.

The maintenance phase is continued probably indefinitely.

In patients who are already being treated with immunosuppressive agents, the pulse therapy should not be started until 2 weeks has elapsed off cytotoxics to avoid severe marrow toxicity. All patients should be prescribed ordinary (not enteric coated) prednisolone in view of potential problems of variable absorption of active compound. During the pulse week, all patients are prescribed ranitidine, 150 mg at night, and amphotericin lozenges, 10 mg four times daily as prophylaxis. Further use is at the discretion of the physician. If patients develop infections during the course of treatment, then immunosuppression will be stopped until the infection has been adequately treated. If there is evidence of renal impairment or marrow suppression, or the patient is over the age of 70 years, the dose of cyclophosphamide and steroid will be adjusted accordingly (see dose adjustments below).

Phase	Time (week)	Pulse number	Route	Dose	
				Cyclophosphamide	Prednisolone
Induction	0, 2, 4,	1–3	IV	15 mg/kg × 1	10 mg/kg
	7, 10, 13, 17, 21, 25	4–9	Oral	5 mg/kg × 3	1.3 mg/kg × 3
Remission	30, 35, 40, 48, 52	10–14	Oral	5 mg/kg × 3	1.3 mg/kg × 3
Consolidation	58, 64, 70, 78	15–21	Oral	5 mg/kg × 3	1.3 mg/kg × 3
Maintenance	Oral prednisolone, 0.15 mg/kg alternate days				

Table 2 Protocol for the treatment of vasculitis

Tests before each pulse

Full blood count should be measured before each pulse is given, and the values should be in the normal range. If they are not, the pulse should be deferred until they are. Creatinine should be measured before each pulse, and if abnormal the dose of the pulse should be adjusted according to Table 3.

Serum creatinine	Cyclophosphamide dose	Prednisolone dose
< 150	15 mg/kg	10 mg/kg
150–250	10 mg/kg	10 mg/kg
251–500	7.5 mg/kg	10 mg/kg
> 500	5 mg/kg	7 mg/kg

Table 3 Bolus therapy

Escalation protocol

At any stage as decided by the clinician, the following may be given:

1. Additional bolus methylprednisolone intravenously, 1 g/day × 3 days;
2. Additional bolus cyclophosphamide intravenously, 5 mg/kg.day × 3 days;
3. Plasma exchange (optional).

Mesna is to be given in three oral doses totalling 75 per cent of the dose of cyclophosphamide for every dose of intravenous cyclophosphamide.

Dose adjustments are devised for the following reasons:

1. Maximum doses: the maximum bolus dose of cyclophosphamide, regardless of weight, will be 1000 mg. The maximum bolus dose of prednisolone regardless of weight will be 1000 mg.
2. Cytopenia prior to bolus therapy: delay bolus until count restored to above lower limit of normal (white cell count > 3.5 or neutrophil count > 2.0, platelets > 140). If cytopenia recurs, reduce cyclophosphamide bolus by 25 per cent.
3. Renal failure on bolus therapy: reduce as shown in Table 3.
4. Bolus therapy in the elderly (defined as over the age of 70 years): reduce the cyclophosphamide dose to 10 mg/kg, steroid dose unchanged.

Chapter References

Adu, D., Howie, A.J., Scott, D.G.I., Bacon, P.A., McGonigle, R.J.S., and Michael, J. (1987). Polyarteritis and the kidney. *Quarterly Journal of Medicine*, **62**, 221–37.

Adu, D., Pall, A., Luqmani, R.A., *et al*. (1993). Controlled trial of treatment of vasculitis. *Clinical and Experimental Immunology*, **93** (Suppl.1), 38.

Arkin, A. (1930). A clinical and pathological study of periarteritis nodosa. *American Journal of Pathology*, **6**, 401–26.

Austin, H.A., Klippel, J.H., Balow J.E., *et al*. (1986). Therapy of lupus nephritis: controlled trial of prednisolone and cytotoxic drugs. *New England Journal of Medicine*, **314**, 614–19

Bacon, P.A. (1987). Vasculitis—clinical aspects and therapy. *Acta Medica Scandinavica*, **715** (Suppl.), 157–63.

Bacon, P.A., Luqmani, R.A., Scott, D.G.I., and Adu, D. (1992). Immunopharmacology of vasculitic syndrome. In *Immunopharmacology in autoimmune disease and transplants* (ed. H.E. Rugstad), pp. 273–89. Plenum Publishing Corporation, New York.

Bacon, P.A., Moots, R.J., Exley, A., Luqmani, R.A., and Rasmussen, N. (1995). Vital assessment of vasculitis. *Clinical and Experimental Rheumatology*, **13**, 275–8.

Borleffs, J.C., Derksen, R.H., and Hene, R.J. (1987). Treatment of Wegener's granulomatosis with cyclosporin. (Letter.) *Annals of the Rheumatic Diseases*, **46**, 175.

Brouwer, E., Huitema, M.G., Mulder, A.H.L., *et al*. (1994). Neutrophil activation in vitro and in vivo in Wegener's geanulomatosis. *Kidney International*, **45**, 1121–311.

Bruijn, J.A. and Dinklo, N.J.C.M. (1993). Distinct patterns of expression of intercellular adhesion molecule-1, vascular cell adhesion molecule-1, and endothelial-leukocyte adhesion molecule-1 in renal disease. *Laboratory Investigation*, **69**, 329–35.

Bryant, B.M., Jarman, M., Ford, H.T., and Smith, I.E. (1980). Prevention of isophosphamide-induced urothelial toxicity with 2-mercaptoethane-sulphanate sodium (mesnum) in patients with advanced carcinoma. *Lancet*, **ii**, 657–9.

Chumbley, L.C., Harrison, R.A., and DeRemee, R.A. (1977). Allergic granulomatous angiitis (Churg–Strauss syndrome). Report and analysis of 30 cases. *Mayo Clinic Proceedings*, **52**, 477–84.

Churg, J. and Strauss, L. (1951). Allergic granulomatosis, allergic angiitis, and periarteritis nodosa. *American Journal of Pathology*, **27**, 277–301.

Clutterbuck, E.J., Evans, D.J., and Pusey, C.D. (1990). Renal involvement in Churg–Strauss syndrome. *Nephrology Dialysis Transplantation*, **5**, 161–7.

Cohen, R.D., Conn, D.L., and Ilstrup, D.M. (1980). Clinical features, prognosis and response to treatment in polyarteritis. *Mayo Clinic Proceedings*, **55**, 146–55.

Cohen Tervaert, J.W., Van Der Wonde, F.J., Fauci, A.S., *et al*. (1989). Association between active Wegener's granulomatosis and anti-cytoplasmic antibodies. *Archives of Internal Medicine*, **149**, 2461–5.

Cohen Tervaert, J.W., Limburg, P.C., Elema, J.D., *et al*. (1991). Detection of autoantibodies against myeloid lysosomal enzymes: a useful adjunct to classification of patients with biopsy-proven necrotizing arteritis. *American Journal of Medicine*, **91**, 59–66.

Cooper, J.A.D., White, D.A., and Matthay, R.A. (1986). Drug-induced pulmonary disease. 1. Cytotoxic drugs. *American Review of Respiratory Diseases*, **133**, 321–40.

Coward, R.A., Handy, N.A.T., Shortland, J.S., and Brown, C.B. (1986). Renal microscopic polyarteritis: a treatable condition. *Nephrology Dialysis Transplantation*, **1**, 31–7.

Cupps, T.R. and Fauci, A.S. (1981). *The vasculitides*. Saunders, Philadelphia.

D'Agati, V., Chander, P., Mash, M., and Mancilla-Jimenez, R. (1986). Idiopathic microscopic polyarteritis nodosa: ultrastructural observations on the renal, vascular and glomerular lesions. *American Journal of Kidney Diseases*, **8**, 95–110.

D'Amico, G. and Sinico, R.G. (1990). Treatment and monitoring of systemic vasculitis. *Nephrology Dialysis Transplantation*, **1** (Suppl.), 53–7.

Davies, D.J., Moran, J.E., Niale, J.F., and Ryan, G.B. (1982). Segmental necrotising glomerulonephritis with antineutrophil antibody: possible arbovirus aetiology. *British Medical Journal*, **285**, 606.

Davson, J., Ball, J., and Platt, R. (1948). The kidney in periarteritis nodosa. *Quarterly Journal of Medicine*, **17**, 175–202.

Ewert, B.H., Jennette, J.C., and Falk, R.J. (1992). Anti-myeloperoxidase antibodies stimulate neutrophils to damage human endothelial cells. *Kidney International*, **41**, 375–83.

Fairley, K.F., Barrie, J.U., and Johnson, W. (1972). Sterility and testicular atrophy related to cyclophosphamide therapy. *Lancet*, **i**, 568–9.

Falk, R.L. and Jennette, J.C. (1988). Antineutrophil cytoplasmic auto-antibodies with specificity for myeloperoxidase in patients with systemic vasculitis and idiopathic necrotising and crescentic glomerulonephritis. *New England Journal of Medicine*, **318**, 1651–7.

Falk, R.J., Terrell, R.S., Charles, L.A., and Jennette, J.C. (1990 *a*). Anti-neutrophil cytoplasmic autoantibodies induce neutrophils to degranulate and produce oxygen radicals in vitro. *Proceedings of the National Academy of Sciences USA*, **87**, 4115–19.

Falk, R.J., Hogan, S., Carey, T.S., and Jennette, C. (1990 *b*). Clinical course of anti-neutrophil cytoplasmic autoantibody-associated glomerulonephritis and systemic vasculitis. *Annals of Internal Medicine*, **113**, 656–63.

Fauci, A.S. and Wolff, S.M. (1973). Wegener's granulomatosis: studies in eighteen patients and a review of the literature. *Medicine (Baltimore)*, **52**, 535–61.

Fauci, A.S., Katz, P., Haynes, B.F., and Wolff, S.M. (1979). Cyclophosphamide therapy of severe systemic necrotizing vasculitis. *New England Journal of Medicine*, **301**, 235–8.

Fauci, A.S., Haynes, B.F., Katz, P., and Wolff, S.M. (1983). Wegener's granulomatosis: prospective clinical and therapeutic experience with 85 patients for 21 years. *Annals of Internal Medicine*, **98**, 76–85.

Ferrari, E. (1903). Ueber Polyarteritis acuta nodosa (sogenannte Periarteritis nodosa) und ihre Beiziehungen zur Polymyositis und Polyneuritis acuta. *Beitraege zur Pathologischen Anatomie und Allgemeinen Pathologie*, **34**, 1–25.

Frampton, G., Jayne, D.R.W., Perry, G.J., Lockwood, C.M., and Cameron, J.S. (1990). Auto-antibodies to endothelial cells and neutrophil cytoplasmic antigens in systemic vasculitis. *Clinical and Experimental Immunology*, **82**, 227–32.

Frohnert, P.P. and Sheps, S.G. (1967). Long term follow up study of periarteritis nodosa. *American Journal of Medicine*, **43**, 8–14.

Fuiano, G., Cameron, J.S., Raftery, M., Hartley, B.H., Williams, D.G., and Ogg, C.S. (1988). Improved prognosis of microscopic polyarteritis in recent years. *Nephrology Dialysis Transplantation*, **3**, 383–91.

Goldschmeding, R., Van Der Shoot, C.E., and ten Bokkel Huinink, D. (1989). Wegener's granulomatosis autoantibodies identify a novel diisopropylfluorophosphate-binding protein in the lysosomes of normal human neutrophils. *Journal of Clinical Investigation*, **84**, 1577–87.

Gordon, M., Luqmani, R.A., Adu, D., *et al*. (1993). Relapses in patients with a systemic vasculitis. *Quarterly Journal of Medicine*, **86**, 779–89.

Green, M.H., Harris, E.L., Gershenson, D.M., and Malkasian, G.D. (1986). Melphalan may be a more potent leukemogen than cyclophosphamide. *Annals of Internal Medicine*, **105**, 360–7.

Gremmel, F., Druml, W., Schmidt, P., and Graninger, W. (1988). Cyclosporin in Wegener's granulomatosis. *Annals of Internal Medicine*, **108**, 491.

Guillevin, L., Le Thi Huong Du, Godeau, P., Jais, P., and Wechsler, B. (1988). Clinical findings and prognosis of polyarteritis nodosa and Churg–Strauss angiitis: a study of 165 patients. *British Journal of Rheumatology*, **27**, 258–64.

Guillevin, L., Jarousse, B., Lok, C., *et al*. (1991). Longterm followup after treatment of polyarteritis nodosa and Churg–Strauss angiitis with comparison of steroids, plasma exchange and cyclophosphamide to steroids and plasma exchange. A prospective randomized trial of 71 patients. *Journal of Rheumatology*, **18**, 567–74.

Guillevin, L., Visser, H., Noel, L.H., *et al*. (1993). Antineutrophil cytoplasm antibodies in systemic polyarteritis nodosa with and without hepatitis B virus infection and Churg–Strauss syndrome. *Journal*

of Rheumatology, **20**, 1345–9.

Hall, T.B., Wadham, B., McN., and Wood, C.J. (1984). Vasculitis and glomerulonephritis—a subgroup with an antineutrophil cytoplasmic antibody. *Australian and New Zealand Journal of Medicine*, **14**, 277–8.

Haubitz, M., Frei, U., Rother, U., Brunkhorst, R., and Koch, K.M. (1991). Cyclophosphamide pulse therapy in Wegener's granulomatosis. *Nephrology Dialysis Transplantation*, **6**, 531–5.

Haworth, S.J., Savage, C.O.S., Carr, D., Hughes, J.M.B., and Rees, A.J. (1985). Pulmonary haemorrhage complicating Wegener's granulomatosis and microscopic polyarteritis. *British Medical Journal*, **290**, 1775–8.

Hind, C.R.K., Paraskevakou, H., Lockwood, C.M., Evans, D.J., Peters, D.K., and Rees, A.J. (1983). Prognosis after immunosuppression of patients with crescentic nephritis requiring dialysis. *Lancet*, **1**, 263–5.

Hind, C.R.K., Winearls, C.G., Lockwood, C.M., Rees, A.J., and Pepys, M.B. (1984). Objective monitoring or activity in Wegener's granulomatosis by measurement of C-reactive protein concentration. *Clinical Nephrology*, **21**, 341–5.

Hoffman, G.S., Leavitt, R.Y., Fleischer, T.A., Minor, J.R., and Fauci, A.S. (1990). Treatment of Wegener's granulomatosis with intermittent high-dose intravenous cyclophosphamide. *American Journal of Medicine*, **89**, 403–10.

Hoffman, G.S., Kerr, G.S., Leavitt, R.Y., Hallahan, C.W., Lebovics, R.S., and Travis, W.D. (1992). Wegener's granulomatosis: an analysis of 158 patients. *Annals of Internal Medicine*, **116**, 488–98.

Jayne, D.R.W., Davies, M.J., Fox, C.J.V., Black, C.M., and Lockwood, C.M. (1990). Treatment of systemic vasculitis with pooled intravenous immunoglobulin. *Lancet*, **337**, 1137–9.

Jayne, D.R.W., Gaskin, G., Pusey, C.D., and Lockwood, C.M. (1995). ANCA and predicting relapse in systemic vasculitis. *Quarterly Journal of Medicine*, **88**, 127–33.

Jennette, J.C., Falk, R.J., Andrassy, K., *et al*. (1994). Nomenclature of systemic vasculitides: the proposal of an international consensus conference. *Arthritis and Rheumatism*, **37**, 187–92.

Juma, F.D., Rogers, H.J., and Trounce, J.R. (1981). Effect of renal insufficiency on the pharmacokinetics of cyclophosphamide and some of its metabolites. *European Journal of Clinical Pharmacology*, **19**, 443–51.

Kallenberg, C.G.M., Cohen Tervaert, J.M., and Stegeman, C.A. (1990). Criteria for disease activity in Wegener's granulomatosis: a requirement for longtitudinal studies. *APMIS Suppl*, **19**, 37–9.

Kallenberg, C.G.M, Brouwer, E., Weening, J.J., and Cohen Tervaert, J.W. (1994). Anti-neutrophil cytoplasmic antibodies: current diagnostic and pathophysiological potential. *Kidney International*, **46**, 1–15.

Kerr, G.S., Fleisher, T.A., Hallahan, C.W., Leavitt, R.Y., Fauci, A.S., and Hoffman, G.S. (1993). Limited prognostic value of changes in antineutrophil cytoplasmic antibody titer in patients with Wegener's granulomatosis. *Arthritis and Rheumatism*, **36**, 365–71.

Kussmaul, A. and Maier, R. (1866). Über eine bisher nicht beschreibene eigenthümliche Arteriener Krankung (periarteritis nodosa): die mit Morbus Brightii und rapid fortschreitender allgemeiner Muskellahmung einhergeht. *Deutsches Archiv für Klinische Medizin*, **1**, 484–517.

Lanham, J.G., Elkon, K.B., Pusey, C.D., and Hughes, G.R.V. (1984). Systemic vasculitis with asthma and eosinophilia: a clinical approach to the Churg–Strauss syndrome. *Medicine (Baltimore)*, **63**, 65–81.

Lee, S.S., Adu, D., and Thompson, R. (1990). Anti-myeloperoxidase antibodies in systemic vasculitis. *Clinical and Experimental Immunology*, **79**, 41–6.

Leib, E.S., Restivo, C., and Paulus, H.E. (1979). Immunosuppressive and corticosteroid therapy of periarteritis nodosa. *American Journal of Medicine*, **67**, 941–7.

Lie, J.T. (1994). Nomenclature and classification of vasculitis: plus ça change, plus c'est la mme chose. *Arthritis and Rheumatism*, **37**, 181–6.

Lockwood, C.M., Thiru, S., Isaacs, J.D., Hale, G., and Waldmann, H. (1993). Long term remission of intractable systemic vasculitis with monoclonal antibody therapy. *Lancet*, **341**, 1620–2.

Luqmani, R.A., Bacon, P.A., Moots, R.J., *et al*. (1994). Birmingham Vasculitis Activity Score in systemic necrotizing vasculitis. *Quarterly Journal of Medicine*, **87**, 671–8.

Masi, A.T., Hunder, G.G., Lie, J.T., Michel, B.A., Bloch, D.A., and Arend, W.P. (1990). The American College of Rheumatology 1990 criteria for the classification of Churg–Strauss syndrome (allergic granulomatosis and angiitis). *Arthritis and Rheumatism*, **33**, 1094–110.

Mathieson, P.W., Cobbold, S.P., Hale, G., *et al*. (1990). Monoclonal antibody therapy in systemic vasculitis. *New England Journal of Medicine*, **323**, 250–4.

Miller, J.J., III, Williams, G.F., and Leissring, J.C. (1971). Multiple late complications of therapy with cyclophosphamide including ovarian destruction. *American Journal of Medicine*, **50**, 530–5.

Minkowitz, G., Smoller, B.R., and McNutt, N.S. (1991). Benign cutaneous polyarteritis nodosa. *Archives of Dermatology*, **127**, 1520–3.

Nolle, B., Specks, V., Ludeman, J., Rohrbach, M.S., De Remee, R.A., and Gross, W.L. (1989). Anticytoplasmic autoantibodies: their immunodiagnostic value in Wegener's granulomatosis. *Annals of Internal Medicine*, **111**, 28–40.

Noronha, I.L., Krüger, C., Andrassy, K., Ritz, E., and Waldherr, R. (1993). *In situ* production of TNFa, IL-1b and IL-2R in ANCA-positive glomerulonephritis. *Kidney International*, **43**, 682.

Novack, S.N. and Pearson, C.M. (1971). Cyclophosphamide therapy in Wegener's granulomatosis. *New England Journal of Medicine*, **285**, 1493–6.

Olsen, T.L., Whiting O'Keefe, Q.E., and Hellman, D.B. (1992). Validity and precision of a vasculitis activity index (VAI). *Arthritis and Rheumatology*, **35**, S164.

Pall, A.A., Adu, D., Drayson, M., Taylor, C.M., Richards, N.T., and Michael, J. (1994). Circulating soluble adhesion molecules in systemic vasculitis. *Nephrology Dialysis Transplantation*, **9**, 770–4.

Pall, A.A., Howie, A.J., Adu, D., *et al*. (1996). Glomerular VCAM-1 expression in renal vasculitis. *Journal of Clinical Pathology*, 49, 238–42.

Porges, A.J., Redecha, P.B., Kimberly, W.T., Csernok, E., Gross, W.L., and Kimberly, R.T. (1994). Anti-neutrophil cytoplasmic antibodies engage and activate human neutrophils via FcgR11a. *Journal of Immunology*, **153**, 1271–80.

Pusey, C.D., Rees, A.J., Evans, D.J., Peters, D.K., and Lockwood, C.M. (1991). Plasma exchange in focal necrotizing glomerulonephritis without anti-GBM antibodies. *Kidney International*, **40**, 757–63.

Ronco, P., Verroust, P., Mignon, F., *et al*. (1983). Immunopathologic studies of polyarteritis nodosa and Wegener's granulomatosis: a report of 43 patients with 51 renal biopsies. *Quarterly Journal of Medicine*, **52**, 212–23.

Sack, M., Cassidy, J.T., and Bole, G.G. (1975). Prognostic factors in polyarteritis. *Journal of Rheumatology*, **2**, 411–20.

Savage, C.O.S., Winearls, C.G., Evans, D.J., Rees, A.J., and Lockwood, S.M. (1985). Microscopic polyarteritis: presentation, pathology and prognosis. *Quarterly Journal of Medicine*, **56**, 467–83.

Savage, C.O.S., Jones, S., Winearls, C.G., Marshall, P.D., and Lockwood, C.M. (1987). Prospective study of radio-immunoassay for antibodies against neutrophil cytoplasm in diagnosis of systemic vasculitis. *Lancet*, **i**, 1389–93.

Savage, C.O.S., Pottinger, B.E., Gaskin, G., Pusey, C.D., and Pearson, J.D. (1992). Autoantibodies developing to myeloperoxidase and proteinase 3 in systemic vasculitis stimulate neutrophil cytotoxicity towards cultured endothelial cells. *American Journal of Pathology*, **141**, 335–42.

Scott, D.G.I. and Bacon, P.A. (1984). Intravenous cyclophosphamide plus methylprednisolone in the treatment of systemic rheumatoid vasculitis. *American Journal of Medicine*, **76**, 377–84.

Scott, D.G.I., Bacon, P.A., Elliott, P.J., Tribe, C.R., and Wallington, T.B. (1982). Systemic vasculitis in a district general hospital 1972–1980: clinical and laboratory features, classification and prognosis of 80 cases. *Quarterly Journal of Medicine*, **51**, 292–311.

Serra, A. and Cameron, J.S. (1985). Clinical and pathological aspects of renal vasculitis. *Seminars in Nephrology*, **5**, 15–33.

Serra, A., Cameron, J.S., Turner, D.R., *et al*. (1984). Vasculitis affecting the kidney: presentation, histopathology and long term outcome. *Quarterly Journal of Medicine*, **53**, 181–207.

Spargo, B.H., Seymour, A.E., and Ordenez, N.G. (1980). Vasculitis. In *Renal biopsy pathology with diagnostic and therapeutic implications*, pp. 205–18. Wiley, New York.

Specks, U. and De Remee, R.A. (1990). Granulomatous vasculitis: Wegener's granulomatosis and Churg–Strauss syndrome in vasculitic syndromes. *Rheumatic Disease Clinics of North America*, **16**, 377–97.

Stillwell, T.J. and Benson, R.C., Jr. (1988). Cyclophosphamide-induced haemorrhage cystitis. A review of 100 patients. *Cancer*, **61**, 451–7.

Stillwell, T.J., Benson, R.C., Jr., De Remee, R.A., McDonald, T.J., and Weiland, L.H. (1988). Cyclophosphamide-induced bladder toxicity in Wegener's granulomatosis. *Arthritis and Rheumatism*, **31**, 465–70.

Travers, R.L., Allison, D.J., Brettle, R.P., and Hughes, G.R.V. (1979). Polyarteritis nodosa: a clinical and angiographic analysis of 17 cases. *Seminars in Arthritis and Rheumatism*, **8**, 184–99.

Van der Woude, F.J., Rasmussen, N., Lobatto, S., *et al*. (1985). Auto-antibodies against neutrophils and monocytes: tool for diagnosis and marker of disease activity in Wegener's granulomatosis. *Lancet*, **ii**, 425–9.

Van der Woude, F.J., Daha, M.R., and Vanes, L.A. (1989). Review: the current status of neutrophil cytoplasmic antibodies. *Clinical and Experimental Immunology*, **78**, 143–8.

Varagunam, M., Adu, D., Taylor, C.M., *et al*. (1992*a*). Endothelium myeloperoxidase-antimyeloperoxidase interaction in vasculitis. *Nephrology Dialysis Transplantation*, **7**, 1077.

Varagunam, M., Nwosu, Z., Adu, D., *et al*. (1992*b*). Little evidence for anti-endothelial cell antibodies in microscopic polyarteritis and Wegener's granulomatosis. *Nephrology Dialysis Transplantation*, **8**, 113.

Walton, E.W. (1958). Giant cell granuloma of the respiratory tract (Wegener's granulomatosis). *British Medical Journal*, **1**, 265–70.

Wegener, F. (1936). Über generalisierte, septische Gefasserkrankungen. *Verhandlungens der Deutschen Gesellschaft für Pathologie*, **29**, 202–9.

Woolf, A.D., Wakerley, J., Wallington, T.B., Scott, D.J.I., and Dieppe, P.A. (1987). Factor VIII related antigen in the assessment of vasculitis. *Annals of the Rheumatic Diseases*, **46**, 441–4.

Wuthrich, R.P. (1992). Intercellular adhesion molecules and vascular cell adhesion molecule-1 and the kidney. *Journal of the American Society of Nephology*, **3**, 1201–11.

Zeek, P.M. (1952). Perarteritis nodosa: a critical review. *American Journal of Clinical Pathology*, **22**, 777–90.

5.11.4 Small vessel vasculitis

Clive E. H. Grattan and Victoria A. Jolliffe

Introduction

Small vessel vasculitis is characterized by inflammation centred on damaged capillaries and postcapillary venules. Muscular arteries are not usually involved although, confusingly, a recent international consensus on the classification of vasculitis (Jennette *et al*. 1994) includes patterns of vasculitis involving small and medium-sized arteries (Wegener's granulomatosis, Churg–Strauss syndrome, and microscopic polyangiitis) within the spectrum of small vessel vasculitis. It is however clear that there is a range of clinical and pathological features within defined diagnostic groups depending on the organ and size of vessel affected. The cutaneous lesions of Wegener's granulomatosis, for instance, may show the typical clinical and histological features of small vessel vasculitis as part of the wider spectrum of this systemic granulomatous disorder (Daoud *et al*. 1994; Francès *et al*. 1994). Small vessel vasculitis usually presents in the skin although the microvasculature of any tissue may be affected, especially the joints and kidneys. The most widely recognized clinicopathological patterns of small vessel vasculitis are listed in Table 1. Henoch–Schönlein purpura is regarded by many as a special form of allergic vasculitis. The division into leucocytoclastic and non-leucocytoclastic patterns of vasculitis is not absolute since skin biopsies may show a spectrum of pathological changes ranging from predominantly neutrophilic to predominantly mononuclear cell infiltrates in Sjögren's syndrome, drug-induced vasculitis, urticarial vasculitis, and nodular vasculitis. There is some evidence that cellular infiltrates may evolve from neutrophilic to mononuclear as lesions mature, but in many patients the pattern of inflammation remains unchanged throughout the course of their illness.

1. Leucocytoclastic vasculitis
   - Allergic vasculitis (synonym: hypersensitivity angiitis)
     - Drugs, infections, autoimmune diseases, inflammatory diseases, malignancy, idiopathic
     - Henoch–Schönlein purpura (synonym: anaphylactoid purpura)
   - Urticarial vasculitis (synonym: hypocomplementaemic vasculitis)
   - Cryoglobulinaemia
   - Hypergammaglobulinaemic purpura (of Waldenström)
   - Erythema elevatum diutinum and granuloma faciale
2. Non-leucocytoclastic vasculitis
   - Drug-related vasculitis
   - Nodular vasculitis
   - Livedo vasculitis
   - Pityriasis lichenoides acuta and chronica

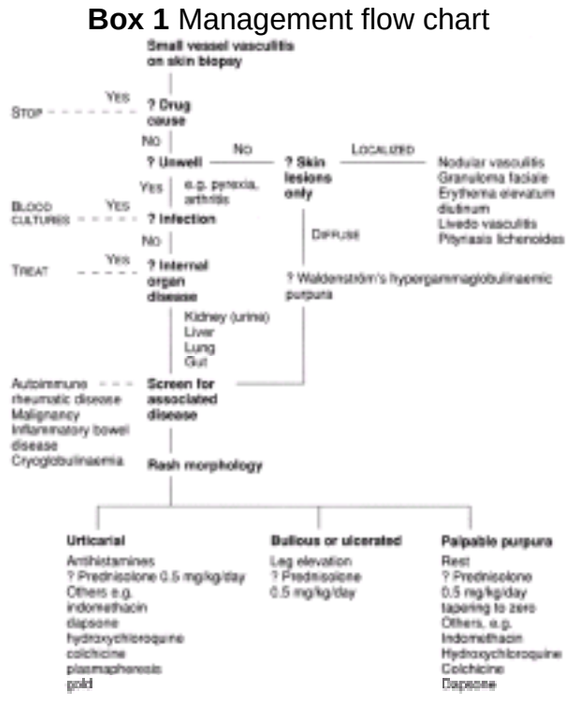
Table 1 Clinicopathological classification of small vessel vasculitis

The definition of small vessel vasculitis is open to different interpretations. Some authors restrict it to lesions with a clearly defined leucocytoclastic histology showing neutrophils within and around blood vessel walls, leucocytoclasis (fragmentation of neutrophils), fibrin in and around the damaged blood vessels, and swelling, hypertrophy, or necrosis of endothelial cells. Others will accept the presence of inflammatory cells within vessel walls and at least one of the above features as being sufficient for diagnosis. Evidence of vessel wall damage should be the key feature for diagnosis whether the predominant cell type is neutrophilic, eosinophilic, mononuclear, or mixed. Red cells may be seen in the surrounding tissue and intraluminal fibrin thrombi may form as a result of it. Of 98 cases of cutaneous vasculitis diagnosed on skin biopsy over a 5-year period in an adult population of around 400 000, 49 had definite leucocytoclastic vasculitis, 13 showed some features of it, and 12 had a predominantly lymphocytic vasculitis (unpublished data).

The clinical presentation of cutaneous leucocytoclastic vasculitis can vary considerably even though the underlying pathological changes may be indistinguishable. Conversely, the appearance of skin lesions may be similar even though the nature of the inflammatory infiltrate may differ. The full-blown picture of small vessel vasculitis with systemic features can be very similar to other patterns of vasculitis, including polyarteritis nodosa. This overlap may cause difficulty in distinguishing one form of vasculitis from another with complete certainty.

A management flow chart is given in Box 1. Although skin biopsy is essential to confirm the diagnosis of small vessel vasculitis, clinical decisions should not be delayed until the result is available in an unwell patient.

Box 1 Management flow chart



Leucocytoclastic vasculitis

Pathogenesis

Immune complex deposition appears to be important in the pathogenesis of most, if not all, forms of leucocytoclastic vasculitis. Experimental immune complex disease in animals is accompanied by leucocytoclastic vasculitis. Immune complex deposits were demonstrated by tissue immunofluorescence as early as 20 min after the onset of experimental Arthus reactions in guinea-pig skin and had disappeared by 18 h (Cream *et al*. 1971). Cochrane (1971) showed that soluble factors from sensitized rabbit basophils caused platelet clumping when reacted with antigen, this in turn caused release of vasoactive amines leading to permeability of the blood vessel and subsequent deposition of immune complexes. Immunofluoresence studies have demonstrated immunoglobulins and C3 within small blood vessels of lesional and adjacent normal skin in patients with leucocytoclastic vasculitis (Sams *et al*. 1975). Immune complexes can be demonstrated in a subendothelial location by electron microscopy and direct immunofluorescence of histamine-induced vasculitis lesions and clinically normal skin of patients with leucocytoclastic vasculitis (Braverman and Yen 1975; Gower *et al*. 1977). Serial biopsies of cold-induced urticaria in a patient with cold urticaria and vasculitis demonstrated C3 deposition within 5 min of ice application, preceding the deposition of fibrin, immunoglobulin, or obvious mast cell degranulation (Eady *et al*. 1981). An antibody raised against the terminal components of the complement pathway (C5b to C9) stained 13 of 15 lesional biopsies from patients with leucocytoclastic vasculitis (Boom *et al*. 1987) indicating that activation of the terminal complement components may contribute to local tissue damage. Antigens of candida, *Mycobacterium tuberculosis*, and streptococci have been identified in skin lesions of patients with neutrophilic vasculitis, together with IgG (Parish 1971), suggesting that bacterial antigen–antibody complexes formed *in vivo* may trigger the subsequent inflammatory events. In a series of 53 patients with cutaneous vasculitis 77 per cent showed two or more positive serological tests for circulating immune complexes (Andrews *et al*. 1979).

The following sequence of pathogenetic events may occur: (1) the endothelium of postcapillary venules becomes more permeable due to release of vasoactive factors, including histamine; (2) circulating immune complexes, formed in response to infections, foreign haptens, or self-antigens lodge beneath the endothelium; (3) local complement activation initiates adherence of neutrophils to endothelium through adhesion molecule expression; (4) they release lysosomal enzymes that damage the vessel wall and surrounding tissues; (5) newly generated C3a and C5a anaphylatoxins degranulate mast cells, releasing neutrophil and eosinophil chemotactic factors and vasoactive mediators, thereby amplifying the tissue reaction; (6) the acute response is followed by an influx of mononuclear cells and macrophages (Zax *et al*. 1990) which promote resolution of the reaction through recruitment of suppressor T lymphocytes and phagocytosis of the cellular debris.

Clinical

Clinical features, laboratory abnormalities, treatment, and prognosis are considered below.

Allergic vasculitis

Allergic vasculitis corresponds approximately to 'hypersensitivity angiitis' in Pearl Zeek's original classification of necrotizing vasculitis (Zeek 1952). The term allergic is a little contentious since it implies an immunological aetiology, which may be an oversimplification. Leucocytoclastic vasculitis is sometimes used synonymously as a clinicopathological term but this usage is better avoided because it may lead to confusion with the histological reaction pattern.

Allergic vasculitis is the commonest pattern of leucocytoclastic vasculitis in adults. It is characterized by purpuric or necrotic skin lesions, with or without systemic features. Men and women are affected equally. The onset may be from childhood to old age. Skin lesions tend to be distributed maximally over the lower legs. The buttocks, trunk, and upper extremities may also be affected. There is often a mixture of haemorrhagic papules that do not blanche on pressure (palpable purpura) (Fig. 1), erythematous or purpuric macules, urticarial plaques, vesicles, pustules, haemorrhagic bullae (Fig. 2), erosions, and ulcers. One or more systemic features occur in at least 50 per cent of patients. Arthralgia and arthritis, microscopic haematuria, abdominal symptoms (pain, nausea, diarrhoea, or bleeding), low-grade fever or malaise are the commonest. Pulmonary involvement (cough, dyspnoea, or haemoptysis) and neurological manifestations (peripheral neuropathy, benign intracranial hypertension, aseptic meningitis, and uveitis) have been reported.

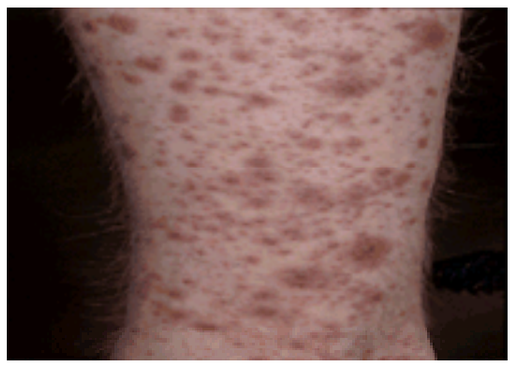


Fig. 1 Palpable purpura on the lower leg of a patient with allergic vasculitis.

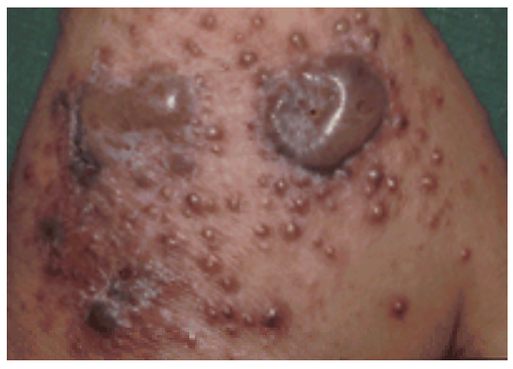


Fig. 2 Haemorrhagic blisters on the back of the hand of the same patient shown in Fig. 1.

The erythrocyte sedimentation rate (**ESR**) is usually raised, in the region of 50 mm/h, and there may be mild to moderate anaemia. Complement consumption and immune complex formation tend to reflect disease severity. Hypocomplementaemia was found in patients showing neutrophilic rather than predominantly mononuclear infiltrates in lesional skin biopsies (Soter *et al*. 1976). Non-organ-specific autoantibodies, including antinuclear antibody, rheumatoid factor, and extractable nuclear antigen antibodies may be detectable. It is important to perform frequent urinalysis. Renal function should be monitored carefully if persistant haematuria or proteinuria are found and a renal biopsy performed where indicated.

Some important precipitants and disease associations are summarized in Table 2. Allergic vasculitis has been reported with many classes of drugs including antibiotics, thiazide diuretics, non-steroidal anti-inflammatory drugs, and recombinant human granulocyte colony stimulating factor (Jain 1994). There have been case reports of many others, including additives in a drug formulation (Lowry *et al*. 1994). Human immunodeficiency virus (**HIV**) antigen has been found in lesional skin biopsies of symptomatic HIV-infected individuals with granular immune deposits in small vessel walls (Gherardi *et al*. 1993). Embolic lodging of bacteria and microthrombi in the skin microvasculature may account for the vasculitic lesions seen in infective endocarditis and septicaemia. Small vessel vasculitis may occur in autoimmune rheumatic disorders (rheumatoid arthritis, Sjögren's syndrome, and systemic lupus erythematosus). Rheumatoid arthritis was the commonest association in a series of 88 patients presenting with cutaneous leucocytoclastic vasculitis (Ekenstam and Callen 1984). The presence of antibodies to Ro and rheumatoid factor in Sjögren's syndrome was associated closely with skin lesions showing leucocytoclastic vasculitis on histology (Molina *et al*. 1985). Various other inflammatory disorders have been associated, including inflammatory bowel disease, chronic active hepatitis, and sarcoidosis (Aractingi *et al*. 1993). Greer *et al*. (1988) have

reviewed the association with malignancy. However, despite thorough clinical evaluation and laboratory investigation over 50 per cent of cases remain unexplained (idiopathic) in most reported series.

Precipitants
Drugs, e.g. sulphonamides, penicillins, thiazides, and many others
Infections
Viral, e.g. hepatitis B, human immunodeficiency virus
Bacterial, e.g. β-haemolytic streptococcus
Foreign protein, e.g. serum sickness
Associations
Autoimmune diseases, e.g. rheumatoid arthritis, Sjögren's syndrome (anti-Ro positive), systemic lupus erythematosus
Inflammatory diseases, e.g. chronic active hepatitis, ulcerative colitis, Crohn's disease, sarcoidosis
Malignancy, e.g. myelo- and lymphoproliferative disorders, solid tumours

Table 2 Causes of allergic vasculitis

The condition is often acute and self-limiting but may pursue a relapsing or chronic course. Occasionally death may occur from a number of causes, including renal failure, gastrointestinal haemorrhage, or perforation. In a review of lesional skin biopsies Hodge *et al*. (1987) found that the overall histological severity correlated with a clinical severity score based on cutaneous and visceral involvement but did not predict the presence or absence of systemic vasculitis. Deeper infiltrates were associated with a higher frequency of pulmonary involvement.

Identifiable causes should be removed or treated whenever possible. Bed rest seems to reduce the appearance of new skin lesions but probably does not alter the long-term outcome. The legs should be elevated and bandaged to reduce oedema. Analgesia may be required. Although encouraging responses have been claimed with indomethacin (Millns *et al*. 1980), hydroxychloroquine (Lopez *et al*. 1984), colchicine (Callen 1985), and dapsone (Fredenberg and Malkinson 1987) for control of the underlying disease, systemic steroids are often required and may have to be used at high doses, especially for progressive renal involvement. Immunosuppressive agents, such as azathioprine, may be appropriate for patients with refractory disease.

Henoch–Schönlein purpura

This tends to be regarded as a special form of allergic vasculitis. It is also known as 'anaphylactoid purpura'. The classical presentation is with purpura, arthritis, haemorrhagic gastrointestinal involvement, and glomerulonephritis. It occurs most often in children but adults of any age may be affected. IgA is usually detectable in biopsies of skin, gut, and kidney. Complement appears to be activated by the alternative rather than the classical pathway. It is not clear whether the cases originally described as peliosis rheumatica by Schönlein (1837) and those of Henoch (1874) would have fitted the current concept of a predominantly IgA-associated small vessel vasculitis, because immunological investigations, including direct immunofluorescence, were not available at the time. The paediatric aspects of Henoch–Schönlein purpura are discussed in Chapter 5.11.8.

The skin lesions are often indistinguishable from those seen in allergic vasculitis. The appearance of palpable purpuric plaques on the lower legs with multifocal areas of haemorrhage or necrosis and reticulated borders has been emphasized as a distinctive sign of Henoch–Schönlein purpura in adults (Piette and Stone 1989). Systemic involvement is common but not invariable. Cream *et al*. (1970) found evidence of renal involvement in 50 per cent of 77 adults, presenting with urinary abnormalities alone, acute nephritis, or as slowly progressive renal failure without an initial acute nephritic syndrome. Gastrointestinal involvement (abdominal pain, melaena stool, haematemesis, diarrhoea, or constipation) occurred in 44 per cent, arthralgia or arthritis in 56 per cent, and oedema of legs and ankles in 57 per cent of their series.

Corticosteroids given during the acute illness appear to relieve abdominal pain and arthralgia but there is little convincing evidence that they prevent progression of renal disease or influence the eventual prognosis (Roth *et al*. 1985).

Urticarial vasculitis

Urticarial skin lesions and arthritis are the commonest presentation of this systemic disorder. Females outnumber males in most reported series by over 2:1. Patients are usually middle-aged. Morphologically, the urticarial wheals resemble those of chronic (ordinary) urticaria (Fig. 3) but may also show central purpura or rarely resemble erythema multiforme. They can occur at sites of pressure and be associated with facial and laryngeal angio-oedema. As a general rule, the wheals of urticarial vasculitis last from 24 to 72 h and tend to have a burning or painful quality, whereas those of chronic urticaria are pruritic and last less than 24 h. The commonest systemic features are musculoskeletal. Malaise and low-grade fever may accompany the attacks of urticaria. Renal damage (evidenced by haematuria, proteinuria), gastrointestinal symptoms (abdominal pain, nausea, vomiting, diarrhoea), and pulmonary disease (cough, dyspnoea, haemoptysis) may occur. Other less common manifestations include lymphadenopathy, uveitis, and benign intracranial hypertension.

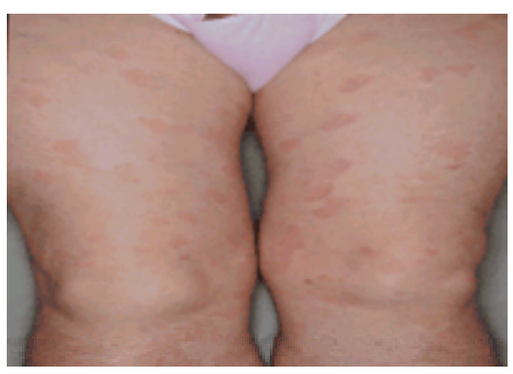


Fig. 3 The wheals of urticarial vasculitis resemble those of chronic urticaria but last longer and may show bruising.

A rare variant of urticarial vasculitis (Schnitzler's syndrome) is characterized by chronic urticaria-like lesions, arthralgia, lymphadenopathy, bone pain, intermittent fever, high ESR, leucocytoclastic vasculitis, and IgM macroglobulinaemia (Schnitzler *et al*. 1974). Interleukin-a antibodies found in some of these patients may play a part in its pathogenesis (Saurat *et al*. 1991).

The term 'hypocomplementaemic vasculitis' has been used for urticarial vasculitis because some early reports described cases in which the early components of complement were reduced (McDuffie *et al*. 1973), but complement abnormalities are by no means invariable and many cases are normocomplementaemic. Hypocomplementaemia and circulating immune complexes were detected in just under half the patients studied in one large series (Sanchez *et al*. 1982). Reduced levels of C1q have been linked with the presence of immunoglobulin G C1q precipitins in some patients (Zeiss *et al*. 1980) but they are not specific for urticarial vasculitis. The most consistent laboratory abnormality is an elevated ESR (Soter *et al*. 1974). Non-organ-specific antibodies are uncommon. Lesional skin biopsies show a range of histological changes from typical leucocytoclastic vasculitis to dense or sparse mixed perivascular infiltrates (Russell Jones *et al*. 1983). Interstitial dermal neutrophilic infiltrates were found in addition to perivascular infiltrates in patients with hypocomplementaemic urticarial vasculitis (Mehregan *et al*. 1992).

Patients with hypocomplementaemia tend to have more systemic manifestations than those with normocomplementaemia. However, the majority of patients with urticarial vasculitis, including those with systemic involvement, follow a chronic but benign course. Glomerulonephritis and chronic obstructive airways disease are

potentially the most serious complications. The relationship between urticarial vasculitis and systemic lupus erythematosus is not entirely clear as urticarial lesions showing histological leucocytoclastic vasculitis have been reported in 7 per cent of patients with systemic lupus erythematosus (Provost *et al*. 1980). Progression of urticarial vasculitis to systemic lupus erythematosus appears to be rare but has been described (Bisaccia *et al*. 1988).

The treatment of urticarial vasculitis has been reviewed by Berg *et al*. (1988). Systemic antihistamines are used widely but tend to be disappointing. Indomethacin, hydroxychloroquine, dapsone, colchicine, plasmapheresis, and gold (Handfield-Jones and Greaves 1991) have been reported to be helpful in small series or anecdotal reports. Systemic steroids often have to be given in high doses to achieve control and tapered to the lowest maintenance level to prevent relapse.

Cryoglobulinaemia

Cryoglobulins are immunoglobulins which precipitate when cold. They have been divided into three types: in type I the cryoglobulin fraction is a single monoclonal immunoglobulin; type II has mixed monoclonal and polyclonal components, usually consisting of IgM rheumatoid-like factor combined with polyclonal IgG; type III has exclusively mixed polyclonal components. Mixed cryoglobulins are associated with autoimmune rheumatic diseases, lymphoproliferative disorders, and infections. The prefix 'essential' is used if no underlying disorder can be found. Hepatitis B virus infection should always be excluded as the virus or its antibody were found in a high proportion of patients previously thought to have essential mixed cryoglobulinaemia when cryoprecipitates were examined as well as sera (Levo *et al*. 1977).

Mixed cryoglobulinaemia presents with purpuric skin lesions showing leucocytoclastic vasculitis on biopsy, polyarthralgia, weakness, and progressive renal disease. It is uncommon. Women are affected about twice as frequently as men, usually in their sixth decade. Gorevic *et al*. (1980) have summarized an 18-year experience of 40 patients with mixed cryoglobulinaemia and reviewed the literature: recurrent palpable purpura were present in all their patients, polyarthralgias in 72 per cent, and renal disease in 55 per cent. Hepatic involvement (hepatomegally, abnormal liver function tests, or mild to severe inflammation on biopsy) was present in 70 per cent of their series. Oedema, hypertension, leg ulcers, Raynaud's phenomenon, abdominal pain, and susceptibility to bacterial pneumonia also occurred in descending order of frequency. A more recent report describes similar skin changes with type I cryoglobulinaemia (Cohen *et al*. 1991). Peripheral neuropathy may occur. Accompanying laboratory abnormalities include reductions in the early components of complement, a raised ESR, and anaemia.

The prognosis is much worse in patients with renal disease. The main causes of death are renal failure, systemic vasculitis, and infection. Necropsy findings showed widespread arteritis in some patients (Gorevic *et al*. 1980) indicating that the vasculitis is not always confined to small vessels.

Treatment of symptomatic mixed cryoglobulinaemia is generally unsatisfactory. High-dose steroids, chemotherapy, plasmapheresis, or a combination of all three (Geltner *et al*. 1981) may bring about limited improvement in renal function, leg ulcers, and purpura. Anecdotal success has been reported with high-dose intravenous gammaglobulin (Boom *et al*. 1988).

Hypergammaglobulinaemic purpura

Waldenström (1943) described three female patients with hyperglobulinaemia, long-standing purpura, and an elevated ESR. This benign disorder should not be confused with Waldenström's macroglobulinaemia which is a lymphoma characterized by a monoclonal IgM paraproteinaemia. Histology of the purpuric skin lesions shows a leucocytoclastic vasculitis. The similarity between hypergammaglobulinaemic purpura and the cutaneous features of some patients with Sjögren's syndrome has been emphasized (Alexander and Provost 1983).

Erythema elevatum diutinum and granuloma faciale

Erythema elevatum diutinum and granuloma faciale are rare but distinctive forms of localized chronic cutaneous leucocytoclastic vasculitis. There is no systemic involvement. The aetiology of these disorders is unknown but erythema elevatum diutinum has been associated with myeloma.

Erythema elevatum diutinum is characterized by slowly enlarging oedematous purplish-brown plaques over the backs of hands, elbows, or knees which heal slowly over months or years with fibrosis. Early lesions may blister (Fig. 4). The clinical and laboratory features have been reviewed by Gibson and Su (1990).

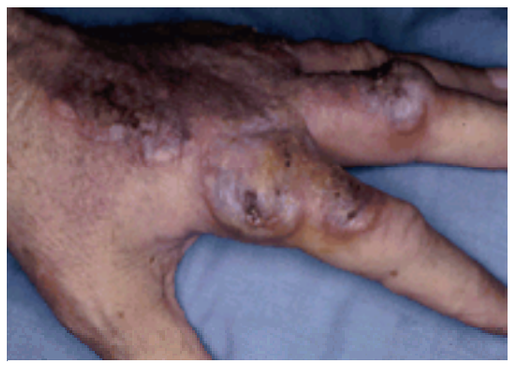


Fig. 4 An early lesion of erythema elevatum diutinum on the back of the hand showing bullous changes.

Granuloma faciale presents with single or multiple pink to brown, well-defined smooth papules and plaques on the face (Fig. 5) which persist for years. It is distinguished histologically from erythema elevatum diutinum by the presence of numerous eosinophils and a zone of normal collagen beneath the epidermis.

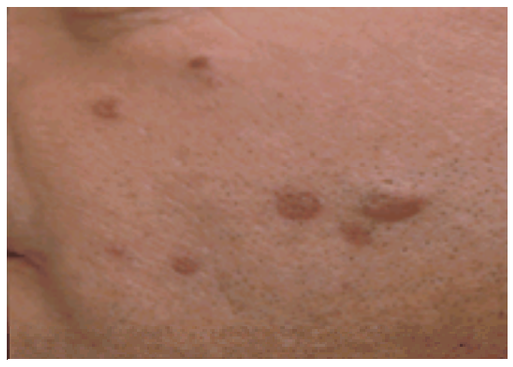


Fig. 5 Reddish-brown papules on the cheek typical of granuloma faciale.

The chronicity of both erythema elevatum diutinum and granuloma faciale is surprising in view of the histology which would suggest an acute pattern of inflammation. Erythema elevatum diutinum may respond well to dapsone. Intralesional steroids can help granuloma faciale.

Non-leucocytoclastic vasculitis

Drug-related vasculitis

Cutaneous and systemic vasculitis showing mononuclear infiltrates and eosinophils in vessel walls without fibrin deposition or necrosis, have been attributed to a variety of drugs, including penicillins and thiazides (Mullick *et al*. 1979).

Nodular vasculitis

The differential diagnosis of nodular forms of cutaneous vasculitis embraces a wide range of disorders, including erythema nodosum and other inflammatory diseases of the subcutaneous fat (Ryan 1992).

Nodular vasculitis is regarded as a distinct subgroup characterized by recurrent subcutaneous nodules usually occurring on the legs of young or middle-aged women. Patients are otherwise healthy. The histological changes range from perivascular lymphocytic infiltrates and granulomatous changes to leucocytoclastic vasculitis with fibrinoid necrosis. The aetiology is uncertain. Associated streptococcal infection may occasionally be found and should be treated. Tuberculosis does not appear to be associated, as with erythema induratum (Bazin's disease). The condition tends to resolve spontaneously but may persist for many years. Individual lesions may respond to intralesional triamcinolone. High doses of potassium iodide have been reported to be beneficial (Schulz and Whiting 1976).

Livedo vasculitis

Ischaemic damage may result from vascular occlusion rather than primary vessel wall inflammation. Livedo vasculitis (segmental hyalinizing vasculitis) is characterized histologically by endothelial proliferation and intraluminal thrombosis. Elevated fibrinopeptide levels, normal complement, and absence of immune complexes on serological studies favour a thrombogenic vasculopathy (McCalmont *et al*. 1992).

Clinically there is a livedo-like pattern of purpura, ulcers, and white atrophic scars known as *atrophie blanche* (Bard and Winkelmann 1967) (Fig. 6). The relevance of antiphospholipid antibodies to the pathogenesis of this disorder and others characterized by vaso-occlusion, including Degos' disease, has been reviewed (Grattan and Burton 1991).

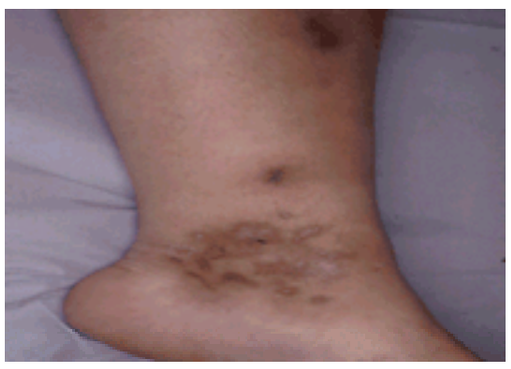


Fig. 6 Acute lesions of livedo vasculitis heal with pigmentation and ' *atrophie blanche*' scarring.

Pityriasis lichenoides

The acute form of this relatively uncommon skin disorder is characterized by crops of oedematous pink papules which enlarge rapidly, and may become haemorrhagic before developing necrotic centres which heal over several weeks with scarring (Fig. 7). Although the eruption may be accompanied by mild constitutional symptoms, such as fever and arthralgia, there are no systemic complications and the disorder is usually self-limiting. Histology shows a lymphocytic perivascular infiltrate. While the endothelial cells are often blurred or swollen, fibrinoid necrosis is seen very rarely and some authorities believe that the changes do not amount to true vasculitis. However, the presence of IgM and C3 deposition in the walls of superficial blood vessels of fresh lesions in both acute and chronic forms of the disease support an immune complex aetiology (Clayton and Haffenden 1978).

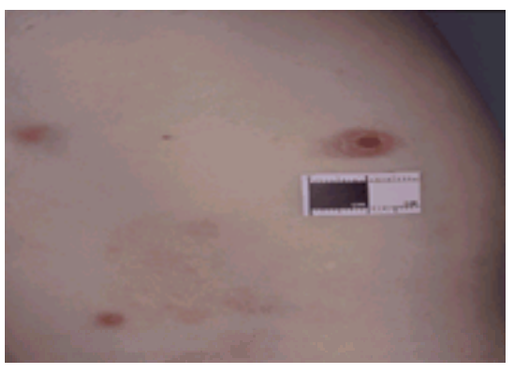


Fig. 7 Oedematous papules on the upper arm of a patient with pityriasis lichenoides acuta showing early central necrosis which will heal with scarring.

Pityriasis lichenoides chronica may succeed the acute form or arise *de novo*. The early pink lesions mature into reddish-brown papules with adherent scale which usually heal without scars. Successive crops of lesions may erupt consecutively for months or years. Treatment with ultraviolet B irradiation is often helpful.

Chapter References

Alexander, E.L. and Provost, E.L. (1983). Cutaneous manifestations of primary Sjögren's syndrome: a reflection of vasculitis and association with anti-Ro (SSA) antibodies. *Journal of Investigative Dermatology*, **80**, 386–91.

Andrews, B.S., Cains G., McIntoch, J., Petts, J., and Penny, R. (1979). Circulating and tissue immune complexes in cutaneous vasculitis. *Journal of Clinical and Laboratory Immunology*, **1**, 311–20.

Aractingi, S., Cadranel, J., Milleron, B., Saiag, P., Malepart, M.J., and Dubertet, L. (1993). Sarcoidosis associated with leucocytoclastic vasculitis. *Dermatology*, **187**, 50–3.

Bard, J.W. and Winkelmann, R.K. (1967). Livedo vasculitis. Segmental hyalinizing vasculitis of the dermis. *Archives of Dermatology*, **96**, 489–99.

Berg, R.E., Kantor, G.R., and Bergfeld, W.F. (1988). Urticarial vasculitis. *International Journal of Dermatology*, **27**, 468–72.

Bisaccia, E., Adamo, V., and Rozan, S.W. (1988). Urticarial vasculitis progressing to systemic lupus erythematosus. *Archives of Dermatology*, **124**, 1088–90.

Boom, B.W., Out–Luiting, C.J., Baldwin, W.M., Westedt, M–L., Daha, M.R., and Vermeer, B-J. (1987). Membrane attack complex of complement in leukocytoclastic vasculitis of the skin. *Archives of Dermatology*, **123**, 1192–5.

Boom, B.W., Brand, A., Bavinck, J–N.B., Eernisse, J.G., Daha, M.R., and Vermeer, B-J. (1988). Severe leukocytoclastic vasculitis of the skin in a patient with essential mixed cryoglobulinaemia treated with high-dose g-globulin intravenously. *Archives of Dermatology*, **124**, 1550–3.

Braverman, I.M. and Yen, A. (1975). Demonstration of immune complexes in spontaneous and histamine-induced lesions and in normal skin of patients with leukocytoclastic vasculitis. *Journal of*

Investigative Dermatology, **64**, 105–12.

Callen, J.P. (1985). Colchicine is effective in controlling chronic cutaneous leukocytoclastic vasculitis. *Journal of the American Academy of Dermatology*, **13**, 193–200.

Clayton, R. and Haffenden, G. (1978). An immunofluoresence study of pityriasis lichenoides. *British Journal of Dermatology*, **99**, 491–3.

Cochrane, C.G. (1971). Mechanisms involved in the deposition of immune complexes in tissues. *Journal of Experimental Medicine*, **134**, 75–89.

Cohen, S.J., Pittelkow, M.R., and Su, W.P.D. (1991). Cutaneous manifestations of cyroglobulinaemia: clinical and histopathological study of seventy two patients. *Journal of American Academy of Dermatology*, **25**, 21–7.

Cream, J.J., Gumpel, J.M., and Peachey, R.D.G. (1970). Schönlein–Henoch purpura in the adult. A study of 77 adults with anaphylactoid or Schönlein–Henoch purpura. *Quarterly Journal of Medicine*, **156**, 461–84.

Cream, J.J., Bryceson, A.D.M., and Ryder, G. (1971). Disappearance of immunoglobulin and complement from the Arthus reaction and its relevance to studies of vasculitis in man. *British Journal of Dermatology*, **84**, 106–9.

Daoud, M.S., Gibson, L.E., DeRemee, R.A., Specks, U., el-Azhary, R.A., and Su, W.P.D. (1994). Cutaneous Wegener's granulomatosis: clinical, histopathologic, and immunopathologic features of thirty patients. *Journal of American Academy of Dermatology*, **31**, 605–12.

Eady, R.A.J., Keahey, T.M., Sibbald, R.G., and Kobza Black, A. (1981). Cold urticaria with vasculitis: report of a case with light and electron microscopic, immunofluorescence and pharmacological studies. *Clinical and Experimental Dermatology*, **6**, 355–66.

Ekenstam, E. and Callen, J.P. (1984). Cutaneous leukocytoclastic vasculitis. *Archives of Dermatology*, **120**, 484–9.

Francès, C. *et al*. (1994). Wegener's granulomatosis. Dermatological manifestations in 75 cases with clinicopathologic correlation. *Archives of Dermatology*, **130**, 861–7.

Fredenberg, M.F. and Malkinson, F.D. (1987). Sulphone therapy in the treatment of leukocytoclastic vasculitis. *Journal of the American Academy of Dermatology*, **16**, 772–8.

Geltner, D., Kohn, R.W., Gorevic, P., and Franklin, E.C. (1981). The effect of combination therapy (steroids, immunosuppressives, and plasmapheresis) on 5 mixed cryoglobulinaemia patients with renal, neurologic, and vascular involvement. *Arthritis and Rheumatism*, **24**, 1121–7.

Gherardi, R. *et al*. (1993). The spectrum of vasculitis in human immunodeficiency virus-infected patients. *Arthritis and Rheumatism*, **36**, 1164–74.

Gibson, L.E. and Su, W.P.D. (1990). Cutaneous vasculitis. *Rheumatic Disease Clinics of North America*, **16**, 309–24.

Gorevic, P.D. *et al*. (1980). Mixed cryoglobulinemia: clinical aspects and long-term follow-up of 40 patients. *American Journal of Medicine*, **69**, 287–308.

Gower, R.G., Sams, W.M., Thorne, E.G., Kohler, P.F., and Claman, H.N. (1977). Leukocytoclastic vasculitis: sequential appearance of immunoreactants and cellular changes in serial biopsies. *Journal of Investigative Dermatology*, **69**, 477–84.

Grattan, C.E.H. and Burton, J.L. (1991). The antiphospholipid syndrome and cutaneous vasocclusive disorders. *Seminars in Dermatology*, **10**, 152–9.

Greer, J.M., Longley, S., Edwards N.L., Elfenbein, G.J., and Panush, R.S. (1988). Vasculitis associated with malignancy. Experience with 13 patients and literature review. *Medicine*, **67**, 220–30.

Handfield-Jones, S.E. and Greaves, M.W. (1991). Urticarial vasculitis-response to gold therapy. *Journal of the Royal Society of Medicine*, **84**, 169–70.

Henoch, E. (1874). Ueber eine eigenthümliche Form von Purpura. *Berliner Klinische Wochenschrift*, **51**, 641–3.

Hodge, S.J., Callen, J.P., and Ekenstam, E. (1987). Cutaneous leukocytoclastic vasculitis: correlation of histopathological changes with clinical severity and course. *Journal of Cutaneous Pathology*, **14**, 279–84.

Jain, K.K. (1994). Cutaneous vasculitis associated with granulocyte colony stimulating factor. *Journal of American Academy of Dermatology*, **31**, 213–15.

Jennette, J.C. *et al*. (1994). Nomenclature of systemic vasculitides. Proposal of an international consensus conference. *Arthritis and Rheumatism*, **37**, 187–92.

Levo, Y., Gorevic, P.D., Kassab, H.J., Zucker-Franklin, D., and Franklin, E.C. (1977). Association between hepatitis B virus and essential mixed cryoglobulinemia. *New England Journal of Medicine*, **296**, 1501–4.

Lopez, L.R., Davis, K.C., Kohler, P.F., and Schocket, A.L. (1984). The hypocomplementaemic urticarial-vasculitis syndrome: therapeutic response to hydroxychloroquine. *Journal of Allergy and Clinical Immunology*, **73**, 600–3.

Lowry, M.D., Hudson, C.F., and Callen J.P. (1994). Leukocytoclastic vasculitis. *Journal of the American Academy of Dermatology*, **30**, 854–5.

McCalmont, C.S., McCalmont, T.H., Jorizzo, J.L., White, W.L., Leshin, B., and Rothberger, H. (1992). Livedo vasculitis: vasculitis or thrombotic vasculopathy? *Clinical and Experimental Dermatology*, **17**, 4–8.

McDuffie, F.C., Sams, W.M., Maldonado, J.E., Andreini, P.H., Conn, D.L., and Samayoa, E.A. (1973). Hypocomplementemia with cutaneous vasculitis and arthritis. Possible immune complex syndrome. *Mayo Clinic Proceedings*, **48**, 340–8.

Mehregan, D.R., Hall, M.J., and Gibson, L.E. (1992). Urticarial vasculitis. *Journal of the American Academy of Dermatology*, **26**, 441–8.

Millns, J.L, Randle, H.W., Solley, G.O., and Dicken, C.H. (1980). The therapeutic response of urticarial vasculitis to indomethacin. *Journal of the American Academy of Dermatology*, **3**, 349–55.

Molina, R., Provost, T.T., and Alexander, E.L. (1985). Two types of inflammatory vascular disease in Sjögren's syndrome. Differential association with seroreactivity to rheumatoid factor and antibodies to Ro (SS-A) and with hypocomplementaemia. *Arthritis and Rheumatism*, **28**, 1251–8.

Mullick, F.G., McAllister, H.A., Wagner, B.M., and Fenoglio, J.J. (1979). Drug related vasculitis. Clinicopathological correlations in 30 patients. *Human Pathology*, **10**, 313–25.

Parish, W.E. (1971). Studies on vasculitis I. Immunoglobulins, b1C, C-reactive protein, and bacterial antigens in cutaneous vasculitis lesions. *Clinical Allergy*, **1**, 97–109.

Piette, W.W. and Stone, M.S. (1989). A cutaneous sign of IgA-associated small dermal vessel leukocytoclastic vasculitis in adults (Henoch–Schönlein purpura). *Archives of Dermatology*, **125**, 53–6.

Provost, T.T., Zone, J.J., Synkowski, D., Maddison, P.J., and Reichlin, M. (1980). Unusual cutaneous manifestations of systemic lupus erythematosus: I. Urticaria-like lesions. Correlation with clinical and serological abnormalities. *Journal of Investigative Dermatology*, **75**, 495–9.

Roth, D.A., Wilz, D.R., and Theil, G.B. (1985). Schönlein–Henoch syndrome in adults. *Quarterly Journal of Medicine*, **55**, 145–52.

Russell Jones, R., Bhogal, B., Dash, A., and Schifferli, J. (1983). Urticaria and vasculitis: a continuum of histological and immunopathological changes. *British Journal of Dermatology*, **108**, 695–703.

Ryan, T.J. (1992). Cutaneous vasculitis. In *Rook/Wilkinson/Ebling Textbook of Dermatology*, Vol. 3, (5th edn) (ed. R.H. Champion, J.L. Burton, and F.J.G. Ebling), Ch. 45, pp. 1893–961. Blackwell Scientific Publications, Oxford.

Sams, W.M., Claman, H.N., Kohler, P.F., McIntosh, R.M., Small, P., and Mass, M.F. (1975). Human necrotizing vasculitis: immunoglobulins and complement in vessel walls of cutaneous lesions and normal skin. *Journal of Investigative Dermatology*, **64**, 441–5.

Sanchez, N.P., Winkelmann, R.K., Schroeter, A.L., and Dicken, C.H. (1982). The clinical and histopathologic spectrums of urticarial vasculitis: study of forty cases. *Journal of the American Academy of Dermatology*, **7**, 599–605.

Saurat, J-H., Schifferli, J., Steiger, G., Dayer, J-M., and Didierjean, L. (1991). Anti-interleukin-1a autoantibodies in humans: characterisation, isotype distribution, and receptor-binding inhibition—higher frequency in Schnitzler's syndrome (urticaria and macroglobulinaemia). *Journal of Allergy and Clinical Immunology*, **88**, 244–56.

Schnitzler, P.L., Schubert, B., Boasson, M., Gardais, J., and Tourmen, A. (1974). Urticaire chronique, lésions osseuses, macroglobulinémie IgM: maladie de Waldenström? *Bulletin de la Societé Française de Dermatologie et Syphiligraphie* **81**, 363.

Schönlein, J.L. (1837). Peliosis rheumatica. *Allgemeine und specielle Pathologie und Therapie, Freyburg*, **2**, 48–9.

Schulz, E.J. and Whiting, D.A. (1976). Treatment of erythema nodosum and nodular vasculitis with potassium iodide. *British Journal of Dermatology*, **94**, 75–8.

Soter, N.A., Austen, K.F., and Gigli, I. (1974). Urticaria and arthralgias as manifestations of necrotizing angiitis (vasculitis). *Journal of Investigative Dermatology*, **63**, 485–90.

Soter, N.A., Mihm, M.C., Gigli, I., Dvorak, H.F., and Austen, K.F. (1976). Two distinct cellular patterns in cutaneous necrotizing angiitis. *Journal of Investigative Dermatology*, **66**, 344–50.

Waldenström, J. (1943). Kliniska metoder för påvisande av hyperproteinämi och deras praktiska värde för diagnostiken. *Nordisk Medicin*, **20**, 2288–95.

Zax, R.H., Hodge, S.J., and Callen, J.P. (1990). Cutaneous leukocytoclastic vasculitis. Serial histopathological evaluation demonstrates the dynamic nature of the infiltrate. *Archives of Dermatology*, **126**, 69–72.

Zeek, P.M. (1952). Periarteritis nodosa: a critical review. *American Journal of Clinical Pathology*, **22**, 777–90.

Zeiss, C.R., Burch F.X., Marder R.J., Furey, N.L., Schmid, F.R., and Gewurz, H. (1980). A hypocomplementemic vasculitic urticarial syndrome. Report of four new cases and definition of the disease. *The American Journal of Medicine*, **68**, 867–75.

5.11.5 Polymyalgia rheumatica

G. S. Panayi

Definitions

Polymyalgia rheumatica

Since both polymyalgia rheumatica and giant cell arteritis are clinical syndromes it is important that agreed definitions of the conditions are employed not only for uniformity of diagnosis but also for epidemiological and investigational studies. The criteria of Jones and Hazleman (Table 1) are succinct and easily applied in practice for the diagnosis of polymyalgia rheumatica. Some comments are necessary. There may be apparent weakness on testing the shoulder and pelvic girdle muscles but this is due to pain rather than intrinsic muscle weakness. Investigations to exclude inflammatory arthritis, malignant disease, or muscle diseases should be carried out objectively but with restraint as discussed under 'diagnosis' below.

1. Shoulder and pelvic girdle pain which is primarily muscular in the absence of true muscle weakness
2. Morning stiffness
3. Duration of at least 2 months unless treated
4. ESR over 30 mm/h or C-reactive protein over 6 μg/ml
5. Absence of rheumatoid or inflammatory arthritis or malignant disease
6. Absence of objective signs of muscle disease
7. Prompt and dramatic response to systemic corticosteroids

From Jones and Hazleman (1981).

Table 1 The diagnostic criteria for polymyalgia rheumatica

Giant cell arteritis

The criteria for diagnosis of giant cell arteritis according to Jones and Hazleman (1981) are shown in Table 2, whilst those by Ellis and Ralston (1983) are shown in Table 3. The advantages of the criteria of Ellis and Ralston (1983) are that they do not include the erythrocyte sedimentation rate (**ESR**) and have a more inclusive description of the clinical features of giant cell arteritis compared with the criteria of Jones and Hazleman (1981). However, the diagnostic criteria of the latter have the advantage of brevity. It should be remembered that although up to one-quarter of patients can present with normal ESR, before being treated with corticosteroids, they are still liable to serious complications including visual loss. This can happen even during treatment when there is a normal ESR. The biopsy is classifically taken from the temporal artery but other sites have included practically all vessels arising from the arch of the aorta.

1. Positive temporal artery biopsy or cranial artery tenderness noted by a physician
2. One or more of the following: visual disturbance, headache, jaw pain, cerebrovascular insufficiency
3. ESR over 30 mm/h or C-reactive protein over 6 μg/ml
4. Response to systemic corticosteroids

Table 2 The Jones and Hazleman (1981) diagnostic criteria for giant cell arteritis

1. Age greater than 55 years
2. Positive response within 48 h to corticosteroid therapy
3. Length of history greater than 2 weeks
4. Positive temporal artery biopsy
5. Proximal, symmetrical girdle, or upper arm muscle pain+ stiffness+tenderness
6. Jaw claudication
7. Clinical abnormality of the temporal artery (tenderness, thickening, redness)
8. Systemic symptoms or signs (malaise, anorexia, weight loss, anaemia, pyrexia)
9. Temporal headache
10. Visual disturbance (loss, diplopia, blurring)

Positive diagnosis present if criteria 1 to 3 present plus any three of criteria 5 to 10 or criterion 4.

Table 3 The Ellis and Ralston (1983) diagnostic criteria for giant cell arteritis

The American College of Rheumatology (**ACR**) has developed criteria for the classification of giant cell arteritis by comparing a group of 214 patients with the condition and 593 controls with other forms of vasculitis (Hunder *et al*. 1990) (Table 4). In the traditional format a patient shall be said to have giant cell arteritis if at least three out of five criteria listed in the Table are present. The presence of three or more criteria yielded a sensitivity of 93.5 per cent and a specificity of 91.2 per cent. A classification tree was also constructed using six criteria. Elevated ESR was excluded in the tree format due to low specificity (48 per cent) and claudication of jaw and/or tongue on deglutition was included. In the classification tree the presence of temporal artery tenderness of decreased pulsation separated cases from controls better than any other criteria. The tree classification defines giant cell arteritis in fairly simple terms as a vasculitis with onset above 50 years of age, abnormal temporal arteries, or claudification of jaw and/or tongue upon deglutition, and arterial biopsy showing vasculitis with predominantly mononuclear cells or granulomatous inflammation. One hundred and ninety-six of 214 patients with giant cell arteritis had biopsies showing arteritis and 18 lacked biopsy proof. Fifteen out of 18 had negative biopsies and in three biopsy was not performed.

1. Age at disease onset ⩾50 years
2. New headache
3. Temporal artery tenderness or decreased pulsation
4. Elevation of ESR ⩾50 mm/h
5. Abnormal artery biopsies showing necrotizing arteritis with mononuclear infiltrate or granulomatous inflammation usually with multinucleated giant cells

Diagnosis of giant cell arteritis, if three of five criteria are present.

Table 4 ACR 1990 criteria for the classification of giant cell arteritis (traditional format)

It should be noted that several controls with other types of vasculitis, such as Wegener's granulomatosis and polyarteritis nodosa, had biopsies of lung or other tissues showing chronic granulomatous arteritis together with enough additional criteria to misclassify them as giant cell arteritis. Biopsies of a temporal artery was not specified because few controls had this procedure. If temporal artery biopsy was specified, few controls would have been misclassified.

The ACR criteria are easy to apply and do not require exclusions other than the presence of a autoimmune rheumatic disease. In the tree classification the use of scalp tenderness and headache as surrogates for temporal artery abnormality and positive biopsy provide flexibility when some clinical data are not available.

Relationship between polymyalgia rheumatica and giant cell arteritis

Patients with features of both polymyalgia and giant cell arteritis are seen and this raises the question of the relationship between the two. Polymyalgia rheumatica is a more common condition than giant cell arteritis. One-half of the patients with giant cell arteritis have polymyalgic symptoms and 15 to 22 per cent of patients with polymyalgia rheumatica have been shown to have giant cell arteritis by either temporal artery biopsy or clinical symptoms (Huston *et al*. 1978; Chuang *et al*. 1982). In other studies the proportion of patients with a myalgic presentation who have positive arterial biopsies but no clinical features of giant cell arteritis varies from 6 per cent (Hunder and Allen 1973), to 40 per cent (Fauchald *et al*. 1972), to 50 per cent, which is the highest recorded (Malmvall and Bengston 1978). It used to be thought that polymyalgia rheumatica was always a manifestation of giant cell arteritis but most patients with the former, even if followed for many years, do not develop clinical or biopsy-proven giant cell arteritis. This was true even in the days before the use of systemic corticosteroids. There are, undoubtedly, patients with polymyalgia rheumatica who have a positive biopsy for giant cell arteritis but whose clinical picture is identical to patients with polymyalgia rheumatica without the lesions of giant cell arteritis (Fauchald *et al*. 1972). Conversely, there are patients with giant cell arteritis who also have clinical evidence of polymyalgia rheumatica. Thus, polymyalgia rheumatica and giant cell arteritis may be considered as components of a single syndrome, the expression of which depends on unknown factors; genetic factors may be the most important of these.

Genetic and epidemiological studies

Genetic studies

A genetic basis for polymyalgia rheumatica/giant cell arteritis has long been suspected because of the occurrence of the disease within families and in sib pairs (Moss and Soukop 1988). However, their mode of inheritance remained unknown until studies on possible HLA associations were undertaken. The study of Hansen *et al*. (1985) has been the only one to show an association with a class I major histocompatibility antigen, HLA Cw3. The increased occurrence of HLA A31, HLA B40 and, of greater importance, HLA DR4 was ascribed as being secondary and due to their linkage disequilibria with HLA Cw3. Armstrong *et al*. (1983) showed a link between polymyalgia rheumatica and giant cell arteritis and HLA Cw6 with a relative risk of 9.0. This is the only reported association with a C locus antigen. However, most studies agree that the true link is with HLA DR4. Richardson *et al*. (1987) found an association between polymyalgia rheumatica or polymyalgia rheumatica plus giant cell arteritis with HLA DR4, whilst giant cell arteritis alone did not show this association. By contrast, Ninet *et al*. (1987) found that it was giant cell arteritis, whether alone or associated with polymyalgia rheumatica, which was linked to HLA DR4. Whether this discrepancy can be accounted for by the racial differences in the two studies, Anglo-Saxon in the former and French in the latter, is not known, but it should be emphasized that Armstrong *et al*. (1983) found a link with polymyalgia rheumatica or giant cell arteritis while Cid *et al*. (1988), in a Spanish population, found a link of HLA DR4 with polymyalgia rheumatica alone. It may be that differences in the population of patients being studied may account for some of the discrepancies. The studies by Sakkas *et al*. (1990) have used restriction fragment length polymorphism (**RFLP**) with DRb, DQa, and DQb probes and appropriate restriction endonucleases in order to investigate the class II MHC association at the molecular level. In this study the link between polymyalgia rheumatica and HLA DR4 was confirmed whilst the DQ specifities DQw7 and DQw8 (previously DQ3.1 and DQ3.2 respectively) were found in similar frequency to the controls. Rheumatoid arthritis is linked to the third hypervariable region of HLA DR4. Whilst polymyalgia rheumatica and giant cell arteritis are also linked to HLA DR4, molecular analysis has shown that this is with the second hypervariable region of the HLA DRB1 gene (Weyand *et al*. 1994a). Two conclusions can be drawn from these findings. First, the reported association of rheumatoid arthritis with polymyalgia rheumatica/giant cell arteritis is probably due to chance or to misdiagnosis. Second, HLA-DRB1 alleles are not predictive for progression of polymyalgia rheumatica to giant cell arteritis.

Demaine *et al*. (1983) have reported that the immunoglobulin allotypic marker G1m(2) was significantly increased in patients with giant cell arteritis but not with polymyalgia rheumatica. The increase in G1m(2) in the group with giant cell arteritis was not accompanied by a corresponding rise in the number of patients homozygous for G1m(2), i.e. all the increase could be attributed to patients with the G1m(1, 2, 3): G3M (5, 10, 21) phenotype. Using a DNA probe for the switch region of Igμ and a1 heavy chain genes and RFLP analysis, Sakkas *et al*. (1990) were unable to pinpoint these Gm polymorphisms further. However, these findings suggest that genes outside the MHC and probably residing in the region of the immunoglobulin heavy chain on chromosome 14q are also involved in the genetics of polymyalgia rheumatica/giant cell arteritis. It is these additional genetic factors which may modify the clinical expression of disease.

The T-cell receptor genes are highly polymorphic and intimately involved in the genetics of the immune response. No RFLP association was found with T-cell receptor a, b, or g genes (Sakkas *et al*. 1990).

Epidemiological studies

The prevalence of giant cell arteritis/polymyalgia rheumatica has strikingly different rates in different racial groups with Scandinavians having the highest and black and Hispanic races the lowest (Gonzalez *et al*. 1989). Furthermore, only 12 to 25 per cent of black patients with giant cell arteritis have polymyalgia rheumatica compared with 40 to 60 per cent in white patients (Love *et al*. 1986) and this may be related to the low frequency of HLA DR4 in the black population. There may be geographical variation in the expression of clinical features as a Japanese study has reported a higher frequency of fever and involvement of the pelvic girdle than the

Western reports (Nishioka *et al*. 1986).

Ninety-six patients with polymyalgia rheumatica were identified in a 10-year survey (1970 to 1979) at the Olmstead County Hospital, Minnesota (Chuang *et al*. 1982) leading to an average annual incidence of 11.1/100 000 persons (53.7/100 000 in persons over 50 years of age). Polymyalgia rheumatica is more common in older age groups with age-specific incidence increasing from 19.8/100 000 in persons aged 50 to 59 years to 112.2/100 000 in those aged 70 to 79 years. Giant cell arteritis was found in 15 out of 96, i.e. 18 per cent of those surveyed. Bengtsson and Malmvall (1982) surveyed the incidence of polymyalgia rheumatica over 3 years at Goteborg, Sweden which has a population of 400 000. The annual incidence was lower with an incidence rate of 6.7/100 000 persons or 20.4/100 000 in those 50 years or older. In the Swedish study the incidence of giant cell arteritis was 41 out of 90 patients with polymyalgia rheumatica, i.e. 46 per cent. The differences in data between the two studies may be partly due to the difference in diagnostic criteria and methods of data collection. A 9-year study of polymyalgia rheumatica/giant cell arteritis in Reggio Emilia, Italy, from 1980 to 1988 revealed an annual incidence of 6.9/100 000 in persons over 50 years of age (Salvarani *et al*. 1991).

The incidence of polymyalgia rheumatica/giant cell arteritis, at least amongst the white races in the northern hemisphere, lies between 1.7 and 7.7 per 1000 of the elderly population. However, it is obvious that these figures derive from studies using different diagnostic criteria, prospective or retrospective design, and different catchment populations. Properly controlled, large, and prospective studies in different racial groups from different parts of the world are urgently required. Such data may enhance our understanding of the genetic and environmental factors contributing to the development and expression of polymyalgia rheumatica/giant cell arteritis.

The frequent acute influenza-like onset has stimulated interest in the search for an environmental agent in polymyalgia rheumatica/giant cell arteritis. Polymyalgia rheumatica has been reported as being preceded by viral infection, vaccination for influenza or typhoid fever, and by *Yersinia enterocolitica* infection. The finding of antibodies to hepatitis B antigen (Bacon *et al*. 1975) has not been verified (Bridgeford *et al*. 1980). The higher prevalence and titre of antibodies to adenovirus and respiratory syncytial virus in patients with polymyalgia rheumatica compared with controls is intriguing (Cimmino *et al*. 1993). Case reports of polymyalgia rheumatica and giant cell arteritis in conjugal pairs also suggest as environmental agent (Kyle *et al*. 1984). Antibodies to intermediate filaments have been described in high concentrations in polymyalgia rheumatica/giant cell arteritis (Dasgupta *et al*. 1987). These antibodies are typically found in viral infections and may suggest a viral aetiology. Polymyalgia rheumatica has been reported following a tick bite in an 84-year-old woman with serological evidence of *Borrelia burgdorferi* injection. Sixty-three per cent of patients (12 of 19) with polymyalgia rheumatica/giant cell arteritis were found to have elevated or borderline IgG antibody titres (Vaith *et al*. 1988) to this pathogen.

Arterial biopsy

In diagnosis

The histological proof of granulomatous change in the temporal artery or in an artery arising from the aorta or in the aorta itself is generally considered diagnostic proof of giant cell arteritis even if the symptoms are those of incomplete giant cell arteritis or even polymyalgia rheumatica alone (Table 5). An adequate biopsy requires a segment of artery 2- to 3-cm long. Many centres recommend biopsy of the opposite artery if histology is negative in the selected artery. Granulomatous giant cell arteritis is seen in approximately 50 per cent of positive biopsies, showing panarteritis with a mixed cellular infiltrate that is predominantly lymphomononuclear (Lie *et al*. 1990) (Fig. 1). Occasionally a circumferential band of fibrinoid necrosis may also be seen. There is no correlation between histological and clinical features in individual patients (Lie 1987). Pretreatment biopsies give the most successful diagnostic rate (80 per cent) which falls to 60 per cent after 1 week of treatment and becomes much lower after longer periods of treatment (Lie 1987), although others have found that up to 14 days of glucocorticoid treatment does not alter the biopsy positivity rate (Achkar *et al*. 1994).

Disease	Artery	Positive biopsy (%)	Year	Author
GCA	Temporal	30/88 (31)	1992	Chmielewski et al.
PMR/GCA	Temporal	14/75 (14)	1989	Stuart
PMR/GCA	Temporal	21/81 (25.9)	1988	Robb-Nicholson et al.
GCA	Temporal	28/107 (27.1)	1988	Fernandez-Herlihy
GCA	Temporal	42/200 (21)	1988	Pongo et al.
GCA	Temporal	5/45 (11.1)	1988	Weyman et al.
GCA	Temporal	43/102 (43.7)	1987	Villaseca et al.

Table 5 Recent studies on the value of temporal artery biopsy in the diagnosis of polymyalgia rheumatica (PMR)/giant cell arteritis (GCA)

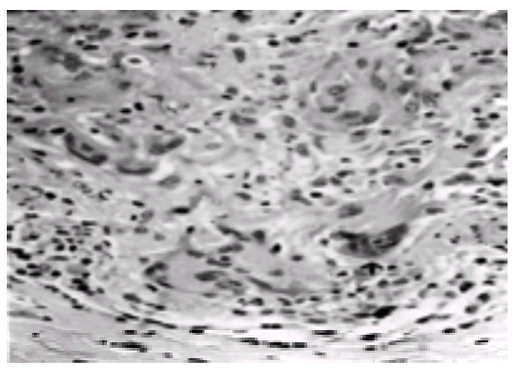


Fig. 1 Typical arterial histology from a patient with giant cell arteritis showing granulomatous inflammation with giant cells (arrow points to a giant cell).

The presence of skip lesions and of differential involvement means that histological proof of giant cell arteritis can vary from 11 to 44 per cent of temporal artery biopsies depending on various selection factors. Efforts have been made to improve the positive biopsy rate by means of arteriographic or Doppler ultrasound examination of the temporal artery but to no avail. There can be large artery involvement with a variety of symptoms including intermittent claudication of an extremity, paraesthesia, Raynaud's phenomenon, and aortic rupture. Clinical examination may reveal absent or decreased pulsation of large arteries with bruits over them. The angiographic findings are characteristic and help to distinguish it from atheromatous stenosis (Klein *et al*. 1975). Angiography reveals (1) long segments of smooth arterial stenoses alternating with those of normal or an increased calibre, (2) smoothly tapered occlusions of affected large arteries, (3) absence of irregular plaques and ulceration, which is characteristic of atheromatous involvement, and (4) anatomic distribution with major involvement of subclavian, axillary, or brachial arteries.

For understanding the pathogenesis

The pathogenesis of polymyalgia rheumatica/giant cell arteritis is not known. Since there is obvious vascular involvement in giant cell arteritis and since giant cell arteritis and polymyalgia rheumatica frequently coexist, it has been assumed that polymyalgia rheumatica also has a vascular basis but direct proof of this supposition is difficult to find. Although evidence of endothelial damage or activation can be found in elevated levels of circulating von Willebrand factor and factor VIII, this observation is common to many inflammatory conditions. Immune complexes have been conspicuous by their absence. Antineutrophil cytoplasmic antibodies (**ANCA**), which are elevated in many forms of macroscopic and microscopic vasculitides, are not found in polymyalgia rheumatica/giant cell arteritis, although Cats and colleagues have found high titres of ANCA, but of unknown specificity, in the serum of patients with giant cell arteritis (Cats *et al*. 1993). Antibodies to intermediate filaments may reflect underlying disease activity (Monteagudo *et al*. 1994).

However, patients with of polymyalgia rheumatica/giant cell arteritis have low numbers of circulating CD8 T cells which are activated, as shown by their high positivity for HLA-DR antigen, which is normally absent from the surface of resting T cells. The relationship of these findings to the pathogenesis of polymyalgia rheumatica/giant cell arteritis is not known but they are similar to those found in the blood of children with Kawasaki's syndrome in which arterial damage is very prominent. Activated CD8 T cells can be cytotoxic and also produce a variety of proteolytic enzymes, especially granzymes, which have cytotoxic and inflammatory potential. A temporal artery with positive histological appearance shows infiltration of mononuclear cells, T cells, and macrophages, and disruption of the elastic lamina of the artery. The lesion may contain deposits of IgG, IgM, IgA, complement, and fibrinogen and the use of immunofluorescence to detect these deposits may be more sensitive and specific than light microscopy alone, although this has not been adopted for routine clinical use (Wells *et al*. 1989). There is also a marked accumulation of fibronectin, fibrin, fibrinogen, and their degradation products, as well as factor VIII, but none of these findings is characteristic (Chemnitz *et al*. 1987).

The T cells infiltrating the temporal artery in giant cell arteritis are CD4+: these are activated as shown by their positivity for HLA DR and the interleukin-2 receptor but, interestingly, show little or no surface expression of the transferrin receptor which is another T-cell activation marker (Andersson *et al*. 1987; Cid 1989). The reason for this discrepancy is unknown but it is worth remarking that although CD4+ T cells in the rheumatoid synovium are HLA-DR positive, few of them express the interleukin-2 receptor. Thus T-cell activation in different chronic inflammatory foci may be at different stages. The elevated serum levels of soluble interleukin-2 receptor provide systemic evidence for T-cell activation but are not useful for disease monitoring (Salvarani *et al*. 1992). Analysis of the T-cell receptor genes used by CD4 T cells expanded from the temporal artery biopsies suggests that they are of limited heterogeneity. This implies that locally present antigen may be driving the granulomatous reaction (Schaufelberger *et al*. 1993; Weyand *et al*. 1994b).

CD8+ T cells form the minority T-cell population and there are few or no B cells and no natural killer cells. The paucity of B cells in the lesions correlates with the lack of a systemic hypergammaglobulinaemia. Although corticosteroids do not influence the cellular distribution within the lesion, a finding similar to that in the rheumatoid synovium, they nevertheless induce functional changes in T cells since within 4 days of starting corticosteroid therapy there is a dramatic fall in the number of T cells positive for interleukin-2 receptors from 87.5 to 14 per cent (Cid 1989).

Macrophages are found in increased numbers in all lesions and are activated as shown by their high expression of HLA DR and transferrin receptor (Andersson *et al*. 1987). Macrophages in the arterial biopsies have been shown to produce interleukin 6, interleukin 1b, and a 72-kDa type IV collagenase. Monocytes from the peripheral blood of patients with giant cell arteritis or polymyalgia rheumatica express interleukins 6 and 1b. These findings suggest that these diseases have local as well as systemic inflammation (Wagner *et al*. 1994). Studying the bone marrow might be rewarding. Macrophages could also act as antigen-presenting cells. The additional cells present are interdigitating cells, which are a form of antigen-presenting cell, found in 41 per cent of temporal artery biopsies; these patients have a significantly shorter disease duration before presentation (mean 1.5 months compared with 3.8 months) (Cid 1989). Arterial smooth muscles are HLA-DR negative so it is unlikely that they are serving as antigen-presenting cells (Andersson *et al*. 1988).

Clinical features

Polymyalgia rheumatica and giant cell arteritis are rare in patients less than 50 years old and the mean age at onset is approximately 70 years. Women are affected more than men (ratio 2:1). The onset is frequently abrupt with pain and stiffness in the neck and shoulder girdle. Hips and thighs are involved less often. Prolonged and severe morning stiffness is a characteristic feature. There may be asymmetry at onset although symptoms become quickly bilateral. Systematic symptoms such as malaise, anorexia, weight loss, low-grade fever, and depression are present frequently. Active joint mobility is often restricted by pain and stiffness whereas passive movements are full. There is no objective muscle weakness although severe muscle pain and stiffness may give this misleading impression

Arthralgia and even synovitis is not uncommon (Chou and Schumacher 1984). Synovitis of the knees and sternoclavicular joints may be evident clinically. Shoulder and hip joint involvement may be more difficult to detect due to the overlying muscles. Synovial fluid examination, arthroscopic synovial biopsies, and joint scintiscanning (O'Duffy *et al*. 1976) have confirmed joint inflammation. Indeed, it has even been proposed that the prominence of shoulder and pelvic girdle symptoms is due to axial synovitis involving the glenohumoral and hip joints (Koski 1992). Arthroscopic synovial biopsies have revealed a lymphocytic inflammatory infiltrate. However, florid synovitis is uncommon in true polymyalgia rheumatica and a clinically significant, chronic, erosive arthropathy suggests an alternative diagnosis of rheumatoid arthritis as this may have a polymyalgic onset in the elderly.

There are case reports of polymyalgia rheumatica associated with underlying malignancy, although it is unclear whether the incidence of neoplasm is increased truly compared with age and sex-matched controls. A retrospective study of case notes reported a high incidence of hypothyroidism in polymyalgia rheumatica/giant cell arteritis (Wiseman *et al*. 1989) but a subsequent study failed to confirm these findings (Dasgupta *et al*. 1990).

Giant cell arteritis

Presenting symptoms may vary widely but headache is the most common initial symptom, being present in two-thirds of patients. The pain is often severe and usually localized to arteries of the scalp (superficial, temporal, and occipital) though it may be more diffuse. The artery involved may be tender or exhibit reduced pulsations. Frequently there is diffuse scalp tenderness.

Visual symptoms are said to occur in 25 to 50 per cent of cases and include diplopia, ptosis (transient or permanent), and partial or total blindness. The symptoms are caused by ischaemia of the optic nerve secondary to involvement of branches of the ophthalmic or posterior ciliary arteries. Recent reports suggest that visual complications may be less frequent than previously reported in pure polymyalgia rheumatica (Myles *et al*. 1992). Unfortunately there are no predicative features for when these may occur and such complications have occurred in patients with quiescent disease and normal erythrocyte sedimentation rates.

Laboratory features

A marked acute-phase response is characteristic of both polymyalgia rheumatica and giant cell arteritis. There is a mild to moderate normochromic anaemia, thrombocytosis is common, and the white cell count is usually normal. The erythrocyte sedimentation rate (ESR) is elevated and frequently much raised, over 100 mm/h, although in occasional cases it may be completely normal. C-reactive protein is often raised and some studies have claimed it is more sensitive than ESR. Long-term studies have shown that ESR is as good an indication of the acute-phase response (Kyle and Hazleman 1989). There may be a decrease in albumin concentration and increase in a_2-globulins, fibrinogen, and other acute-phase reactants.

Serum immunoglobulins are usually normal although they may be raised in severe cases. Complement levels are normal. Tests for autoantibodies and rheumatoid factor are negative. Antibodies to intermediate filaments and to neutrophil cytoplasmic antigens have been described (see above).

Increases in alkaline phosphatase can occur and less frequently there is increase in transaminases with prolonged prothrombin time. Liver biopsies are generally normal although granulomatous hepatitis has been described. Muscle enzymes and electromyography are normal. Muscle histology is normal, although ultrastructural studies of muscle have shown crystalline inclusions in the mitochondria. Thyroid function tests are usually normal although they may reflect changes due to non-thyroidal illness. Synovial fluid analysis is compatible with an inflammatory state with poor mucin clot and leucocyte counts ranging from 1000 to 20 000/mm^3 with a majority of polymorphonuclear leucocytes. Biopsies have shown a lymphocytic synovitis.

Levels of factor VIII/von Willebrand factor have been found to be elevated in polymyalgia rheumatica/giant cell arteritis. Values are highest in giant cell arteritis but do not closely parallel the ESR and may reflect endothelial damage. Nordberg *et al*. (1991) studied 53 patients with giant cell arteritis by serial analysis of von Willebrand factor antigen and plasminogen activator inhibitor. The concentration of von Willebrand factor slowly decreased and reached control range 18 months after diagnosis. However, levels failed to correlate with clinical features and results of temporal biopsy, and failed to predict flare-up of disease and vascular complications. The activity of plasminogen activator inhibitor was no different from levels in age-matched controls. Von Willebrand factor antigen is present along the lamina elastica of the arterial wall in giant cell arteritis but absent from arterial biopsy in pure polymyalgia rheumatica (Olsson *et al*. 1990). Fibrinolysis has also been described in these conditions (Grau *et al*. 1984). Plasma fibronectin was not significantly different from controls in a study by Puccetti *et al*. (1987).

Interleukin 6 (**IL-6**) is a multifunctional cytokine responsible for secretion of acute-phase proteins by hepatocytes. The level of serum IL-6 was raised in 15 patients with untreated polymyalgia rheumatica/giant cell arteritis and declined following therapy (Dasgupta *et al*. 1990; Roche *et al*. 1993). However, in seven patients it remained elevated above baseline at 6 months despite optimum treatment and remission of symptoms. This suggests ongoing disease activity despite suppression of acute-phase response and symptoms. As stimulated endothelial cells are a potent source of IL-6, elevated IL-6 levels may reflect endothelial damage. However, immunohistochemistry and in situ hybridization have shown that IL-6 is produced predominantly by macrophages and some fibroblasts in the media of the inflamed

artery and by endothelial cells (Emilie *et al*. 1994). Paradoxically, the very effectiveness of glucocorticoids in returning the acute-phase response to normal pose a problem in deciding how quickly and by how much steroids should be decreased. Serial estimation of a $_1$-antichymotrypsin may be helpful as it does not return to normal until some 18 months after institution of glucocorticoid therapy (Pountain *et al*. 1994).

A study of T-cell subsets in peripheral blood showed selective depletion of CD8+ cells, which remained depressed after 1 year of treatment and only becomes normal after 2 years of treatment (Dasgupta *et al*. 1989).

Blood flow has been studied in temporal arteries (Brunholz and Mullen 1988) and in the central retinal, short posterior ciliary, and ophthalmic arteries (Ho *et al*. 1994) by Doppler ultrasonography. Abnormal flow (reduced, reversed, or alternating flow) was seen. This procedure may be useful in diagnosis and management as some of the waveforms were not seen in non-arteritic optic neuropathy.

Diagnosis

The diagnosis of polymyalgia rheumatica is based essentially on clinical symptoms and signs and it is significant that on follow-up in two large series the diagnosis was revised in 20 to 25 per cent of cases. Thus, even after the diagnosis has been made and treatment started, the possibility of other diseases should be kept in mind. Table 6 lists the conditions which can present with polymyalgic symptoms. The following conditions cause greatest difficulty.

1. Rheumatic disease in the elderly
   Rheumatoid arthritis
   Systemic lupus erythematosus
2. Inflammatory myopathies
3. Endocrinopathy
   Hypothyroidism
   Hyperthyroidism
4. Neoplasia
   Carcinoma
   Multiple myeloma
5. Occult sepsis
6. Bilateral shoulder capsulitis especially with diabetes
7. Osteoarthritis
8. Depressive illness
9. Parkinsonism

Table 6 Conditions that can present with polymyalgic symptoms

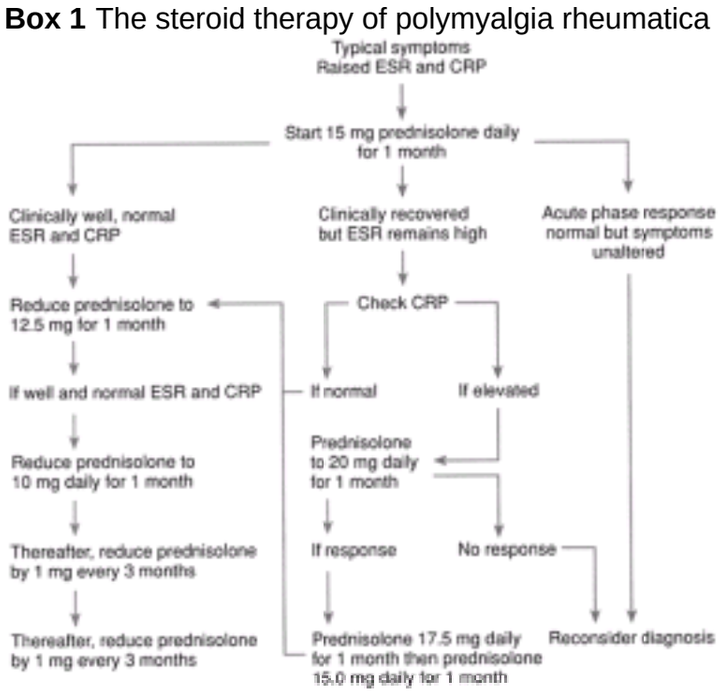
1. Rheumatoid arthritis. It may be impossible to distinguish polymyalgic onset of rheumatoid arthritis from true polymyalgia rheumatica. High titres of rheumatoid factor, prominent articular symptoms, and partial response to low-dose steroids may distinguish the former, while polymyalgia rheumatica is associated with low-grade synovitis and low titre or, more commonly, absent rheumatoid factor.
2. Neoplasia. Although it is still debatable whether there is true association of neoplasia with polymyalgia rheumatica, it is important not to mistake the generalized musculoskeletal aching, systemic symptoms, weight loss, and raised ESR of occult cancer for polymyalgia rheumatica. A thorough physical examination is mandatory and although it is not necessary to embark on detailed investigation, tests such as a protein electrophoretic strip and a chest radiograph should be carried out.
3. Inflammatory myopathies may present with muscle pains. However, in polymyalgia rheumatica there should be no objective muscle weakness. In addition creatine kinase levels and electromyography are normal.
4. Chronic sepsis may simulate polymyalgia rheumatica especially if associated with low-grade pyrexia. Blood cultures should be obtained if fever is present.
5. Hypothyroidism and less frequently hyperthyroidism may present with muscle pain and generalized stiffness. Thyroid function should be checked although abnormalities may merely reflect changes from a non-thyroidal illness.
6. Parkinsonian rigidity can occasionally mislead the unwary, although the condition is easily distinguished once it is considered and sought.
7. Bilateral capsulitis of shoulders, such as is seen in diabetes, can mimic polymyalgia rheumatica but should easily be differentiated by the restriction of passive movements. Similarly, osteoarthritis of hips and shoulders can cause difficulty but can be excluded by radiographs.
8. Depressive illnesses can frequently present with myalgic symptoms and should therefore be considered.

The limitations to the diagnosis of polymyalgia rheumatica are that there are no specific clinical features or any specific laboratory tests for this disease. The acute-phase response and its response to steroids is helpful but should be interpreted with caution as a raised ESR, whatever the cause, is likely to decrease with high-dose steroids (e.g. 20 mg or more daily). What distinguishes polymyalgia rheumatica is a complete and quick response (within 4 days) to low-dose steroids (e.g. 10 to 15 mg daily).

CD8+ lymphocytes in the peripheral blood may help in providing a marker of disease activity that is independent of the acute-phase response. We have found that even after a year CD8 cell numbers were low, probably reflecting disease activity. Most other conditions with similar presentations are accompanied by normal levels of CD8 so this test may help in distinguishing the condition, although further comparative studies are required to establish the specificity and sensitivity. A study of CD8 levels in 108 patients undergoing temporal artery biopsies suggests a positive predictive value of 85 per cent for either biopsy-positive giant cell arteritis or polymyalgia rheumatica while the likelihood of absence of giant cell arteritis or giant cell arteritis syndrome with normal CD8+ values was 60 per cent (Elling *et al*. 1990).

Treatment (Box 1,Box 2)

Box 1 The steroid therapy of polymyalgia rheumatica



Box 2 Summary

Elevated ESR in the face of a normal CRP suggests that inflammation is controlled

Give calcium and vitamin D supplements to prevent osteoporosis

Only some 50% of patients have ceased prednisolone by 2 years

Follow same schedule for giant cell arteritis but start therapy at 30 mg prednisolone daily

Polymyalgia rheumatic and giant cell arteritis are indications for corticosteroid treatment. Dramatic symptomatic response usually occurs within 48 h of introducing steroids. Despite this early gratifying response, continuing skilled supervision is necessary for several reasons, including the need to monitor for disease complications and the complications of steroid therapy. Correct assessment of disease activity is paramount in preventing steroid-related morbidity.

Polymyalgia rheumatica responds to low-dose steroids and treatment should be started with prednisolone at 15 mg daily as a single morning dose. There are no universal guidelines to tapering steroid dosage although, according to a recent report, prednisolone at 10 mg/day is required in the second month to prevent an unacceptable relapse rate (Kyle and Hazleman 1990). Further reductions below 10 mg are recommended in 1 mg decrements every 4 to 6 weeks.

Behn *et al*. (1983) in a prospective study of 176 patients reported that 10 mg of prednisolone is sufficient for control of symptoms of polymyalgia rheumatica. Others (Healey and Wilske 1977; Spiera and Davison 1978) have concurred but Kyle and Hazleman (1989), in a controlled prospective study, found that 13 of 20 patients with polymyalgia rheumatica relapsed on an initial dose of 10 mg/day, and 15 to 20 mg of prednisolone was required for adequate disease control.

The rationale for a lower dose of steroids than used hitherto in polymyalgia rheumatica/giant cell arteritis is based on several studies showing that major and minor steroid side-effects are quite common and their occurrence is related to initial dose, treatment duration, and, most critically, to cumulative dose. It is important to use the lowest satisfactory dose which contains symptoms and not merely be guided by the ESR. It is wrong to use the ESR as a chief guide to therapy since active disease may exist with a normal ESR and, on the other hand, it may be impossible to attain a normal ESR without an excessively high steroid dosage. Laboratory measures independent of the acute-phase response, such as the percentage of circulating CD8+ T cells and serum a $_1$-antichymotrypsin, may be helpful in long-term disease monitoring.

Other routes of steroid therapy may also be useful in polymyalgia rheumatica. In a pilot study, 120 mg of intramuscular methylprednisolone acetate administered initially for 3 weeks and then monthly at reducing doses was successful in induction and maintenance of disease remission (Dasgupta *et al*. 1991). Cumulative steroid dosage was 40 to 60 per cent lower than in a conventional regimen. This route has the advantage of ensuring compliance, and may not suppress the hypothalamic–pituitary axis. Further controlled studies are needed to compare the efficacy of this steroid regimen with that of oral prednisolone, as well as incidence of steroid side-effects with particular emphasis on bone mineral density.

Other agents such as dapsone, azathioprine, and d-penicillamine have been used in polymyalgia rheumatica/giant cell arteritis. Use of dapsone has been associated with very serious haematological complications such as agranulocytosis and haemolytic anaemia. Azathioprine has been anecdotally reported to have been of benefit and a controlled study showed a small but significant response (De Silva 1986).

Non-steroidal anti-inflammatory drugs have been used in polymyalgia rheumatica but should not be the first choice unless there is a major contraindication to steroid therapy.

Chapter References

Achkar, A.A., Lie, J.T., Hunder, G.G., O'Fallon, W.M., and Gabriel, S.E. (1994). How does previous corticosteroid treatment affect the biopsy findings in giant cell (temporal) arteritis? *Annals of Internal Medicine*, **120**, 987–92.

Andersson, R., Jonsson, R., Tarkowski, A., Bengtsson, B., and Malmvall, B.E. (1987). T cell subsets and expression of immunological activation markers in the arterial walls of patients with giant cell arteritis. *Annals of the Rheumatic Diseases*, **46**, 915–23.

Andersson, R., Hansson, G.K., Soderstrom, T., Jonsson, R., Bengtsson, B.A., and Hordborg, E. (1988). HLA-DR expression in the vascular lesion and circulating T lymphocytes of patients with giant cell arteritis. *Clinical and Experimental Immunology*, **73**, 82–7.

Armstrong, R.D., Behn, A., Myles, A., Panayi, G.S., and Welsh, K. (1983). Histocompatibility antigens in polymyalgia rheumatica and giant cell arteritis. *Journal of Rheumatology*, **10**, 659–61.

Bacon, P.A., Doherty, S.M., and Zuckerman, A.J. (1975). Hepatitis B antibody in polymyalgia rheumatica. *Lancet*, **ii**, 476.

Behn, A.R., Perera, T., and Myles, A.B. (1983). Polymyalgia rheumatica and corticosteroids: how much for how long? *Annals of the Rheumatic Diseases*, **42**, 374–8.

Bengtsson, B.A. and Malmvall, B.E. (1982). Giant cell arteritis. *Acta Medica Scandinavica*, Suppl. **658**, 1–102.

Bridgeford, P.H., Lowenstein, M., Bocanegra, T.S., Vasey, F.B., Germain, B.F., and Espinoza, L.R. (1980). Polymyalgia rheumatica and giant cell arteritis: histocompatibility typing and hepatitis B infection studies. *Arthritis and Rheumatism*, **23**, 516–8.

Brunholz, L.C. and Mullen, H.R. (1988). Doppler sonography findings in arteritis temporalis. *Ultraschall in der Medizin*, **9**, 232–6.

Cats, H.A., Tervaert, J.W., van Wijk, R., Limburg, P.C., and Kallenberg, C.G. (1993). Anti-neutrophil cytoplasmic antibodies in giant cell arteritis and polymyalgia rheumatica. *Advances in Experimental Medicine and Biology*, **336**, 363–6.

Chemnitz, J., Christensen, B.C., Christoffersen, P., Garbarsch, C., Hansen, T.M., and Lorenzen, I. (1987). Giant cell arteritis. Histological, immunohistochemical and electron microscopic studies. *Acta Pathologica et Microbiologica Immunologia Scandinavica* (A), **95**, 251–62.

Chmelewski, W.L., McKnight, K.M., Agudelo, C.A., Wise, C.M. (1992). Presenting features and outcome in patients undergoing temporal artery biopsy. A review of 98 patients. *Annals of Internal Medicine*, **152**, 1690–5.

Chou, C.T. and Schumacher, H.R. (1984). Clinical and pathological studies of synovitis in polymyalgia rheumatica. *Arthritis and Rheumatism*, **27**, 1107–17.

Chuang, T.Y., Hunder, G.G., Ilstrup, D.M., and Kurland, L.T. (1982). Polymyalgia rheumatica: a 10 year epidemiologic and clinical study. *Annals of Internal Medicine*, **97**, 672–80.

Cid, M.C. (1989). Immunohistochemical analysis of lymphoid and macrophage cell subsets and their immunological activation markers in temporal arteritis. Influence of corticosteroid treatment. *Arthritis and Rheumatism*, **32**, 884–93.

Cid, M.C. *et al*. (1988). Polymyalgia rheumatica: a syndrome associated with polymyalgia syndrome. *Arthritis and Rheumatism*, **31**, 678–81.

Cimmino, M.A., Grazi, G., Balistreri, M., and Accardo, S. (1993). Increased prevalence of antibodies to adenovirus and respiratory syncytial virus in polymyalgia rheumatica. *Clinical and Experimental Rheumatology*, **11**, 309–13.

Dasgupta, B., Duke, O., Kyle, V., Macfarlane, D.G., Hazleman, B.L., and Panayi, G.S. (1987). Antibodies to intermediate filaments in polymyalgia rheumatica and giant cell arteritis: a sequential

study. *Annals of the Rheumatic Diseases*, **46**, 746–9.

Dasgupta, B., Duke, O., Timms, A.M., and Panayi, G.S. (1989). Selective depletion and activation of CD8+ lymphocytes from peripheral blood of patients with polymyalgia rheumatica and giant cell arteritis. *Annals of the Rheumatic Diseases*, **48**, 307–11.

Dasgupta, B., Gruindy, E., and Stainer, E. (1990). Hypothyroidism in polymyalgia rheumatica and giant cell arteritis: lack of any association. *British Medical Journal*, **301**, 96–7.

Dasgupta, B., Gray, J., Fernandes, L., and Olliff, C. (1991). Treatment of polymyalgia rheumatica with intramuscular depot methylprednisolone. *Annals of the Rheumatic Diseases*, **50**, 942–5.

Demaine, A.G., Vaughan, R.W., Behn, A.R., Mules, A.B., Panayi, G.S., and Welsh, K.I. (1983). Immunoglobulin (Gm) allotype frequencies in patients with giant cell arteritis and polymyalgia rheumatica. *Journal of Immunogenetics*, **10**, 343–8.

De Silva, M. and Hazleman, B.L. (1986). Azathiprine in giant cell arteritis: a double-blind study. *Annals of the Rheumatic Diseases*, **45**, 136–8.

Elling, P., Olson, A., and Elling, H. (1990). CD8+ T lymphocyte subset in giant cell arteritis and related disorders. *Journal of Rheumatology*, **17**, 225–7.

Ellis, M.E. and Ralston, S. (1983). The ESR in the diagnosis and management of polymyalgia rheumatica/giant cell arteritis syndrome. *Annals of the Rheumatic Diseases*, **43**, 163–70.

Emilie, D. *et al*. (1994). Production of interleukin 6 by granulomas of giant cell arteritis. *Human Immunology*, **39**, 17–24.

Fauchald, P., Rygvold, I., and Oystese, B. (1972). Temporal arteritis and polymyalgia rheumatica: clinical and biopsy findings. *Annals of Internal Medicine*, **77**, 842–52.

Fernandez-Herlihy, L. (1988). Temporal arteritis: clinical aids to diagnosis. *Journal of Rheumatology*, **15**, 1797–801.

Gonzalez, E.B., Varner, W.T., Lisse, J.R., Daniels, J.C., and Hokanson, J.A. (1989). Giant cell arteritis in the southern United States. An 11 year retrospective study from the Texas Gulf coast. *Archives of Internal Medicine*, **149**, 1561–5.

Grau, R.G., Kassan, S.S., Franks, J.J., Kaplan, H., Walker, S.H., and Tan, E.M. (1984). Fibrin(ogen)olysis in polymyalgia rheumatica and temporal arteritis: preliminary findings of association with disease activity. *Annals of the Rheumatic Diseases*, **43**, 721–4.

Hansen, J.A., Healey, L.A., and Wilske, K.R. (1985). Association between giant cell (temporal arteritis) and HLA-Cw3. *Human Immunology*, **13**, 193–8.

Healey, L.A. and Wilske, K.R. (1977). Manifestations of giant cell arteritis. *Medical Clinics of North America*, **61**, 621–70.

Hunder, G.G. and Allen, G.L. (1973). The relationship between polymyalgia and temporal arteritis. *Geriatrics*, **28**, 134–42.

Ho, A.C., Sergott, R.C., Regillo, C.D., Savino, P.J., Lieb, W.E., Flaharty, P.M., and Bosley, T.M. (1994). Color Doppler haemodynamics of giant cell arteritis. *Archives of Ophthalmology*, **112**, 938–45.

Hunder, G.G. *et al*. (1990). The American College of Rheumatology 1990 criteria for the classification of giant cell arteritis. *Arthritis and Rheumatism*, **33**, 1122–8.

Huston, K.A. *et al*. (1978). Temporal arteritis: a 25 year epidemiologic, clinical and pathologic study. *Annals of Internal Medicine*, **88**, 162–7.

Jones, J.G. and Hazleman, B.L. (1981). The prognosis and management of polymyalgia rheumatica. *Annals of the Rheumatic Diseases*, **40**, 1–5.

Klein, R.G., Hunder, G.G., Stanson, A.W., and Sheps, S.G. (1975). Large artery involvement in giant cell (temporal) arteritis. *Annals of Internal Medicine*, **83**, 806.

Koski, J.M. (1992). Ultrasonographic evidence of synovitis in axial joints in patients with polymyalgia rheumatica. *British Journal of Rheumatology*, **31**, 201–3.

Kyle, V. and Hazleman, B.L. (1989). Treatment of polymyalgia rheumatica and giant cell arteritis. I. Steroid regimens in the fist two months. *Annals of the Rheumatic Diseases*, **48**, 658–61.

Kyle, V. and Hazleman, B.L. (1990). Stopping steroids in polymyalgia rheumatica and giant cell arteritis. *British Medical Journal*, **300**, 344–5.

Kyle, V., Hazleman, B.L., and King, R.H. (1984). Polymyalgia rheumatica giant cell arteritis in husband and wife. *Clinical Rheumatology*, **3**, 395–6.

Lie, J.T. (1987). The classification and diagnosis of vasculitis in large and medium sized blood vessels. *Pathologische Annalen*, **22**, 125–62.

Lie, J.T. *et al*. (1990). Illustrated histopathologic classification criteria for selected vasculitis syndromes. *Arthritis and Rheumatism*, **33**, 1074–87.

Love, D.G. *et al*. (1986). Temporal arteritis in blacks. *Annals of Internal Medicine*, **105**, 387–9.

Malmvall, B.E. and Bengston, B.A. (1978). Giant cell arteritis: clinical feature and involvement of different organs. *Scandinavian Journal of Rheumatology*, **7**, 154–8.

Mashiah, A., Mashiah, T., and Hod, I. (1988). Quick diagnosis of temporal arteritis. *Israel Journal of Medical Science*, **24**, 188–9.

Monteagudo, I., Corrigall, V., Panayi, G.S., and Hazleman, B.I. (1994). Antibodies to intermediate filaments in polymyalgia rheumatica/giant cell arteritis: do they reflect the underlying disease activity rather than the acute phase response? *Annals of the Rheumatic Diseases*, **53**, 150.

Moss, R.R. and Soukop, M. (1988). Polymyalgia rheumatica in a sibling pair. Two case reports and a brief review of the literature. *Scottish Medical Journal*, **33**, 342–3.

Myles, A.B., Pereira, P., and Ridley, M.G. (1992). Prevention of blindness in giant cell arteritis by corticosteroids treatment. *British Journal of Rheumatology*, **31**, 103–5.

Ninet, J. *et al*. (1987). Distribution of HLA-DR antigens in unrelated giant cell arteritis. *Presse Medicale*, **16**, 1725–8.

Nishioka, J. *et al*. (1986). Symptomatological review of polymyalgia rheumatica in Japan: induction of diagnostic criteria for Japanese patients. *Japanese Journal of Rheumatology*, **1**, 113–21.

Nordberg, E., Anderson, R., Tengborn, L., Eden, S., and Bingtsson, B.A. (1991). Von Willebrand factor antigen and plasminogen activator inhibitor in giant cell arteritis. *Annals of the Rheumatic Diseases*, **50**, 316–20.

O'Duffy, J.D., Wahner, H.W., and Hunder, G.G. (1976). Joint imaging in polymyalgia rheumatica. *Mayo Clinic Proceedings*, **51**, 519–24.

Olsson, A., Elling, P., and Elling, H. (1990). Serological and immunohistochemical determination of von Willebrand factor antigen in serum and biopsy specimens from patients with arteritis temporalis and polymyalgia rheumatica. *Clinical and Experimental Rheumatology*, **8**, 55–8.

Ponge, T., Barrier, J.H., Grolleau, J.-Y., Ponge, A., Vlasak, A.M., and Cottin, S. (1988). The efficacy of selective unilateral temporal artery biopsy versus bilateral biopsies for diagnosis of giant cell arteritis. *Journal of Rheumatology*, **15**, 997–1000.

Pountain, G.D., Calvin, J., and Hazleman, B.L. (1994). a $_1$-Anticymotrypsin, C-reactive protein and erythrocyte sedimentation rate in polymyalgia rheumatica and giant cell arteritis. *British Journal of Rheumatology*, **33**, 550–4.

Puccetti, L. *et al*. (1987). Plasma levels of fibronectin in polymyalgia rheumatica/giant cell arteritis. *Rheumatology International*, **7**, 249–54.

Richardson, J.E., Gladman, D.D., Fam, A., and Keystone, E.C. (1987). HLA-DR4 in giant cell arteritis: association with polymyalgia syndrome. *Arthritis and Rheumatism*, **30**, 1293–7.

Robb-Nicholson, C. *et al*. (1988). Diagnostic value of the history and examinations in giant cell arteritis: a clinical pathological study of 81 temporal artery biopsies. *Journal of Rheumatology*, **15**, 1793–6.

Roche, N.E., Fulbright, J.W., Wagner, A.D., Hunder, G.G., Goronzy, J.J., and Weyand, C.M. (1993). Correlation of interleukin-6 production and disease activity in polymyalgia rheumatica and giant cell arteritis. *Arthritis and Rheumatism*, **36**, 1286–94.

Sakkas, L.J., Loqueman, N., Panayi, G.S., Myles, A.B., and Welsh, K.I. (1990). Immunogenetics of polymyalgia rheumatica. *British Journal of Rheumatology*, **29**, 331–4.

Salvarani, C. *et al*. (1991). Epidemiologic and immunogenetic aspects of polymyalgia rheumatica and giant cell arteritis in Northern Italy. *Arthritis and Rheumatism*, **34**, 351–6.

Salvarani, C. *et al*. (1992). Soluble interleukin 2 receptors in polymyalgia rheumatica/giant cell arteritis. Clinical and laboratory correlations. *Journal of Rheumatology*, **19**, 1100–6.

Schaufelberger, C., Stemme, S., Andersson, R., and Hansson, G.K. (1993). T lymphocytes in giant cell arteritic lesions are polyclonal cells expressing ab type antigen receptors and VLA-1 integrin

receptors. *Clinical and Experimental Immunology*, **91**, 421–8.

Spiera, T.T. and Davison, S. (1978). Long term follow up of polymyalgia rheumatica. *Mount Sinai Journal of Medicine*, **45**, 225–9.

Stuart, R.A. (1989). Temporal artery biopsy in suspected temporal arteritis: a five year study. *New Zealand Medical Journal*, **102**, 431–3.

Vaith, P., Rother, E., Vogt, A., and Peter, H.H. (1988). Polymyalgia rheumatica following *Borrelia* infection. *Immunitaet und Infekt.*, **16**, 71–3.

Vilaseca, J., Gonzalez, A., Cid, M.C., Lopez-Vivancos, J., and Ortega, A. (1987). Clinical usefulness of temporal artery biopsy. *Annals of the Rheumatic Diseases*, **46**, 282–5.

Wagner, A.D., Goronzy, J.J., and Weyand, C.M. (1994). Functional profile of tissue-infiltrating and circulating CD68+ cells in giant cell arteritis. Evidence for two components of the disease. *Journal of Clinical Investigation*, **94**, 1134–40.

Wells, K.K., Folberg, R., Goeken, J.A., and Kemp, J.D. (1989). Temporal artery biopsies. Correlation of light microscopy and immunofluorescence microscopy. *Ophthalmology*, **96**, 1058–64.

Weyand, C.M., Hunder, N.N., Hicok, K.C., Hunder, G.G., and Goronzy, J.J. (1994 *a*). HLA-DRB1 alleles in polymyalgia rheumatica, giant cell arteritis and rheumatoid arthritis. *Arthritis and Rheumatism*, **37**, 514–20.

Weyand, C.M., Schonberger, J., Oppitz, U., Hunder, N.N., Hicok, K.C., and Goronzy, J.J. (1994 *b*). Distinct vascular lesions in giant cell arteritis share identical T cell clonotypes. *Journal of Experimental Medicine*, **179**, 951–60.

Wiseman, P., Stewart, K., and Rai, G.S. (1989). Hypothyroidism in polymyalgia rheumatica and giant cell arteritis. *British Medical Journal*, **298**, 647–9.

5.11.6 Large vessel vasculitis

Richard A. Watts and David G. I. Scott

Introduction

Systemic vasculitis may be classified by the size of the predominant vessel involved and presumed aetiology (Scott and Watts 1994). Large vessels may be defined as the aorta and its largest branches to the major body regions—subclavian, carotid, and femoral arteries (Jeanette *et al*. 1994). The primary vasculitides are those in which there is no known precipitating cause, whereas the secondary vasculitides occur in a setting of infection or pre-existing autoimmune disease such as rheumatoid arthritis. The classification of large vessel vasculitis is given in Table 1. The major causes of primary large vessel vasculitis—Takayasu's arteritis and giant cell arteritis involve predominately large vessels, occasionally medium-sized vessels (i.e. renal, hepatic, coronary, mesenteric arteries), but almost never involve vessels smaller than arteries. Similarly medium and small vessel diseases such as Wegener's granulomatosis rarely involve large vessels (Davenport *et al*. 1994) By contrast some causes of secondary large vessel vasculitis such as rheumatoid arthritis have a broader spectrum of vessel involvement and frequently involve medium and small vessels.

Primary	Secondary
Takayasu's arteritis	Infection
Giant cell arteritis	Bacterial
Buerger's disease	Fungal
Cogan's syndrome	Mycobacterial
Wegener's granulomatosis	Spirochaetal
Isolated CNS angiitis	Rheumatoid arthritis
	Seronegative spondylarthropathy
	Systemic lupus erythematosus
	Juvenile chronic arthritis
	Kawasaki disease
	Sarcoidosis
	Rheumatic fever
	Relapsing polychondritis

CNS=central nervous system

Table 1 Causes of large vessel vasculitis

Giant cell arteritis is probably the most common of the primary systemic vasculitides in North America and Western Europe (Churg 1993). The estimated annual incidence in the United Kingdom is 13 per million adults and 42 per million individuals over the age of 60 (Jonasson *et al*. 1979). In Sweden, the annual incidence of giant cell arteritis proved by biopsy is 55 per million adults and 168 per million adults aged over 50 years (Bengtsson and Malmvall 1981), whilst in Iceland the annual incidence in individuals aged over 50 years is 270 per million (Baldursson *et al*. 1994). Takayasu's arteritis is much less common with an estimated annual incidence in Sweden of 0.8 per million (Waern *et al*. 1983). Systemic rheumatoid vasculitis which can involve large vessels has an annual incidence in the United Kingdom of 12.6 per million adults (Watts *et al*. 1994). However, large vessel involvement is uncommon occurring in less than 5 per cent of cases (Luqmani *et al*. 1994). There is no data on the incidence of other types of large vessel vasculitis. Data from the Norwich Vasculitis Register suggests that in the United Kingdom large vessel arteritis is rare. Over a six-year period (1988 to 1994) only 6 cases of large vessel vasculitis have been registered out of a total of 200 cases of systemic vasculitis (Scott and Watts, unpublished data). These 6 cases comprise 3 with arteritis complicating giant cell arteritis, 2 with brachial arteritis and one with infective aortitis.

The aorta is of mesenchymal origin, any injury to the aorta, whether of infectious, toxic, traumatic or immunological aetiology, results in similar morphological appearances (Lande and Berkman 1976). In the acute phase of most types of aortitis, inflammation is restricted to the media and adventitia without involvement of the intima. During the healing or chronic phase of aortitis, the cellular infiltrate resolves and is replaced by collagen, which retracts resulting in a 'wrinkled' or characteristic 'tree-bark' appearance of the aorta. Fibrosis leads to smooth tapering stenosis or even occlusion of the lumen. Sites of inflammation and intimal damage act as focal points for the development of accelerated atherosclerosis.

P>In Takayasu's arteritis the large elastic arteries are primarily involved, whilst in giant cell arteritis the changes are seen most often in the medium-sized muscular arteries which possess well developed internal and external elastic laminae (e.g. temporal artery). In 10 to 15 per cent of cases of giant cell arteritis there is involvement of the large elastic arteries (Hunder 1986).

The clinical features of large vessel arteritis are protean. Arch aortitis (e.g. Takayasu's arteritis) presents with diminished or absent peripheral pulses, claudication, bruits, hypertension and heart failure, often in the presence of a systemic illness and acute phase response. Occlusion or stenosis of the carotid or vertebral arteries leads to transient or permanent neurological deficits. Infective arteritis by contrast usually presents with aneurysm formation rather than stenosis.

The differential diagnosis of aortitis includes atherosclerosis, inflammatory aneurysms of the abdominal aorta, and aortic dissection. These conditions occur predominantly in the elderly and are not associated with an acute phase response. In younger patients abdominal aneurysms may be inflammatory and associated with a raised erythrocyte sedimentation rate. They are often only recognized at surgery (Rijbroek *et al*. 1994). Inflammatory abdominal aortic aneurysms occur in the infrarenal aorta and are associated with atherosclerosis. There is no evidence of infection or other inflammatory process. In peripheral arteries the differential

diagnosis includes atherosclerosis, fibromuscular dysplasia and thromboangiitis obliterans (Buerger's disease). Fibromuscular dysplasia affects typically the renal arteries, although other muscular arteries can be involved. Buerger's disease attacks middle and small arteries of the extremities leading to peripheral gangrene. It occurs in smokers. Inflammatory or infective angiitis must be considered in patients with clinical evidence of arteritis and an acute phase response. Histology of resected arteries will show evidence of arteritis or infection. Resected specimens should be cultured whenever an infective aetiology is considered.

The assessment of patients with suspected large vessel arteritis should include: assessment of acute phase response, disease extent and aetiology (Table 2). Total aortography is essential to establish the extent of disease in the aorta and major vessels. Increasingly, magnetic resonance imaging is being substituted for conventional arteriography. Temporal artery biopsy is useful in older patients with a large vessel arteritis, even if there is no clinical evidence of temporal arteritis.

Non-specific	
	Full blood count
	Acute phase response (erythrocyte sedimentation rate, plasma viscosity, C-reactive protein)
Disease extent	
	Chest radiograph
	Aortography (including major branches)
	Magnetic resonance imaging
	Echocardiography
Organ involvement	
	Creatinine
	Urinalysis
Aetiology	
	Blood cultures
	VDRL test
	Temporal artery biopsy
	Rheumatoid factor
	Antinuclear antibody
	Antineutrophil cytoplasmic antibody
	Anticardiolipin antibody

Table 2 Assessment of patients with suspected large vessel arteritis

Antineutrophil cytoplasmic antibodies (ANCA) are associated with systemic vasculitis—in particular Wegener's granulomatosis and microscopic polyangiitis. They are not associated with large vessels vasculitis of either inflammatory or infective origin (Churg 1993).

Treatment of inflammatory large vessel arteritis is with corticosteroids and in resistant cases immunosuppressive agents. Infectious arteritis should be treated with appropriate antibiotics. Surgical resection of aneurysms or bypass of stenotic lesions may be required.

Takayasu's arteritis

Takayasu's arteritis is a chronic granulomatous panarteritis affecting large elastic arteries such as the aorta and its major branches, and less frequently the pulmonary arteries.

In 1908 Takayasu presented to a clinical meeting the case of a 21-year-old woman with an unusual arteriovenous malformation in the retina associated with blindness due to cataract formation (Takayasu 1908). Colleagues pointed out the association between ocular abnormalities and the absence of radial pulses. Earlier descriptions include those by Savory in 1856 and Morgagni in 1761 (DiGiacomo 1984). Morgagni described a 40-year-old woman with a 6-year history of absent radial pulses who died in pulmonary oedema with an ectatic proximal aorta, lower thoracic aortic stenosis, and cardiac hypertrophy secondary to aortic incompetence.

There are a number of synonyms that have previously been widely used to describe patients presenting with evidence of aortitis. These include: aortic arch syndrome, pulseless disease, idiopathic aortitis, stenosing aortitis, aortoarteritis and occlusive thromboarteriopathy.

Epidemiology

The epidemiology of Takayasu's arteritis is poorly documented. In the United States (Olmstead County) three cases were seen in the period 1971 to 1983, equivalent to an annual incidence of 2.6 per million (Hall *et al*. 1985). The annual incidence was estimated in Sweden (1969 to 1975) to be 0.8 per million (Waern *et al*. 1983). In Birmingham (United Kingdom) during a 6 year period the annual incidence was estimated to be 0.15 per million (Scott unpublished ata), whilst in the Norwich (United Kingdom) epidemiological study of systemic vasculitis no cases were seen in the period 1988 to 1994 (Watts *et al*. 1995). The prevalence is, however, higher in the Far East, Central and South America and the India subcontinent, but the annual incidence is unknown.

Takayasu's arteritis predominately affects woman (up to 90 per cent of cases; Lupi-Herrera *et al*. 1977; Ishikawa 1988; Kerr *et al*. 1994). The age of onset is in the second to third decade (Lupi-Herrera *et al*. 1977; Ishikawa 1988; Kerr *et al*. 1994). European series have reported a higher median age of onset (41 years) (Waern *et al*. 1983). Takayasu's arteritis does occur in children (Sharma *et al*. 1991) and in the recent American series they accounted for one third of cases (Kerr *et al*. 1994).

In rare cases Takayasu's arteritis has been associated with rheumatoid arthritis (Rush *et al*. 1986), seronegative spondylarthropathy (Magaro *et al*. 1988), systemic lupus erythematosus (Saxe and Altman 1992), juvenile chronic arthritis (Hall and Nelson 1986), adult Still's disease (Wilson *et al*. 1979) and inflammatory bowel disease (Achar and Al-Nahib 1986); whether these are true associations or chance occurrences is unknown.

Definition

Ishigawa proposed diagnostic criteria for Takayasu's arteritis in 1988 and suggested that an age of less than 40 years should be an obligatory criterion for the diagnosis (Ishikawa 1986). The diagnosis of the disease was based on one obligatory criterion (age of less than 40 years), two major criteria (left and right mid-subclavian artery lesions) and nine minor criteria (high erythrocyte sedimentation rate, common carotid tenderness, hypertension, aortic regurgitation or annuloaortic ectasia, and lesions of the pulmonary artery, left mid-carotid artery, distal brachiocephalic trunk, thoracic aorta and abdominal aorta). In addition to the obligatory criterion, the presence of two major or one major plus two or more minor criteria gave a high probability of Takayasu's arteritis. In a study of 96 patients with the disease, there was 84 per cent sensitivity. None of 12 patients with other aortic diseases met these criteria for the diagnosis of Takayasu's arteritis.

The American College of Rheumatology 1990 classification criteria for the diagnosis of Takayasu's arteritis are given in Table 3 (Arend *et al*. 1990). The disease is clearly distinguished from giant cell arteritis by age of onset. In this study, 63 patients with Takayasu's arteritis were compared with 744 patients with other forms of systemic vasculitis. A high erythrocyte sedimentation rate, carotid artery tenderness and/or hypertension lacked the specificity and sensitivity to differentiate accurately patients with Takayasu's arteritis from other forms of arteritis. The sensitivity of the criteria in the traditional table format (Table 3) was 90.5 per cent with a specificity of 97.8 per cent. A classification tree was also constructed with five of the same six criteria omitting limb claudication. This gave a sensitivity of 92.1 per cent and a specificity of 97.0 per cent.

Criterion	Definition
1. Age at disease onset	Development of symptoms or findings related to Takayasu's arteritis at age < 40 years
2. Claudication of extremities	Development and worsening of fatigue and discomfort in muscles of one or more extremity while in use, especially the upper extremities
3. Decreased brachial artery pulse	Decreased pulsation of one or both brachial arteries
4. BP difference	Difference of > 10 mmHg in systolic blood pressure between arms
5. Bruit over subclavian arteries or aorta	Bruit audible on auscultation over one or both subclavian arteries or abdominal aorta
6. Arteriogram abnormality	Arteriographic narrowing or occlusion of the entire aorta, its proximal branches, or large arteries in the proximal upper or lower extremities, not due to atherosclerosis, fibromuscular dysplasia, or similar causes; changes usually focal or segmental

Reproduced from Arend *et al*. (1990) with permission.
For purposes of classification, a patient shall be said to have Takayasu arteritis if at least three of these six criteria are present. The presence of any three or more criteria yields a sensitivity of 90.5 per cent and a specificity of 97.8 per cent. BP = blood pressure.

Table 3 American College of Rheumatology 1990 criteria for the classification of Takayasu's arteritis (traditional format)

The definition of Takayasu's arteritis was considered by the Chapel Hill Consensus conference on the nomenclature of the systemic vasculitides (Jeanette *et al*. 1994). It was defined as a 'granulomatous inflammation of the aorta and its major branches, usually occuring in patients aged less than 50 years'. They recognized that age is a useful discriminator between Takayasu's arteritis and giant cell arteritis.

Clinical features

The clinical features of Takayasu's arteritis have been reviewed recently in a series of patients from the United States (Table 4) and are broadly similar to previous series (Lupi-Herrera *et al*. 1977). The spectrum of presentation, disease severity, and pace of progression can often lead to a delay in diagnosis. The classic description of Takayasu's arteritis is of a triphasic illness characterized by pre-pulseless inflammation, followed by painful ischaemic vessels, and ending in 'burnt out disease' (Lupi-Herrera *et al*. 1977; Hall *et al*. 1985). More recent studies suggest that 20 per cent of patients have a monophasic illness not requiring therapy and that 57 per cent of patients never have constitutional symptoms (Kerr *et al*. 1994).

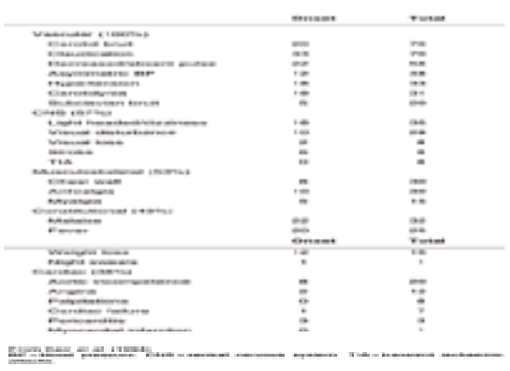


Table 4 Clinical features of Takayasu's arteritis

The hallmark of the disease is vascular ischaemic symptoms with bruit, claudication, or diminished pulses. Disease is often extensive involving the aorta and its branches. The most common vascular abnormality (more than 90 per cent of cases) is a bruit occurring most frequently in the carotid vessels (70 per cent) and in multiple sites in a third of cases. Claudication and a diminished or absent pulse is more common in the arms than legs. Carotidynia occurs in a third of patients and a similar proportion complain of lightheadedness or dizzyness. Hypertension occurs in 30 to 70 per cent of patients as a consequence of renal artery stenosis, which may be either unilateral or bilateral (Lupi-Herrera *et al*. 1977; Ishikawa 1988; Kerr *et al*. 1994).

Musculoskeletal symptoms occur in half of patients, with arthralgias in one third to one half of patients. Peripheral synovitis is less common occurring in 18 per cent of patients. Myalgia is found infrequently but can be severe.

Cardiac lesions occur in 40 per cent with aortic incompetence (20 per cent) resulting from aortitis, which may also involve the proximal coronary arteries. Congestive cardiac failure may result from aortic incompetence or systemic hypertension and accounts for the majority of fatal outcomes. Hypertension contributes to left ventricular hypertrophy and myocardial failure, but coronary arteritis results in myocardial infarction and fibrosis. Pulmonary artery involvement is common (50 per cent) with lesions localized to large and medium pulmonary vessels (Lupi-Herrera *et al*. 1977).

Cutaneous manifestations include erythema nodosum (8 per cent) which responds to treatment of the disease and probably reflects vasculitis rather than panniculitis. Leg ulcers, hypocomplementaemic vasculitis and pyoderma gangrenosum have all been described occurring in association with Takayasu's arteritis.

Neurological features occur in all stages of the disease and reflect carotid artery involvement. Non-specific symptoms such as headache and vertigo are seen in up to 90 per cent of cases (Churg 1993), but sensory loss and aphasia occur infrequently. Strokes are an important contributor to mortality. Visual disturbance occurs in 30 per cent of patients with either blurring of vision, which may be posturally dependent, diplopia or amaurosis fugax. Retinal disease with arteriovenous anastomosis occurs late in the disease. Visual disturbances are often related to systemic hypertension and 40 per cent of cases have fundal changes consistent with hypertension. Cataracts and neovascularization as described by Takayasu is now uncommon (Hall *et al*. 1985; Kerr *et al*. 1994).

Takayasu's arteritis may develop during childhood. In Eastern countries presentation is often with cardiac failure and hypertension in association with renal artery stenosis. Comparison of paediatric with adult onset patients in the United States suggests that systemic features are more common, whilst claudication is less common in childhood cases (Kerr *et al*. 1994).

Takayasu's arteritis should be considered in any young patient with limb or organ ischaemia, hypertension, or systemic illness. These patients should be investigated with full angiography (see below).

Assessment

Objective assessment of disease activity in systemic vasculitis is difficult and differentiation must be made between new lesions and scars from previous active disease. Generic assessment schemes such as the Birmingham Vasculitis Assessment Score (BVAS) may be utilized (Luqmani *et al*. 1994). Active disease was defined by Kerr and colleagues as new or worsening of two or more characteristic features. (Kerr *et al*. 1994). Complete resolution of all clinical features or their stabilization, in the setting of fixed vascular lesions, was indicative of remission. Clinical assessment of involved vessels may be difficult as symptoms of ischaemia will depend on the degree of stenosis and the adequacy of any collateral circulation. Measurement of blood pressure may be difficult in patients with diminished or absent pulses in several limbs.

Laboratory investigation

The erythrocyte sedimentation rate is often elevated in active disease. Serial measurements may correlate with disease activity, but not with histological extent (Kerr et al., 1994). Other surrogate measures such as von Willebrand factor and endothelin-1 have been reported to reflect disease activity, but have not been established as reliable markers (Woolf *et al*. 1987). Antineutrophil cytoplasmic antibodies are not associated with Takayasu's arteritis (Churg 1993). Levels of soluble adhesion molecules such as soluble intracellular adhesion molecule-1 (sICAM-1), soluble vascular adhesion molecule-1 (sVCAM-1), and soluble E-selectin may reflect endothelial activation and damage, and hence be used as markers of vascular damage. Levels of sICAM-1, sVCAM-1 and sELAM are elevated in other forms of systemic vasculitis (e.g. Wegener's granulomatosis and giant cell arteritis) (Carson *et al*. 1993) but there have not yet been any systematic studies in Takayasu's arteritis.

The diagnosis of Takayasu's arteritis has depended on arteriography to demonstrate the characteristic changes in the aorta and its major branches; long stenotic lesions occur in almost all patients (Fig. 1 and Fig. 2). Total arteriography is essential for angiographic diagnosis of the extent of the disease. Irregularity of the vessel wall occurs in one-third and aneurysms in 27 per cent of patients (Kerr *et al*. 1994). The distribution of lesions has been classified to reflect disease above and below the diaphragm—type I disease is localized to the aortic arch and its branches; type II involves the thoracic descending aorta and abdominal aorta but not the aortic arch; type III is a mixture of types I and II; type IV features aspects of types I, II, or III together with involvement of the pulmonary artery (Lupi-Herrera *et al*. 1977).

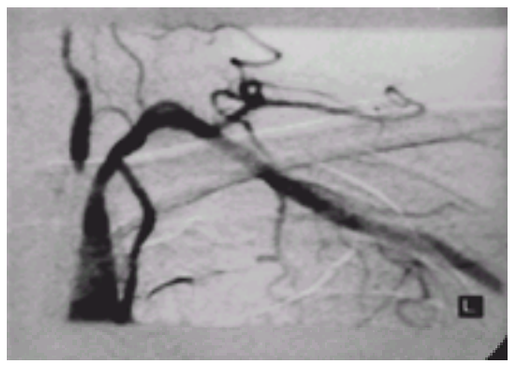


Fig. 1 Arch aortogram from a patient with aortic arch syndrome showing a long tapering stenosis of the left subclavian artery.

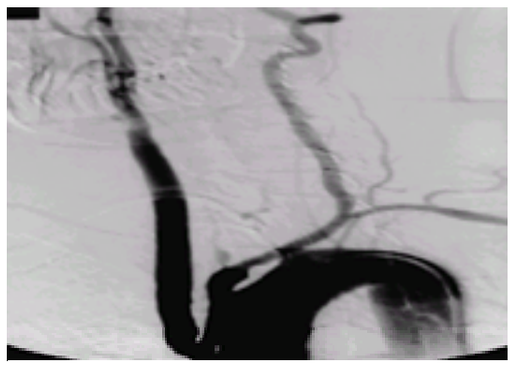


Fig. 2 Arch aortogram from a patient with Takayasu's arteritis. This shows (i) occlusion of the left subclavian at its origin, (ii) absence of the left common carotid artery, (iii) grafts *in situ* between the origin of the left common carotid and the left subclavian/vertebral artery and also between the aortic arch and right common carotid artery.

High resolution B-mode ultrasonography is more sensitive in detecting carotid lesions than angiography and correlates with the presence of bruits (Maeda *et al*. 1991). The ultrasonographic hallmark of Takayasu's arteritis is a diffuse thickening of the intima–media complex, the 'Macaroni sign'. Ultrasonography is limited by its inability to image the thoracic aorta and pulmonary arteries.

The role of magnetic resonance angiography (MRA) is still being evaluated. MRA is able to image the aorta and its major branches and demonstrate aortic wall thickening, stenosis and aneurysms. Improving resolution and newer scanning techniques are improving rapidly the ability of MRA to evaluate blood vessels and it is likely that it will become the technique of choice, replacing conventional contrast angiography (Link *et al*. 1993).

Aetiology and immunopathogenesis

The aetiopathogenesis of Takayasu's arteritis is not well understood. The association with mycobacterial infection has not been confirmed (Lupi-Herrera *et al*. 1977). Occasional examples of multicase families have been published, suggesting a common genetic or environmental factor. Immunopathological findings include hypergammaglobulinaemia, the presence of rheumatoid factors, anti-aorta antibodies, immune complexes (Lupi-Herrera *et al*. 1977) and anti-endothelial cell antibodies in peripheral blood (Sima *et al*. 1994). Antineutrophil cytoplasmic antibodies are not present. Elevated numbers of CD4+ T cells in peripheral blood, together with a decrease in CD8+ T cells and B cells have been reported but the significance of this finding is uncertain (Sagar *et al*. 1992). The inflammatory infiltrate conversely has an excess of CD8+ T cells, which are cytotoxic to endothelial cells in vitro (Scott *et al*. 1986).

In Japanese patients there is an association with the HLA haplotypes Bw52, Dw12, DR2, and DQW 1 (Dong *et al*. 1992; Kasuya *et al*. 1992) but this association has not been found in American patients (Hall *et al*. 1992). HLA-Bw52 positive Japanese patients have an association with heart disease and a worse prognosis with a higher incidence of aortic incompetence and left ventricular perfusion abnormalities (Kasuya *et al*. 1992)

A single case of Takayasu's arteritis following hepatitis B vaccination has been reported (Castrenasa-Isla *et al*. 1993). The patient developed erythema nodosum, arthralgia and upper limb claudication after receiving plasma-derived hepatitis B vaccine. Hepatitis B infection has not otherwise been associated with Takayasu's arteritis, unlike the well recognized association with polyarteritis nodosa.

Management

Medical treatment

Initial medical treatment for patients with active Takayasu's arteritis is with corticosteroids (prednisolone, 60 to 80 mg/day) and response rates of 20 to 100 per cent have been reported, with subsequent resolution of symptoms and stabilization of vascular abnormalities (Sen *et al*. 1963; Fraga *et al*. 1972; Lupi-Herrera *et al*. 1977; Kerr *et al*. 1994). However, 40 per cent of patients require additional cytotoxic agents, and 23 per cent have chronic unremitting disease (Kerr *et al*. 1994). The median time to remission in this study was 22 months (Kerr *et al*. 1994). Our experience of 7 patients confirms this data, with 3 patients requiring cytotoxic therapy in addition to corticosteroids (2, cyclophosphamide; 1, azathioprine). Continuous oral cyclophosphamide has been shown to induce remission in patients with corticosteroid-resistant disease in 4 of 6 patients treated (Shelhamer *et al*. 1985). The toxicity of cyclophosphamide in the treatment of systemic vasculitis can be reduced with the use of pulse intravenous regimens. There is little experience in the use of pulse cyclophosphamide in Takayasu's arteritis, although we have used such a regimen with good results in a small number of patients. Methotrexate is being used increasingly as a steroid sparing agent in the vasculitides. Hoffman and colleagues recently reported their experience of low-dose oral methotrexate in 18 patients with Takayasu's arteritis (Hoffman *et al*. 1994). Sixteen patients were followed up for an average of 2.8 years. The mean dose of methotrexate was 17.1 mg per week, which together with corticosteroids, resulted in remission in 13 patients; however, 7 patients relapsed on reduction of corticosteroid dose. Reintroduction of methotrexate resulted in a further remission. Sustained remissions of 4 to 34 months have been observed. Methotrexate is an effective method of inducing disease remission, permits reduction in the total corticosteroid dose and is therefore an alternative to cyclophosphamide in patients with disease that is difficult to control with corticosteroids alone. A flow diagram illustrating the managment of Takayasu's arteritis is given in Fig. 3.

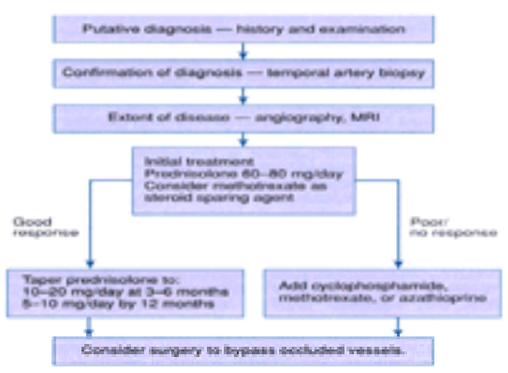


Fig. 3 Flow diagram illustrating the immunosuppresive treatment of large vessel involvement in Takayasu's arteritis and giant cell arteritis.

Hypertension can be difficult to manage in patients with Takayasu's arteritis, particularly in those individuals with involvement of subclavian and femoral arteries, making measurement of blood pressure difficult. Wherever possible hypertension should be controlled by angioplasty or surgery. Medical control of hypertension must be cautious because of the risk to organs perfused by stenosed arteries from a fall in blood pressure. Angiotensin converting enzyme (ACE) inhibitors must be used with caution in patients with renal artery stenosis, because such patients have elevated renin levels and may be especially sensitive to the first dose of an ACE inhibitor. Beta-blockers may be used as an alternative. Vasodilators are potentially dangerous in the presence of fixed stenotic lesions (Ito 1992).

Anticoagulation is not currently recommended for Takayasu's arteritis. The role of anticoagulation and/or antiplatelet drugs in patients with no evidence of a hypercoagulable state has not been evaluated, but low dose aspirin may be sensible.

Surgical treatment

Surgery has an important adjunctive role in the management of patients with Takayasu's arteritis. Hypertension secondary to renal artery stenosis is a common mode of presentation, particularly in South-East Asia. Percutaneous transluminal angioplasty has been reported to restore patency in up to 80 per cent of cases (Sharma *et al*. 1992), but restenosis can occur within 1 to 2 years (Kumar *et al*. 1990). This can be managed with repeated angioplasty but formal bypass surgery may be necessary.

Elective bypass procedures are best performed during the inactive phase of the disease. Operative procedures include bypass of stenosed segments of arteries, resection of aneurysms and replacement of aortic valves (Ohtecki *et al*. 1992; Robbs *et al*. 1994). Anastomotic sites should be chosen in areas unaffected by inflammation, and autologous grafts perform better than synthetic grafts (Kerr *et al*. 1994). Overall operative mortality is 3 to 4 per cent and is associated with surgery for aneurysm rupture (Robbs *et al*. 1994).

Prognosis

The majority of patients (74 per cent) have some impairment of activities of daily living, and 47 per cent are permanently disabled (Kerr *et al*. 1994). Mortality is, however, low and 5- and 10-year survival rates of 80 to 90 per cent have been reported (Hall *et al*. 1985; Ishikawa 1988; Subramayan *et al*. 1989). Hypertension, cardiac involvement, aortic or arterial aneurysms, and severe functional disability predict greater morbidity and mortality in these studies. Ishikawa recently reported the long-term outcome in 120 patients followed for a median of 13 years (Ishikawa and Maetani 1994). The survival rate at 15 years was 82.9 per cent. The major determinants of outcome were the presence of complications (retinopathy, hypertension, aortic regurgitation and aneurysm) and the pattern of previous disease.

Pregnancy does not appear to exacerbate the inflammatory process, but if pregnancy is considered then it may be safer in a quiescent phase (Kerr *et al*. 1994). Hypertension, aneurysms and extent of disease are associated with an increased risk of maternal or fetal death (Wong *et al*. 1983). The increases in intravascular volume during pregnancy may exacerbate hypertension, aortic regurgitation and congestive cardiac failure. Moderate doses of corticosteroids have not had a detrimental effect on the fetus. Normal vaginal delivery is possible and caesarean section should be considered on obstetric merits and not because of coexistence of Takayasu's arteritis (Kerr *et al*. 1994).

Giant cell arteritis

Giant cell arteritis is a granulomatous arteritis of the aorta and its major branches, with a predilection for the extracranial branches of the carotid artery (Jeanette *et al*. 1994). The main features of the disease and its relation to polymyalgia rheumatica (PMR) are described elsewhere (Chapter 5.11.5). This section will review large vessel involvement in giant cell arteritis.

The inflammatory process of giant cell arteritis may involve the aorta and its major branches; this involvement may be minimal and clinically silent, however, extensive inflammation may lead to aortic incompetence, aneurysm formation, aortic rupture and death. Postmortem studies have demonstrated the extensive nature of the inflammatory process in giant cell arteritis, but many lesions are clinically silent (Cooke *et al*. 1946; Wilkinson and Russell 1972; Ostberg 1973). In a prospective study of 889 postmortem cases in which the temporal arteries and two transverse sections of the aorta were examined, giant cell arteritis was found in 1.7 per cent (Ostberg 1973).

Hamrin found that the aortic arch syndrome was present in 14 out of 93 cases of giant cell arteritis (Hamrin *et al*. 1972), whilst Klein and colleagues reviewed 248 patients with this disease and found that 34 (13.7 per cent) had definite or possible clinical evidence of involvement of the aorta and its major branches (Klein *et al*. 1975). At presentation of giant cell arteritis, 5 per cent of patients have large vessel involvement (Klein *et al*. 1975). Upper limb involvement is present in 6.3 per cent of patients with giant cell arteritis/polymyalgia rheumatica (Ninet *et al*. 1990). Our experience suggests that large vessel involvment in giant cell arteritis is less common. We have seen 3 cases over a 6 year period compared with approximately 100 cases of uncomplicated giant cell arteritis/polymyalgia rheumatica.

Clinical features

Systemic manifestations are present in a majority of patients with large vessel involvement, with malaise, fever, anorexia, sweats and weight loss. There is no difference in the frequency of polymyalgic symptoms and permanent visual loss between patients with large artery involvement and those without (Klein *et al*. 1975).

Intermittent claudication is the most common symptom occurring in 75 per cent of patients (Table 5) (Hunder 1990). Raynaud's phenomenon occurs in 20 per cent and may be unilateral. Physical examination reveals bruits over large arteries and/or decreased pulses. Bruits may be audible over long sections of artery. The arms alone were involved in 50 per cent of cases, legs alone in 12 per cent and arms and legs in 24 per cent (Klein *et al*. 1975). The remaining 12 per cent were either asymptomatic or had symptoms related to the chest or abdomen. Lie reviewed 72 cases with histologically proven extracranial giant cell arteritis: the ascending and aortic arch were most frequently involved (39 per cent), subclavian and axillary arteries (26 per cent), and femoropopliteal arteries (18 per cent) (Lie 1995). Evans and colleagues from the Mayo clinic identified 41 patients with giant cell arteritis and thoracic aortic aneurysms seen between 1950 and 1991 (Evans *et al*. 1994). During this period 1330 cases of biopsy-proven giant cell arteritis were seen. Laboratory investigations were similar in patients with and without aortic involvement. In 8 patients thoracic aorta involvement was diagnosed before or simultaneously with giant cell arteritis. The median interval in the remaining 33 patients was 7 years (3 months to 20 years). In only half of the patients was there evidence of active or recurrent giant cell arteritis at the time of development of thoracic involvement.

Symptom/sign	%
Intermittent claudication	74
Upper limb	40
Lower limb	13
Both	22
Raynaud's phenomenon	22
Paraesthesia	9
Decreased/absent pulses	91
Upper limb	61
Lower limb	17
Both	13

From Hunder (1990).

Table 5 Clinical features of large vessel arteritis occurring in giant cell arteritis

Large vessel involvement can become clinically apparent at any stage of the disease, either at initial presentation, when corticosteroid dose is being reduced or after corticosteroids have been withdrawn (Hunder 1990). Development of intermittent claudication of the arms may be the first manifestation of giant cell arteritis in a patient without symptoms suggestive of either polymyalgia rheumatica or temporal artery tenderness.

Diagnosis of large vessel involvement requires careful examination of arteries for tenderness, decreased pulses and bruits. Angiography is necessary to demonstrate the extent of arterial disease.

The differential diagnosis of patients with large vessel involvement due to giant cell arteritis includes Takayasu's arteritis and atherosclerosis. Takayasu's arteritis involves a younger age group (less than 50 years), different ethnic communities (giant cell arteritis is rare in non-Caucasians, whilst Takayasu's arteritis is common in Asians), the disease is usually more extensive and temporal artery biopsy is negative. Atherosclerosis, like giant cell arteritis, affects the population over 60 years of age, but is a more common cause of claudication than giant cell arteritis; however, the presence of claudication with evidence of a systemic illness (fever, malaise, weight loss, raised erythrocyte sedimentation rate) should suggest the possibility of an underlying vasculitic process. The diagnosis can be confirmed by angiography or temporal artery biopsy.

Laboratory investigations

The erythrocyte sedimentation rate is usually elevated at the time of diagnosis of large vessel involvement, but cases have been reported of large vessel disease developing at a time of apparent disease quiescence. The diagnosis should be considered in patients with an elevated erythrocyte sedimentation rate and limb, especially upper limb, ischaemia. In these patients temporal artery biopsy should be performed even if the temporal arteries are clinically normal.

Histology of large arteries is similar to that seen in the temporal artery with disruption of the elastic lamina, mononuclear cell infiltrate, giant cells and granulomata. Fibrosis occurs in longstanding lesions and results in luminal narrowing.

The demonstration of the extent of large vessel involvement in giant cell arteritis is dependent on imaging of the vessels involved. Doppler ultrasonography will demonstrate diminution of blood flow in an extremity and allow localization of lesions. Angiography has been the key to diagnosis of large vessel involvement (Fig. 4). The angiographic features most suggestive of arteritis are: (i) long segments of smooth arterial stenosis alternating with areas of normal or increased calibre, (ii) smooth tapered occlusions of affected large arteries, (iii) absence of irregular plaques and ulceration, and (iv) anatomical distribution of these changes with major involvement of subclavian, axillary and brachial arteries (Klein *et al*. 1975). The role of magnetic resonance imaging is uncertain at present, but as with Takayasu's arteritis will undoubtedly become more important.

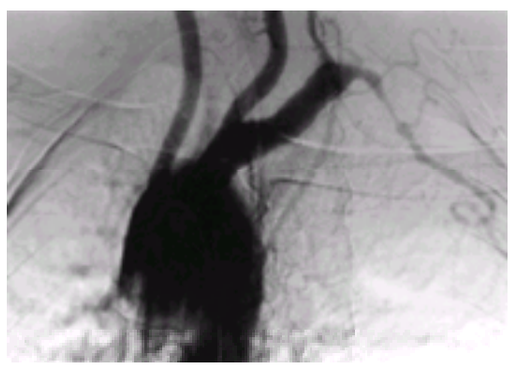


Fig. 4 Arch aortogram from a patient with giant cell arteritis showing dilatation of the right subclavian artery terminating in a long stricture distal to the origin of the vertebral artery (from Watts *et al*. 1989 with permission).

Treatment and prognosis

Improvement in blood flow through extremities rapidly occurs following treatment with high-dose corticosteroids (45 to 60 mg/day), rendering reconstructive surgery unnecessary. Evolution of arterial stenosis can be serially monitored using Doppler ultrasonography. Vessels that are occluded on the angiogram rarely recanalize and in these cases clinical improvement is due to collateral formation. Surgery should be considered in patients with persistent limb ischaemia. This should be after the inflammatory phase has subsided to prevent early thrombosis of the graft (Ninet *et al*. 1990).

Large vessel involvement in giant cell arteritis is potentially fatal following rupture of an inflammatory aortic aneurysm, stroke or myocardial infarction (Säve-Söderburgh *et al*. 1986). All patients with giant cell arteritis should be assessed carefully for large artery lesions.

Cytotoxic agents (methotrexate or cyclophosphamide) should be considered for those patients with persistent disease activity.

The vasculitides in general, although well recognized as a cause of coronary ischaemia, are rarely considered in the assessment of patients with ischaemic heart disease.

Infection-related arteritis

Infection-related aortitis was first described by Ambrose Paré in the sixteenth century, who associated syphilis with aortic aneurysms. During the nineteenth century, syphilis was recognized as the most common cause of aortitis, but pyogenic organisms and tuberculosis were increasingly recorded. A wide variety of organisms have since been recognized as causing a large vessel arteritis, including bacteria, fungi and spirochaetes (Table 6). Viral infections are not known to be associated with a large vessel arteritis.

Bacterial	*Staphylococcus* *Streptococcus* *Pneumococcus* *Klebsiella* *Pseudomonas* *Brucella* *Salmonella* *Haemophilus* *Serratia*
Mycobacterial	*Mycobacterium tuberculosis*
Fungal	*Candida* *Cryptococcus* *Aspergillus* *Coccidioides* *Histoplasma* *Mucor* *Blastomyces*
Spirochaetal	Syphilis

Table 6 Causes of infective large vessel arteritis

The vascular reaction to direct infection depends on the organism and the site of infection. Two mechanisms for infection-related vasculitis have been proposed: first, direct microbial toxicity either by endothelial invasion or the effect of microbial toxins on endothelium; or second, immune-mediated either via immune complexes or

cellular responses. Large vessel vasculitis secondary to infection occurs as a result of direct microbial endothelial invasion and leads to an erosive arteritis with mycotic aneurysm formation (Fig. 5).

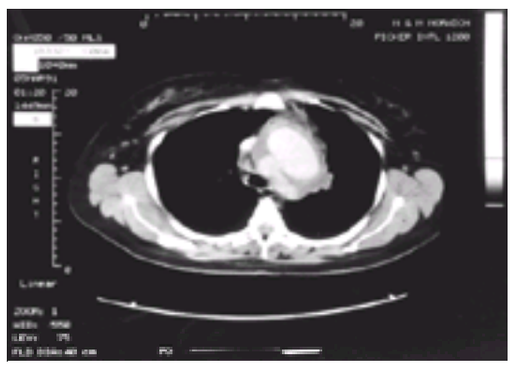


Fig. 5 Thoracic CT scan showing an aneurysm of the descending aorta. *Pneumococcus pneumoniae* was cultured from the resected vessel (from Chakravarty and Scott 1992 with permission).

Bacterial infective arteritis

Infective arteritis occurring as a result of direct endothelial invasion by bacteria or fungi results in formation of a mycotic aneurysm. Such aneurysms may develop in several ways, by septic embolization from a focus of infection such as a cardiac valve, by direct extension from a neighbouring abscess or by haematogenous spread of organisms from a portal of entry. Previously damaged or atherosclerotic vessel walls are very susceptible to bacterial seeding. Predisposing factors for development of a mycotic aneurysm include immunodeficiency, malignancy, diabetes mellitus, intravenous drug abuse and malignancy.

A wide variety of bacteria have been reported as causing infective large vessel arteritis including *Staphylococcus*, *Streptococcus*, *Klebsiella*, *Pneumococcus*, and *Pseudomonas* spp. (Table 6). Recently vasculitis has been described in association with more unusual organisms: *Salmonella* (Oskoui *et al*. 1993), *Haemophilus* (Chakravarty and Scott 1992) and *Brucella* (Aguado *et al*. 1987). *Staphylococcus* and *Salmonella* were the most frequently identified organisms in a 10-year study of 21 patients (Oz *et al*. 1989).

Aortic mycotic aneurysms may be clinically silent until they reach a massive size and rupture, with fatal haemorrhage occurring before a clinical diagnosis is made. Clinical symptoms are usually related to the systemic manifestations of infection (fever, leucocytosis), and if the aneurysm is large enough it may cause pressure effects—cough, chest pain, hoarseness and dysphagia with thoracic aneurysms, and epigastric or lumbar pain and a palpable mass with abdominal aneurysms. In the limbs mycotic aneurysms usually present with pain, tenderness and fever associated with a pulsatile mass, which may be palpable when present in the extremities. They are associated with a high mortality from sepsis and haemorrhage. The source of the infection is often from infected heart valves (80 per cent), though rarely it may be the result of direct microbial arteritis. Transient bacteraemia occurs frequently in intravenous drug abusers and results in direct vessel wall infection. *Salmonella*, *Staphylococcus*, and *E. coli* are the usual organisms. The presentation may be insidious and diagnosis delayed until haemorrhage occurs.

Diagnosis is based on a high level of suspicion particularly in patients with fever, leucocytosis, abdominal pain and positive blood cultures. The aorta may appear normal on plain radiographs and a CT scan may be required to demonstrate aortic infection

The mortality of bacterial aortitis is still high. In a series of 21 cases seen over 10 years, there were 8 disease-related deaths, but no graft infections were seen in the survivors (Oz *et al*. 1989).

Spirochaetal infection

Treponema pallidum is a well recognized cause of aortitis, with aneurysm formation and development of aortic incompetence. The primary histological changes are in the vasa vasorum with endarteritis, and perivascular infiltration with lymphocytes and plasma cells. Infection occurs early in the disease and organisms lie dormant in the aortic wall. Spirochaetes can, however, only rarely be detected in tissue. Symptomatic aortic disease occurs as a feature of tertiary disease. The majority of syphilitic aneurysms are located in the thoracic aorta with the following distribution: sinus of Valsalva less than 1 per cent, ascending aorta 36 per cent, aortic arch 24 per cent, descending aorta 24 per cent, and multiple locations 4 per cent (Berkman 1986). Presentation is with aortic incompetence (60 per cent of cases) or coronary ostial stenosis. Coronary arteritis may occur independently of aortitis. Diagnosis is based on the radiographic appearances and serological tests. Treatment is with antibiotics and surgical resection. The differential diagnosis of syphilitic aortitis includes Takayasu's arteritis, extracranial giant cell arteritis, and the aortitis of rheumatic diseases.

Borrelia burgdorferi infection causes a small vessel vasculitis which does not involve large vessels.

Fungal infections

Fungal infections of the endothelium are rare. *Candida*, *Aspergillus*, and *Histoplasma* are the most common organisms; however, *Blastomyces*, *Coccidioides*, *Cryptococcus*, and *Mucor* have been described rarely (Berkman 1986; Leavitt and Kauffman 1988). Infection results in formation of a mycotic aneurysm, similar to that seen with bacterial arterial infections. Direct infection is rare. Most cases of fungal aortitis occur in patients who are already severely immunocompromised from other diseases, those who have received multiple courses of intravenous antibiotics, or intravenous drug abusers. These patients have a high mortality (70 per cent) and require aggressive antifungal therapy.

Mycobacterial infection

Mycobacterial infection may involve blood vessels of any size, with veins being more vulnerable than arteries. Large vessel tuberculous arteritis is uncommon, with an equal division between thoracic and abdominal aneurysms (Silbergleit *et al*. 1965; Berkman 1986). The aorta may be infected by the tubercle bacillus by haematogenous dissemination in miliary disease, lymphangitic spread and direct extension from contiguous structures. Aneurysm formation occurs in half of the cases. Clinical presentation is non-specific with fever, chest pain, haemoptysis, dysphagia, abdominal pain and a palpable mass. Tuberculous vasculitis results in granuloma formation and a panarteritis. Acid-fast bacilli may be detectable in macrophages of the inflammatory infiltrate. Treatment is with antituberculous drugs, combined with surgical resection.

Mycobacterium leprae may cause vasculitis but this typically involves small arteries. Other atypical mycobacteria have not been described as causing a large vessel aortitis.

Rheumatoid arthritis

Systemic rheumatoid vasculitis was first described by Bywaters in 1949 and the clinical features delineated over the next decade (Bywaters 1957). The annual incidence of this disease is 12.6 per million, making it one of the most common forms of systemic vasculitis (Watts *et al*. 1994). All sizes of vessel from large arteries to capillaries may be involved. Large vessel involvement with aortitis or aortic regurgitation is rare in patients with rheumatoid arthritis and was only first clearly described by Zvaifler and Weintraub (1963) . In a large comparative series seen between 1970 and 1994 from three centres in the United Kingdom comprising nearly 150 cases of systemic rheumatoid vasculitis, large vessel involvement was seen in 5 per cent (Luqmani *et al*. 1995). Aortitis was identified in 10 cases in a series of 180 postmortems performed on patients with rheumatoid vasculitis (Gravallese *et al*. 1989). As with other forms of systemic rheumatoid vasculitis, the typical patient with large vessel disease has long-standing seropositive erosive rheumatoid vasculitis. In most cases there is multiorgan involvement (Gravallese *et al*. 1989) including a coronary arteritis. Aortitis is rarely diagnosed antemortem, but may be the cause of death, particularly in cases with associated coronary arteritis. Aortitis

and aortic incompetence may result in haemodynamic compromise with congestive cardiac failure.

Histological features of rheumatoid aortitis include necrosis of medial smooth muscle and elastica, with an inflammatory infiltrate comprising lymphocytes and plasma cells. A panmural aortitis can occur. Rheumatoid granulomata are seen in half of the cases (Gravallese *et al*. 1989).

Systemic rheumatoid vasculitis responds well to corticosteroids and cyclophosphamide (Luqmani *et al*. 1995). There is however no data on the response of rheumatoid aortitis or aortic incompetence to immunosuppressive therapy, however this treatment should be considered in patients presenting with active aortitis since prevention or delay in the necessity for surgery may be achieved.

Spondylarthopathies

Aortitis may complicate the seronegative spondylarthropathies in particular ankylosing spondylitis (Buckley and Roberts 1973) and Reiter's syndrome (Morgan *et al*. 1986). The aortic ring and ascending aorta are the typical sites of involvement but distal aortitis has been described (Morgan *et al*. 1984). Aortic incompetence occurs in up to 5 per cent of cases of ankylosing spondylitis, 2.5 per cent of cases of Reiter's syndrome and probably less frequently in the other seronegative spondylarthropathies (Townend *et al*. 1991).

The seronegative spondylarthropathies are associated with the HLA-B27 haplotype. Inflammatory diseases associated with this haplotype have been found in 15 to 20 per cent of patients with lone aortic regurgitation, suggesting that lone aortic regurgitation may reflect an inflammatory valvulitis or aortitis (Bergfeldt *et al*. 1988)

Echocardiography shows thickening of the aortic leaflets, subaortic echodense bumps and aortic root densities (Labresh *et al*. 1985). Histologically there is thickening of the aortic valve cusps and aorta, together with lymphocytic infiltration in the aortic wall and fibrosis of the aortic root (Buckley and Roberts 1973). These changes can be difficult to distinguish from lesions occurring in syphilitic aortitis.

Treatment of the inflammatory process with immunosuppressive agents has been suggested, but no controlled data exists as to their benefit in this situation, in particular whether the need for aortic valve replacement can be delayed or prevented (Townend *et al*. 1991).

Systemic lupus erythematosus

The typical vascular lesion of systemic lupus erythematosus involves small vessels, large vessels are rarely involved. Saxe and Altman reviewed 19 reported cases of Takayasu's arteritis associated with systemic lupus erythematosus, and pointed out that the diseases have a similar age of onset and female preponderance (Saxe and Altman 1992). However, the absence of specific markers for systemic lupus erythematosus in patients with Takayasu's arteritis who subsequently develop systemic lupus erythematosus suggests that the coexistence of these two conditions may be coincidental. Some of these cases are incompletely reported and would not be considered necessarily to be either of the diseases using current diagnostic criteria. Patients with systemic lupus erythematosus who develop large vessel thrombosis as part of the antiphospholipid antibody syndrome, may mimic the obstructive vasculopathy seen in Takayasu's arteritis.

Relapsing polychondritis

Relapsing polychondritis is a systemic disease of unknown aetiology with inflammation of cartilaginous structures. Vasculitis occurs in 11 to 56 per cent of cases with involvement of large and medium-sized vessels (Michet 1990). Large artery involvement is uncommon (less than 15 per cent of cases) and present as an aortic arch syndrome, thoracic or abdominal aortic aneurysm with rupture, or aortic regurgitation. Small vessel vasculitis and a segmental necrotizing glomerulonephritis may occur simultaneously. The vasculitis may present at the same time as other manifestations of relapsing polychondritis or many years afterwards. The aortitis of relapsing polychondritis results in loss of glycosaminoglycans from the aortic wall, particularly in the media with loss of elastic tissue.

Treatment is with corticosteroids, and the role of cytotoxic agents is uncertain. Surgical repair is required for valvular and artery occlusions.

Behçet's syndrome

Behçet's syndrome is a chronic relapsing condition characterized by a triad of aphthous stomatitis, genital ulceration, and uveitis. A large vessel arteritis is well described (Little and Zarins 1982; Lakhanpal *et al*. 1985). Major arterial occlusions were present in 36 per cent and there was evidence of vasculitis in 18 per cent of cases at autopsy (Lakhanpal *et al*. 1985). Shimizu and colleagues described 81 patients with Behçet's disease in whom occlusion of large arteries was present in 17 and aneurysms in 24 patients (Shimizu *et al*. 1979). Aortitis is uncommon and is represented histologically by inflammation in the media and adventitia, particularly around the vasa vasorum. Multiple pulmonary aneurysms occur which may communicate with the bronchi and present as massive haemoptysis.

Cogan's syndrome

Cogan's syndrome is a rare condition characterized by recurrent episodes of acute sensorineural hearing loss and non-syphilitic keratitis (Cogan 1945). Systemic features occur and vascular involvement is common (Vollertsen *et al*. 1986).

Vasculitis occurs early in the disease, within the first year in two-thirds of patients, but can occur up to 8 years after onset (Vollertsen *et al*. 1986). The clinical manifestations will depend on the vessel involved.

Cardiac involvement occurs in 15 to 25 per cent of patients (Cheson *et al*. 1976; Vollertsen *et al*. 1986), manifesting mainly as aortic regurgitation which may be rapidly progressive over a period of months. It is associated with either a valvulitis or an aortitis. Aortic regurgitation may occur at diagnosis or up to 12 years after diagnosis (Vollertsen *et al*. 1986). Aortitis may lead to narrowing of the coronary ostia and arteries arising from the aortic arch; a distal coronary arteritis may also occur.

Angiography demonstrates arterial occlusion of large and medium vessels with scattered stenosis, and a diffuse irregular narrowing can also be seen (Vollertsen *et al*. 1986).

Histological examination of the aorta reveals neutrophils, mononuclear cells, giant cells, destruction of the elastic lamina, neovascularization, necrosis, scarring and fibrotic hypertrophy (Cheson *et al*. 1976). Similar findings are observed in other vessels.

Treatment of Cogan's syndrome is with corticosteroids, and clinical improvement occurs in most cases. The role of cytotoxic drugs in undetermined. Treatment of aortic regurgitation is by valve replacement (Cochrane and Tatoulis 1991). Arterial bypass grafting may be required.

Kawasaki disease

Kawasaki disease is a inflammatory vascular disease of childhood which typically results in the formation of coronary artery aneurysms. Systemic artery aneurysms occur in 4 per cent of patients with Kawasaki disease, the most common site being the axillary and iliac arteries (Inioue *et al*. 1988). Aortitis and consequent aneurysm formation is rare. Fibrosis and arterial luminal stenosis is a late sequelae of acute Kawasaki disease. Treatment of acute disease is with aspirin and intravenous immunoglobulin.

Wegener's granulomatosis

Wegener's granulomatosis is a disease of predominately small and medium vessels. The clinical features and management of Wegener's granulomatosis is described in Chapter 5.11.2. Coronary and large artery involvement has only rarely been described in patients with Wegener's granulomatosis (Davenport *et al*. 1994; Logar *et al*. 1994). Two patients had clinically significant aortic valve incompetence. Histology of the resected valve showed myxoid degeneration due to previous vasculitis affecting the vessels of the aortic wall, together with valvular necrosis (Davenport *et al*. 1994). Carotid arteritis has been described in a patient with ANCA-positive

crescentic glomerulonephritis (Logar *et al*. 1994).

Rheumatic fever

Rheumatic fever occurs as part of the immune response to Group A streptococci. During acute rheumatic fever arteritis can develop in different sized arteries, and Aschoff bodies may develop in the adventitia of the aorta and other large arteries. An acute vasculitis may involve the vasa vasorum, with endothelial cell swelling, luminal leucocytes, and a perivascular infiltrate of lymphocytes, macrophages, and leucocytes (Virmani and McAllister 1986).

Sarcoidosis

Sarcoidosis is a systemic granulomatous disease of unknown aetiology. Cardiac involvement is well recognized in patients who die suddenly. Sarcoid granulomata have been described rarely in the aorta and its large branches. The granulomata are found in the adventitia (Virmani and McAllister 1986).

Central nervous system angiitis

Primary angiitis of the central nervous system is a rare disease of unknown aetiology, occasionally it involves large vessels but more typically medium or small arteries (Calabrese and Mallek 1988; Abu-Shakra *et al*. 1994). The clinical features are non-specific and entirely neurological. Presentation may be acute with severe headache, stroke, neurocognitive defect, or seizure. There may be evidence of systemic illness with elevation of acute phase proteins. The diagnosis is made by the presence of characteristic histological findings, with a granulomatous vasculitis in the leptomeninges or cortex. Angiography demonstrates beading or ectasia alternating with stenosis. These changes are typically distal to the termination of the carotid arteries. Treatment is with corticosteroids with or without cyclophosphamide.

Chapter References

Abu-Shakra, M., Khraisha, M., Grosman, H., Lewtas, J., Cividino, A., and Keystone, E.C. (1994). Primary angiitis of the CNS diagnosed by angiography. *Quarterly Journal of Medicine*, **87**, 351–8.

Achar, K.N. and Al-Nahib, B. (1986). Takayasu's arteritis and ulcerative colitis. *American Journal of Gastroenterology*, **81**, 1215–7.

Aguado, J.M., Barros, C., Gomez Garces, J.L., and Fernandez-Geurrero, M.L. (1987). Infective aortitis due to *Brucella melitensis*. *Scandinavian Journal of Infectious Disease*, **19**, 483–4.

Arend, W.P. *et al*. (1990). The American College of Rheumatology 1990 criteria for the classification of Takayasu arteritis. *Arthritis and Rheumatism*, **33**, 1129–34.

Baldursson, O., Steinsson, K., Bornsson, L., and Lie, J.T. (1994). Giant cell arteritis in Iceland. An epidemiologic and histopathologic analysis. *Arthritis and Rheumatism*, **37**, 1007–12.

Bengtsson, B-Å. and Malmvall, B.E. (1981). The epidemiology of giant cell arteritis including temporal arteritis and polymyalgia rheumatica: incidences of different clinical presentations and eye complications. *Arthritis and Rheumatism*, **24**, 899–904.

Bergfeldt, L., Insulander, P., Lindblom, D., Möller, E., and Edhag, O. (1988). HLA-B27: an important genetic risk factor for lone aortic regurgitiation and severe conduction abnormalities. *American Journal of Medicine*, **85**, 12–18.

Berkman, Y. (1986). Medical aspects of infectious aortitis. In *Aortitis* (ed. A. Lande, Y.M. Berkman, and H.A. McAllister), pp. 161–72. Raven Press, New York.

Buckley, B.H. and Roberts, W.C. (1973). Ankylosing spondylitis and aortic regurgitation. Description of the characteristic lesion from study of eight necropsy cases. *Circulation*, **48**,1014–27.

Bywaters, E.G.L. (1957). Peripheral vascular obstruction in rheumatoid arthritis and its relationships to other vascular lesions. *Annals of the Rheumatic Diseases*, **16**, 84–103.

Calabrese, L. and Mallek, J.A. (1988). Primary angiitis of the central nervous system. *Medicine*, **67**, 20–39.

Carson, C.W., Beall, L.D., Hunder, G.G., Johnson, C.M., and Newman, W. (1993). Serum ELAM-1 is increased in vasculitis, scleroderma, and systemic lupus erythematosus. *Journal of Rheumatology*, **20**, 809–14.

Castrenasa-Isla, C.J., Herrera-Martinez, G., and Vega-Molin, J. (1993). Erythema nodosum and Takayasu's arteritis after immunisation with plasma derived hepatitis B vaccine. *Journal of Rheumatology*, **20**, 1417–8.

Chakravarty, K.C. and Scott, D.G.I. (1992). Mycotic aneurysm of the aortic arch masquerading as systemic lupus erythematosus. *Annals of the Rheumatic Diseases*, **51**, 1079–81.

Cheson, B.D., Bluming, A.Z., and Alroy, J. (1976). Cogan's syndrome and systemic vasculitis. *American Journal of Medicine*, **60**, 549–55.

Churg, J. (1993). Large vessels arteritis. *Clinical and Experimental Immunology*, **93** (Suppl.), 11–12.

Cochrane, A.D. and Tatoulis, J. (1991). Cogan's syndrome with aortitis, aortic regurgitation, and aortic arch vessel stenoses. *Annals of Thoracic Surgery*, **52**, 1166–7.

Cogan, D.G. (1945). Syndrome of nonsyphilitic interstitial keratitis and vestibulo-auditory symptoms. *Archives of Ophthalmology*, **33**, 144–9.

Cooke, W.T., Cloake, P.C.P., and Govan, A.D.T. (1946). Temporal arteritis: a generalised disease. *Quarterly Journal of Medicine*, **15**, 47–76.

Davenport, A., Goodfellow, J., Goel, S., MacIver, A.G., and Walker, P. (1994). Aortic valve disease in Wegener's granulomatosis. *American Journal of Kidney Diseases*, **24**, 205–8.

DiGiacomo, V. (1984). A case of Takayasu's disease occurred over two hundred years ago. *Angiology*, **35**, 750–4.

Dong, R-P., Kimur, A., Numamno, F., Nishimura, Y., and Sasazuki, T. (1992). HLA-linked susceptibility and resistance to Takayasu arteritis. *Heart Vessels*, **7**, 73–81.

Evans, J.M., Bowles, C., Bjornsson, J., Mullany, C.G., and Hunder, G.G. (1994). Thoracic aortic aneurysm and rupture in gaint cell arteritis. *Arthritis and Rheumatism*, **37**, 1539–47.

Fraga, A., Mintz, G., Valle, L., and Flores-Izquierd, G. (1972). Takayasu's arteritis: frequency of systemic manifestations (study of 22 patients) and favourable response to maintenance steroid therapy with adrenocorticosteroids. *Arthritis and Rheumatism*, **15**, 617–24.

Gravallese, E.M. *et al*. (1989). Rheumatoid aortitis: a rarely recognised but clinically significant entity. *Medicine (Baltimore)*, **68**, 95–106.

Hall, S. and Nelson, A.M. (1986). Takayasu's arteritis and juvenile rheumatoid arthritis. *Journal of Rheumatology*, **13**, 431–3.

Hall, S. *et al*. (1985). Takayasu arteritis: a study of 32 North American patients. *Medicine (Baltimore)*, **64**, 89–99.

Hall, S., Saga, S., Ganguly, N.K., Koicha, M., and Sharma, B.K. (1992). Immunopathogenesis of Takayasu's arteritis. *Heart Vessels*, **7**, 85–90.

Hamrin, B. (1972). Polymyalgia arteritica. *Acta Medica Scandinavica*, **533** (Suppl.), 1–131.

Hoffman, G.S., Leavitt, R.Y., Kerr, G.S., Rottem, M., Sneller, M.C., and Fauci, A.S. (1994). Treatment of gluco-corticoid resistant or relapsing Takayasu arteritis with methotrexate. *Arthritis and Rheumatism*, **37**, 78–82.

Hunder, G.G. (1986). Giant cell arteritis and giant cell aortitis. In *Aortitis* (ed. A. Lande, Y.M. Berkman, and H.A. McAllister), pp. 193–204. Raven Press, New York.

Inioue, O. *et al*. (1988). Systemic artery involvement in Kawasaki disease. In *Proceedings of the 3rd International Kawasaki disease symposium*, Tokyo, 29 November to 2 December 1988, p. 23.

Ishikawa, K. (1986). Pattern of symptoms and prognosis in occlusive thromboaortopathy (Takayasu's disease). *Journal of the American College of Cardiology*, **8**, 1041–6.

Ishikawa, K. (1988). Diagnostic approach and proposed criteria for the clinical diagnosis of Takayasu's arteriopathy. *Journal of the American College of Cardiology*, **12**, 964–72.

Ishikawa, K. and Maetani, S. (1994). Long term outcome for 120 Japanese patients with Takayasu's disease. Clinical and statistical analyses of related prognostic factors. *Circulation*, **90**, 1855–60.

Ito, I. (1992). Medical treatment of Takayasu's arteritis. *Heart Vessels*, **7** (Suppl.), 133–7.

Jeanette, J.C. *et al*. (1994). Nomenclature of systemic vasculitides. Proposal of an international consensus conference. *Arthritis and Rheumatism*, **37**, 187–92.

Jonasson, F., Cullen, J.F., and Elton, R.A. (1979). Temporal arteritis: a 14 year epidemiological, clinical and prognostic study. *Scottish Medical Journal*, **24**, 111–7.

Kasuya, K., Hashimoto, Y., and Numano, F. (1992). Left ventricular dysfunction and HLA Bw5z antigen in Takayasu arteritis. *Heart Vessels*, **7**, 116–9.

Kerr, G.S. *et al*. (1994). Takayasu's arteritis. *Annals of Internal Medicine*, **120**, 919–29.

Klein, R.G., Hunder, G.G., Stanson, A.W., and Sheps, S.G. (1975). Large artery involevement in giant cell (temporal) arteritis. *Annals of Internal Medicine*, **83**, 806–12.

Kumar, S. *et al*. (1990). Percutaneous transluminal angioplasty in non-specific aortoarteritis (Takayasu's disease): experience of 16 cases. *Cardiology Interventional Radiology*, **12**, 321–5.

Labresh, K.A., Lally, E.V., Sharma, S.C., and Ho, G. (1985). Two dimensional echocardiographic detection of preclinical aortic root abnormalities in rheumatoid variant diseases. *American Journal of Medicine*, **78**, 908–12.

Lakhanpal, S. *et al*. (1985). Pathological features of Behçet's disease. A review of Japanese autopsy registry data. *Human Pathology*, **16**, 790–5.

Lande, A. and Berkman, M.Y. (1976). Aortitis: pathological, clinical and arteriographic review. *Radiological Clinics of North America*, **14**, 219–40.

Leavitt, A.D. and Kauffman, C.A. (1988). Cryptococcal aortitis. *American Journal of Medicine*, **85**, 108–10.

Lie, J.T. (1995). Aortic and extra-cranial large vessel giant cell arteritis. *Seminars in Arthritis and Rheumatism*, **24**, 422–31.

Link, K.M., Loehr, S.P., Baker, D.M., and Lesko, N.M. (1993). Magnetic resonance imaging of the thoracic aorta. *Seminars in Ultrasound, CT and MRI*, **14**, 91–105.

Little, A.G. and Zarins, C.K.(1982). Abdominal aortic aneurysm and Behçet's disease. *Surgery*, **91**, 359–62.

Logar, D., Rozman, B., Vizjak, A., Ferluga, D., Mulder, A.H., and Kallenberg, C.G. (1994). Arteritis of both carotid arteries in a patient with focal crescentic glomerulonephritis and anti-neutrophil cytoplasmic antibodies. *British Journal of Rheumatology*, **33**, 167–9.

Lupi-Herrera, E., Sanchez-Torres, G., Marcushamer, J., Mispireta, J., Horwitz, S., and Vela, J.E. (1977). Takayasu's arteritis. Clinical study of 107 cases. *American Heart Journal*, **93**, 94–103.

Luqmani, R.A. *et al*. (1994). Birmingham vasculitis activity score (BVAS) in systemic vasculitis. *Quarterly Journal of Medicine*, **87**, 671–8.

Luqmani, R.A., Watts, R.A., Scott, D.G.I., and Bacon, P.A. (1994). Treatment of vasculitis in rheumatoid arthritis. *Annals de Medicine Interne*, **145**, 566–76.

Maeda, H. *et al*. (1991). Carotid lesions detected by B-mode ultrasonography in Takayasu's arteritis: 'macaroni sign' as an indicator of disease. *Ultrasound Medicine and Biology*, **17**, 695–701.

Magaro, M., Altomonte, L., Mirone, L., Zoli, A., and Corvino, G. (1988). Seronegative spondarthritis associated with Takayasu's arteritis. *Annals of the Rheumatic Diseases*, **47**, 595–7.

Michet, C.J. (1990). Vasculitis and relapsing polychondritis. *Rheumatic Disease Clinics of North America*, **16**, 441–4.

Morgan, S.H., Asherson, R.A., and Hughes, G.R.V. (1984). Distal aortitis complicating Reiter's syndrome. *British Heart Journal*, **52**, 115–6.

Ninet, J.P., Bachet. P., Dumontet, C.M., Du Colombier, P.B., Stewart, M.D., and Pasquier, J.H. (1990). Subclavian and axillary involvement in temporal arteritis and polymyalgia rheumatica. *American Journal of Medicine*, **88**, 13–20.

Ohtecki, H. *et al*. (1992). Aortic valve replacement for Takayasu's arteritis. *Journal of Thoracic Cardiovascular Surgery*, **104**, 482–6.

Oskoui, R., Davis, R.A., and Gomes, M.N. (1993). Salmonella aortitis. *Archives of Internal Medicine*, **153**, 517–25.

Ostberg, G. (1972). Morphological changes in the large arteries in polymyalgia arteritica. *Acta Medica Scandinavia*, **533** (Suppl.), 135–64.

Oz, M.C. *et al*. (1989). A ten year experience with bacterial aortitis. *Journal of Vascular Surgery*, **10**, 439–49.

Rijbroek, A., Moll, F.L., van-Dijk, H.A., Meijer, R., and Jansen, J.W (1994). Inflammation of the abdominal aortic aneurysm wall. *European Journal of Vascular Surgery*, **8**, 41–6.

Robbs, J.V., Abdool-Carrim, A.T., and Kadwa, A.M. (1994) Arterial reconstruction for non-specific arteritis (Takayasu's disease). *European Journal of Vascular Surgery*, **8**, 401–7.

Rush, P.J., Inman, R., and Reynolds, W.J. (1986). Rheumatoid arthritis after Takayasu's arteritis. *Journal of Rheumatology*, **13**, 427–30.

Sagar, S., Ganguly, N.K., Koicha, M., and Sharma, B.K. (1992) Immunopathogenesis of Takayasu arteritis. *Heart Vessels*, **7**, 85–90.

Säve-Söderburgh, J., Malmvall, B-E., Andersson, R., and Bengtsson, B-Å. (1986). Giant cell arteritis as a cause of death. *Journal of the American Medical Association*, **255**, 493–6.

Saxe, P.A., and Altman, R.D. (1992). Takayasu's arteritis syndrome associated with systemic lupus erythematosus. *Seminars in Arthritis and Rheumatism*, **21**, 295–305.

Scott, D.G.I. and Watts, R.A. (1994). Classification and epidemiology of systemic vasculitis. *British Journal of Rheumatology*, **33**, 897–9.

Scott, D.G.I. *et al*. (1986). Takayasu's arteritis: a pathogenic role for cytotoxic T lymphocytes? *Clinical Rheumatology*, **5**, 517–22.

Sen, P.K., Kinare, S.G., Engineer, S.D., and Parulkar, G.B. (1963). The middle aortic arch syndrome. *British Heart Journal*, **35**, 610–8.

Sharma, S. *et al*. (1991). Non-specific aorto-arteritis (Takayasu's disease) in children. *British Journal of Radiology*, **64**, 690–8.

Sharma, S. *et al*. (1992). Renal artery stenosis caused by non-specific arteritis (Takayasu disease): results of treatment with percutaneous transluminal angioplasty. *American Journal of Roentgenology*, **158**, 417–22.

Shelhamer, J.H, Volkman, D.J., Parrillo, J.E., Lawley, T.J., Johnston, M.R., and Fauci, A.S. (1985). Takayasu's arteritis and its therapy. *Annals of Internal Medicine*, **103**, 121–6.

Shimizu, T., Ehrlich, G.E., and Hayashi, K. (1979). Behçet disease (Behçet syndrome). *Seminars in Artritis and Rheumatism*, **8**, 223–60.

Silbergleit, A. *et al*. (1965). Tuberculous aortitis—surgical resection of ruptured false aneurysms. *Journal of the American Medical Association*, **193**, 333–5.

Sima, D. *et al*. (1994). Anti-endothelial cell antibodies in Takayasu arteritis. *Arthritis and Rheumatism*, **37**, 441–2.

Subramayan, R., Joy, J., and Balakrishan, K.G. (1989). Natural history of aortoarteritis (Takayasu's disease). *Circulation*, **80**, 429–37.

Takayasu, M. (1908). Case with unusual changes of the central vessels of the retina. *Acta Societatis Ophthalmologica Japonica*, **21**, 554.

Townend, J.N., Emery, P., Davies, M.K., and Littler, W.A. (1991). Acute aortitis and aortic incompetence due to systemic rheumatological disorders. *International Journal of Cardiology*, **33**, 253–8.

Virmani, R. and McAllister, H.A. (1986). Pathology of the aorta and major arteries. In *Aortitis* (ed. A. Lande, Y.M. Berkman, and H.A. McAllister), pp. 7–53. Raven Press, New York.

Vollertsen, R.S., McDonald, T.J., Younge, B.R., Banks, P.M., Stanson, A.W., and Ilstrup, D.M. (1986). Cogan's syndrome: 18 cases and a review of the literature. *Mayo Clinic Proceedings*, **61**, 344–61.

Waern, A.U., Andersson, P., and Hemmingsson, A. (1983). Takayasu's arteritis: a hospital- region based study on occurrence, treatment and prognosis. *Angiology*, **34**, 311–20.

Watts, R.A,. Bhalla, A.K., Binder, A.I., and Coppen, M. (1989). Giant cell arteritis presenting as limb claudication. *Journal of the Royal Society of Medicine*, **82**, 51–2.

Watts, R.A., Carruthers, D.M., Symmons, D.P.M., and Scott, D.G.I. (1994). The incidence of rheumatoid vasculitis in the Norwich Health Authority. *British Journal of Rheumatology*, **33**, 832–3.

Watts, R.A., Carruthers, D.M., and Scott, D.G.I. (1995). Epidemiology of systemic vasculitis changing incidence or definition? *Seminars in Arthritis and Rheumatism*, **25**, 28–34.

Wilkinson, I.M.S. and Russell, R.W.R. (1972). Arteritis of the head and neck in giant cell arteritis. *Archives of Neurolology*, **27**, 378–91.

Wilson, W.A., Morgan, O.S., Bain, B., and Taylor, J.E. (1979). Takayasu's arteritis association with Still's disease in an adult. *Arthritis and Rheumatism*, **22**, 684–8.

Wong, V.C.W., Yang, R.Y.C., and Tse, T.F. (1983). Pregnancy and Takayasu's arteritis. *American Journal of Medicine*, **75**, 597–601.

Woolf, A.D., Wakerley, G., and Wallington, T.B. (1987). Factor VIII related antigen in the assessment of vasculitis. *Annals of the Rheumatic Diseases*, **46**, 441–7.

Zvaifler, N.J. and Weintraub, A.M. (1963). Aortiitis and aortic insufficiency in the chronic rheumatic disorders—a reappraisal. *Arthritis and Rheumatism*, **6**, 241–5.

5.11.7 Behçet's syndrome

Hasan Yazici and Sebahattin Yurdakul and Vedat Hamuryudan

Introduction

Behçet's syndrome is a systemic vasculitis (Lie 1992) of unknown aetiology with a definite and peculiar geographic distribution. Most cases are clustered around the countries of the Mediterranean basin, the Middle East, and the Far East. Its most dreaded complication, eye disease, is one of the leading causes of blindness in these areas.

Beginning with Hippocrates many have written about patients with disorders that today are considered elements of Behçet's syndrome. In 1937 Hulusi Behçet, professor of dermatovenereology in Istanbul, described in detail three patients with oral and genital ulceration and hypopyon uveitis, and proposed that this was a distinct entity. Subsequently it was realized that many other clinical manifestations were part of this syndrome (Shimizu *et al*. 1979). Table 1 gives the most important of these manifestations.

Lesion	Prevalence (%)
Aphthous ulcerations	97–100
Genital lesions	80–90
Skin lesions	80
Eye lesions	50
Arthritis	40–50
Thrombophlebitis	25
Neurological involvement	1–15
Gastrointestinal involvement	0–25

Table 1 Clinical findings in Behçet's syndrome

Epidemiology

The usual onset of the syndrome is in the third or fourth decade. The onset is rare in children and after the age of 45. Recently there has been an increased awareness of childhood cases (Özdoğan 1994).

The male:female ratio is approximately equal but the syndrome has a more severe course in men. Large scale epidemiological studies are lacking. Based on case registries, the prevalence is about 1:300 000 in Northern Europe and 1:10 000 in Japan. The prevalence may be higher in Mediterranean countries. In Turkey, based on two spot surveys among the adult population, the prevalence rates were 8 and 38:10 000 (Yurdakul *et al*. 1988). For unexplained reasons, few cases are reported from the United States and Australia.

Genetics

The syndrome is associated with HLA B5, specifically with HLA B51 (Ohno *et al*. 1982), however there is geographical variation. Patients from the Mediterranean countries and Japan show this association whereas those from the United States and United Kingdom, perhaps with the exception of those with eye disease, do not (Yazici *et al*. 1980). The association with HLA B5 is primarily with those patients attending hospital and therefore with more severe disease (Yurdakul *et al*. 1988).

In a syndrome associated with an HLA allele, one would expect a more pronounced familial occurrence than one ordinarily sees in Behçet's syndrome. In general there is no consistent inheritance pattern (Bird Stewart 1986). We have seen two monozygotic brothers concordant for the syndrome (Hamuryudan *et al*. 1991a). There is also an increased prevalence of lymphocytotoxic antibodies among the healthy children of patients with Behçet's yndrome (Günaydin *et al*. 1992).

Clinical features

Skin and mucosal involvement

Oral aphthae (Fig. 1)

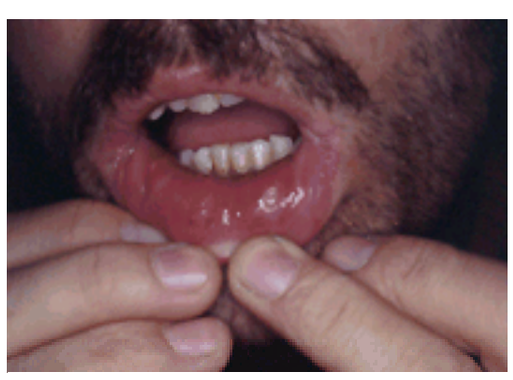


Fig. 1 Oral aphthae.

These are almost always present. However, 1 to 3 per cent of patients can have several of the other features of the syndrome without ever having aphthae. Aphthae are frequently the first manifestation of the syndrome. It is not uncommon for some patients to have only aphthae for many years before other features appear. The majority of oral ulcers in Behçet's syndrome are indistinguishable from those seen in recurrent oral ulceration, but tend to be multiple and occur more frequently. Large (major) ulcers are less frequent and herpetiform ulcers are rare. Major ulcers, however, can be very troublesome because they heal with scarring, which can even occlude the oropharynx. The minor ulcers do not as a rule leave scars. The histology reveals non-specific ulceration with necrotic material. There is evidence of vasculitis with an increase in mast cells.

Genital ulceration (Fig. 2)

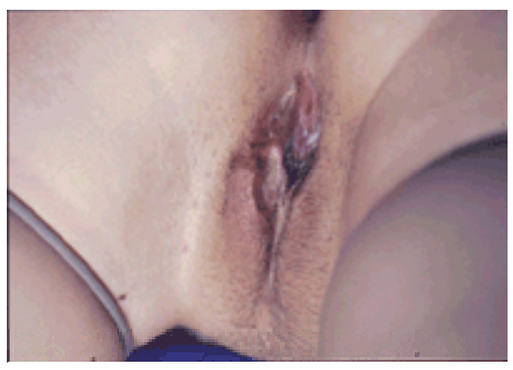


Fig. 2 Fresh genital ulceration.

In the male, 90 per cent of the genital ulcers occur on the scrotum and almost always leave scars. They are less frequent on the shaft and on the glans penis. Urethritis is not observed unless there is an associated meatal ulcer.

In the female the labia (major and minor) are affected commonly. Ulcers are less frequent in the more proximal part of the genital tract and cervical lesions are rarely seen. Histologically, they are indistinguishable from oral aphthae.

Skin lesions (Fig. 3)

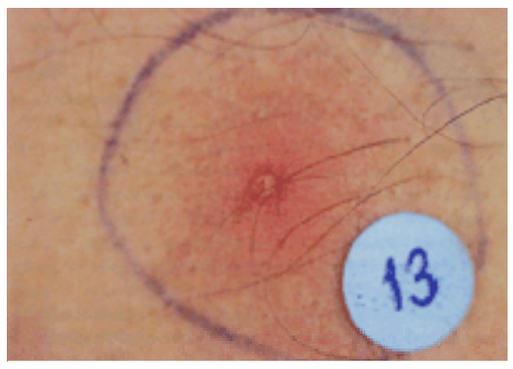


Fig. 3 The pathergy reaction in the form of a pustule.

The skin lesions of Behçet's syndrome can be divided into three main types: (i) nodular lesions resembling erythema nodosum, (ii) papulopustular lesions also called acneiform or simply acne, (iii) others (palpable purpura, ulcerations, Sweet's syndrome, etc.) all representing various forms of vasculitis. The lesions resembling erythema nodosum are similar to idiopathic erythema nodosum and those due to other conditions (e.g. sarcoidosis). Their tendency to leave a pigmented area after the acute period is a somewhat distinguishing feature, and one which is explained by the erythrodiapedesis observed histologically. Sometimes superficial thrombophlebitis can be clinically indistinguishable from erythema nodosum. The papulopustular lesions, when present in intertrigenous areas like axillae and the inguinal folds, tend to erode and ulcerate. Most papulopustular lesions are histologically very similar to ordinary acne. However, they differ from the latter in their propensity to occur also in the extremities. The other forms of skin lesions are leucocytoclastic vasculitis, necrotizing arteritis of the small and medium arteries, superficial thrombophlebitis, and unclassifiable papules and pustules (Azizlerli *et al*. 1992). Painful papules, which histologically show a neutrophilic infiltration of the skin without fibrinoid necrosis (Sweet's syndrome) are also seen (Lakhanpal *et al*. 1988).

The pathergy reaction (Fig. 3), a curious hyperreactivity of the skin to a needle prick, is peculiar to this syndrome (Gilhar *et al*. 1989). The only other condition in which it is known to be positive with any consistency is pyoderma gangrenosum. After a skin puncture with a needle a papule or a pustule forms in 24 to 48 h.

There is some debate as to the true prevalence of this reaction. It is seldom found among patients in Northern Europe or the United States. In patients from Japan and Turkey it is positive in around 60 to 70 per cent when tested repetitively. It is, however, possible that its prevalence may be decreasing in recent years (DilŠen *et al*. 1993). Whether this is a matter of patient selection, the recent use of disposable needles, or a true decline in prevalence is a matter of debate.

The mechanism of the pathergy reaction is still obscure. Surgical cleaning of the skin considerably dampens this reaction (Fresko *et al*. 1993), which suggests that more than disrupting the integrity of the epidermis and dermis is required.

The reaction shows an initial accumulation of neutrophils followed by mast cells and mononuclear cells. The pathergy phenomenon in Behçet's syndrome is not confined to the skin: various tissues are known to be hyperreactive to surgical trauma, and it is not uncommon to have attacks of uveitis after eye surgery and synovitis after an arthrocentesis.

The propensity for inflammation in Behçet's syndrome can also be observed in the response these patients have to intradermal injections of sodium monourate crystals (Çakir *et al*. 1991). Patients with Behçet's syndrome have an augmented response to these crystals and it is interesting to note that British patients show a similar heightened response even though they are, as noted above, generally pathergy negative.

Eye involvement (Fig. 4)

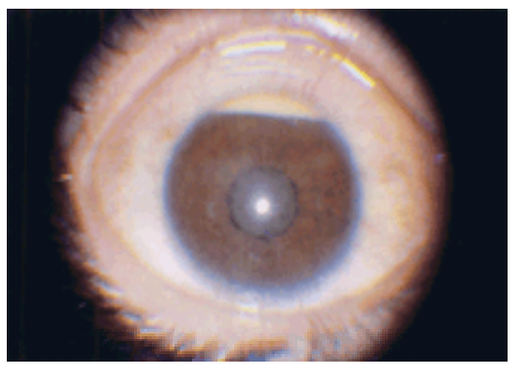


Fig. 4 Hypopyon uveitis.

This is one of the most serious manifestations. Males and those with younger age of onset, i.e. less than 25 years of age, have an increased prevalence. While the overall prevalence is about 50 per cent, in the younger male it is approximately 70 per cent. Females are affected less severely. Disease is bilateral in 90 per cent of the patients with ocular involvement. The onset of eye disease is usually within 2 to 3 years of the development of the syndrome.

Eye disease in Behçet's syndrome consists of a chronic relapsing posterior and anterior uveitis. Isolated anterior uveitis is found in only 10 per cent of those with ocular involvement. Conjunctivitis is rare. Occasionally episcleritis and conjunctival ulceration are seen.

Hypopyon uveitis (Fig. 4) is very typical of Behçet's syndrome, although occasionally it can be observed in Reiter's syndrome. It is an accumulation of white cells and debris in the anterior chamber that precipitates to form a layer due to gravity. Hypopyon is seen in 20 per cent of patients with eye disease and as rule is almost always associated with severe retinal disease.

The basic retinal lesion is a vasculitis, which can lead to exudates, haemorrhages, venous thrombosis, papilloedema, and macular disease that frequently results in a hole. The pars plana is also involved. During an acute flare there is a marked influx of fibrin, inflammatory cells, and cellular debris into the vitreous. After each flare there is usually some residual structural damage in the form of retinal changes, vitreal opacities, posterior synechiae, and cataracts. Secondary glaucoma frequently develops. The extent of these structural changes determinates the cours e of eye disease in Behçet's syndrome.

Musculoskeletal system

Behçet himself, as early as 1938, described 'rheumatoid' involvement in Behçet's syndrome. It can occasionally be the initial manifestation. Involvement of the joints (Table 2) is seen in about one-half of the patients in the form of arthritis or arthralgia (Yurdakul *et al*. 1983). It is mono- or oligoarticular but can be symmetrical. It is quite common to have symmetrical disease of the wrist or elbow, which can be confused with rheumatoid arthritis. Usually lasting a few weeks it seldom leads to deformity. Erythema of the overlying skin is not seen. Erosions are uncommon. Chronic synovitis lasting months to years with deformity and erosions are seen, but rare. The synovial fluid is inflammatory (see below). The histological changes are non-specific.

Seen in 50 per cent of the patients either as arthritis or arthralgia
Non-deforming and short lasting (few weeks)
Knees, ankles, wrists, and elbows are involved most commonly
Sacroiliac joint involvement is not prevalent in controlled studies
Usually mono- or oligoarticular but can be symmetrical with a potential for confusion with rheumatoid arthritis
Synovial fluid is commonly inflammatory but a good mucin clot is usual
Synovial histology is non-diagnostic; paucity of plasma cells and superficial-layer ulceration are the most outstanding features
Aseptic necrosis of the bone not related to steroid use can be seen

Table 2 Features of arthritis in Behçet's syndrome

Knees are the most commonly affected joints, followed in frequency by ankles, wrists, and elbows. There has been much debate about the involvement of the sacroiliac joints in Behçet's syndrome, but an increased prevalence has not been found in controlled studies. Furthermore there is considerable interobserver variation in interpretation of the sacroiliac joint on a plain radiograph of the pelvis (Yazici *et al*. 1987a). Regardless of the presence or absence of involvement of this joint, back pain is quite rare in Behçet's syndrome. The main debate about sacroiliac involvement centres around the erroneous inclusion of Behçet's syndrome among the seronegative spondylarthropathies (Moll *et al*. 1974). Other reasons why Behçet's syndrome is not a spondylarthritis include the lack of familial association with the spectrum of spondylarthritides, the association with HLA B51 and, most importantly, the presence of widespread vasculitis. In addition, the nature of the 'shared' clinical features in Behçet's syndrome is very different from that in this group of diseases. The genital ulceration is usually scrotal, urogenital infection is absent, nail changes are not seen, and the nature and course of eye involvement are totally different (Yazici 1987).

Myositis is occasionally found in Behçet's syndrome. It is usually local but generalized forms can also be seen. The muscle enzymes are not raised in the local forms and the histological features are indistinguishable from those of polymyositis.

Another musculoskletal manifestation associated with Behçet's syndrome is aseptic necrosis of the bone. This is possibly related to vasculitis and not necessarily to steroid use.

Neurological involvement (Fig. 5)

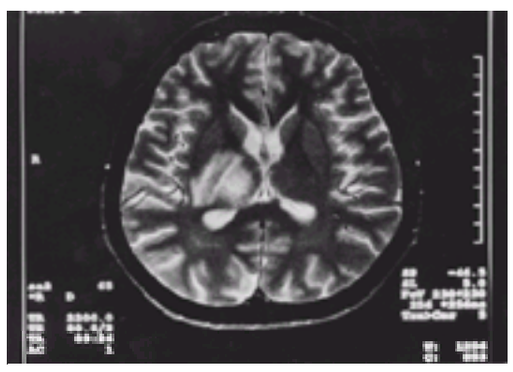


Fig. 5 Involvement of the central nervous system (T_2 -weighted MRI showing a hyperintense lesion in the right thalamus, capsula interna, and nucleus lentiformis in a 28-year-old male patient).

There is much variation in the reported prevalence rates of neurological involvement. In a prospective survey a prevalence rate of 5 per cent was found (Serdaroğlu *et al*. 1989). Pyramidal signs are the most common, followed by cerebellar and sensory symptoms and signs. Symptoms of increased intracranial pressure and meningeal irritation are also observed. Dementia can develop in an occasional patient. In the majority of the cases, headaches can be related to the basic pathology only if there are other associated signs and symptoms in the central nervous system. Papilloedema usually indicates occlusion of a venous sinus (Wechsler *et al*. 1992). As is true with eye involvement, the most severe forms of central nervous involvement are seen in the male.

The findings in cerebrospinal fluid are non-specific. Opening pressures and protein content are usually increased. Pleocytosis is present. Glucose can be low. Computed tomographic scans are of limited value unless localizing symptoms are present. Magnetic resonance imaging may prove to be more sensitive and specific.

The most common site of involvement is the brainstem. Hemispheric, meningeal, and spinal cord lesions are also seen. In contrast to the other vasculitides, peripheral nerve disease is unusual.

The histological appearances of the central nervous lesions in Behçet's syndrome are non-specific. There are inflammatory and degenerative changes. Frank vasculitis is not seen. The usual vascular change is a perivascular infiltration, sometimes with perivascular lamellar fibrosis. Microabscesses can also be observed.

Some patients develop psychiatric problems. Reliable prevalence rates are lacking. As in systemic lupus some of the psychiatric problems in Behçet's syndrome may be the result of a situational response to a chronic disease and occasionally of steroid use.

Cardiovascular and pulmonary involvement

Cardiac involvement

Endocarditis, myocarditis, and pericarditis can all occur but are rare. Cases with coronary vasculitis and ventricular aneurysms have also been documented. In a prospective controlled survey among 64 patients we failed to find an increased prevalence of cardiac disease (Özkan *et al*. 1992).

Venous lesions

Involvement of the veins is one of the main manifestations of Behçet's syndrome. In fact, Behçet's syndrome, together with systemic lupus and Buerger's disease, is one of the few vasculitides that can involve the venous side of the circulatory system together with the arterial side. Furthermore, in contrast to the other two, it can involve the vena cavae.

Thrombophlebitis occurs in 25 per cent of all patients. More frequent in the calf, it is also seen in veins of the upper extremity. There is a propensity to develop venous thromboses after venepunctures. A frequent outcome of thrombosis in the lower extremities is dermatitis of chronic stasis associated with recurring skin ulcers. Thrombosis of the iliac veins and both vena cavae is less frequent and seen almost only in male patients. This is usually less associated with the formation of collateral vessels. Occlusion of the suprahepatic veins can cause a Budd–Chiari syndrome, which carries a high mortality.

In the common variety of thrombophlebitis, for example of the pelvic veins in the postoperative state, only a short segment of the vessel wall is diseased. The thrombus is adherent to the wall at this site and has a long, non-adherent tail. This is not so in Behçet's syndrome. Here large segments of the vessel wall are diseased and consequently the thrombus formed is adherent to the wall at all places. Sometimes the entire length of the inferior vena cava is thus affected. This explains, we believe, why pulmonary embolism is rare in Behçet's syndrome even though thrombophlebitis is seen in one-quarter of the patients.

Arterial lesions (Fig. 6 and Fig. 7)

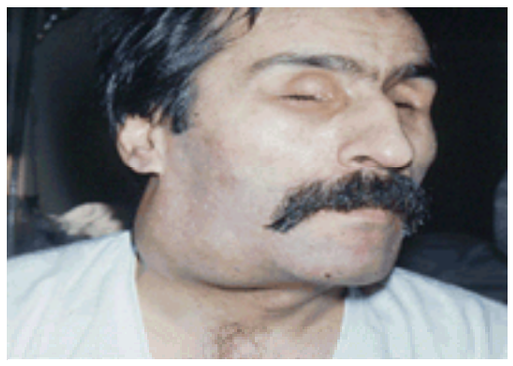


Fig. 6 Carotid artery aneurysm.

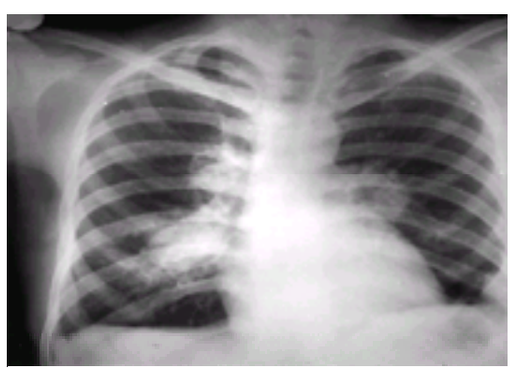


Fig. 7 Pulmonary artery aneurysms.

Starting with the aorta, all the arterial tree can be afflicted (Hamza 1987; Koç *et al*. 1992). Aneurysms of the abdominal aorta, carotid, femoral, popliteal, and less commonly coronary arteries can be seen (Fig. 6). The rupture of these aneurysms is frequently fatal. The basic pathology is thought to be a vasculitis of the vasa vasorum. There is a high recurrence rate after reconstructive surgery of the aneurysms of the peripheral vessels. This complication is thought to be related to the pathergy phenomenon.

The basic pulmonary pathology in Behçet's syndrome is also related to vasculitis. Pulmonary vascular changes in the form of arterial aneurysms, arterial and venous thromboses, and pulmonary infarcts are found (Efthimiou *et al*. 1986; Hamuryudan *et al*. 1994a). The classic radiographic finding is the non-cavitating mass lesion (Fig. 7). Computed tomographic scans are usually adequate for diagnosis.

Pulmonary arterial aneurysms in Behçet's syndrome carry a grave prognosis. Among our patients the mortality was 50 per cent despite treatment, the terminal event usually being the rupture of an aneurysm into a bronchus.

Pleural effusions are seldom in Behçet's syndrome. They may be associated with the pulmonary vascular disease or may be present independently.

Gastrointestinal involvement

While gastrointestinal involvement is seen in about one-third of patients from Japan (Shimizu *et al*. 1979), it is quite rare among patients reported from the Mediterranean basin. The basic pathology is that of mucosal ulceration. This is seen most commonly in the ileum, followed by the caecum and other parts of the colon. Occasionally oesophageal ulceration can occur. Histologically the ulcers are indistinguishable from those found in ulcerative colitis. Intestinal lymphangiectasia have also been reported. The course of intestinal Behçet's syndrome is that of exacerbations and remissions. The usual symptoms are abdominal pain and melaena. A mass is often palpable in the abdomen. The ileocaecal ulcers have the worst prognosis, with a distinct tendency to perforate.

Hepatic problems are not common in Behçet's syndrome except when an associated Budd–Chiari syndrome is present. A slightly enlarged spleen is noticed in 20 per cent of the male patients (Soysal *et al*. 1990), not necessarily related to portal hypertension secondary to venous thrombosis.

Other clinical features

Renal involvement is seen much less frequently than one would expect in a systemic vasculitis. There are occasional reports of glomerulonephritis (Hamuryudan *et al*. 1991b). Epididymitis is seen around 5 per cent of the patients.

Amyloidosis is seen sporadically and when present usually presents with a nephrotic syndrome. It is of the AA type and carries a grave prognosis. In a rectal biopsy survey of non-symptomatic patients without any proteinuria, no cases of silent amyloidosis could be identified (Yurdakul *et al*. 1990).

Diagnosis

The full-blown syndrome is easy to identify; the so-called incomplete forms sometimes cause problems. Reiter's and Stevens–Johnson syndromes can also be problems in differential diagnosis for the uninitiated. However, an occasional patient with inflammatory bowel disease associated with oral ulcers, skin lesions, and episcleritis may be impossible to differentiate from Behçet's syndrome. Another difficult aspect is a patient with multiple sclerosis with one or two features of Behçet's syndrome.

Many criteria have been proposed to diagnose Behçet's syndrome. Recently a set of classification criteria have been proposed by a computer analysis of clinical features of 914 cases collected around the globe (International Study Group for Behçet's Disease 1990). Diseased controls with features that may be confused with Behçet's syndrome were also included in the analysis. Using the presence of recurrent aphthous ulceration as mandatory, together with proper exclusions, these criteria require involvement of two other organ systems. In this scheme a positive pathergy test can replace involvement of an organ system.

Laboratory investigations

There are no laboratory findings specific for Behçet's syndrome. A moderate anaemia of chronic disease and leucocytosis are seen in 15 per cent of the patients. Neither reflects the clinical activity. The erythrocyte sedimentation rate is only mildly elevated, as is the C-reactive protein. Again neither correlates well with disease activity (Müftüoğlu *et al*. 1986).

The synovial fluid is usually inflammatory. The cell count is between 5 000 and 50 000/mm ^{3} with neutrophils predominating. Despite this high cell count the mucin clot is usually good (Yurdakul *et al*. 1983).

Serum immunoglobulins are sometimes elevated, while autoantibodies are absent. Complement levels may be high. Despite the pauci-immune nature of the basic disease process, antineutrophilic antibodies are not a feature of Behçet's syndrome.

Pathogenesis

The prevailing opinion is that Behçet's syndrome is caused by a defect in immunoregulation and this defect in turn is triggered by an infectious agent(s) (Lehner *et al*. 1991). Among the candidates for the infectious agents are herpes simplex type 1 (Denman *et al*. 1980; Eglin *et al*. 1982) and some strains of streptococcus (Mizushima *et al*. 1988). Microbial and human heat shock proteins (**HSP**) show marked amino acid homology. Behçet's syndrome patients have both lymphocytes and antibodies which recognize certain synthetic peptides derived from mycobacterial HSP. Likewise a similar recognition exists to human mitochondrial HSP peptides that are homologous to mycobacterium HSP peptides. Finally these peptides are uveitogenic in rats (Stanford *et al*. 1994). These observations indicate that multiple agents, through molecular mimicry of their HSP peptides to the human tissues, can initiate an immune response and organ disease in Behçet's syndrome.

The evidence for immunological aberration is abundant (Arbesfeld and Kurban 1988). Antibodies to oral mucosa, circulating immune complexes, lymphocytotoxins in serum, a decrease in the OKT4:OKT8 ratio and in natural killer cells during disease exacerbations are among these. Increased levels of soluble interleukin-2 receptors, a common finding in autoimmune disease, also occur in Behçet's syndrome (Hamzoui and Ayed 1989). There is also an increased *in vivo* accumulation of activated lymphocytes expressing g T-cell receptors at the sites of inflammation (Hamzoui *et al*. 1994).

In contrast, the lack of autoantibodies, the rarity of associated Sjögren's syndrome (Günaydin *et al*. 1994), the lack of association with other autoimmune diseases both among the patients and their relatives, the rather poor response to steroids, and finally the more severe disease expression among males are points against an autoimmune mechanism in Behçet's syndrome (Yazici 1987).

The heightened inflammatory response, best manifested by the pathergy reaction (Tüzün *et al*. 1979), prompted research into leucocyte activity. Polymorphonuclear leucocyte activity is certainly increased (Matsumura and Mizushima 1975). A transgenic mice model has also been developed which attempts to provide an explanation for this heightened inflammatory activity (Takeno *et al*. 1995).

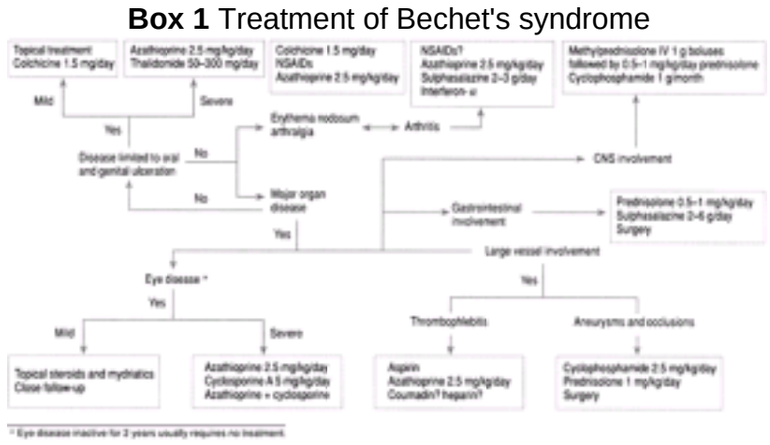
The widespread vascular disease in Behçet's syndrome has justifiably turned attention to endothelial pathology, and the increase in thromboses to problems with coagulation. Increased levels of factor VIII-related antigen are found only in patients with major (i.e. pulmonary artery and its branches; vena cavae) involvement of vessels (Yazici *et al*. 1987b). There is also a defect in prostacyclin production from the vessel wall (Kansu *et al*. 1986). Two other stigmas of endothelial injury in this syndrome are the presence of the antiendothelial cell antibodies (Aydintuğ *et al*. 1993) and increased serum levels of endothelin (Koç *et al*. 1993). Both observations might obviously be secondary events. Fibrinolytic activity is defective in Behçet's syndrome (Hampton *et al*. 1991). However, no single abnormality has been shown to account for this. The initial reports of increased levels of antiphospholipid antibodies have not been confirmed in subsequent studies.

Any theory that tries to explain the pathogenesis of Behçet's syndrome should also take into consideration the reasons for the more severe disease seen in the male and the young (Yazici *et al*. 1984). The pathergy reaction is also more strongly positive in the male (Yazici *et al*. 1985). This brings sex hormones into consideration. Serum concentrations of the major sex hormones are normal among the patients; however, this does not negate a possible influence of hormonal effects via metabolic products or alterations of end-organ responsiveness. In this respect the so-called acneiform skin lesion of Behçet's syndrome is clinically and histologically indistinguishable from ordinary acne, an androgen-dependent lesion.

Management (Yazici *et al*. 1995)

There are several important features of Behçet's syndrome that have to be taken into consideration when planning the management: (i) the usual course of the syndrome in any organ system is that of exacerbations and remissions with the overall activity generally abating with the passage of time; thus the principal aim is to prevent irreversible structural damage, which is the outcome of the early stormy course; (ii) being young and male are separate and additive negative prognostic factors; (iii) eye disease usually occurs, if at all, either initially or within the first few years; (iv) the syndrome can be fatal, especially in the young male; and (v) there are many patients with Behçet's syndrome who do not need any treatment but reassurance. Box 1 contains a scheme for the treatment of Behçet's syndrome.

Box 1 Treatment of Bechet's syndrome



Immunosuppressive drugs are the main line of treatment for eye involvement. Although corticosteroids have been used for a long time, there is no formal evidence that they are effective. In fact there is suspicion that they may be harmful. Their short-term use over a few months, however, may shorten the duration of an attack. Large centres treating Behçet's syndrome are now using steroids less frequently.

In the only controlled, double-blind study of cytotoxic immunosuppressive agents, azathioprine at 2.5 mg/kg per day was shown to be superior to placebo in maintaining visual acuity and perhaps more importantly in preventing the emergence of new eye disease (Yazici *et al*. 1990). It was not useful in restoring the already compromised vision. It would be unrealistic to expect restoration of impaired visual acuity when this is due to structural changes like synechiae. Thus it is important that treatment is begun well before these changes appear. Other cytotoxic, immunosuppressive agents such as chlorambucil and cyclophosphamide are also used, and extrapolating from the controlled experience with azathioprine their use is justified (O'Duffy 1990).

Cyclosporin is an effective and rapidly acting drug in the uveitis of Behçet's syndrome (Masuda *et al*. 1989). Problems with cyclosporin are the potential nephrotoxicity, especially at doses greater than 5 mg/kg per day, and the very frequent relapses after cessation of therapy. The high cost is another problem. Eye disease in remission for 2 years or more needs no further treatment. Young males usually need to be treated more vigorously. Once initiated, the usual course for cytotoxic or cyclosporin therapy is a minimum of 2 years, after which attempts at discontinuation are made. In some patients, after a course for 6 to 8 months of combined azathioprine and cyclosporin, treatment is continued with azathioprine only. In more resistant cases azathioprine is used in combination with cyclosporin (both at conventional doses) for extended periods of time. Our uncontrolled experience with this mode of therapy in severe eye disease is quite favourable.

Structural damage to the eye can be managed surgically (i.e. vitrectomy) at specialized centres. However, the results are not uniformly satisfactory. There is always the problem of new attacks of inflammation in surgically handled tissue in Behçet's syndrome. Also the already established disease in the retina can not be helped by surgery. Local mydriatics should be used in the acute stage to prevent synechiae.

The oral and genital ulcers are usually well controlled by immunosuppressives and steroids. However, these should be reserved for more severe cases. Most of the time, a local steroid preparation that adheres to fresh ulceration (such as triamcinolone acetonide oral paste) is all that is required. Thalidomide in doses ranging from 50 mg on 3 nights a week to 300 mg daily seems to be highly effective in the treatment of the orogenital ulcers of Behçet's syndrome. Limitations for its use are teratogenicity and peripheral neuropathy. The incidence and relationship of the neuropathy to the dose used is a matter of debate (Denman *et al*. 1993).

Colchicine has been claimed to be beneficial for almost every manifestation of Behçet's syndrome, but without any formal trials. The rationale for its use is the increased chemotaxis of the polymorphonuclear leucocytes found in this syndrome. In the only controlled study with colchicine it was superior to placebo only in the treatment of erythema nodosum and arthralgias (Aktulga *et al*. 1980). When started the usual practice is to use the colchicine for several years.

Cyclophosphamide, either 2 to 2.5 mg/kg per day orally or 500 to 1500 mg weekly or monthly intravenous boluses, is required to treat those patients with severe cutaneous, arterial, or pulmonary vasculitis or those with arterial aneurysms or vena caval involvement. There is no formal experience with any therapy for central nervous disease; however, steroids and immunosuppressives are again used. Steroids are usually employed in the form of pulsed intravenous methylprednisolone at a dosage of 1 g for 3 to 7 days followed by oral prednisolone at 0.5 to 1 mg/kg per day for a few months.

Gastrointestinal involvement is initially managed by sulphasalazine at a dosage of 2 to 6 g/day. Sometimes surgery is required, with resection of large segments of bowel. Usually a good portion of uninvolved area should also be removed to prevent recurrences. Surgery is also mandatory for aneurysms of peripheral vessels.

Controlled studies have not shown either the viral agent acyclovir, or transfer factor to be effective. In an open, self-controlled study a-interferon was shown to be effective in controlling some of the extraocular manifestations, especially the arthritis, of the syndrome (Hamuryudan *et al*. 1994b).

There is debate whether to use heparin or oral anticoagulants for the thrombophlebitis of Behçet's syndrome. Pulmonary embolism is seldom observed, as explained above. In this setting antiplatelet drugs (i.e. aspirin) are probably sufficient. We also use azathioprine in thrombophlebitis of Behçet's syndrome as an agent to suppress the disease activity in general.

Prognosis

We had in the past thought that Behçet's syndrome caused significant morbidity but not appreciable mortality. Our recent experience shows otherwise. Mention has already been made of those patients with pulmonary arterial aneurysms. We have also now assessed the 10-year mortality of 152 patients registered in our clinic in 1982. With 81 per cent follow-up information the minimum mortality was 6.5 per cent among our male patients. This was statistically more than expected for the 15 to 24 age bracket when compared with standard mortality statistics. In contrast, there was no observed mortality among the females (Yazici *et al*. 1996).

Apart from the bleeding pulmonary arterial aneurysms, the other causes of mortality in Behçet's syndrome are ruptured peripheral aneurysms, severe central nervous system disease, Budd–Chiari syndrome, and finally intestinal ulceration and perforation especially prominent among the Japanese.

Chapter References

Aktulga, E. *et al*. (1980). A double blind study of colchicine in Behçet's disease. *Haematologica*, **65**, 399–402.

Arbesfeld, S.J. and Kurban, A.K. (1988). Behçet's disease. New perspectives on an enigmatic syndrome. *Journal of the American Academy of Dermatology*, **19**, 767–79.

Aydintuğ, A.O. *et al*. (1993). Antibodies to endothelial cells in patients with Behçet's disease. *Clinical Immunology and Immunopathology*, **67**, 157–62.

Azizlerli, G., Özarmağan, G., Övül, C., Sarica, R., and Mustafa, S.O. (1992). A new kind of skin lesion in Behçet's disease: extragenital ulcerations. *Acta Dermalogica et Venereologica* (*Stockholm*), **72**, 286.

Bird Stewart, J.A. (1986). Genetic analysis of families of patients with Behçet's syndrome: data incompatible with autosomal recessive inheritance. *Annals of the Rheumatic Diseases*, **45**, 265–8.

Çakir, N. *et al*. (1991). Response to intradermal injection of monosodium urate crystals in Behçet's syndrome. *Annals of the Rheumatic Diseases*, **50**, 634–6.

Denman, A.M., Fialkow, P.J., Pelton, B.K., Salo, A.C., Appleford, D.J., and Gilchrist, C. (1980). Lymphocyte abnormalities in Behçet's syndrome. *Clinical and Experimental Immunology*, **42**, 175–85.

Denman, A.M., Graham, E., Howe, L., Denman, E.J., and Lightman, S. (1993). Low dose thalidomide treatment of Behçet's syndrome. In *Behçet's disease* (ed. B. Wechsler and P. Godeau), pp. 649–52, Excerpta Medica, Elsevier, Amsterdam.

DilŠen, N. *et al*. (1993). Comparative study of the skin pathergy test with blunt and sharp needles in Behçet's disease. Confirmed specificity but decreased sensitivity with sharp needles. *Annals of the Rheumatic Diseases*, **52**, 823–5.

Efthimiou, J., Johnston, C., Spiro, S.G., and Turner-Warwick, M. (1986). Pulmonary disease in Behçet's syndrome. *Quarterly Journal of Medicine*, **58**, 259–80.

Eglin, R.P., Lehner, T., and Subak Sharpe, J.H. (1982). Detection of RNA complementary to herpes-simplex virus in mononuclear cells from patients with Behçet's syndrome and recurrent oral ulcers. *Lancet*, **ii**, 1356–61.

Fresko, I., Yazici, H., Bayramiçli, M., Yurdakul, S., and Mat, C. (1993). Effect of surgical cleaning of the skin on the pathergy phenomenon in Behçet's syndrome. *Annals of the Rheumatic Diseases*, **52**, 619–20.

Gilhar, A., Winterstein, G., Turani, H., Landau, J., and Etzioni, A. (1989). Skin hyperreactivity response (pathergy) in Behçet's disease. *Journal of the American Academy of Dermatology*, **21**, 547–52.

Günaydin, I., Yazici, H., Özbakir, F., Hamuryudan, V., and Yurdakul, S. (1992). Lymphocytotoxic antibodies in the healthy children of patients with Behçet's syndrome. *Journal of Rheumatology*, **19**, 1996–8.

Günaydin, I;. *et al*. (1994). The prevalence of Sjögren's syndrome in Behçet's syndrome. *Journal of Rheumatology*, **21**, 1662–4.

Hampton, K.K., Chamberlain, M.A., Menon, D.K., and Davies, J.A. (1991). Coagulation and fibrinolytic activity in Behçet's disease. *Thrombosis and Haemostasis*, **66**, 292–4.

Hamuryudan, V., Yurdakul, S., Özbakir, F., Yazici, H., and Hekim, N. (1991 *a*). Monozygotic twins concordant for Behçet's syndrome. *Arthritis and Rheumatism*, **34**, 1071–2.

Hamuryudan, V., Yurdakul, S., Kural, A.R., Ince, Ü., and Yazici, H. (1991 *b*). Diffuse proliferative glomerulonephritis in Behçet's syndrome. *British Journal of Rheumatology*, **30**, 63–4.

Hamuryudan, V. *et al*. (1994*a*). Pulmonary arterial aneurysms in Behçet's syndrome: a report of 24 cases. *British Journal of Rheumatology*, **33**, 48–51.

Hamuryudan, V. *et al*. (1994*b*). Systemic interferon-a 2b treatment in Behçet's syndrome. *Journal of Rheumatology*, **21**, 1098–100.

Hamza, M. (1987). Large artery involvement in Behçet's disease. *Journal of Rheumatology*, **14**, 554–9.

Hamzoui, K. and Ayed, K. (1989). Soluble interleukin 2 receptors in patients with Behçet's disease. *Journal of Rheumatology*, **16**, 852–3.

Hamzoui, K. *et al*. (1994). Phenotype and functional profile of T cells expressing ld receptor from patients with Behçet's disease. *Journal of Rheumatology*, **21**, 2301–6.

International Study Group for Behçet's Disease (1990). Criteria for diagnosis of Behçet's disease. *Lancet*, **335**, 1078–80.

Kansu, E., Iahin, G., Iahin, F., Sivri, B., Sayek, I., and Batman, F. (1986). Impaired prostacyclin synthesis by vessel walls in Behçet's disease. *Lancet*, **ii**, 1154.

Koç, Y. *et al*. (1992). Vascular involvement in Behçet's disease. *Journal of Rheumatology*, **19**, 402–10.

Koç, Y. *et al*. (1993). Endothelin 1–2 levels in Behçet's disease. In *Behçet's disease* (ed. B. Wechsler and P. Godeau), pp. 97–101. Excerpta Medica, Elsevier, Amsterdam.

Lakhanpal, S., O'Duffy, J.D., and Lie, J.T. (1988). Pathology. In *Behçet's disease. A contemporary synopsis* (ed. G.R. Plotkin, J.J. Calabro, and J.D. O'Duffy), pp. 101–42. Futura, New York.

Lehner, T., Fortune, F., and Studd, M. (1991). Recent advances in T cell immunoregulation and in the microbial causes of Behçet's disease. In *Behçet's disease* (ed. J.D. O'Duffy and E. Kökmen), pp. 463–73. Marcel Dekker, New York.

Lie, J.T. (1992). Vascular involvement in Behçet's disease: arterial and venous and vessels of all sizes. *Journal of Rheumatology*, **19**, 341–3.

Masuda, K., Nakajima, A., Urayama, A., Nakae, K., Kogure, M., and Inaba, G. (1989). Double-masked trial of cyclosporin versus colchicine and long-term open study of cyclosporin in Behçet's disease. *Lancet*, **i**, 1093–6.

Matsumura, N. and Mizushima, Y. (1975). Leucocyte movement and colchicine treatment in Behçet's disease. *Lancet*, **ii**, 813.

Mizushima, Y., Matsuda I., Hoshi, K., and Ohno, S. (1988). Induction of Behçet's disease symptoms after dental treatment and streptococcal antigen skin test. *Journal of Rheumatology*, **15**, 1029–30.

Moll, J.M., Haslock, I., Macrae, I.F., and Wright, V. (1974). Associations between ankylosing spondylitis, psoriatic arthritis, Reiter's disease, the intestinal arthropathies, and Behçet's. *Medicine* (*Baltimore*), **53**, 343–64.

Müftüoğlu, A.U. *et al*. (1986). Behçet's disease. Relation of serum C-reactive protein and erythrocyte sedimentation rates to disease activity. *International Journal of Dermatology*, **25**, 235–9.

O'Duffy, J.D. (1990). Behçet's syndrome. *New England Journal of Medicine*, **322**, 326–8.

Ohno, S., Ohguchi, M., Hirose, S., Matsuda, H., Wakisaka, A., and Aizawa, M. (1982). Close association of HLA-Bw51 with Behçet's disease. *Archives of Ophthalmology*, **100**, 1455–8.

Özdoğan, H. (1994). Behçet's syndrome in children. *Rheumatology in Europe*, **23** (Suppl. 12), 34.

Özkan, M. *et al*. (1992). M-mode, 2-D and Doppler echocardiographic study in 65 patients with Behçet's syndrome. *European Heart Journal*, **13**, 638–41.

Serdaroğlu P., Yazici, H., Özdemir, C., Yurdakul, S., Bahar, S., and Aktin, E. (1989). Neurologic involvement in Behçet's syndrome. A prospective study. *Archives of Neurology*, **46**, 265–9.

Shimizu, T., Ehrlich, G.E., Inaba, G., and Hayashi, K. (1979). Behçet disease (Behçet syndrome). *Seminars in Arthritis and Rheumatism*, **8**, 223–60.

Soysal, M. *et al*. (1990). Splenic size in Behçet's syndrome. *British Journal of Rheumatology*, **29**, 497–8.

Stanford, M.R. *et al*. (1994). Heat shock protein peptides reactive in patients with Behçet's disease are uveitogenic in Lewis rats. *Clinical and Experimental Immunology*, **97**, 226–31.

Takeno, M. *et al*. (1995). Excessive function of peripheral blood neutrophils from patients with Behçet's disease and from HLA B51 transgenic mice. *Arthritis and Rheumatism*, **38**, 426–43.

Tüzün, Y., Yazici, H., Pazarli, H., Yalçin, B., Yurdakul, S., and Müftüo ğlu, A. (1979). The usefulness of the nonspecific skin hyperreactivity (the pathergy test) in Behçet's disease in Turkey. *Acta Dermatolgica et Venereologica* (*Stockholm*), **59**, 77–9.

Wechsler, B. *et al*. (1992). Cerebral venous thrombosis in Behçet's disease: clinical study and long-term follow-up of 25 cases. *Neurology*, **42**, 614–18.

Yazici, H. (1987). Behçet's syndrome, a personal view. *Clinical and Experimental Rheumatology*, **5**, 1–3.

Yazici, H., Chamberlain, M.A., Schreuder, I., D'Amaro, J., and Müftüo ğlu, A. (1980). HLA antigens in Behçet's disease: a reappraisal by a comparative study of Turkish and British patients. *Annals of the Rheumatic Diseases*, **39**, 344–8.

Yazici, H. *et al*. (1984). Influence of age of onset and patient's sex on the prevalence and severity of manifestations of Behçet's syndrome. *Annals of the Rheumatic Diseases*, **43**, 783–9.

Yazici, H. *et al*. (1985). Male patients with Behçet's syndrome have stronger pathergy reactions. *Clinical and Experimental Rheumatology*, **3**, 137–41.

Yazici, H., Turunç, M., Özdogan, H., Yurdakul, S., Akinci, A., and Barnes, C.G. (1987 *a*). Observer variation in grading sacroiliac radiographs might be a cause of 'sacroiliitis' reported in certain disease states. *Annals of the Rheumatic Diseases*, **46**, 139–45.

Yazici, H. *et al*. (1987*b*). Von Willebrand factor in Behçet's syndrome. *Journal of Rheumatology*, **14**, 305–6.

Yazici, H. *et al*. (1990). A controlled trial of azathioprine in Behçet's syndrome. *New England Journal of Medicine*, **322**, 281–5.

Yazici, H., Yurdakul, S., and Hamuryudan, V. (1995). Behçet's syndrome: how should we treat it? *Clinical Immunotherapy*, **3** (2), 102–7.

Yazici, H. *et al*. (1996). The ten-year mortality in Behçet's syndrome. *British Journal of Rheumatology*, **35**, 139–41.

Yurdakul, S. *et al*. (1983). The arthritis of Behçet's disease: a prospective study. *Annals of the Rheumatic Diseases*, **42**, 505–15.

Yurdakul, S. *et al*. (1988). The prevalence of Behçet's syndrome in a rural area in northern Turkey. *Journal of Rheumatology*, **15**, 820–2.

Yurdakul, S., Tüzüner, N., Yurdakul, I., and Hamuryudan, V. (1990). Amyloidosis in Behçet's syndrome. *Arthritis and Rheumatism*, **33**, 1586–9.

5.11.8 Primary vasculitis in children

Michael J. Dillon

Vasculitis occurs in many different diseases and syndromes in childhood (Hicks 1988). It is the predominant manifestation of the disorder in some but in others may be one aspect of a more widespread autoimmune rheumatic disease (Ansell 1980; Cassidy and Petty 1995). As a result classification is difficult and is not entirely satisfactory (Dillon 1989). A classification that has proved useful, based on that of Fink (1986), is outlined in Table 1. There are other classifications including those of the American College of Rheumatology Subcommittee on Classification of Vasculitis (Hunder *et al*. 1990) and the Chapel Hill Conference on the Nomenclature of Systemic Vasculitis (Jennette *et al*. 1994). The latter is reproduced in Chapter 5.11.1. However, for the purposes of this chapter the classification shown in Table 1 will be utilized.

Polyarteritis
Macroscopic
Microscopic
Kawasaki disease
Mucocutaneous lymph node syndrome
Granulomatous vasculitis
Wegener's granulomatosis
Churg–Strauss syndrome
Leucocytoclastic vasculitis
Henoch–Schönlein purpura
Hypersensitivity angiitis
Cutaneous polyarteritis
Vasculitis associated with connective tissue disease
Systemic lupus erythematosus
Juvenile chronic arthritis
Mixed connective tissue disease
Dermatomyositis
Scleroderma
Giant cell arteritis
Takayasu disease
Miscellaneous vasculitides

Table 1 Childhood vasculitis

Polyarteritis (macroscopic and microscopic)

The classic form of polyarteritis, usually referred to as polyarteritis nodosa, is a necrotizing vasculitis associated with aneurysmal nodules along the walls of medium-sized muscular arteries and is akin to that described initially by Kussmaul and Maier (1866). Although there is some overlap with smaller vessel disease, polyarteritis nodosa appears to be a distinct entity and despite being rare can and does occur in childhood (Blau *et al*. 1977; Magilavy *et al*. 1977; Reimold *et al*. 1977). It is of interest that this type of vasculitis is less common in adults in whom microscopic polyangiopathy and Wegener's granulomatosis predominate.

The average age of patients with polyarteritis at the Hospital for Sick Children in London, was 7.5 years (ranging from 17 days to 15 years) with a male:female ratio of 1.6:1 (Dillon 1989; Dillon 1990). Many children on presentation were initially misdiagnosed with infections, autoimmune rheumatic disorders and malignant diseases often considered before the correct diagnosis was established. Although in the majority of cases reported genetic factors do not play a part, the occasional familial case has been described (Mason *et al*. 1994).

Microscopic polyarteritis differs from classic polyarteritis nodosa by the presence of extensive glomerular involvement and is defined as small vessel vasculitis with focal segmental glomerulonephritis, but without granulomatous disease of the respiratory tract (Savage *et al*. 1985). It can be difficult to distinguish it from Wegener's granulomatosis, especially if the latter presents atypically and often presents as idiopathic crescentic glomerulonephritis which is considered currently to be a manifestation of this form of polyarteritis (Jardim *et al*. 1992).

Clinical manifestations

The main clinical features of the macroscopic disease are malaise, fever, skin rash, abdominal pain and arthropathy (Reimold *et al*. 1977; Dillon 1989; Dillon 1990; Cassidy and Petty 1995). In addition myalgia, ischaemic heart and testicular pain, renal manifestations such as haematuria and hypertension, neurological features such as focal deficits, hemiplegia, visual loss, mononeuritis multiplex and organic psychosis can occur. Skin lesions are variable and may masquerade as those of Henoch–Schönlein purpura or multiform erythema, but can also be necrotic and associated with peripheral gangrene (Fig. 1). System involvement in the Hospital for Sick Children patients was as follows: renal 63 per cent; cardiac 26 per cent; neurological 26 per cent; respiratory 21 per cent; and gut 15 per cent (although a much higher percentage of patients with abdominal pain implied that gut disease was more common) (Dillon 1990). This spectrum of disease activity and distribution in the referred patients to the above hospital is not dissimilar to other published data, even to the extent of renal involvement, which is reported in 50 to 60 per cent of children with polyarteritis nodosa (Blau *et al*. 1977; Magilavy *et al*. 1977; Reimold *et al*. 1977; Cassidy and Petty 1995). Similar features may also be seen occasionally in the microscopic variety since extra renal manifestations are reported (Gaskin and Pusey 1992). However, the predominant presentation in this type of vasculitis is renal impairment with proteinuria, haematuria and hypertension (Coward *et al*. 1986; Adu *et al*. 1987).

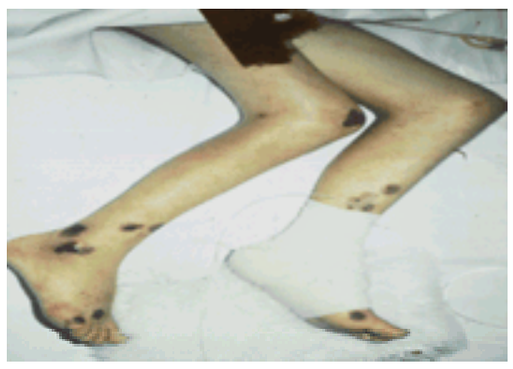


Fig. 1 Necrotic skin lesions in polyarteritis nodosa

Laboratory and radiological investigations

Investigation usually demonstrates anaemia, polymorphonuclear leucocytosis, thrombocytosis, an increased erythrocyte sedimentation rate and C reactive protein (Blau *et al*. 1977; Magilavy *et al*. 1977; Reimold *et al*. 1977; Dillon 1989; Cassidy and Petty 1995). Platelets are hyperaggregable and there are circulating immune complexes in plasma (Levin *et al*. 1985). The plasma of some patients with active disease has a greenish colour due to a caeruloplasmin-like acute reactant that can be diagnostically helpful. Antineutrophil cytoplasmic antibodies (ANCA) have been detected in certain forms of polyarteritis (Jennette and Falk 1990; Dillon and Tizard 1991) and the indirect immunofluorescence appearance, the immunoglobulin class and the antigen to which the antibody is raised all have important bearings on the diagnositic value of the observations. In macroscopic polyarteritis nodosa ANCA have been considered to be absent but recently, in adults, both cytoplasmic and perinuclear patterns of ANCA staining by indirect immunofluorescence have been reported (Gross *et al*. 1991) and this has been confirmed in children (Wong *et al*. 1992). In microscopic polyarteritis ANCA findings are heterogeneous, with predominantly a perinuclear pattern of antibody staining using the indirect immunofluorescence technique and evidence that myeloperoxidase is the antigen involved (Jennette and Falk 1990; Dillon and Tizard 1991; Wong *et al*. 1992).

Biopsy material may prove to be diagnostically important, especially if an affected area or lesion in skin, muscle or other tissue is available. It is worth emphasizing, however, that the characteristic histopathological changes may not be present in sections examined and this would not necessarily exclude the diagnosis, since vasculitis is very variable in terms of the extent that tissues are affected. Renal biopsy is more valuable in diagnosing microscopic polyarteritis when features of crescentic glomerulonephritis may be revealed (Coward *et al*. 1986). Macroscopic vasculitis affecting the renal vessels may be suggested by patchy areas of decreased isotope uptake on technetium-99 DMSA (dimercaptosuccinic acid) scanning (Dillon 1990). However, the most valuable investigative procedure for macroscopic disease is renal and hepatic angiography (Fig. 2(a and b)) (McLain *et al*. 1972; Yousefzadeh *et al*. 1981; Dillon 1989; Dillon 1990).

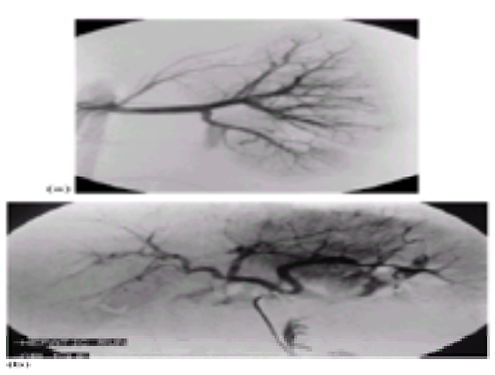


Fig. 2 (a) Renal angiogram showing small peripheral aneurysms in polyarteritis nodosa. (b) Coeliac angiogram showing abnormal hepatic vessels with aneurysms in polyarteritis nodosa.

Treatment and prognosis

Treatment for macroscopic and microscopic polyarteritis includes the use of steroids and oral cyclophosphamide, although pulsed intravenous injections of the latter are advocated by some. There is also a case for administering some antiplatelet therapy such as dipyridamole, but aspirin is usually contraindicated since the patients initially are on high-dose steroids and there may be a risk of gastrointestinal bleeding. Plasma exchange (Tizard and Dillon 1991) and intravenous prostacyclin have roles in life-threatening situations. Recently high-dose intravenous gammaglobulin has been introduced with some success (Jayne *et al*. 1991). Azathioprine and cyclosporin appear to have a role in maintaining remission but are not that helpful in inducing remission. A number of reviews deal with the therapy of childhood vasculitis (Cameron 1988; Larsson and Baum 1988; Dillon 1990; Roberti *et al*. 1993).

Prognosis is variable and mortality is at times over 20 per cent in spite of aggressive therapy (Fink 1977; Dillon 1989; Dillon 1990). However, in some series of childhood polyarteritis the outcome has been substantially better, at times with steroid alone, raising the question as to the precise diagnosis in and the severity of the disease in the affected children (Blau *et al*. 1977; Magilavy *et al*. 1977; Reimold *et al*. 1977).

Kawasaki disease (mucocutaneous lymph node syndrome)

This systemic childhood vasculitis was first described by Tomisaku Kawasaki in Japan in 1967 (Kawasaki 1967). Over 80 000 cases have subsequently been reported from that country alone, but the disease is of worldwide distribution and mainly affects infants and young children under 5 years of age. There is a male preponderance, an ethnic bias towards oriental or Afro-Caribbean children, some seasonality, and occasional epidemics (Hicks and Mellish 1986; Rowley *et al*. 1988; Kawasaki 1989; Shulman 1987). The incidence in Japan is 150 per 100 000 children under 5 years of age per year, but elsewhere it is substantially less: 10.3 in the United States, 2.9 in Germany and 3.4 in the United Kingdom (Shulman 1987; Kawasaki 1989; Dhillon *et al*. 1993).

Clinical features

The principal symptoms are outlined in Table 2 (Japan Kawasaki Disease Research Committee 1984), and illustrated in Fig. 3 and Fig. 4. For the diagnosis to be made at least five of the six items should be satisfied. Patients with four items can be diagnosed if coronary artery aneurysms are seen on two-dimensional echocardiography or coronary angiography. The incidence of cardiovascular complications can be up to 35 per cent if transient coronary artery dilatation, pericardial effusion, ECG abnormalities, pericarditis, myocardial infarction, ventricular aneurysms, mitral incompetence and cardiac failure are included (Table 3) (Rowley *et al*. 1988; Susuki *et al*. 1990; Tizard *et al*. 1991b). The incidence of coronary artery aneurysms (Fig. 5) ranges from 20 to 30 per cent. Cases of incomplete Kawasaki disease, however, are being recognized and the criteria may need to be reviewed in the future. Other significant symptoms or findings in Kawasaki disease are shown in Table 3 (Japan Kawasaki Disease Research Committee 1984).

Fever persisting for 5 days or more
Changes in the peripheral extremities (reddening of palms and soles, indurative oedema in initial stage, and membranous desquamation of the skin of the hands and feet in the convalescent phase)
Polymorphous exanthema
Bilateral conjunctival injection
Changes of lips and oral cavity (reddening of lips, strawberry tongue, diffuse injection of oral and pharyngeal mucosa)
Acute non-purulent cervical lymphadenopathy

Table 2 Principal symptoms of Kawaski disease

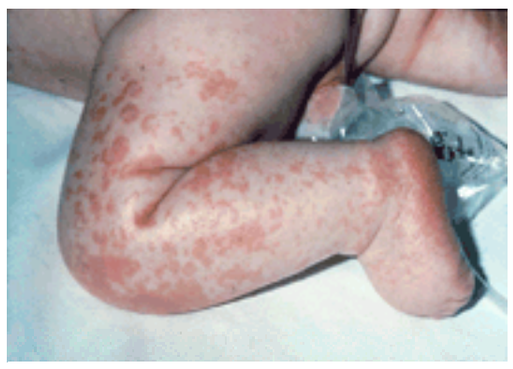


Fig. 3 Kawasaki disease showing characteristic rash, redness of the sole and oedema of the dorsum of the foot.

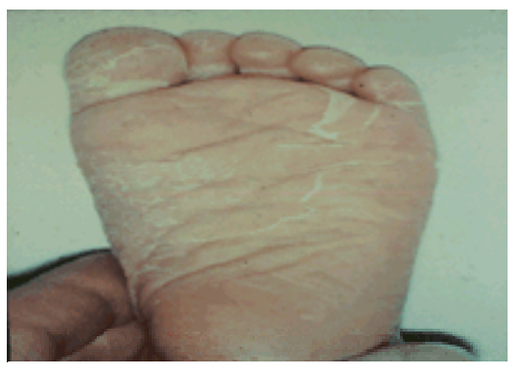


Fig. 4 Kawasaki disease showing membranous desquamation of skin of the foot in convalascent phase.

The cardiovascular system: heart murmurs, gallop rhythm, ECG changes, cardiomegaly, two-dimensional echo findings of pericardial effusion, coronary artery aneurysms, aneurysms of peripheral arteries, angina pectoris and myocardial infarction
The gastro-intestinal tract: diarrhoea, vomiting, abdominal pain, hydrops of the gallbladder, ileus and jaundice
The blood: leucocytosis, thrombocytosis, increased erythrocyte sedimentation rate, increased C reactive protein, hypo-albuminaemia, and anaemia
The urine: proteinuria, increased leucocytes in sediment
The skin: transverse furrows of finger nails
The respiratory tract: cough and rhinorrhoea
The joints: pain and swelling
The neurological system: pleocytosis in cerebrospinal fluid, convulsions and facial palsy

Table 3 Other significant symptoms and findings in Kawasaki disease

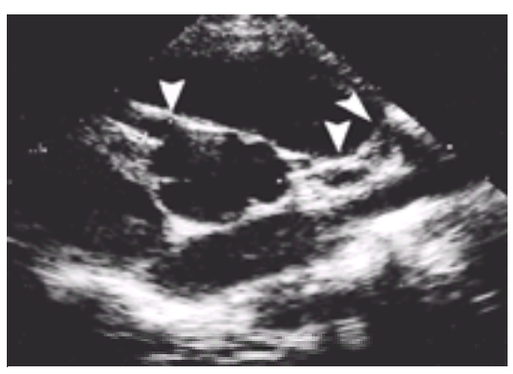


Fig. 5 Two-dimensional echocardiogram (parasternal short-axis view) showing medium-sized aneurysm (6.8 mm) of right coronary artery and smaller aneurysms of left main and left arterior descending coronary arteries (arrows) in child with Kawasaki disease.

Laboratory investigations

In patients with Kawasaki disease a polymorphonuclear leucocytosis and thrombocytosis is seen and circulating platelet-aggregating factors and immune complexes are demonstrable (Levin *et al*. 1985). Both ANCA and anti-endothelial cell antibodies (AECA) have also been determined in acute sera from patients (Leung *et al*. 1986; Savage *et al*. 1989; Tizard *et al*. 1991c; Kaneko *et al*. 1994). The ANCA findings on immunofluorescence were thought to be characteristic with a diffuse cytoplasmic staining and possibly distinct antigens involved (Kaneko *et al*. 1993). However, the diagnostic value of ANCA and AECA in Kawasaki disease has recently been questioned (Guzman *et al*. 1994; Nash *et al*. 1995) in studies showing that these antibodies did not differentiate early Kawasaki disease from other childhood illnesses.

Aetiology

Most workers agree that Kawasaki disease is likely to have an infective basis but the nature of the infective agent and the mechanisms involved remain in doubt. A number of organisms have been considered including *Streptococcus sanguis*, *Propionibacterium acnes*, Epstein–Barr virus, human herpes virus 6, chlamydia, rickettsia and retroviruses (Shulman 1987; Kawasaki 1989; Takahashi and Taubert 1993). Presently there has been considerable interest in the possibility that the condition is caused by new clones of staphylococci that cause toxic shock syndrome or streptococci that produce pyrogenic exotoxin (Leung *et al*. 1993).

Treatment and prognosis

Therapeutically aspirin and high-dose intravenous gammaglobulin, either as four or five daily doses of 400 mg/kg per day (Newburger *et al*. 1986) or one dose of 2 g/kg (Newburger *et al*. 1991), are recommended. Dipyridamole has also been used by some groups in addition to aspirin. When giant coronary artery aneurysms are present, especially with evidence of ischaemia, intravenous prostacyclin has been utilized (Tizard *et al*. 1991b) and intra-arterial or intravenous urokinase has been advocated in circumstances of coronary artery occlusion with thrombus (Terai *et al*. 1985). Revascularization surgery of the coronary arteries may be indicated for stenotic lesions after the acute phase of the illness is over (Susuki *et al*. 1985). The use of steroids remains controversial. There has been a policy contraindicating their use because of increased risks of coronary artery complications compared with aspirin, but this may need to be reconsidered since they certainly have been of value, given with antiplatelet therapy, in children with severe disease who do not respond to gammaglobulin treatment at The Hospital for Sick Children in London.

The overall outlook for children with Kawasaki disease is good. A 1 to 2 per cent acute mortality rate from myocardial infarction has been reduced further in many countries by alertness of clinicians to the diagnosis and early use of gammaglobulin with antiplatelet therapy. Nonetheless, there is a late morbidity and occasional mortality caused by stenotic lesions of coronary arteries in later life and it has been postulated that adult atheromatous coronary disease in some cases may have its origins in childhood and be due to covert or overt Kawasaki disease (Brecker *et al*. 1988).

Wegener's granulomatosis

Wegener's granulomatosis (Wegener 1936) is a necrotizing granulomatous vasculitis of the upper and lower respiratory tract, associated with glomerulonephritis and variable degrees of small vessel vasculitis elsewhere. It can be a generalized disorder or present in a limited local form, involving, initially at least, the respiratory tract. It is rare in childhood but has been reported (Moorthy *et al*. 1977; Baliga *et al*. 1978; Orlowski *et al*. 1978; Halstead *et al*. 1986; Dillon 1990; Singer *et al*. 1990; Rottem *et al*. 1993) and the association with ANCA showing a cytoplasmic immunofluorescence appearance has created another means of diagnosis as well as a method of monitoring disease activity (Van der Woude *et al*. 1985; Savage *et al*. 1987).

Clinical and laboratory findings

Clinical manifestations are similar to those reported in polyarteritis nodosa. There are, however, particular features that might lead clinicians to the diagnosis including subglottic stenosis due to granulomatous involvement of the trachea (Fig. 6), at times requiring tracheostomy (Halstead *et al*. 1986; Dillon 1990; Tizard *et al*. 1991a; Rottem *et al*. 1993), and other upper airway findings such as sinus opacity, nasal septum disease and lower respiratory tract features frequently masquerading as infection (Fig. 7). Some of these findings are also seen in relapsing polychondritis, an even rarer condition of childhood, which can cause diagnostic confusion (Blau 1976). Glomerulonephritis, at times taking the form of an aggressive crescentic glomerulonephritis, is seen but is not a prerequisite for the diagnosis. A proportion of children have disease that is limited to the upper and lower respiratory tract (Hall *et al*. 1985b; Rottem *et al*. 1993). Patients' serum usually contains circulating ANCA with the characteristic granular cytoplasmic distribution (C-ANCA) on indirect immunofluorescence and an antibody directed against proteinase 3 (Jennette and Falk 1990; Dillon and Tizard 1991; Gross *et al*. 1991; Wong *et al*. 1992).

Fig. 6 Subglottic stenosis due to granulomatous involvement of the trachea in Wegener's granulomatosis. Note tracheostomy tube.

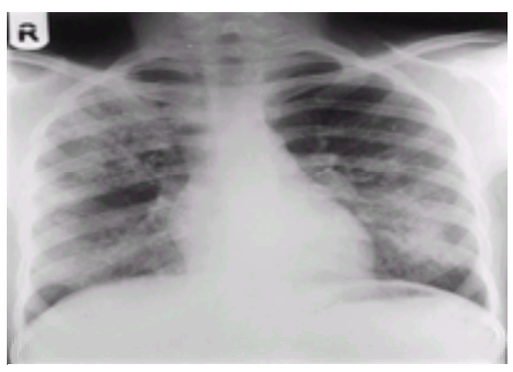


Fig. 7 Pulmonary involvement in Wegener's granulomatosis.

Treatment and prognosis

Steroids and cyclophosphamide (orally or by pulsed intravenous injection) are the mainstays of treatment, but additionally some use antiplatelet medication and antimicrobials such as cotrimoxazole and, in life-threatening situations, plasma exchange (Gaskin and Pusey 1992; Roberti *et al*. 1993; Rottem *et al*. 1993). Response to therapy is variable but the majority of patients need to remain on long-term steroid and immunosuppressive agents. Once remission has been induced, it can be maintained in many children on a combination of steroid and azathioprine or cyclosporin (Dillon 1990; Gaskin and Pusey 1992). Mortality rate amongst The Hospital for Sick Children patients has been of the order of 15 per cent (Tizard *et al*. 1991a).

Churg–Strauss syndrome

Churg–Strauss syndrome or allergic granulomatosis is extremely rare in childhood (Churg and Straus 1951; Frayha 1982). The clinical picture consists of variable vasculitic features with asthma, eosinophilia, infiltrates on chest radiographs, and extravascular granulomata on biopsy (Chumbley *et al*. 1977). ANCA may be present with perinuclear or cytoplasmic staining on indirect immunofluorescence. Steroids are the mainstay of therapy but cytotoxic agents such as cyclophosphamide may need to be introduced to control disease activity (Chumbley *et al*. 1977; Roberti *et al*. 1993).

Henoch–Schönlein purpura (anaphylactoid purpura)

Henoch–Schönlein purpura is the most common form of systemic vasculitis of childhood and comes into the general category of leucocytoclastic vasculitides, involving small vessels especially in the skin (Gairdner 1948; Allen *et al*. 1960). Within this general group of conditions there is often an identifiable initiating factor such as a drug or a microorganism, hence the alternative descriptive terms of allergic or hypersensitivity vasculitis. Henoch–Schönlein purpura is chiefly a disease of childhood, more common in males, often preceded by an upper respiratory infection and more common in the winter months (Meadow *et al*. 1972; Levy *et al*. 1976).

Clinical characteristics

The salient features are a non-thrombocytopenic palpable purpura, arthritis or arthralgia, abdominal pain, gastrointestinal haemorrhage and glomerulonephritis. The palpable purpura is prominent on dependent and pressure-bearing areas such as the lower limbs and buttocks, but in young children facial involvement and subcutaneous oedema are features. Glomerulonephritis occurs in 50 per cent of patients and in 10 per cent of these it is serious (Kobayashi *et al*. 1977; Koskimies *et al*. 1974). The spectrum of renal disease varies from isolated microscopic haematuria to a nephritic/nephrotic syndrome with renal failure.

Laboratory investigations

The majority of affected children have a moderate leucocytosis and thrombocytosis. Thrombocytopenia would raise serious doubts about the diagnosis. Serum IgA and IgM are increased in 50 per cent of patients and a proportion have circulating IgA containing immune complexes and cryoglobulins (Trygstad and Stiehm 1971; Levinsky and Barratt 1979; Knight 1990). In some patients IgA ANCA have been demonstrated (Van den Wall Bake *et al*. 1987). Haematuria and proteinuria are found frequently. Renal function impairment, with or without hypertension in the presence of nephrotic range proteinuria, would be an indication for a renal biopsy. Histological examination of biopsy material may reveal proliferative glomerulonephritis of variable severity from focal segmental lesions to extensive crescentic glomerulonephritis. Immunofluorescence will often demonstrate mesangial IgA deposits and in some cases C3, fibrin, and IgM.

Treatment and prognosis

Treatment is supportive. Corticosteroids given over a few weeks may help severe gut disease but are not thought to benefit the renal disease except in the context of aggressive rapidly progressive glomerulonephritis when steroid, immunosuppressive and antiplatelet therapy have a therapeutic role. If renal function is rapidly deteriorating pulsed methyl prednisolone and/or plasmapheresis appear to have some beneficial effects.

The prognosis is usually good, but there is some morbidity and occasional mortality associated with gut and kidney disease (Allen *et al*. 1960). Amongst patients who present with a nephritic, nephrotic or nephritic/nephrotic syndrome, 44 per cent have hypertension or impaired renal function on long-term follow-up, whereas 82 per cent who present with haematuria (with or without proteinuria) are normal (Goldstein *et al*. 1992). Overall, 2 to 5 per cent of children with Henoch–Schönlein purpura progress to endstage renal failure accounting for approximately 10 per cent of renal failure in childhood (Kobayashi *et al*. 1977; Koskimies *et al*. 1981). Renal transplantation is possible and successful in such patients and recurrence in the graft rare.

Hypersensitivity angiitis

The commonest small vessel vasculitis other than Henoch–Schönlein purpura is hypersensitivity vasculitis. This is usually a drug-induced condition but can be associated with other antigens including infectious agents (Kunnamo *et al*. 1986). The predominant clinical manifestations involve the skin, with palpable nodules or purpura (Fig. 8), although other organs can be affected and it is important to realise that the condition can with time evolve into one of the systemic necrotizing vasculitides (Fauci *et al*. 1978). Removal of the precipitating agent is usually followed by resolution, but sometimes non-steroidal anti-inflammatory drugs, and steroids may be indicated.

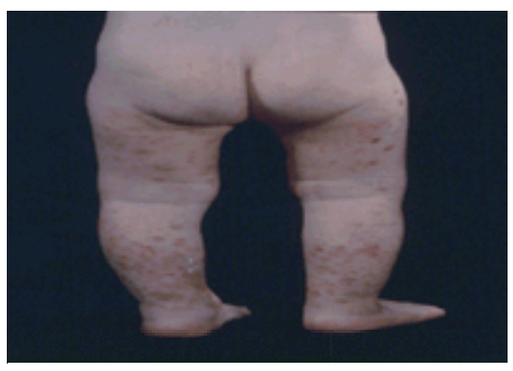


Fig. 8 Palpable purpuric nodules in skin of patient with hypersensitivity vasculitis.

Cutaneous polyarteritis

This is another syndrome with some similarities to Henoch–Schönlein purpura with crops of painful skin nodules and livido reticularis often with a story of preceding upper respiratory tract infection (Diaz-Perez and Winklemann 1980; Fink 1991; David *et al*. 1993). The condition responds to penicillin, non-steroidal anti-inflammatory drugs and steroids. However, the concern is that the condition may be a manifestation of a systemic necrotizing vasculitis. Should this be the case, additional immunosuppressive therapy will be necessary. David *et al*. (1993) reported that all their patients had an erythematous, nodular, painful rash, often on the medial aspect of the feet, plus an evanescent arthritis affecting the knees and ankles. More than 50 per cent of patients also had brawny oedema of muscles and periorbital oedema. Two of their 12 patients went on to develop angiographically-confirmed polyarteritis nodosa (David *et al*. 1993). Cutaneous vasculitis usually runs a benign course, but relapses, particularly in association with recurrent streptococcal infection, are seen in up to 25 per cent of cases. Clinicians seeing this condition usually advise continuing penicillin prophylaxis throughout childhood to prevent relapses since it is amongst the relapsing group that systemic vasculitis tends to occur.

Vasculitis associated with autoimmune rheumatic disorders

Vasculitis is seen at times in systemic lupus erythematosus, juvenile chronic arthritis, mixed connective tissue disease, dermatomyositis and scleroderma and the concept that some of these disorders may be forms of systemic vasculitis has a degree of support. This is particularly relevant to juvenile dermatomyositis in which necrotizing vasculitis in arterioles, capillaries and venules of striated muscle, skin, subcutaneous tissue and gastrointestinal tract has been identified (Banker and Victor 1966; Dillon and Ansell 1995). Crowe *et al*. (1982) reported a group of children with dermatomyositis in which there was muscle infarction, lymphocytic vasculitis and a non-inflammatory vasculopathy associated with extensive erythematous and ulcerative cutaneous disease and who tended to develop calcinosis. Clinically, within this group of patients, periungual erythema and telangiectasia of the nail beds are characteristic, as is telangiectasia along the eyelids which may be accompanied by oedema. Occlusive endarteropathy is likely to be responsible for the cutaneous lesions with ulceration and this, at times, occurs in a livido reticular pattern. Infarction of the palate is associated with weakening of palatal movement and retinal exudates can occur as a result of retinal vascular involvement. Vasculitis can cause acute gastrointestinal ulceration with bleeding or perforation and occasionally there can also be myocardial involvement. Therapeutically, bolus intravenous methylprednisolone is recommended for severe gastrointestinal involvement or vasculitic ulcers. Intravenous gammaglobulin may also have a role and plasma exchange has been utilized. Persistence of widespread ulceration in spite of therapy with severe changes in the capillary loops would be an indication for additional immunosuppressive therapy. Methotrexate in recalcitrant dermatomyositis has been generally accepted (Miller *et al*. 1992); however, there is relatively little information as to how effective it is in vasculitis. As with other vasculitides cyclophosphamide may prove effective when other regimens have failed.

Giant cell arteritis (Takayasu disease)

Giant cell arteritis (Takayasu disease) is a segmental inflammatory vasculitis causing stenosis and aneurysm formation in large arteries, especially the aorta and its major branches (Hall *et al*. 1985a). After Henoch–Schönlein purpura and Kawasaki disease it is the third commonest form of childhood vasculitis if considered worldwide (Lee *et al*. 1967; Wiggelinkhuizen and Cremin 1978). Its aetiology is unclear but genetic and infective factors, for example, tuberculosis, may play a part (Wiggelinkhuizen and Cremin 1978).

Clinical features and management

Early disease manifestations include fever, night sweats, anorexia, weight loss, and arthritis (Lupi-Herrera *et al*. 1977). Subsequently features of hypertension, heart failure and pulse deficits become apparent (Lee *et al*. 1967; Wiggelinkhuizen and Cremin 1978 ; Hall *et al*. 1985a). Leg length inequality can also occur as a result of the aortic pathology compromising lower limb blood supply (Morales *et al*. 1991). The erythrocyte sedimentation rate is increased as are other acute-phase reactants and there may be widening and calcification of the aorta on a plain abdominal radiograph. Doppler ultrasonography, plus magnetic resonance or standard angiography are usually required to establish the diagnosis (Southwood *et al*. 1988). Corticosteroids, plus other immunosuppressives, have been utilized therapeutically in the acute phase (Sunamori *et al*. 1976; Lupi-Herrera *et al*. 1977). A positive tuberculin test may justify antituberculous therapy and hypotensive agents and antiplatelet therapy also have a place. Reconstructive surgery and transluminal angioplasty have been undertaken in older children with inactive disease (Shelhamer *et al*. 1985).

Miscellaneous vasculitides

A number of other conditions in which there is a vasculitic component exist that are too numerous to deal with individually in this chapter. However, there are three that deserve special mention: Behçet's syndrome, familial Mediterranean fever, and Cogan's syndrome.

Behçet's syndrome, first reported by Behçet in 1937, consists of the triad of aphthous stomatitis, genital ulceration, and iritis. It is relatively rare in childhood, but can and does occur and it is well recognised that a vasculitic component is an important feature (Hamza 1993; Koné-Paut and Bernard 1993; Shafaie *et al*. 1993; Özdagan 1994). Systemic therapy is usually necessary for patients with severe cutaneous, systemic or pulmonary vasculitis or those with arterial aneurysms or vena caval involvement. Agents that have been used include steroids and cytotoxic drugs, such as azathioprine or cyclophosphamide, but other medications for treating various aspects of childhood Behçet's syndrome, such as colchicine, levamizole, cyclosporin, and thalidomide may have roles. The prognosis is variable, as in adults, but vascular involvement is related to a poor outcome (Özdagan 1994).

Familial Mediterranean fever often manifests itself for the first time in childhood (Sohar *et al*. 1967; Pras *et al*. 1994). Although not normally considered to be associated with vasculitis there are now a number of reports describing both polyarteritis nodosa and Henoch–Schönlein purpura in affected patients (Flatau *et al*. 1982; Schlesinger *et al*. 1985; Sachs *et al*. 1987; Glikson *et al*. 1989; Pras *et al*. 1994). It is therefore important to consider the possibility of an associated vasculitis in patients with familial Mediterranean fever that may need treating in its own right. Patients with Henoch–Schönlein purpura symptomatically appear to respond to systemic steroid treatment (Pras *et al*. 1994) and those with polyarteritis nodosa with steroid and cyclophosphamide (Sachs *et al*. 1987; Glikson *et al*. 1989).

Non-syphilitic interstitial keratitis with vestibuloauditory dysfunction was first described by Cogan (1945). Although rare and usually affecting young adults it has been reported in children (Cheson *et al*. 1976; Kundell and Ochs 1980) and in addition to the main features there can be evidence of a widespread vasculitis that may require treatment with corticosteroids (Kundell and Ochs 1980; Cassidy and Petty 1995)

Conclusion

As can be seen there is a wide spectrum of vasculitic disease affecting children. The distinction from other conditions, especially those associated with infection, is often difficult, but it is usually possible to categorize the disorders into clinicopathological entities even though there is a substantial degree of overlap. The current range of investigative tools usually allows the correct diagnosis to be established and modern therapeutic approaches have made a major impact in achieving control of disease activity, but in spite of this the level of morbidity and mortality is not inconsequential.

Chapter References

Adu, D. *et al*. (1987). Polyarteritis and the kidney. *Quarterly Journal of Medicine*, **62**, 221–37.

Allen, D.M., Diamond, L.K., and Howell, D.A. (1960). Anaphylactoid purpura in children (Schönlein–Henoch syndrome): review with follow-up of the renal complications. *American Journal of Diseases of Children*, **99**, 833–54.

Ansell, B.A. (1980). *Rheumatic disorders in childhood*. Butterworths, London.

Baliga, R. *et al*. (1978). A case of generalized Wegener's granulomatosis in childhood: successful therapy with cyclophosphamide. *Pediatrics*, **61**, 286–90.

Banker, B.Q. and Victor, M. (1966). Dermatomyositis (systemic angiopathy) of childhood. *Medicine*, **45**, 261–89.

Behçet, H. (1937). Uber rezidivierende Aphthose, durch ein Virus verursachte Geschwure am Mund, am Auge und an den Genitalien. *Dermatologische Wochenschrift*, **105**, 1152–7.

Blau, E.B. (1976). Relapsing polychondritis and retroperitoneal fibrosis in an 8 year old boy. *American Journal of Disease in Children*, **113**, 1149–52.

Blau, E.B., Morris, R.F., and Yunis, E.J. (1977). Polyarteritis nodosa in older children. *Pediatrics*, **60**, 227–34.

Brecker, S.J.D., Gray, H.H., and Obedershaw, P.J. (1988). Coronary artery aneurysms and myocardial infarction: adult sequelae of Kawasaki disease. *British Heart Journal*, **59**, 509–12.

Cameron, J.S. (1988). Renal disease and vasculitis. *Pediatric Nephrology*, **2**, 490–7.

Cassidy, J.T. and Petty, R.E. (1995). *Textbook of pediatric rheumatology*, (3rd edn). W.B. Saunders, Philadelphia.

Cheson, B.D., Bluming, A.Z., and Alroy, J (1976). Cogan's syndrome: a systemic vasculitis. *American Journal of Medicine*, **60**, 549–55.

Chumbley, L.C., Harrison, E.G., and De Remee, R.A. (1977). Allergic granulomatosis and angiitis (Churg–Strauss syndrome). Report and analysis of 30 cases. *Mayo Clinic Proceedings*, **52**, 477–84.

Churg, J., and Strauss, L. (1951). Allergic granulomatosis, allergic angiitis and periarteritis nodosa. *American Journal of Pathology*, **27**, 277–301.

Cogan, D.G. (1945). Syndrome of non-syphilitic interstitial keratitis and vestibuloauditory symptoms. *Archives of Ophthalmology*, **33**, 144–9.

Coward, R.A. *et al*. (1986). Renal micro polyarteritis: a treatable condition. *Nephrology, Dialysis, Transplantation*, **1**, 31–7.

Crowe, W.E., Bove, K.E., and Levinson, J.E. (1982). Clinical and pathogenetic implication of histopathology in childhood polydermatomyositis. *Arthritis and Rheumatism*, **25**, 126–39.

David, J., Ansell, B.M., and Woo, P. (1993). Polyarteritis nodosa associated with streptococcus. *Archives of Disease in Childhood*, **69**, 685–8.

Dhillon, R. *et al*. (1993). Management of Kawasaki disease in the British Isles. *Archives of Disease in Childhood*, **69**, 631–6.

Diaz-Perez, J.L. and Winklemann, R.K. (1980). Cutaneous periarteritis nodosa: a study of 33 cases. In *Vasculitis* (ed. R.K. Winklemann and K. Woolf), pp. 273–84. Lloyd-Luke, London.

Dillon, M.J. (1989). Classification and pathogenesis of arteritis in children. *Toxicologic Pathology*, **17**, 214–18.

Dillon, M.J. (1990). Vasculitic syndromes. In *Paediatric rheumatology update* (ed. P. Woo, P.H. White, and B.M. Ansell), pp. 227–42. Oxford University Press.

Dillon, M.J. and Ansell, B.M. (1995). Vasculitis in children and adolescents. *Rheumatic Disease Clinics of North America*, **21**, 1115–36.

Dillon, M.J. and Tizard, E.J. (1991). Anti-neutrophil cytoplasmic antibodies and anti- endothelial cell antibodies. *Pediatric Nephrology*, **5**, 256–9.

Fauci, A.S., Haynes, B.F., and Katz, P. (1978). The spectrum of vasculitis: clinical, pathologic, immunologic and therapeutic considerations. *Annals of Internal Medicine*, **89**, 660–76.

Fink, C.W. (1977). Polyarteritis and other diseases with necrotising vasculitis in childhood. *Arthritis and Rheumatism*, **20**, 378–84.

Fink, C.W. (1986). Vasculitis. *Pediatric Clinics of North America*, **33**, 1203–19.

Fink, C.W. (1991). The role of streptococcus in post-streptococcal reactive arthritis and childhood polyarteritis nodosa. *Journal of Rheumatology*, **18** (Suppl 29), 14–20.

Flatau, E. *et al*. (1982). Schönlein–Henoch syndrome in patients with familial Mediterranean fever. *Arthritis and Rheumatism*, **25**, 42–47.

Frayha, R.A. (1982). Churg–Strauss syndrome in a child. *Journal of Rheumatology*, **9**, 807–9.

Gairdner, D. (1948). The Schönlein–Henoch syndrome (anaphylactoid purpura). *Quarterly Journal of Medicine*, **17**, 95–122.

Gaskin, G. and Pusey, C.D. (1992). Systemic vasculitis. In *Oxford textbook of clinical nephrology* (ed. S. Cameron *et al*.), pp. 612–35. Oxford Medical Publications, Oxford.

Glikson, M. *et al*. (1989). Polyarteritis nodosa and familial Mediterranean fever. A report of two cases and review of the literature. *Journal of Rheumatology*, **16**, 536–9.

Goldstein, A.R. *et al*. (1992). Long term follow up of children with Henoch–Schönlein nephritis. *Lancet*, **339**, 280–82.

Gross, W.L., Schmitt, W.H., and Csernok, E. (1991). Antineutrophil cytoplasmic autoantibody-associated diseases: a rheumatologist's perspective. *American Journal of Kidney Disease*, **18**, 175–9.

Guzman, J., Fung, M., and Petty, R.E. (1994). Diagnostic value of anti-neutrophil cytoplasmic and anti-endothelial cell antibodies in early Kawasaki disease. *Journal of Pediatrics*, **124**, 917–20.

Hall, S. *et al*. (1985*a*). Takayasu arteritis. A study of 32 North American patients. *Medicine*, **64**, 89–99.

Hall, S.L. *et al*. (1985*b*). Wegener's granulomatosis in pediatric patients. *Journal of Pediatrics*, **106**, 739–44.

Halstead, L.A., Karmody, C.S., and Wolff, S.M. (1986). Presentation of Wegener's granulomatosis in young patients. *Otolaryngology, Head and Neck Surgery*, **94**, 368–71.

Hamza, M. (1993). Juvenile Behçet's disease. In *Behçet's disease: Proceedings of the 6th International Conference* (ed. B. Wechsler and P. Godeau), p. 381. Excerpta Medica International Conference Series 1037, Amsterdam.

Hicks, R.V. (ed.) (1988). Vasculopathies of childhood. PSG Publishing Co., Littleton, MA.

Hicks, R.V. and Melish, M.E. (1986). Kawasaki syndrome. *Pediatric Clinics of North America*, **33**, 1151–75.

Hunder, G.G. *et al*. (1990). The American College of Rheumatology 1990 criteria for the classification of vasculitis: introduction. *Arthritis and Rheumatism*, **33**, 1065–7.

Japan Kawasaki Disease Research Committee (1984). *Diagnostic guidelines of Kawasaki disease* (4th revised edn.). Japan Red Cross Medical Centre, Tokyo.

Jardim, H.M.P.F. *et al*. (1992). Crescentic glomerulonephritis in children. *Pediatric Nephrology*, **6**, 231–35.

Jayne, D.R.W. *et al*. (1991). Treatment of systemic vascultis with pooled intravenous immunoglobulin. *Lancet*, **337**, 1137–9.

Jennette, J.C. and Falk, R.J. (1990). Antineutrophil cytoplasmic autoantibodies and associated diseases: A review. *American Journal of Kidney Diseases*, **15**, 517–29.

Jennette, J.C. *et al*. (1994). Nomenclature of systemic vasculitides. Proposal of an International Consensus Conference. *Arthritis and Rheumatism*, **37**, 187–92.

Kaneko, K. *et al*. (1993). Kawasaki disease has distinct anti neutrophil cytoplasmic antibodies. In *Proceedings of the 4th International Kawasaki Disease Symposium* (ed. M. Takahashi and K. Taubert), p. 192. American Heart Association, Dallas.

Kaneko, K. *et al*. (1994). Antiendothelial cell antibodies can be cytotoxic to endothelial cells without cytokine pre-stimulation and correlate with ELISA antibody measurement in Kawasaki disease. *Clinical and Experimental Immunology*, **98**, 264–9.

Kawasaki, T. (1967). Acute febrile mucocutaneous syndrome with lymphoid involvement with specific desquamation of the fingers and toes in children. *Japanese Journal of Allergy*, **16**, 178–222.

Kawasaki, T. (ed.) (1989). *Proceedings of the 3rd International Kawasaki Disease Symposium*. Japan Heart Foundation, Tokyo.

Knight, J.F. (1990). The rheumatic poison: a survey of some published investigations of the immunopathogenesis of Henoch–Schönlein purpura. *Pediatric Nephrology*, **4**, 533–41.

Kobayashi, O. *et al*. (1977). Schönlein–Henoch's syndrome in children. *Contributions to Nephrology*, **4**, 48–71.

Koné-Paut, I. and Bernard, J.L. (1993). Behçet's disease in children: a French nationwide survey. In *Behçet's disease: Proceedings of the 6th International Conference* (ed. B. Wechsler and P. Godeau), p. 385. Excerpta Medica International Conference Series 1037, Amsterdam.

Koskimies, O. *et al*. (1974). Renal involvement in Schönlein–Henoch purpura. *Acta Pediatrica Scandinavia*, **63**, 357–63.

Koskimies, O. *et al*. (1981). Henoch–Schönlein nephritis: long term prognosis of unselected patients. *Archives of Disease in Childhood*, **56**, 482–4.

Kundell, S.P. and Ochs, H.D. (1980). Cogan syndrome in childhood. *Journal of Pediatrics*, **97**, 96–8.

Kunnamo, I. *et al*. (1986). Serum-sickness-like disease is a common cause of acute arthritis in children. *Acta Paediatrica Scandinavia*, **75**, 964–9.

Kussmaul, A. and Maier, K. (1866). Über eine bischer nicht beschreibene eigenthümliche arteriener krankung (periarteritis nodosa), die mit Morbus Brightii und rapid fortschreitender allgemeiner Muskallahmung einhergeht. *Deutsche Archiv für Klinische Medizin*, **1**, 484–517.

Larsson, L-G., and Baum, J. (1988). Treatment of the child with vasculitis syndromes. In *Vasculopathies of childhood* (ed. R.V. Hicks), pp. 347–68. PSG Publishing Co., Littleton, MA.

Lee, K.S. *et al*. (1967). Primary arteritis (pulseless disease) in Korean children. *Acta Paediatrica Scandinavia*, **56**, 526–36.

Leung, D.Y.M., Collins, T., Lapierre, L.A., Geha, R.S., and Pober, J.S. (1986). Immunoglobulin M antibodies present in the acute phase of Kawasaki syndrome lyse cultured vascular endothelial cells stimulated by gamma interferon. *Journal of Clinical Investigation*, **77**, 1428–35.

Leung, D.Y.M. *et al*. (1993). Toxic shock syndrome toxin-secreting *Staphylococcus aureus* in Kawasaki syndrome. *Lancet*, **342**, 1385–8.

Levin, M. *et al*. (1985). Platelet immune complex interaction in pathogenesis of Kawasaki disease and childhood polyarteritis. *British Medical Journal*, **290**, 1456–60.

Levinsky, R.J. and Barratt, T.M. (1979). IgA immune complexes in Henoch–Schönlein purpura. *Lancet*, **ii**, 1100–103.

Levy, M. *et al*. (1976). Anaphylactoid purpura nephritis in childhood: natural history and immunopathology. *Advances in Nephrology*, **6**, 183–228.

Lupi-Herrera, E. *et al*. (1977). Takayasu arteritis: clinical study of 107 cases. *American Heart Journal*, **93**, 94–103.

Magilavy, D.B. *et al*. (1977). A syndrome of childhood polyarteritis. *Journal of Pediatrics*, **91**, 25–30.

Mason, J.C. *et al*. (1994). Familial polyarteritis nodosa. *Arthritis and Rheumatism*, **37**, 1249–53.

McLain, L.G., Bookstein, J.J., and Kelsch, R.C. (1972). Polyarteritis nodosa diagnosed by renal arteriography. *Journal of Pediatrics*, **80**, 1032–5.

Meadow, S.R. *et al*. (1972). Schönlein–Henoch nephritis. *Quarterly Journal of Medicine*, **41**, 241–58.

Miller, L.C. *et al*. (1992). Methotrexate treatment of recalcitrant childhood dermatomyositis. *Arthritis and Rheumatism*, **35**, 1143–9.

Moorthy, A.V. *et al*. (1977). Wegener's granulomatosis in childhood: prolonged survival following cytotoxic therapy. *Journal of Pediatrics*, **91**, 616–18.

Morales, E., Pineda, C., and Martinez-Lavin, M. (1991). Takayasu's arteritis in children. *Journal of Rheumatology*, **18**, 1081–4.

Nash, M.C. *et al*. (1995). Anti-neutrophil cytoplasmic antibodies and anti-endothelial cell antibodies are not increased in Kawasaki disease. *British Journal of Rheumatology*, **34**, 882–7.

Newburger, J.W. *et al*. (1986). The treatment of Kawasaki syndrome with intravenous gammaglobulin. *New England Journal of Medicine*, **315**, 341–7.

Newburger, J.W. *et al*. (1991). A single intravenous infusion of gammaglobulin as compared with four infusions in the treatment of acute Kawasaki syndrome. *New England Journal of Medicine*, **324**, 1633–9.

Orlowski, J.P., Clough, J.D., and Dyment, P.G. (1978). Wegener's granulomatosis in the pediatric age group. *Pediatrics*, **61**, 83–90.

Özdagan, H. (1994). Behçet's syndrome in children. *Rheumatology in Europe*, **23** (Suppl. 2) 34.

Pras, M. *et al*. (1994). Vasculitis in familial Mediterranean fever. *Rheumatology in Europe*, **23** (Suppl. 2), 33–4.

Reimold, E.W. *et al*. (1977). Polyarteritis in children. *American Journal of Diseases of Children*, **130**, 534–41.

Roberti, I., Reisman, L., and Churg, J. (1993). Vasculitis in childhood. *Pediatric Nephrology*, **7**, 479–89.

Rottem, M. *et al*. (1993). Wegener granulomatosis in children and adolescents: clinical presentation and outcome. *Journal of Pediatrics*, **122**, 26–31.

Rowley, A.H., Gonzalez-Crussi, F., and Shulman, S.T. (1988). Kawasaki syndrome. *Reviews of Infectious Diseases*, **10**, 1–15.

Sachs, D. *et al*. (1987). Polyarteritis nodosa and familial Mediterranean fever. *British Journal of Rheumatology*, **26**, 139–41.

Savage, C.O.S. *et al*. (1985). Microscopic polyarteritis: presentation, pathology and prognosis. *Quarterly Journal of Medicine*, **56**, 467–83.

Savage, C.O.S. *et al*. (1987). Prospective study of radioimmunoassay for antibodies against neutrophil cytoplasm in diagnosis of systemic vasculitis. *Lancet*, **i**, 1389–93.

Savage, C.O.S. *et al*. (1989). Antineutrophil cytoplasm antibodies in Kawasaki disease. *Archives of Disease in Childhood*, **64**, 360–63.

Schlesinger, M., Vardy, P.A., and Rubinow, A. (1985). Henoch–Schönlein purpura and familial Mediterranean fever. *Israel Journal of Medical Science*, **21**, 83–5.

Shafaie, N. *et al*. (1993). Behçet's disease in children. In *Behçet's disease. Proceedings of the 6th International Conference* (ed. B. Wechsler and P. Godeau), p. 381. Excerpta Medica International Conference 1037, Amsterdam.

Shelhamer, J.H. *et al*. (1985). Takayasu's arteritis and its therapy. *Annals of Internal Medicine*, **103**, 121–6.

Shulman, S.T. (ed.) (1987). *Kawasaki disease. Proceedings of the 2nd International Kawasaki Disease Symposium*. Alan R. Liss, New York.

Singer, J., Suchet, I., and Horwitz, T. (1990) Wegener's granulomatosis: two case histories and a review of the literature. *Clinical Radiology*, **42**, 50–51.

Sohar, M. *et al*. (1967). Familial Mediterranean fever. A survey of 470 cases and review of literature. *American Journal of Medicine*, **43**, 227–53.

Southwood, T.R. *et al*. (1988). New techniques for detection of large vessel arteritis. *Arthritis and Rheumatism*, **31**, 1218.

Sunamori, M. *et al*. (1976). Aortitis syndrome due to Takayasu's disease. A guideline for the surgical indications. *Journal of Cardiovascular Surgery*, **17**, 423–56.

Suzuki, A. *et al*. (1985). Indication of aortocoronary by-pass for coronary arterial obstruction due to Kawasaki disease. *Heart Vessels*, **1**, 94–100.

Suzuki, A. *et al*. (1990). Kawasaki disease: echocardiographic features in 91 cases presenting in the United Kingdom. *Archives of Disease in Childhood*, **65**, 1142–6.

Takahashi, H. and Taubert, K. (ed.) (1993). *Proceedings of the 4th International Symposium on Kawasaki Disease*. American Heart Association, Dallas.

Terai, M. *et al*. (1985). Coronary arterial thrombi in Kawasaki disease. *Journal of Pediatrics*, **106**, 76–8.

Tizard, E.J. and Dillon, M.J. (1991). Plasmapheresis in childhood. *Care of the Critically Ill*, **7**, 51–5.

Tizard, E.J., Barratt, T.M., and Dillon, M.J. (1991 *a*). Wegener's granulomatosis in six children. *American Journal of Kidney Disease*, **18**, 212.

Tizard, E.J. *et al*. (1991*b*). Clinical aspects of 100 patients with Kawasaki disease. *Archives of Disease in Childhood*, **66**, 185–8.

Tizard, E.J. *et al*. (1991*c*). Anti endothelial cell antibodies detected by a cellular based ELISA in Kawasaki disease. *Archives of Disease in Childhood*, **66**, 189–92.

Trygstad, C.W. and Stiehm, E.R. (1971). Elevated serum IgA globulin in anaphylactoid purpura. *Pediatrics*, **47**, 1023–8.

Van den Wall Bake, A.W.L. *et al*. (1987). IgA antibodies directed against cytoplasmic antigens of polymorphonuclear leucocytes in patients with Henoch–Schönlein purpura. In *Advances in experimental medicine and biology*, 216B (ed. J.R. McGee and J. Mestecky), pp. 1593–8. Plenum Press, New York.

Van der Woude, F.J. *et al*. (1985). Autoantibodies against neutrophils and monocytes: tool for diagnosis and marker of disease activity in Wegener's granulomatosis. *Lancet*, **i**, 425–9.

Wegener, F. (1936). Über generalisiert, Septische Gefasser Krankurgen. *Verhandlungen der Deutschen Gesellschaft für Pathologie*, **29**, 202–9.

Wiggelinkhuizen, J. and Cremin, J. (1978). Takayasu arteritis and renovascular hypertension in childhood. *Pediatrics*, **62**, 209–17.

Wong, S-N., Shah, V., and Dillon, M.J. (1992). Anti-neutrophil cytoplasmic antibodies (ANCA) in childhood systemic vasculitis. *Journal of the American Society of Nephrology*, **3**, 668.

Yousefzadeh, D.K., Chow, K.C., and Benson, C.A. (1981). Polyarteritis nodosa: regression of arterial aneurysms following immunosuppressive and corticosteroid therapy. *Pediatric Radiology*, **10**, 139–41.

5.12 Overlap syndromes in adults and children

Enrique Roberto Soriano and Neil John McHugh

Introduction

Autoimmune rheumatic diseases (connective tissue diseases) are an overlapping group of disorders of unknown aetiology. Their classification depends on identifying clusters of clinical and laboratory features. At least three problems arise when attempting to classify individual patients into one of the defined autoimmune rheumatic diseases early on in their disease or when they present with clinical overlaps: (1) most of the clinical or laboratory features are not exclusive to one disease; (2) many of the symptoms and signs that define the autoimmune rheumatic diseases do not occur concurrently, but rather occur sequentially; (3) as many as 25 per cent of patients with autoimmune rheumatic disease present with an overlap syndrome with typical features of more than one disorder.

The terms undifferentiated connective tissue disease, overlap syndrome, and even mixed connective tissue disease (**MCTD**), have been used for patients who are not comfortably placed within any one of the defined autoimmune rheumatic diseases. These terms are not interchangeable, but unfortunately are often applied loosely. This chapter will describe some of the more common overlap conditions, and what is known concerning environmental triggers. Mixed connective tissue disease is discussed at some length reflecting the large proportion of studies in this area, although there is debate as to whether this disease is a distinct entity or a disease in evolution. Mention will be made of attempts to define more homogeneous subsets of autoimmune rheumatic disease by knowledge of immunogenetic and serological profiles.

Undifferentiated connective tissue disease

The diagnosis of undifferentiated connective tissue disease is best applied to those patients with features strongly suggestive of an autoimmune rheumatic disease but who do not fulfil criteria for any one disorder. The features usually include Raynaud's phenomenon, polyarthritis, rash, and myalgia (Alarcon *et al*. 1991). Undifferentiated connective tissue disease may develop into a well defined autoimmune rheumatic disease, may persist unchanged over time, or the symptoms may even disappear. Outcomes from large multicentre prospective studies are awaited (Alarcon *et al*. 1991). Raynaud's phenomenon is probably the most frequent clinical feature of undifferentiated connective tissue disease. Other factors that may be predictive of the development of later autoimmune rheumatic disease are listed in Table 1.

Scleroderma-like nailfold capillary abnormalities
Positive antinuclear antibody (titre > 1/100)
Anticentromere or antitopoisomerase 1 (Scl-70) antibody
Digital pitting scars
Abnormal erythrocyte sedimentation rate

Adapted from LeRoy and Medsger (1992).

Table 1 Possible prognostic factors in Raynaud's phenomenon

Overlap syndrome

The term overlap syndrome has been used when two or more autoimmune rheumatic diseases, or some of their unique manifestations, occur in the same individual simultaneously. Features that are characteristic of certain defined autoimmune rheumatic diseases are given in Table 2. The overlap may consist of full expression of the features of two or more conditions, or more commonly may be limited to one or more manifestations of each disease. In the latter case, the diagnosis of undifferentiated connective tissue disease may equally apply.

Systemic lupus erythematosus	Systemic sclerosis	Polydermatomyositis	Rheumatoid arthritis
Malar rash	Sclerodermа	Myositis	Erosive arthritis
Photosensitivity	Oesophageal hypomotility	Heliotrope rash	Subcutaneous nodules
Glomerulonephritis		Gottron's papules	Rheumatoid factor
Central nervous system disease	Telangiectasia	Antisynthetase antibodies	
Anti-Sm antibodies		Anti-Mi-2 antibodies	
Anti-double-stranded DNA	Antitopoisomerase I antibodies		

Table 2 Features usually restricted to one connective tissue disease, and less common in others

Serological subsets

A characteristic feature of autoimmune rheumatic diseases is the presence of autoantibodies. As autoantibodies are associated with particular clinical features, knowledge of the autoantibody profile may help in diagnosis and prognosis. The best known example is the association of anti-U1RNP antibodies with mixed connective tissue disease which will be further discussed. Other examples include the association of anti-Ro (SS-A) and anti-La (SS-B) antibodies with Sjögren's syndrome, antiphospholipid antibodies with the antiphospholipid syndrome, and antibodies to tRNA synthetases with the antisynthetase syndrome. There are less common autoantibodies such as anti-Pm-Scl (Fig. 1) which may identify patients with other overlap features. Also, in most cases there are strong associations between autoantibody-defined subsets of disease and HLA genes that may explain the genetic basis for disease susceptibility (Table 3).

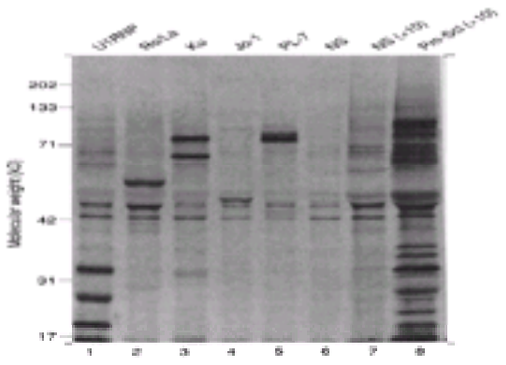


Fig. 1 Proteins immunoprecipitated using sera containing autoantibodies that are associated with overlap syndromes; anti-U1RNP with MCTD, anti-Ro/La antibodies with systemic lupus erythematosus (**SLE**)/Sjögren's syndrome, anti-Ku with systemic lupus erythematosus or scleroderma, anti-Jo-1 (histidyl-tRNA synthetase) and anti-PL-7 (threonyl- tRNA synthetase) with the antisynthetase syndrome, and anti-Pm-Scl with polymyositis/scleroderma/arthritis. NS, normal serum.

Overlap syndrome	Autoantibody specificity	HLA class II allele
Mixed connective tissue disease	U1RNP	HLA-DR4 (DRB1*0401) HLA-DQ3 (DQB1*0301) (Kawano, et al. 1995)
Polymyositis, interstitial lung disease, Raynaud's, arthritis	Jo-1 (histidyl-tRNA synthetase) and less often other tRNA synthetases	HLA-DR3, DR52 (Goldstein et al. 1992)
Systemic sclerosis, polymyositis	PM-Scl Ku U2-RNP	HLA-DR3 (Marguerie et al. 1992)

Table 3 Recognized clusters of clinical overlaps, autoantibodies, and HLA class II alleles

Mixed connective tissue disease

Definition

Mixed connective tissue disease was first described by Sharp and co-workers (Sharp *et al*. 1971; Sharp *et al*. 1972) as an 'apparently distinct rheumatic disease syndrome' in which the clinical characteristics included a combination of features similar to those of systemic lupus erythematosus, scleroderma, and polymyositis. The spectrum has been broadened by some to include features of rheumatoid arthritis (Hench *et al*. 1975; Sullivan *et al*. 1984). Mixed connective tissue disease was considered unique as it was associated with autoantibodies to a ribonuclease-sensitive component of extractable nuclear antigen, now known as U1RNP.

There remains debate as to whether mixed connective tissue disease is a distinct disease, a disease in evolution, or a subset of another autoimmune rheumatic disease such as systemic lupus erythematosus (LeRoy *et al*. 1980; Lazaro *et al*. 1989; McHugh *et al*. 1990; Black and Isenberg 1992). Diagnosis that is dependent on a single serological finding may suffer from ascertainment bias. Many patients with the clinical features 'characteristic' of mixed connective tissue disease do not have anti-U1RNP antibodies (Ginsburg *et al*. 1983; Lazaro *et al*. 1989), and conversely anti-U1RNP antibodies may appear in other conditions (McHugh *et al*. 1990). None the less, mixed connective tissue disease is considered by some to be a syndrome that has a core of manifestations associated with a serological marker (Alarcon-Segovia 1994). Three substantially revised and distinct sets of diagnostic criteria have been proposed (Table 4).

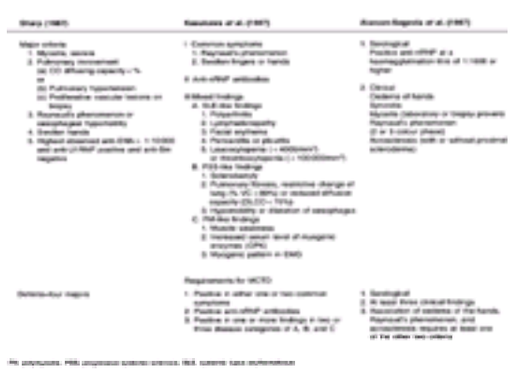


Table 4 Diagnostic criteria for mixed connective tissue disease (MCTD)

Epidemiology

Mixed connective tissue disease is a disease of the second to fourth decades of life with a mean age of onset of 35 years in adults (Sharp *et al*. 1972; Prakash *et al*. 1985), and 10 years in children (Singsen *et al*. 1977; Savouret *et al*. 1983). Women are affected more often than men at all ages, and represent around 80 per cent of patients. The incidence of mixed connective tissue disease in adults remains unknown. A hospital-based retrospective study of children under 15 in Sweden, reported an incidence of 0.13/100 000 per year, compared with an incidence of 0.43/100 000 per year for systemic lupus erythematosus and 0.28/100 00 per year for polymyositis/dermatomyositis (Magnusson 1993). In a nationwide, prospective, hospital-based study for 4 years in Finland in children aged 0 to 15 years, the annual incidence rate for mixed connective tissue disease (diagnosed according to Sharp's criteria) was 0.10/100 000 compared with 0.37 for systemic lupus erythematosus, 0.30 for polymyositis/dermatomyositis, and 0.05 for scleroderma (Pelkonen *et al*. 1994).

Clinical features (Table 5)

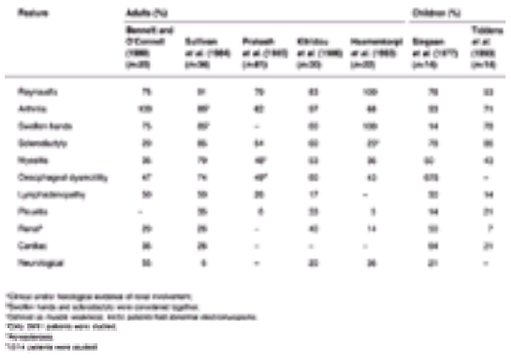


Table 5 Frequent clinical features in adults and children with mixed connective tissue disease

General symptoms

The typical patient is a 20- to 30-year-old female presenting with Raynaud's phenomenon, arthralgias/arthritis, swollen hands and/or puffy fingers, in association with a high titre, speckled pattern, antinuclear antibody (**ANA**) (Fig. 2). Other presenting symptoms may include fatigue, fever, serositis, mild myositis, aseptic meningitis or unexplained lymphadenopathy.

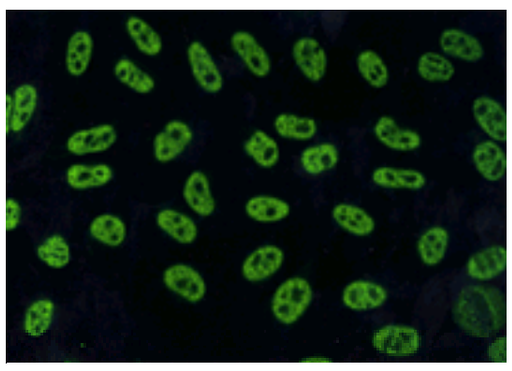


Fig. 2 Typical coarse speckled pattern of ANA detected by indirect immunofluorescence in serum from a patient with mixed connective tissue disease. Note the relative sparing of staining of the nucleolus within the cell nucleus.

Vascular involvement

Raynaud's phenomenon is present in 75 to 100 per cent of adults and children (Table 5), and is often the first symptom to appear. Ischaemic necrosis and ulcerations of the fingertips are rare but may occur (Sharp *et al*. 1972; Gilliam and Prystowsky 1977; Peller *et al*. 1985). Nailfold capillaroscopy shows capillary dilatation and capillary loss (avascular areas) in 50 to 90 per cent of patients with mixed connective tissue disease (Maricq *et al*. 1980; Peller *et al*. 1985; Granier *et al*. 1986) and may be associated with pulmonary disease (Sullivan *et al*. 1984; Pallis *et al*. 1991). Dystrophic, branched 'bushy' capillaries may be especially characteristic for mixed connective tissue disease (Granier *et al*. 1986). Angiographic studies have shown evidence of vasculopathy of both small and medium-sized vessels in the hands (Peller *et al*. 1985).

Skin

Swelling of the hands, particularly of the fingers leading to a sausage appearance (Fig. 3), occurs more frequently in adults than in children. The skin of the hands may be taut and thick and histologically resembles scleroderma. Scleroderma extending more proximally is not usually a feature of mixed connective tissue disease. Other less frequent cutaneous manifestations are alopecia, depigmentation, telangiectasia, erythema nodosum, and chronic discoid lesions. A malar rash, suggestive of systemic lupus erythematosus, and dermatomyositis-like skin are more frequent in children than in adults.

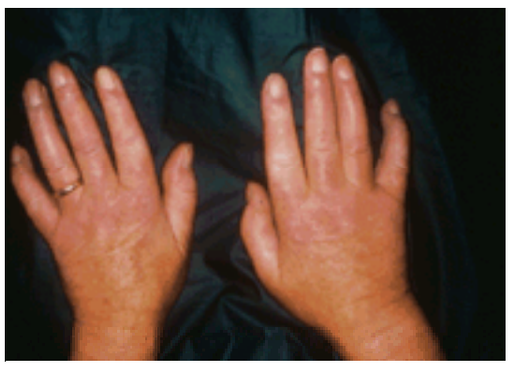


Fig. 3 Typical 'sausage-shaped' swollen fingers of a patient with mixed connective tissue disease.

Joints

Polyarthralgia is an early symptom and occurs in most patients. A symmetrical polyarthritis most often involving hands and wrists may mimic rheumatoid arthritis but is less deforming and erosive. None the less ulnar deviation, swan neck changes, and flexion contractures are not rare and atlantoaxial subluxation has been reported (Stuart and Maddison 1991). Erosions are found in about 60 per cent of the patients (Udoff *et al*. 1977; Bennett and O'Connell 1980; Catoggio *et al*. 1983), and are usually small, punched-out, and asymetrically distributed (O'Connell and Bennett 1977). Prominent periarticular calcification is not uncommon (Fig. 4) and arthritis induced by hydroxyapatite crystal has been reported (Hutton *et al*. 1988). Avascular necrosis of bone may occur in children (Tiddens *et al*. 1993) and adults (O'Connell and Bennett 1977). Minute, multiple peritendinous nodules may be found adjacent to the flexor tendons of the forearms and the extensor tendons of the hands (Babini *et al*. 1985). Arthritis may be more frequent and destructive in children in whom the initial diagnosis may be juvenile chronic arthritis.

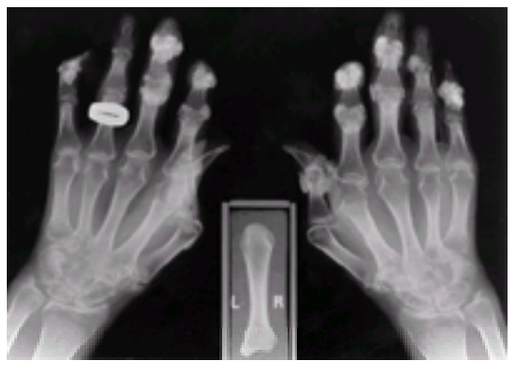


Fig. 4 Prominent periarticular calcification in a patient with mixed connective tissue disease.

Muscle

Myalgias are common and about two-thirds of patients develop an inflammatory myopathy that is identical clinically and histologically to polymyositis (Sharp *et al*. 1972; Bennett and O'Connell 1980). Very often myositis may occur acutely in a patient who has other mild features (Bennett and O'Connell 1980; Lundberg *et al*. 1992a), and prompts the diagnosis of mixed connective tissue disease. The prognosis seems more favourable than in polymyositis/dermatomyositis with less corticosteroid treatment needed (Nimelstein *et al*. 1980; Lundberg *et al*. 1992a). In children, vasculitis and inflammatory infiltration of skeletal muscle has been reported at autopsy, without clinical or laboratory evidence of muscle disease (Singsen *et al*. 1977).

Gastrointestinal tract

Gastrointestinal involvement is similar to that in systemic sclerosis, with heartburn and dysphagia the most common symptoms (Table 6). Oesophageal manometry is abnormal in up to 85 per cent of patients (Marshall *et al*. 1990; Doria *et al*. 1991). Corticosteroid treatment may improve oesophageal dysfunction (Marshall *et al*. 1990) but is needed rarely. Extensive gastrointestinal tract involvement with malabsorption (Norman and Fleischmann 1978), colonic and small bowel perforations due to vasculitis, acute pancreatitis (Marshall *et al*. 1990), duodenal haemorrhage (Hirose *et al*. 1993), pneumatosis intestinalis and pneumoperitoneum (Bennett and O'Connell 1980; Pun *et al*. 1991), haemobilia due to vasculitis of the gallbladder (Kuipers *et al*. 1991), protein-losing enteropathy (Furuya *et al*. 1992), and secretory diarrhoea (Thiele and Krejs 1985) have all been reported. Hepatomegaly and splenomegaly are found in about 25 per cent of patients but major liver involvement is uncommon. The spectrum of gastrointestinal involvement is similar in children with mixed connective tissue disease.

Symptom	Number of patients	Frequency
Asymptomatic	16	26
Heartburn or regurgitation	29	48
Dysphagia	23	38
Dyspepsia	12	20
Diarrhoea	2	8
Constipation	3	5
Vomiting	2	3

Adapted from Marshall *et al*. (1990).

Table 6 Gastrointestinal symptoms in 61 patients with mixed connective tissue disease

Keratoconjunctivitis sicca

Sicca symptoms are not uncommon in adults and children with mixed connective tissue disease, depending on the bias of patients selected for study. Sialectasia was found in 82 per cent of 39 adults with mixed connective tissue disease studied by parotid sialography (Ohtsuka *et al*. 1992), although 50 per cent of these patients also fulfilled criteria for Sjögren's syndrome.

Neurological involvement

Trigeminal sensory neuropathy may occur in up to 10 per cent of patients and may be an early manifestation of the disease (Bennett *et al*. 1978; Searles *et al*. 1978; Sullivan *et al*. 1984; Hagen *et al*. 1990). Involvement of the central nervous system is rare in adults although an aseptic meningitis-like syndrome has been reported (Bennett and O'Connell 1980). Other occasional findings are headaches, seizures, peripheral neuropathy (Bennett *et al*. 1978), spinal cord involvement (Kappes and Bennett 1982), transverse myelitis (Weiss *et al*. 1978), hypertrophic cranial pachymeningitis (Fujimoto *et al*. 1993), and demyelination (Nitsche *et al*. 1991). Three of the 14 children reported by Singsen *et al*. had cerebral involvement (Singsen *et al*. 1977).

Heart

Major cardiovascular abnormalities are uncommon (Table 7). Pericarditis may occur in up to 20 per cent of patients and pericardial tamponade has been reported (Bennett and O'Connell 1980; Langley and Treadwell 1994). Electrocardiogram abnormalities are more common (Alpert *et al*. 1983; Oetgen *et al*. 1983), and cardiac conduction defects have been described (Emlen 1979; Rakovec *et al*. 1982). Echocardiographic changes including pericardial effusions, mitral valve prolapse, and right ventricle enlargement have been found in 38 to 60 per cent of patients (Alpert *et al*. 1983; Oetgen *et al*. 1983). Abnormal left ventricular diastolic filling has been documented by Doppler studies in the absence of myocardial disease (Leung *et al*. 1990). Intimal hyperplasia of coronary arteries and inflammatory cell infiltrates of the myocardium may be found in both adults and children (Singsen *et al*. 1977; Alpert *et al*. 1983).

Symptom	Alpert *et al.* (1983) (n=38) (%)	Cetgen *et al.* (1983) (n=16) (%)	Prakash *et al.* (1985) (n=81) (%)
Dyspnoea	63[a]	38[a]	16
Pericardial chest pain	6	19	7[b]
Angina pectoris	8	19	7[b]
Palpitations	16	25	–

[a]Dyspnoea could be attributed to pulmonary origin in all but one patient in each series.
[b]Cause of chest pain not stated.

Table 7 Cardiac involvement in adults with mixed connective tissue disease

Lungs

Pleuropulmonary involvement is common and may be clinically inapparent. The most common clinical findings are dyspnoea, pleuritic pain, and bibasilar rales (Table 8). Children may present with reduced exercise tolerance. Chest radiograph abnormalities include basal interstitial fibrosis, pleural effusion, pneumonic infiltrates, and pleural thickening (Sullivan *et al*. 1984; Prakash *et al*. 1985). Abnormalities of pulmonary function test are very common and may include a restrictive pattern, small airway involvement and respiratory muscle weakness (Izumiyama *et al*. 1993; Lazaro *et al*. 1993). Interstitial lung disease usually responds to corticosteroids but rapidly progressive cases have been reported (Weiner-Kronish *et al*. 1981).

Feature	Sullivan *et al.* (1984) (n=34) (%)	Prakash *et al.* (1985) (n=81) (%)
Asymptomatic	33	75
Dyspnoea	58	16
Pleuritic pain	40	7
Bibasilar rales	42	–
Cough	24	5
Abnormal chest radiograph	30	21
Restrictive pulmonary function tests	41	69
Decreased diffusing capacity	72	69

Table 8 Pulmonary involvement in adults with mixed connective tissue disease

Pulmonary hypertension is frequent in mixed connective tissue disease, may be difficult to detect and is a major cause of death (Table 9). The presence of scleroderma-type capillary changes on nailfold capillary microscopy may be predictive of the development of pulmonary hypertension (Sullivan *et al*. 1984). Anticardiolipin antibodies and/or a lupus anticoagulant may be more frequent in this group of patients (Hainaut *et al*. 1986; Miyata *et al*. 1992). Histological findings show proliferative vascular abnormalities resembling those found in the **CREST** (**c**alcinosis, **R**aynaud's phenomenon, (o)**e**sophageal dysmotility, **s**clerodactly, **t**elangiectasia) syndrome. Patients with post-sternal pain, pulmonary diastolic murmur, right ventricular hypertrophy on ECG, and higher mean pulmonary artery pressure have a poor prognosis (Kasukawa *et al*. 1990). Pulmonary hypertension seems less frequent in children.

Presumed cause of death	Death (n=31)	Frequency (%)
Pulmonary hypertension	11	35
Septicaemia	3	10
Sudden death	3	10
Myocardial infarction	2	6
Pulmonary embolus	1	3
Pulmonary vasculitis	1	3
Scleroderma-like renal crisis	1	3
Others, possibly unrelated	8	26
Unknown	2	6

Data taken from Bennett and O'Connell 1980, Nimelstein *et al*. 1980, Grant *et al*. 1981, Sullivan *et al*. 1984, Prakash *et al*. 1985, and Kitridou *et al*. 1986.

Table 9 Reported causes of death in mixed connective tissue disease

Kidney

In their initial report Sharp *et al*. (1972) stressed the paucity of renal involvement but later reports have found a prevalence of up to 50 per cent in adults and children (Table 5, and Hench *et al*. 1975; Bennett and Spargo 1977; Grant *et al*. 1981). Nephrotic syndrome associated with membranous nephropathy is the most common presentation and may respond to high-dose corticosteroid therapy (Kitridou *et al*. 1986). The characteristic renal lesion is an immune- deposit nephritis. Renal crisis is a rare manifestation but should be suspected in a patient with accelerated hypertension as the treatment with angiotensin-converting enzyme inhibitor may be effective (Satoh *et al*. 1994). Glomerular and vascular deposition of amyloid material have been reported (Piirainen 1989; Kessler *et al*. 1992). As in systemic lupus erythematosus, renal lesions may be found in adults and children without clinically evident renal disease (Bennett and Spargo 1977; Singsen *et al*. 1977; Singsen *et al*. 1980).

Laboratory findings

The most common laboratory features in adults and children are listed in Table 10. A moderate anaemia of chronic inflammation is usually present. Frank haemolytic anaemia and severe thrombocytopenia requiring splenectomy may occur but are rare complications. Thrombocytopenia may be associated with anticardiolipin antibodies (Doria *et al*. 1992a). Hypergammaglobulinaemia is more frequent than in systemic lupus erythematosus. Hypocomplementaemia and cryoglobulins are present in about a third of patients but neither seem specifically associated with renal or other organ involvement. An elevated creatine phosphokinase may be associated with aseptic meningitis and trigeminal neuropathy in addition to myositis (Bennett and O'Connell 1980).

Feature	Adults (%)			Children (%)	
	Bennett and O'Connell (1993) (*n*=22)	Sullivan *et al.* (1994) (*n*=34)	Prahash *et al.* (1992) (*n*=81)	Singsen *et al.* (1977) (*n*=14)	Tolaino *et al.* (1992) (*n*=14)
Anaemia	75	24	–	[#1/45]	64
Leucopenia	75	21	–	64	43
High ESR	100	–	–	100	–
Hypergammaglobulinaemia	75	53	73	86	–
High creatine kinase	43	–	–	–	50
Positive direct Coombs'	60	–	–	–	–
Positive rheumatoid factor	25	59	49	64	57
Low complement	–	32	5	–	–
Positive LE cell	–	18	–	–	–
Thrombocytopaenia	–	–	–	58	21
Positive ANA	100[a]	100	93	100	100
Anti-U1RNP	100[a]	100	100	100	100
Anti-dsDNA	0	23	4	0	21

[a]By definition patients had positive ANA, RNPase-sensitive ENA haemagglutination test, negative anti-Sm and negative anti-dsDNA.
ANA antinuclear antibodies, ESR erythrocyte sedimentation rate, LE lupus erythematosus.

Table 10 Laboratory features in adults and children with mixed connective tissue disease

Autoantibodies

The presence of autoantibodies to the ribonuclease-sensitive component of extractable nuclear antigen (**ENA**) by a haemagglutination test is a characteristic serological feature of mixed connective tissue disease. The autoantibodies are now termed anti-U1RNP antibodies. The presence of these autoantibodies is suggested by a high titre, coarse speckled, nuclear pattern on an indirect immunofluorescence screen for antinuclear antibodies (Fig. 2). Immunodiffusion tests are now more commonly used than haemagglutination to confirm the presence of anti-U1RNP antibodies. The ribonuclease-resistant component of ENA is recognized by anti-Sm antibodies which frequently coexist with anti-U1RNP antibodies. The absence of anti-Sm antibodies and anti-DNA antibodies in sera positive for anti-U1RNP is also felt to be an important discriminatory finding (Table 4) as anti-Sm and anti-DNA antibodies are more specific for systemic lupus erythematosus.

Anti-U1RNP and anti-Sm antibodies recognize polypeptides on small ribonucleoproteins which participate in the formation of spliceosomes and process RNA. Anti-U1RNP antibodies selectively immunoprecipitate the U1RNA by recognition of 70 kDa, A and C polypeptides which are unique to the U1RNP particle, whereas anti-Sm antibodies immunoprecipitate the abundant uridine-rich small RNAs U1, U2, U5, and U4/U6 by recognition of core polypeptides found on all these RNAs. These core proteins common to each snRNP are B' (29 kDa), B (28 kDa), D1 to D3 (16 kDa), E (12 kDa), F (10 kDa) and G (9 kDa), of which B'/B and D are the major targets for anti-Sm antibodies. Precise characterization of the autoantibody specificity may be obtained by the technique of western blotting (Fig. 5). Although there is no clear-cut association between the titre of anti-U1RNP antibodies and activity of disease, the profile of autoantibody recognition to the U1RNP particle may change over time, and in some cases has been related to changes in features of the disease (Fisher *et al*. 1985).

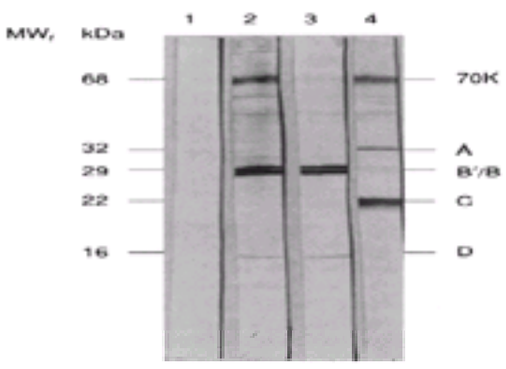


Fig. 5 Immunoblot showing the U1RNP and Sm antigenic peptides. Lane 1, normal serum; Lane 2, anti-U1RNP and anti-Sm serum; Lane 3, anti-Sm serum; Lane 4, anti-U1 RNP and anti-Sm serum.

Pathogenesis

It is likely that environmental factors trigger mixed connective tissue disease in genetically susceptible individuals. Various environmental agents have been reported including drugs such as procainamide, toxins such as polyvinyl chloride (Kahn *et al*. 1989), and silicone implants (Kumagai *et al*. 1984). Familial cases are rare but autoimmune conditions including mixed connective tissue disease may cluster in families, suggesting the inheritance of a common autoimmune diathesis. There appears to be a link with MHC class II genes (Table 3).

Histopathology

Proliferative vascular changes may be widespread and involve organs not clinically affected (Singsen *et al*. 1980). The vascular changes are similar to systemic sclerosis but accompanied by less fibrosis. There is a predilection for intimal thickening of large arteries including coronary, pulmonary, renal, and aortic arteries (Singsen *et al*. 1980; Alpert *et al*. 1983; Sullivan *et al*. 1984). Skin biopsies show dermal thickening with hypertrophy of collagen bundles, thickening of blood vessel walls and a perivascular mononuclear cell infiltration. Immunofluorescence studies of non-involved skin have shown a particulate epidermal nuclear staining pattern and granular dermo–epidermal junction deposits of immunoglobulin (Gilliam and Prystowsky 1977; Reimer *et al*. 1983). High levels of circulating anti-endothelial antibodies (Bodolay *et al*. 1989), factor VIII-related antigen (James *et al*. 1990) and ristomycin-cofactor activity (Udvardy *et al*. 1991) point to endothelial cell injury and alteration of platelet function.

Immunoregulatory abnormalities

B-cell hyperactivity is evident by polyclonal hypergammaglobulinaemia, and spontaneous *in vitro* production of immunoglobulins (Kallenberg *et al*. 1988). However the autoantibody response to the U1RNP particle appears antigen-driven rather than as a result of non-specific polyclonal activation. Recognition of one protein on the snRNP particle may allow processing and presentation of other proteins on the same complex by B cells (Fatenjad *et al*. 1993). Under conditions in which T cell tolerance to snRNPs is overcome, such a mechanism may lead to affinity maturation of a high-titre autoantibody response to the U1RNP particle. The initial autoantibody response may be initiated by molecular mimicry. Indeed, there are several regions of homology between U1RNP proteins and infectious agents such as human retroviral proteins (Query and Keene 1987). Antibodies to native HIV-1 and HTLV-I have been reported in mixed connective tissue disease (Ranki *et al*. 1992), although retroviral genomic material was not isolated in these patients.

T cells reactive to snRNPs have been isolated from patients with mixed connective tissue disease as well as healthy individuals (Hoffman *et al*. 1993). T-cell clones reactive to the snRNP A protein show restricted usage of TCR-b chain genes (Okubo *et al*. 1994). Therefore, snRNP-reactive oligoclonal T cells may accumulate in patients with mixed connective tissue disease. Rapidly dividing cells randomly accumulate gene mutations including mutations in the hypoxanthine-guanine phosphoribosyl transferase (**HPRT**) gene. An increased frequency of mutations in the HPRT gene in T cells isolated from patients with mixed connective tissue disease has been found, which may suggest a role for activated T cells in pathogenesis (Holyst *et al*. 1994). Prolonged survival of autoreactive cells by defects in apoptosis accounts for accelerated autoimmunity including the generation of autoantibodies to snRNPs in certain inbred strains of mice (Watanabe-Fukunaga *et al*. 1992). The role of molecules which regulate apoptosis such as Fas receptor and Fas ligand in human autoimmune disease is currently unknown.

Diagnosis

The presence of anti-U1RNP antibodies is a major criteria for diagnosis (Table 4), but should not be used alone to make the diagnosis. Only the criteria proposed by Sharp (1987) (Table 4) allow for the diagnosis of mixed connective tissue disease without the presence of anti-U1RNP antibodies. All sets of criteria had similar sensitivity when applied to a large population with autoimmune rheumatic disease (Alarcon-Segovia and Cardiel 1989), and although the criteria of Alarcon-Segovia

(Alarcon-Segovia and Villarreal 1987) appeared more specific, the findings need validation in a less selected series of patients. An advantage of Alarcon-Segovia's criteria is their simplicity, but a disadvantage is the need for a haemagglutination test which is now seldom used in clinical practice. The Japanese diagnostic criteria for mixed connective tissue disease (Kasukawa *et al*. 1987) was highly specific and sensitive in a group of Italian patients (Doria *et al*. 1991), and seems useful in children (Tiddens *et al*. 1993).

Longitudinal studies in mixed connective tissue disease

Long-term follow-up studies have shown that after many years of disease, features of only one autoimmune rheumatic disease (mainly scleroderma or systemic lupus erythematosus) predominate over the 'overlap' features present earlier in the course of the disease, which has cast doubt on mixed connective tissue disease being a separate entity. However, many of the patients that fulfil carefully selected criteria for mixed connective tissue disease at the time of diagnosis, will still fulfil them at long-term follow-up. Most of the patients originally described by Sharp *et al*. (1972) developed either systemic sclerosis or systemic lupus erythematosus of variable severity on long-term follow-up, although 5 of 25 patients were virtually asymptomatic (Nimelstein *et al*. 1980). In a series of 14 children with mixed connective tissue disease followed for nearly 10 years, features of systemic lupus erythematosus and polymyositis became less prominent, but scleroderma-like symptoms and joint abnormalities persisted (Tiddens *et al*. 1993).

Scleroderma-like skin abnormalities in adults and children are less responsive to treatment than initially thought and in most cases will persist, or may develop in patients who are initially diagnosed as having seronegative rheumatoid arthritis (Bennett and O'Connell 1980; Catoggio *et al*. 1993). Therefore the clinical features of mixed connective tissue disease do appear to evolve with time; skin and pulmonary lesions persist in adults whereas skin and joint involvement are prominent in children. The differentiation of mixed connective tissue disease into systemic lupus erythematosus or systemic sclerosis may be determined by the immunogenetic background (Gendi *et al*. 1995).

Studies in patients with anti-U1RNP antibodies

About 50 per cent of patients with anti-U1RNP antibodies do not have mixed connective tissue disease, even with extended follow-up (Table 11), although sometimes long periods of time are needed for the development of overlap manifestations (Lemmer *et al*. 1982). High titres of anti-U1RNP antibodies are associated more strongly with the development of mixed connective tissue disease than low titres (Lundberg and Hedfors 1991). Virtually all patients with mixed connective tissue disease will fulfil criteria for another autoimmune rheumatic disease at some stage in their illness. In our experience, patients with anti-U1RNP antibodies who fulfil criteria for mixed connective tissue disease without fulfilling criteria for systemic lupus erythematosus or systemic sclerosis are extremely rare (McHugh *et al*. 1990). None the less, studies that exclude mixed connective tissue disease once another autoimmune rheumatic disease is diagnosed (Calderon *et al*. 1984; Van Den Hoogen *et al*. 1994) may be misleading, as such patients may still be said to have the former disease. In the context of systemic lupus erythematosus, the presence of anti-U1RNP antibodies is associated with Raynaud's phenomenon, swollen fingers, myositis, pulmonary involvement, and less renal disease (Reichlin and Mattioli 1972; Maddison *et al*. 1978; McHugh *et al*. 1990). In patients with an overlap syndrome, the presence of anti-U1RNP antibodies may be associated with Raynaud's phenomenon and pulmonary hypertension (Ginsburg *et al*. 1983; Lazaro *et al*. 1989).

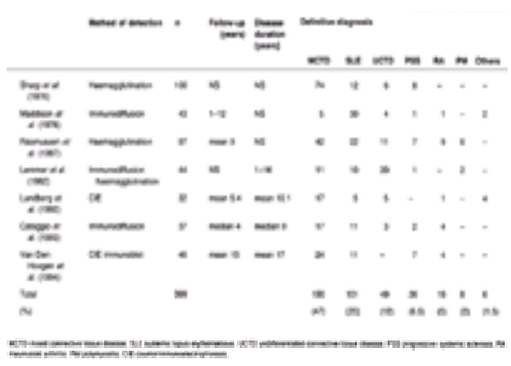
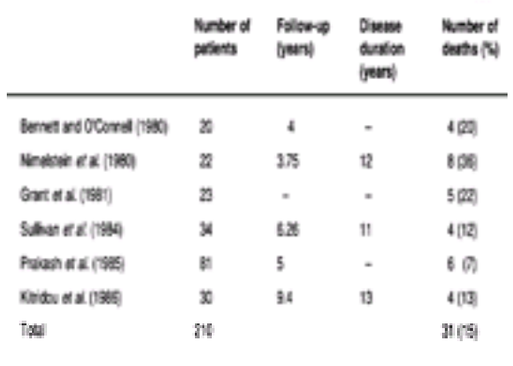


Table 11 Outcome in patients with anti-U1RNP antibodies

Prognosis

The prognosis is not as favourable as originally reported. The major cause of mortality appears to be pulmonary hypertension (Table 9). One-third of the 34 patients followed by Sullivan *et al*. (1984) had very severe disease requiring repeated courses or sustained high doses of corticosteroids, and 4 patients died. The mortality rate in some of the larger series is shown in Table 12. The 5- and 10-year survival rate was 90.5 per cent and 82.1 per cent in a Japanese series of patients, which was similar to that for systemic lupus erythematosus (Miyawaki and Onodera 1987). In children the presence of reactivity to the Sm D polypeptide is associated with a poorer prognosis (Hoffman *et al*. 1993).



	Number of patients	Follow-up (years)	Disease duration (years)	Number of deaths (%)
Bennett and O'Connell (1980)	20	4	–	4 (20)
Nimelstein *et al.* (1980)	22	3.75	12	8 (36)
Grant *et al.* (1981)	23	–	–	5 (22)
Sullivan *et al.* (1984)	34	6.26	11	4 (12)
Prakash *et al.* (1985)	81	5	–	6 (7)
Kitridou *et al.* (1986)	30	9.4	13	4 (13)
Total	210			31 (15)

Table 12 Reported mortality in mixed connective tissue disease

Management

Management of mixed connective tissue disease depends on manifestations of the disease. A suggested approach to management is outlined in Fig. 6. The scleroderma-like features are not as responsive to corticosteroids as originally reported. Raynaud's phenomenon is managed as in scleroderma, with emphasis on the patient keeping warm, avoiding trauma to the fingers, and discouragement of cigarette smoking. Vasodilators such as calcium channel blockers, angiotensin-converting enzymes, pentoxifylline and ketanserin are effective in some patients. Intravenous prostacyclin may be used for acute ischaemic digital lesions.

Fig. 6 Management of overlap syndrome.

Oesophageal involvement may respond to corticosteroids, although such treatment is rarely justified. More often, antacids, H_2-antagonists and omeprazole are required. Fatigue, fever, lymphadenopathy, arthralgia, and myalgia may respond to non-steroidal anti-inflammatory drugs in conventional doses. The addition of antimalarial agents early in the course of the disease is becoming a more widespread practice. Low-dose corticosteroids (less than 20 mg of prednisolone per day) may be needed if general symptoms do not respond to other measures. The development of an erosive polyarthritis may warrant the introduction of methotrexate or azathioprine. Myositis in mixed connective tissue disease may respond to lower doses of corticosteroids than polymyositis (Lundberg *et al*. 1992).

The management of renal disease remains controversial, although knowledge of histology may help in the decision. High-dose corticosteroids were effective for nephrotic syndrome and proteinuria in about two-thirds of patients described by Kitridou *et al*. (1986). However, cytotoxic drugs are probably indicated for those patients with more proliferative lesions, or those patients that require large corticosteroid doses for long periods of time.

Antiphospholipid antibodies should be sought in patients with pulmonary hypertension, and anticoagulation therapy considered when these antibodies are present (Ueda *et al*. 1991). Corticosteroids and cytotoxic drugs should be considered at an early phase because of the poor prognosis associated with pulmonary hypertension in mixed connective tissue disease and its usually rapid progression (Weiner-Kronish *et al*. 1981; Alpert *et al*. 1992). There are reports of patients responding to chlorambucil (Weiner-Kronish *et al*. 1981), cyclophosphamide (Friedman *et al*. 1992), and cyclophosphamide and cyclosporin A (Dahl *et al*. 1992). Recurrent episodes of pulmonary vasoconstriction may contribute to the development of pulmonary hypertension in patients with diffuse systemic sclerosis, the CREST syndrome and mixed connective tissue disease (Alpert *et al*. 1992). Nifedipine, hydralazine, prazosin, phenotolamine, ketanserin, prostacyclin, and captopril are all capable of acute and sustained reduction of pulmonary vascular resistance in such patients and therefore may be of additional therapeutic value.

Polymyositis overlap syndromes

Anti-tRNA synthetases syndrome

Autoantibodies are found in the majority of patients with polymyositis and dermatomyositis (Reichlin and Arnett 1984). The most frequent autoantibody is anti-Jo-1 which is directed against histidyl-tRNA synthetase, one of the 20 different enzymes that attaches the tRNA with its specific amino acid (Mathews and Bernstein 1983; Wasicek *et al*. 1984). Antibodies to five of these synthetases have been identified in patients with inflammatory myopathy (Targoff 1992). Anti-Jo-1 is found in 15 to 20 per cent of all myositis patients, and is found very rarely in dermatomyositis (Nishikai and Reichlin 1980).

A characteristic group of clinical features is associated with this group of autoantibodies. The 'antisynthetase syndrome' is characterized by myositis (100 per cent), interstitial lung disease (50 to 75 per cent), arthritis (60 to 100 per cent), and Raynaud's phenomenon (60 per cent) (Targoff 1992). A particular cutaneous feature is the presence of hyperkeratosis and fissuring along the lateral aspects of the fingers, so-called 'mechanic's (or machinist's) fingers' (Fig. 7).

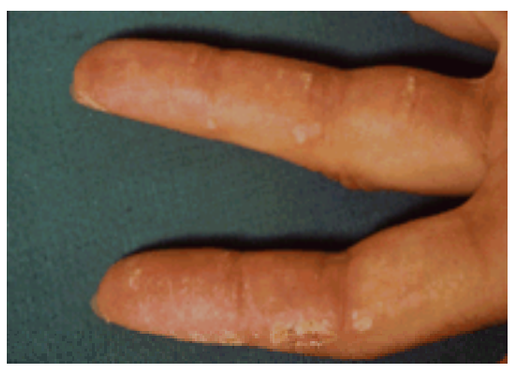


Fig. 7 Mechanic's finger.

Interstitial lung disease is one of the characteristic findings in the antisynthetase syndrome. Interstitial lung disease is found in 50 to 75 per cent of myositis patients with anti-Jo-1, compared with only 10 to 20 per cent of myositis overall (Bernstein *et al*. 1984; Hochberg *et al*. 1984; Love *et al*. 1991; Targoff 1992). The interstitial lung disease is indistinguishable from idiopathic interstitial lung disease, and it may persist despite response of the myositis to treatment and become a greater clinical problem than the myositis itself (Targoff and Arnett 1990). Arthritis is more frequent in patients with anti-Jo-1 antibodies (Bernstein *et al*. 1984; Plotz *et al*. 1989; Oddis *et al*. 1990; Targoff 1992). It is usually non-erosive, but deforming arthritis may be seen in one-third of patients. The wrist and finger joints are the most commonly involved joints.

Almost all patients with anti-Jo-1 antibodies have myositis at some stage in the course of their disease. However, antisynthetase antibodies may be found in patients with 'idiopathic' lung disease without clinical myositis (Bernstein *et al*. 1984; Targoff and Arnett 1990). Subclinical or absent myositis is more common with other non-Jo-1 antisynthetase antibodies (Targoff 1992). The myositis is indistinguishable from that of patients without antisynthetase antibodies, although patients with the autoantibodies may develop myositis at a younger age and have a more rapid form of onset (Plotz *et al*. 1989).

Rhupus

The existence of patients with a clinical picture resembling both rheumatoid arthritis and systemic lupus erythematosus has long been recognized. The description of lupus erythematosus cells by Hargraves in 1948 allowed a more precise segregation of patients with systemic lupus erythematosus, although a number of patients with 'true' rheumatoid arthritis, usually with more severe disease and extra-articular manifestations, were positive for lupus erythematosus cells. Conversely, there are patients with systemic lupus erythematosus who have a prominent synovitis associated with rheumatoid factor.

The appearance of nodules or an erosive arthropathy in patients with systemic lupus erythematosus (Cohen and Webb 1987) suggests a genuine overlap between the two conditions. The term 'rhupus' has been used to describe such patients (Panush *et al*. 1988). Although numerous case reports have appeared documenting a true coexistence of diffuse glomerulonephritis with nodules or erosive polyarthritis, 'rhupus' does not seem to occur more frequently than expected than the chance concurrence of systemic lupus erythematosus and rheumatoid arthritis (Panush *et al*. 1988). Some of the patients described in these earlier studies had overlap features and if serological techniques had been available for detecting anti-U1RNP antibodies, they may have fulfilled criteria for mixed connective tissue disease. Immunogenetic factors may determine which clinical features become most prominent (Brand *et al*. 1992).

Scleroderma overlap syndromes

Scleroderma and primary biliary cirrhosis

The association between systemic sclerosis and primary biliary cirrhosis is well recognized (Murray-Lyon *et al*. 1970; Reynolds *et al*. 1971). Scleroderma occurs in 4 to 17 per cent of patients with primary biliary cirrhosis according to large series (Clarke *et al*. 1978). Scleroderma in primary biliary cirrhosis is usually the CREST syndrome, now better referred to as limited cutaneous systemic sclerosis. Conversely, up to 5 per cent of patients with scleroderma have clinical evidence of primary biliary cirrhosis (Reimer 1990).

Antimitochondrial antibodies, the serological hallmark of primary biliary cirrhosis, are found in 15 to 25 per cent of sera from patients with systemic sclerosis, with a minority of the latter patients having overt evidence of primary biliary cirrhosis (Reimer 1990). Anticentromere antibodies, the serological hallmark of the CREST syndrome, are found in up to 30 per cent of sera from patients with primary biliary cirrhosis (Reynolds *et al*. 1971; Bernstein *et al*. 1982), and more than half of these patients have at least some features of CREST (Bernstein *et al*. 1982). The serological overlap between primary biliary cirrhosis and systemic sclerosis is more prevalent than the clinical overlap, although subclinical hepatic involvement may not be evident without liver biopsy. The serological overlap is not because of cross-reactivity between mitochondrial and centromere-associated antigens, which are independent antigenic targets (Whyte *et al*. 1994).

Patients with primary biliary cirrhosis also have a high prevalence of Sjögren's syndrome or keratoconjunctivitis sicca (Culp 1985). While evaluating 17 patients with primary biliary cirrhosis keratoconjunctivitis sicca was found in 76 per cent and a positive lip biopsy compatible with Sjögren's in 62 per cent of patients (Soriano *et al*. unpublished data)

Scleroderma–polymyositis overlap

Muscle involvement was recognized early in the description of the systemic sclerosis. Two different types of myopathy may occur. The more common myopathy is characterized by mild or no proximal muscle weakness, mild elevation of muscle enzymes, polyphasic motor unit potentials of normal amplitude and duration on electromyograms, and interstitial fibrosis and variation in diameter of muscle fibres without active inflammation on biopsy. This form has been classified as 'simple myopathy' (Clements *et al*. 1978) and does not appear to change significantly over long periods of time. A less frequent form resembles polymyositis with more proximal muscle weakness, very high muscle enzyme concentrations, inflammatory changes on biopsy, and polyphasic motor unit potentials of short duration and small amplitude on electromyography. The latter form requires active therapy and has been labelled inflammatory myopathy or 'complicated myopathy' (Clements *et al*. 1978).

In the series of patients with polymyositis reported by Bohan and co-workers (Bohan *et al*. 1977), 21 per cent of 153 patients had another autoimmune rheumatic disorder and were classified as type V polymyositis. The most frequently associated autoimmune rheumatic disease was scleroderma (36 per cent). The diagnosis of mixed connective tissue disease was not considered, especially as detailed serological investigation was not available. However, anti-U1RNP antibodies were present in 25 to 33 per cent of a later series of similar patients (Hochberg *et al*. 1986; Love *et al*. 1991). Of 34 patients studied in Argentina with polymyositis/dermatomyositis, 14 (41 per cent) had type V polymyositis. Anti-U1RNP was present in 67 per cent of cases, and one- half fulfilled Alarcon Segovia's criteria for mixed connective tissue disease (Imamura *et al*. 1993).

Other autoantibody specificities may also be associated with the scleroderma–polymyositis overlaps. Anti-Ku antibodies are associated with polymyositis–scleroderma overlap syndrome in Japan, but are less frequent in the United States; conversely anti-Pm-Scl antibodies are more frequent in polymyositis–scleroderma overlap in the United States but not in Japan (Targoff 1992). In general anti-Pm-Scl antibodies are found in 8 per cent of myositis patients and 3 per cent of scleroderma patients (Targoff 1992). Among 22 patients with anti-Pm-Scl antibodies, polymyositis/dermatomyositis alone was present in 55 per cent, scleroderma without polymyositis/dermatomyositis in 5 per cent, and myositis–scleroderma overlap in 41 per cent (Reichlin *et al*. 1984).

The prognosis of patients with type V polymyositis is no different from that in polymyositis alone (Bohan *et al*. 1977). It is possible that lower doses of corticosteroids may be sufficient to control myositis in association with mixed connective tissue disease (Lundberg *et al*. 1992).

Eosinophilia–myalgia syndrome, toxic oil syndrome, eosinophilic fasciitis

The eosinophilia–myalgia syndrome was first defined in 1989 (Centers for Disease Control 1989) as a new epidemic in association with the ingestion of L-tryptophan preparations (Belongia *et al*. 1990; Eidson *et al*. 1990), traced to a single Japanese manufacturer. (Belongia *et al*. 1990; Carr *et al*. 1994). The clinical picture and the pathological features, strongly resemble those of the Spanish toxic oil syndrome. Toxic oil syndrome was another epidemic that occurred in Spain in 1981 and was associated with the ingestion of adulterated rapeseed cooking oil. The presence of eosinophilia and fasciitis, link these two epidemics with eosinophilic fasciitis. However, no toxin has been associated with eosinophilic fasciitis.

Eosinophilia–myalgia syndrome and toxic oil syndrome

Epidemiology

The use of L-tryptophan was widespread in the United States in 1989 for the treatment of insomnia, premenstrual syndrome, and depression. Tryptophan was available to consumers without a prescription; thus a very large section of the population were at risk. As of 1 August 1992, the Centers for Disease Control had received a total of 1511 reports, including 38 deaths (Nightingale 1992). The surveillance case definition of eosinophilia–myalgia syndrome was as follows: eosinophil count greater than 1000/mm^3, incapacitating myalgia, and exclusion of other infectious or neoplastic illnesses that could account for the other two findings (Kilbourne *et al*. 1990). However, the actual number of individuals in whom some type of eosinophilia–myalgia syndrome had developed was estimated to be several times higher (Nightingale 1992).

National surveillance data from the United States in 1990 showed that 84 per cent of affected patients were female, 97 per cent were non-Hispanic white, and 87 were aged 35 years or older (median: 48; range: 4 to 85) (Swygert *et al*. 1990). The female preponderance is more likely due to ingestion patterns rather than gender susceptibility (Silver 1993). The usual daily dose ranged from 10 to 15 000 mg (median: 1500), and the median time between beginning use of tryptophan and the onset of symptoms was 127 days (range: 1 to 3668). Twenty-two per cent of those affected with eosinophilia–myalgia syndrome had taken tryptophan for more than 1 year before becoming ill, and 12 per cent had discontinued tryptophan a median of 15 days (range: 1 to 2858) before onset of symptoms (Swygert *et al*. 1990). Two risk factors for developing eosinophilia–myalgia syndrome were older age and the quantity of tryptophan consumed. The latter risk factor was not associated with the syndrome in a recent epidemiological study in Germany (Carr *et al*. 1994), possibly because of batch variations of quantities of the contaminants in the L-tryptophan (Mayeno and Gleich 1994). The occurrence rate of eosinophilia–myalgia syndrome in Germany was estimated as 40 per 100 000 users of the implicated source (Carr *et al*. 1994).

Toxic oil syndrome occurred in Spain in 1981 affecting more than 20 000 individuals with about 500 deaths (Mayeno and Gleich 1994). Epidemiological investigations implicated the use of aniline-denatured rapeseed oil that was sold and reprocessed illegally (Kilbourne *et al*. 1988).

Clinical features and histology

Symptoms at presentation

The more common clinical and laboratory manifestations of eosinophilia–myalgia syndrome and toxic oil syndrome are summarized in Table 13. The majority of patients with eosinophilia–myalgia syndrome presented with a flu-like syndrome characterized by myalgia and arthralgia, associated with profound weakness and fatigue (Kaufman *et al*. 1990). By definition all patients had severe myalgia with an elevated peripheral eosinophil count. Myalgia was generally diffuse, although could be localized, and accompanied by severe episodic muscle cramps (Kaufman *et al*. 1990; Hedberg *et al*. 1992). By contrast, patients with toxic oil syndrome presented as an atypical pneumonia, with non-productive cough, pleuritic chest pain, headache, fever, and bilateral pulmonary infiltrates. Gastrointestinal findings and striking eosinophilia became prominent within the first month.

Feature	EMS (Berygger et al. 1990) n=1075 (%)	TOS (Milesuma et al. 1983) n=181 (%)
Myalgia	100	95.5
Arthralgia	73	8.4
Rash	60	64.7
Peripheral oedema	59	64.7
Cough	59*	68.1
Dyspnoea	59*	42
Fever	36	68.1
Scleroderma like skin	32	–
Alopecia	28	5
Neuropathy	28	30.5
Hepatomegaly	5	37
Leucocytosis	65	71.1
Eosinophilia (> 500 cmm^3)	100*	77.7
Elevated ESR	33	38.6
Elevated aldolase	46	55.4
Elevated creatine kinase	10	–
Abnormal chest radiograph	21	87.6
Abnormal liver function test	49	41.3

*By definition all patients had eosinophilia and myalgia and eosinophil count at most recent visit > 10^3 cells (Berygger et al. 1990).
†Cough and dyspnoea taken together (Berygger et al. 1990).

Table 13 Prevalence of selected clinical and laboratory findings in eosinophilia–myalgia syndrome (EMS) and toxic oil syndrome (TOS)

Neuromuscular involvement

The clinical and pathological manifestations of neuromuscular involvement were qualitatively similar in both diseases, although more severe in toxic oil syndrome (Hertzman and Abaitua Borda 1993). In the acute and intermediate phases myalgia, muscle cramps, weakness, arthralgia, paraesthesia, and dysaesthesia were reported (Table 13). Late severe sensorimotor polyradiculopathy, with progressive weakness and occasional respiratory muscle involvement, and neuropathic contractures of the hands and wrist more commonly occurred in patients with toxic oil syndrome (30 per cent compared with less than 5 per cent) (Hertzman and Abaitua Borda 1993; Bolster and Silver 1994). In toxic oil syndrome the presence of severe initial systemic findings was associated with progression to neuromuscular illness (Kilbourne *et al*. 1983).

The most prominent histological feature in both conditions was a perimysial and epimysial inflammatory infiltrate consisting of mononuclear cells, histiocytes, and eosinophils (Kaufman *et al*. 1990; Varga *et al*. 1993; Lin *et al*. 1992). The degree of inflammatory infiltrate appeared to correlate with the extent of clinical symptoms (Kaufman *et al*. 1990). Myofibril atrophy, regeneration, and necrosis were almost absent in all biopsies, despite the severe and incapacitating myalgia (Lin *et al*. 1992). Later in the course of the disease, muscle fibre atrophy consistent with denervation was observed in both conditions (Hertzman and Abaitua Borda 1993).

Lungs

Pulmonary involvement was a major feature in both conditions. Toxic oil syndrome was characterized by an acute onset of cough and dyspnoea. Radiographic and pathological changes in toxic oil syndrome were consistent with a non-inflammatory non-cardiogenic pulmonary oedema, and 250 deaths occurred from this complication (Hertzman and Abaitua Borda 1993). Acute respiratory involvement in eosinophilia–myalgia syndrome occurred less frequently and was characterized by an interstitial pneumonitis that resolved spontaneously or was highly responsive to corticosteroids (Hertzman and Abaitua Borda 1993). Dyspnoea in eosinophilia–myalgia syndrome was related to parenchymal involvement of the lungs and to respiratory muscle weakness (Read *et al*. 1992).

Pulmonary hypertension developed in the intermediate stage of both conditions (1 to 12 months) in 8 per cent of patients with toxic oil syndrome (Cheg 1993) and in 5 per cent of patients with eosinophilia–myalgia syndrome (Martin *et al*. 1990). Cardiac involvement was rare and usually occurred as a result of pulmonary hypertension. However, primary involvement of the heart and coronary arteries has been reported in eosinophilia–myalgia syndrome (Berger *et al*. 1994) and in toxic oil syndrome (James *et al*. 1991).

Skin

The most frequent cutaneous manifestations in the acute phase of eosinophilia–myalgia syndrome and toxic oil syndrome were diffuse erythematous macules over the trunk and extremities. Oedema associated with cutaneous induration in the arms and legs was found in 28 to 60 per cent of patients with eosinophilia–myalgia syndrome (Kaufman *et al*. 1990; Swygert *et al*. 1990), and in 90 per cent of those with toxic oil syndrome. The oedema was due to fascial inflammation and was often accompanied by venous furrowing and the 'peau d'orange' induration characteristic of eosinophilic fasciitis (Kaufman *et al*. 1990). Histologically the changes could not be distinguished from eosinophilic fasciitis, with an inflammatory infiltrate of the fascial layers, a chronic perivascular infiltrate in the dermis in early cases, and thickening and homogenization of collagen bundles in later cases (Lin *et al*. 1992). Active skin lesions had disappeared in biopsy specimens at late follow-up (Hertzman and Abaitua Borda 1993), with oedema replaced by scleroderma-like changes. Non-necrotizing vascular lesions were more common in toxic oil syndrome than in eosinophilia–myalgia syndrome (Bolster and Silver 1994).

Gastrointestinal tract

Hepatomegaly, abdominal pain, nausea, diarrhoea, odynophagia, and dysphagia were particularly prominent during the first month in patients with toxic oil syndrome. Gastrointestinal symptoms have been reported in 64 per cent of patients who died of causes related to eosinophilia–myalgia syndrome (Swygert *et al*. 1993). Clinical evidence of pancreatic disease was rare in both entities, but 30 per cent of autopsies on patients with toxic oil syndrome showed marked atrophy of the pancreas (Hertzman and Abaitua Borda 1993).

Central nervous system

Cognitive symptoms and significant impairments of verbal and visual memory, conceptual reasoning, and motor speed were reported in 62 per cent of patients with eosinophilia–myalgia syndrome (Krupp *et al*. 1993) and 20 to 30 per cent of patients with toxic oil syndrome (Hertzman and Abaitua Borda 1993). Most patients with eosinophilia–myalgia syndrome were taking L-tryptophan because of insomnia and depression, and the use of antidepressive drugs may have contributed to some of these features. (Krupp *et al*. 1993). Electroencephalograms, computed tomography scans, and magnetic resonance imaging studies were normal (Hertzman and Abaitua Borda 1993; Krupp *et al*. 1993).

Laboratory findings

The laboratory findings are shown in Table 13. The most characteristic finding was the presence of eosinophilia in peripheral blood, usually within the first week. The eosinophilia was responsive to corticosteroid treatment and in eosinophilia–myalgia syndrome to discontinuation of L-tryptophan. The serum IgE levels were characteristically normal in eosinophilia–myalgia syndrome but were elevated in nearly half of patients with toxic oil syndrome (Belongia *et al*. 1990). Antinuclear antibodies were found in 40 per cent of patients with eosinophilia–myalgia syndrome and in a minority of patients with toxic oil syndrome. Normal creatine kinase levels were seen in almost all patients, although isolated elevation of serum aldolase was reported in 58 per cent of patients with eosinophilia–myalgia syndrome (Kaufman *et al*. 1990)

Aetiology and pathogenesis

Eosinophilia–myalgia syndrome

The epidemic nature of eosinophilia–myalgia syndrome, and its association with batches of L-tryptophan originating from a single manufacturer, suggested a contaminant (Varga *et al*. 1993). High-performance liquid chromatography analysis of L-tryptophan revealed more than 60 contaminants (Belongia *et al*. 1990). A single absorbance peak labelled 'peak E', or peak 97, and later identified as 1,1'-ethylidene bis[tryptophan] (**EBT**), was found consistently in L-tryptophan lots associated with eosinophilia–myalgia syndrome (Belongia *et al*. 1990). Lewis rats treated with EBT and/or case-associated L-tryptophan for 6 weeks developed significant increases in the thickness of the myofascia, compared with animals treated with vehicle or L-tryptophan not associated with the case. However, only animals treated with case-associated L-tryptophan (and not those receiving EBT) showed evidence of immune activation. Although these results demonstrated the pathological effect of EBT, they did not rule out the possibility that other impurities in the L-tryptophan associated with eosinophilia–myalgia syndrome may have contributed to some of the features of the syndrome (Love *et al*. 1993). Another murine model of eosinophilia–myalgia syndrome induced by intraperitoneal

administration of EBT was reported recently (Silver *et al*. 1994). A second contaminant (peak UV-5) has been identified as 3-(phenylamino)alanine (**PAA**) and related to L-tryptophan associated with cases of eosinophilia–myalgia syndrome (Mayeno *et al*. 1992). Of interest, PAA is chemically similar to 3-phenylamino-1,2-propanediol, an aniline derivative implicated in the development of toxic oil syndrome. The fact that many patients had ingested tryptophan for long periods (10 years or more) before becoming ill makes the existence of an inborn error of tryptophan metabolism an unlikely cause of eosinophilia–myalgia syndrome (Carr *et al*. 1994; Mayeno and Gleich 1994). However, several studies have found abnormal tryptophan metabolism in patients with eosinophilia–myalgia syndrome and toxic oil syndrome, with a shunting of L-tryptophan metabolism to the kynurenine pathway (Silver *et al*. 1992; Varga *et al*. 1993; Bolster and Silver 1994). These abnormalities could play a role in the pathogenesis of eosinophilia–myalgia syndrome, but more probably are a consequence of inflammation (Varga *et al*. 1993).

Both cellular and humoral autoimmune mechanisms have been implicated in the pathogenesis of eosinophilia–myalgia syndrome (Varga *et al*. 1993). A hypothetical model is that the initial trigger (tryptophan contaminant or metabolite) activates inflammatory cells, which are induced to secrete cytokines (IL-5, GM-CSF, transforming growth factor-b) that cause activation of eosinophils, and fibroblasts. Activated eosinophils may release cytokines and toxic granule proteins (including major basic protein), that may contribute to tissue injury. Activated fibroblasts produce increased amounts of collagen and other extracellular matrix components resulting in the characteristic fibrosis (Varga *et al*. 1993).

Toxic oil syndrome

Fatty acid anilides have been proposed as the aetiological agents in toxic oil syndrome. Fatty acid anilides may alter arachidonic acid metabolism or impair fibrinolytic activity in endothelial cells leading to the early intimal lesions seen in these patients (Bolster and Silver 1994).

Prognosis

The disease-related mortality rate for the first year was 2.7 per cent for eosinophilia–myalgia syndrome and 1.5 to 3.6 per cent for toxic oil syndrome (Alonso-Ruiz *et al*. 1993; Kaufman 1993; Swygert *et al*. 1993). It is likely that the incidence rate of eosinophilia–myalgia syndrome was underestimated, and toxic oil syndrome was more severe than eosinophilia–myalgia syndrome in its acute stages (Kaufman 1993). Progressive polyneuropathy and myopathy accounted for 67 per cent of deaths in patients with eosinophilia–myalgia syndrome. Older age and involvement of more than one organ system, in particular neuromuscular, pulmonary, and cardiovascular sequelae, suggested a poor prognosis (Swygert *et al*. 1993).

Long-term morbidity was frequent in both syndromes. Persistent symptoms such as myalgia (50 to 94 per cent), muscle cramping (43 to 90 per cent), and neuropathy (30 to 91 per cent) were reported up to 36 months after the onset of eosinophilia–myalgia syndrome (Kaufman 1993). The consumption of multivitamin supplements before the appearance of eosinophilia–myalgia syndrome was associated with reduced severity of subacute symptoms (Hatch and Goldman 1993). In an 8-year follow-up study of 332 patients with toxic oil syndrome, only 9 per cent achieved full remission (Alonso-Ruiz *et al*. 1993). The severity of the chronic manifestations was variable, but was mild in most of the cases (Alonso-Ruiz *et al*. 1993). Muscle cramps and chronic musculoskeletal pain were the most common symptoms.

Management

A variety of medical regimens, mainly non-steroidal anti-inflammatory agents and corticosteroids, have been used to treat patients with eosinophilia–myalgia syndrome and toxic oil syndrome. Although acute symptoms such as oedema, acute pulmonary disease, and eosinophilia responded to corticosteroid treatment, such treatment was not associated with long-term improvement in symptoms. Corticosteroids may have been life saving in cases of acute respiratory failure due to non-cardiogenic pulmonary oedema in toxic oil syndrome.

Eosinophilia–myalgia syndrome in children

Eosinophilia–myalgia syndrome occasionally has been identified among children, including a neonate with persistent eosinophilia whose mother ingested tryptophan during pregnancy (Hatch *et al*. 1991). More recently a case of eosinophilia–myalgia syndrome was reported in a child with phenylketonuria who developed the syndrome after ingesting a specialized infant formula containing contaminated tryptophan (Springer *et al*. 1992).

Eosinophilic fasciitis

Eosinophilic fasciitis is a relatively uncommon idiopathic disease, first described by Shulman (1974), and characterized by fasciitis and peripheral blood eosinophilia. It shares similar clinical and histological features with eosinophilia–myalgia syndrome and toxic oil syndrome, but significant internal organ involvement is uncommon. Also, there are differences that distinguish eosinophilic fasciitis from systemic sclerosis: the absence of Raynaud's phenomenon, normal nailfold capillaries, sparing of the dermis, infrequent visceral involvement, absence of the serological features characteristic of systemic sclerosis, and the development of haematological complications, such as aplastic anaemia and thrombocytopenia, seldom reported in typical systemic sclerosis (Michet *et al*. 1981; Maddison 1991). The age of onset is 30 to 40 years (mean: 47; range: 11 to 72), and in half of the patients a history of recent strenuous exertion can be recalled (Lakhampal *et al*. 1988).

Cutaneous manifestations are the most common presenting feature and usually evolve through three stages: pitting oedema, 'peau d'orange', and induration. These stages are often present simultaneously in different areas of the body (Maddison 1991). The arms and legs are affected most commonly and simultaneous involvement of hands and feet is not uncommon. Localized morphoea may occur in other areas, especially in children (Miller 1992). Synovitis is not uncommon and may be the presenting feature. Low-grade myositis may be present, although serum creatine kinase levels are usually normal.

Eosinophilic fasciitis has been associated with malignancy and may present as a paraneoplastic syndrome (Naschitz *et al*. 1994), which remits with successful cancer surgery. Haematological malignancies are overrepresented, and aplastic anaemia in particular has a high associated mortality (Hoffman *et al*. 1982). Eosinophilic fasciitis associated with malignancy has a female predominance and usually fails to respond to corticosteroids (Naschitz *et al*. 1994).

Peripheral eosinophilia, sometimes impressive, is the most striking laboratory feature. Eosinophilia may be transient and the diagnosis should not be dismissed because of its absence. Diagnosis is confirmed by a cutaneous biopsy to include tissue extending from the epidermis to skeletal muscle and the deep fascia. Characteristic histological findings are a widespread inflammatory infiltrate involving the deep fascia and septae of the subdermal fat as well as the dermal layer and a normal epidermis.

The cause of eosinophilic fasciitis is unknown. Of interest is the association with strenuous exertion, especially in men. After the description of eosinophilia–myalgia syndrome associated with L-tryptophan, some retrospective studies have found an association of eosinophilic fasciitis with the consumption of L-tryptophan (Freundlich *et al*. 1990; Martin *et al*. 1991), while others have not (Varga *et al*. 1991). Recently two cases of diffuse fasciitis with peripheral eosinophilia were described in which *Borrelia burgdoferi* was identified on biopsy specimens. The term borrelial fasciitis was used to describe such lesions (Granter *et al*. 1994).

More than half of patients with eosinophilic fasciitis respond to corticosteroids, although complete remission is achieved in only 15 per cent. There may be spontaneous remissions. Other agents such as cimetidine, hydroxychloroquine, colchicine, D-penicillamine, and cyclosporin (Laneuville 1992) have been used with variable results. In children, two-thirds of patients developed cutaneous fibrosis (Farrington *et al*. 1993).

Chapter References

Alarcon, G.S. *et al*. (1991). Early undifferentiated connective tissue disease. I. Early clinical manifestation in a large cohort of patients with undifferentiated connective tissue diseases compared with cohorts of well established connective tissue disease. *Journal of Rheumatology*, **18**, 1332–9.

Alarcon-Segovia, D. (1994). Mixed connective tissue disease and overlap syndromes. *Clinics in Dermatology*, **12**, 309–16.

Alarcon-Segovia, D. and Villarreal, M. (1987) Classification and diagnostic criteria for mixed connective tissue diseases. In *Mixed connective tissue diseases and anti-nuclear antibodies* (ed. R. Kasukawa and G.C. Sharp), pp. 33–40. Excerpta Medica, Amsterdam.

Alarcon-Segovia, D. and Cardiel, M.H. (1989). Comparison between 3 diagnostic criteria for mixed connective tissue disease. Study of 593 patients. *Journal of Rheumatology*, **16**, 328–34.

Alonso-Ruiz, A., Calabozo, M., Perez-Ruiz, F., and Mancebo, L. (1993). Toxic oil syndrome: a long-term follow-up of a cohort of 332 patients. *Medicine (Baltimore)*, **72**, 285–95.

Alpert, M.A. *et al*. (1983). Cardiovascular manifestations of mixed connective tissue disease in adults. *Circulation*, **68**, 1182–93.

Alpert, M.A., Pressly, T.A., Mukerji, V., Lambert, C.R., and Mukerji, B. (1992). Short-and long-term hemodynamic effects of captopril in patients with pulmonary hypertension and selected connective tissue disease. *Chest*, **102**, 1407–12.

Babini, S.M., Maldonado-Cocco, J.A., Barcelo, H.A., and Garcia-Morteo, O. (1985). Peritendinous nodules in overlap syndrome. *Journal of Rheumatology*, **12**, 160–4.

Belongia, E.A. *et al*. (1990). An investigation of the cause of the eosinophilia–myalgia syndrome associated with tryptophan use. *New England Journal of Medicine*, **323**, 357–65.

Bennett, R.M. and O'Connell, D.J. (1980). Mixed connective tissue disease: a clinicopathologic study of 20 cases. *Seminars in Arthritis and Rheumatism*, **10**, 25–50.

Bennett, R.M. and Spargo, B.H. (1977). Immune complex nephropathy in mixed connective tissue disease. *American Journal of Medicine*, **63**, 534–41.

Bennett, R.M., Bong, D.M., and Spargo, B.H. (1978). Neuropsychiatric problems in mixed connective tissue disease. *American Journal of Medicine*, **65**, 955–62.

Berger, P.B., Duffy, J., Reeder, G.S., Karon, B.L., and Edwards, W.D. (1994) Restrictive cardiomyopathy associated with the eosinophilia–myalgia syndrome. *Mayo Clinic Proceedings*, **69**, 162–5.

Bernstein, R.M., Callender, M.E., Neuberger, J.M., Hughes, G.R.V., and Williams, R. (1982). Anticentromere antibody in primary biliary cirrhosis. *Annals of the Rheumatic Diseases*, **41**, 612–14.

Bernstein, R.M. *et al*. (1984). Anti Jo-1 antibody: a marker for myositis with interstitial lung disease. *British Medical Journal*, **289**, 151–3.

Black, C. and Isenberg, D.A. (1992). Mixed connective tissue disease—goodbye to all that. *British Journal of Rheumatology*, **31**, 695–700.

Bodolay, E., Bojan, F., Szegedi, G., Stenszky, V., and Farid, N.R. (1989). Cytotoxic endothelial cell antibodies in mixed connective tissue disease. *Immunology Letters*, **20**, 163–7.

Bohan, A., Peter, J.B., Bowman, R.L., and Pearson, C.M. (1977). A computer-assisted analysis of 153 patients with polymyositis and dermatomyositis. *Medicine (Baltimore)*, **56**, 255–86.

Bolster, M.B. and Silver, R.M. (1994). Eosinophilia–myalgia syndrome, toxic-oil syndrome, and diffuse fasciitis with eosinophilia. *Current Opinion in Rheumatology*, **6**, 642–9.

Brand, C.A., Rowley, M.L, Tate, B.D., Muirden, K.D., and Whittingham, S.F. (1992). Coexistent rheumatoid arthritis and systemic lupus erythematosus. Clinical, serological and phenotypic features. *Annals of the Rheumatic Diseases*, **51**, 173–6.

Calderon, L., Rodriguez-Valverde, V., Sanchez-Andrade, S., Riestra, J.L., and Gomez-Reyno, J. (1984). Clinical profiles of patients with antibodies to nuclear ribonucleoprotein. *Clinical Rheumatology*, **8**, 483–92.

Carr, L., Ruther, E., Berg, P.A., and Lehnert, H. (1994). Eosinophilia–myalgia syndrome in Germany: an epidemiologic review. *Mayo Clinic Proceedings*, **69**, 620–5.

Catoggio, L.J., Evison, G., Harkness, J.A.L., and Maddison, P.J. (1983). The arthropathy of systemic sclerosis (scleroderma); comparison with mixed connective tissue disease. *Clinical and Experimental Rheumatology*, **1**, 101–12.

Catoggio, L.J., Soriano, E.R., Imamura, P., Almeida, V., Gonzales Salas, P., and Mayorga, L.M.(1993). Caracteristicas y evolucion de 52 pacientes con anti U1 RNP. *Revista Argentina de Reumatologia*, **4**, 138 (abstract).

Centers for Disease Control (1989). Eosinophili–myalgia syndrome New Mexico. *Morbidity and Mortality Weekly Report*, **38**, 765–7.

Cheg, T.O. (1993). Pulmonary hypertension in patients with eosinophilia–myalgia syndrome or toxic oil syndrome. *Mayo Clinic Proceedings*, **68**, 823–4.

Clarke, A.K., Galbraith, R.M., Hamilton, E.B.D., and Williams, R. (1978). Rheumatic disorders in primary biliary cirrhosis. *Annals of the Rheumatic Diseases*, **37**, 42–7.

Clements, P.J. *et al*. (1978). Muscle disease in progressive systemic sclerosis. *Arthritis and Rheumatism*, **21**, 62–71.

Cohen, M.G. and Webb, J. (1987). Concurrence of rheumatoid arthritis and systemic lupus erythematosus: report of 11 cases. *Annals of the Rheumatic Diseases*, **46**, 853–8.

Culp, K.S. (1985). Autoimmune associations in primary biliary cirrhosis. *Mayo Clinic Proceedings*, **57**, 365.

Dahl, M., Chalmers, A., Wade, J., Calverley, D., and Munt, B. (1992). Ten year survival of a patient with advanced pulmonary hypertension and mixed connective tissue disease treated with immunosuppressive therapy. *Journal of Rheumatology*, **19**, 1807–9.

Doria, A., Bonavina, L., Anselmino, M., Ruffatti, A., Favaretto, M., and Gambari, P. (1991). Esophageal involvement in mixed connective tissue disease. *Journal of Rheumatology*, **18**, 685–90.

Doria, A., Ghirardello, A., Zambiasi, P., Ruffatti, A., and Gambari, P. (1992). Japanese diagnostic criteria for mixed connective tissue disease in Caucasian patients. *Journal of Rheumatology*, **19**, 259–64.

Doria A. *et al*. (1992). Antiphospholipid antibodies in mixed connective tissue disease. *Clinical Rheumatology*, **11**, 48–50

Eidson, M., Philen, R.M., Sewell, C.M., Voorhees, R., and Kilbourne, E.M. (1990). L-tryptophan and eosinophilia–myalgia syndrome in New Mexico. *Lancet*, **335**, 645–8.

Emlen, W. (1979). Complete heart block in mixed connective tissue disease. *Arthritis and Rheumatism*, **22**, 679–80.

Farrington, M.L., Haas, J.E., Nazar-Stewart, V., and Mellins, E.D. (1993). Eosinophilic fasciitis in children frequently progress to scleroderma-like cutaneous fibrosis. *Journal of Rheumatology*, **20**, 128–32.

Fatenjad, S., Mamula, M.J., and Craft, J. (1993). Role of intermolecular/intrastructural B- and T-cell determinants in the diversification of autoantibodies to ribonucleoprotein particles. *Proceedings of the National Academy of Sciences (USA)*, **90**, 12010–14.

Fisher, D.E., Reeves, W.H., Wisniewolski, R., Lahita, R.G., and Chiorazzi, N. (1985). Temporal shifts from Sm to ribonucleoprotein reactivity in systemic lupus erythematosus. *Arthritis and Rheumatism*, **28**, 1348–55.

Freundlich, B. *et al*. (1990). L-tryptophan ingestion associated with eosinophilic fasciitis but not progressive systemic sclerosis. *Annals of Internal Medicine*, **112**, 758–62.

Friedman, D.M., Mitnick, H.J., and Danilowicz, D. (1992). Recovery from pulmonary hypertension in an adolescent with mixed connective tissue disease. *Annals of the Rheumatic Diseases*, **51**, 1001–4.

Fujimoto, M., Kira, J., Murai, H., Yoshimura, T., Takizawa, K., and Goto, I.(1993). Hypertrophic cranial pachymeningitis associated with mixed connective tissue disease; a comparison with idiopathic and infectious pachymeningitis. *Internal Medicine*, **32**, 510–12.

Furuya, T. *et al*. (1992). Mixed connective tissue disease associated with protein losing enteropathy, successful treatment with intravenous cyclophosphamide therapy. *Internal Medicine*, **31**, 1359–62.

Gendi, N.S.T., Welsh, K.I., Van Venrooij, W.J., Vancheeswaran, R., Gilroy, J., and Black, C. (1995). HLA type as a predictor of mixed connective tissue disease differentiation. Ten-year clinical and immunogenetic followup of 46 patients. *Arthritis and Rheumatism*, **38**, 259–66.

Gilliam, J.N. and Prystowsky, S.D. (1977). Mixed connective tissue disease syndrome. Cutaneous manifestations of patients with epidermal nuclear staining and high titer serum antibody to ribonuclease-sensitive extractable nuclear antigen. *Archives of Dermatology*, **113**, 583–7.

Ginsburg, W.W., Conn, D., Bunch, T.W., and McDuffie, F.C. (1983). Comparison of clinical and serologic markers in systemic lupus erythematosus and overlap syndrome: a review of 247 patients. *Journal of Rheumatology*, **10**, 235–41.

Goldstein, R. *et al*. (1990). HLA-D region genes associated with autoantibody responses to histidyl-transfer RNA synthetase (Jo-1) and other translation-related factors in myositis. *Arthritis and Rheumatism*, **33**, 1240–8.

Granier, F., Vayssairat, M., Priollet, P., and Housset, E. (1986). Nailfold capillary microscopy in mixed connective tissue disease. Comparison with systemic sclerosis and systemic lupus erythematosus. *Arthritis and Rheumatism*, **29**, 189–95.

Grant, K.D., Adams, L.E., and Hess, E.V. (1981). Mixed connective tissue disease—a subset with sequential clinical and laboratory features. *Journal of Rheumatology*, **8**, 587–98.

Granter, S.R., Barnhill, R.L., Hewins, M.E., and Duray, P.H. (1994). Identification of *Borrelia burgdorferi* in diffuse fasciitis with peripheral eosinophilia: borrelial fasciitis. *Journal of the American Medical Association*, **272**, 1283–5.

Hagen, N.A., Stevens, J.C., and Michet, C.J., Jr. (1990). Trigeminal sensory neuropathy associated with connective tissue diseases. *Neurology*, **40**, 891–6.

Hainaut, P., Lavenne, E., Magy, J.M., and Lebacq, E.G. (1986). Circulating lupus type anticoagulant and pulmonary hypertension associated with mixed connective tissue disease. *Clinical Rheumatology*, **5**, 96–101.

Hameenkorpi, R., Ruuska, P., Forsberg, S., Tiilikainen, R., Makitalo, R., and Hakala, M. (1993). More evidence of distinctive features of mixed connective tissue disease. *Scandinavian Journal of Rheumatology*, **22**, 63–8.

Hatch, D.L. and Goldman, L.R. (1993). Reduced severity of eosinophilia–myalgia syndrome associated with the consumption of vitamin-containing supplements before illness. *Annals of Internal Medicine*, **153**, 2368–73.

Hatch, D.L., Garona, J.E., Goldman, L.R., and Walker, K.O. (1991). Persistent eosinophilia in an infant with probable intrauterine exposure to L-tryptophan-containing supplements. *Pediatrics*, **88**, 810–13.

Hedberg, K., Urbach, D., Slutsker, L., Matson, P., and Fleming, D. (1992). Eosinophilia–myalgia syndrome. Natural history in a population-based cohort. *Archives of Internal Medicine*, **152**, 1889–92.

Hench, P.K., Edgington, T.S., and Tan, E.M. (1975). The evolving clinical spectrum of mixed connective tissue disease (MCTD). *Arthritis and Rheumatism*, **18**, 404.

Hertzman, P.A. and Abaitua Borda, I. (1993). The toxic oil syndrome and the eosinophilia–myalgia syndrome: pursuing clinical parallels. *Journal of Rheumatology*, **20**, 1707–10.

Hirose, W. *et al*. (1993). Duodenal hemorrhage and dermal vasculitis associated with mixed connective tissue disease. *Journal of Rheumatology*, **20**, 151–4.

Hochberg, M.C., Feldman, D., Stevens, M.B., Arnett, F.C., and Reichlin, M. (1984). Antibody to Jo-1 in polymyositis/dermatomyositis. Association with interstitial pulmonary disease. *Journal of Rheumatology*, **11**, 663–5.

Hochberg, M.C., Feldman, D., and Stevens, M.B. (1986). Adult onset polymyositis/dermatomyositis: an analysis of clinical and laboratory features and survival in 76 patients with a review of the literature. *Seminars in Arthritis and Rheumatism*, **15**, 168–78.

Hoffman, R., Young, N., Ershler, W.B., Mazur, E., and Gewirtz, A. (1982). Diffuse fasciitis and aplastic anemia: a report of four cases revealing an unusual association between rheumatologic and hematologic disorders. *Medicine*, **61**, 373–81.

Hoffman, R.W. *et al*. (1993). Human T cell clones reactive against U-small nuclear ribonucleoprotein autoantigens from connective tissue disease patients and healthy individuals. *Journal of Immunology*, **151**, 6460–9.

Holyst, M.M., Hill, D.L., Sharp, G.C., and Hoffman, R.W. (1994). Increased frequency of mutations in the hprt gene of T cells isolated from patients with anti-U1–70 kD-autoantibody-positive connective tissue disease. *International Archives of Allergy and Immunology*, **105**, 234–7.

Hutton, C.W., Maddison, P.J., Collins, A.J., and Berriman, J.A. (1988). Intra-articular apatite deposition in mixed connective tissue disease: crystallographic and technetium scanning charcacteristics. *Annals of the Rheumatic Diseases*, **47**, 1027–30.

Imamura, P.M., Catoggio, L.J., Soriano, E.R., Baruzzo, C., and Mayorga, L.M. (1993). Clinical and serological characteristics of 34 Argentine patients with dermato-polymyositis. *Revista Argentina de Reumatologia*, **4**, 90–8.

Izumiyama, T. *et al*. (1993). Small airway involvement in mixed connective tissue disease. *Tokyo Journal of Experimental Medicine*, **170**, 173–83.

James, J.P. *et al*. (1990). Factor VIII-related antigen in connective tissue disease patients and relatives. *British Journal of Rheumatology*, **29**, 6–9.

James, T.N. *et al*. (1991). Cardiac abnormalities in the toxic oil syndrome, with comparative observations on the eosinophilia–myalgia syndrome. *Journal of the American College of Cardiology*, **18**, 1367–79.

Kahn, M.F., Bourgeois, P., Aeschlimann, A., and de Truchis, P. (1989). Mixed connective tissue disease after exposure to polyvinyl chloride. *Journal of Rheumatology*, **16**, 533–5.

Kallenberg, C.G., Van-Dissel Emiliani, F., Huitema, M.G., Limburg, P.C., and The, T.H. (1988). B-Cell proliferation and differentiation in systemic lupus erythematosus and mixed connective tissue disease. *Journal of Clinical Laboratory and Immunology*, **26**, 55–61.

Kappes, J. and Bennett, R.M. (1982). Cauda equina syndrome in a patient with high titer anti-RNP antibodies. *Arthritis and Rheumatism*, **25**, 349–52.

Kasukawa, R. *et al*. (1987). Preliminary diagnostic criteria for classification of mixed connective tissue diseases. In *Mixed connective tissue diseases and anti-nuclear antibodies* (ed. R. Kasukawa and G.C. Sharp), pp. 41–7. Excerpta Medica, Amsterdam.

Kasukawa, R., Nishimaki, T., Takagi, T., Miyawaki, S., Yokohari, R., and Tsunematsu, T. (1990). Pulmonary hypertension in connective tissue disease. Clinical analysis of sixty patients in a multi-institutional study. *Clinical Rheumatology*, **9**, 56–62.

Kaufman, L.D. (1993). Eosinophilia–myalgia syndrome: morbidity and mortality. *Journal of Rheumatology*, **20**, 1644–6.

Kaufman, L.D., Seidman, R.J., and Gruber, B.L. (1990). L-tryptophan-associated eosinophilic perimyositis, neuritis, and fasciitis. A clinicopathologic and laboratory study of 25 patients. *Medicine (Baltimore)*, **69**, 187–99.

Kessler, E., Halpern, M., Chagnac, A., Zevin, D., Hammel, I., and Ben-Bassat, M. (1992). Unusual renal deposit in mixed connective tissue disease. *Archives of Pathology and Laboratory Medicine*, **116**, 261–4.

Kilbourne, E.M. *et al*. (1983). Clinical epidemiology of toxic-oil syndrome. *New England Journal of Medicine*, **309**, 1408–14.

Kilbourne, E.M. *et al*. (1988). Chemical correlates of pathogenicity of oils related to the toxic oil syndrome epidemic in Spain. *American Journal of Epidemiology*, **127**, 1210–27.

Kilbourne, E.M. *et al*. (ed.) (1990). Interim guidance on the eosinophilia–myalgia syndrome. *Annals of Internal Medicine*, **112**, 85–6.

Kitridou, R.C., Akmal, M., Turkel, S.B., Ehresmann, G.R., Quismorio, F.P., Jr., and Massry, S.G. (1986). Renal involvement in mixed connective tissue disease: a longitudinal clinicopathologic study. *Seminars in Arthritis and Rheumatism*, **16**, 135–45.

Krupp, L.B., Masur, D.M., and Kaufman, L.D. (1993). Neurocognitive dysfunction in the eosinophilia–myalgia syndrome. *Neurology*, **43**, 931–6.

Kuipers, E.J., Van-Leeuwen, M.A., Nillels, P.G., Jager, J., and Van Rijswijk, M.H. (1991). Hemobilia due to vasculitis of the gall bladder in a patient with mixed connective tissue disease. *Journal of Rheumatology*, **18**, 617–18.

Kumagai Y., Shiokawa Y., Medsger T.A., and Rodnan G.P.(1984). Clinical spectrum of connective tissue disease after cosmetic surgery. *Arthritis and Rheumatism*, **27**, 1–12.

Kuwana, M., Okano, Y., Kaburaki, J., Tsuji, K., and Inoko, H. (1995). Major histocompatibility complex class II gene associations with anti-U1 small nuclear ribonucleoprotein antibody. *Arthritis and Rheumatism*, **38**, 396–405.

Lakhampal, S., Ginsburg, W.W., Michet, J.J., Doyle, J.A., and Breanndan Morre, S. (1988). Eosinophilic fasciitis: clinical spectrum and therapeutic response in 52 cases. *Seminars in Arthritis and Rheumatism*, **17**, 221–31.

Laneuville, P. (1992). Cyclosporin A induced remission of CD4+ T CLL, associated with eosinophilia and fasciitis. *British Journal of Haematology*, **80**, 252–4.

Langley, R.L. and Treadwell, E.L. (1994). Cardiac tamponade and pericardial disorders in connective tissue diseases: case report and literature review. *Journal of the National Medical Association*, **86**, 149–53.

Lazaro, M.A. *et al*. (1989). Clinical and serologic characteristics of patients with overlap syndrome: is mixed connective tissue disease a distinct clinical entity? *Medicine (Baltimore)*, **68**, 58–65.

Lazaro, M.A., Gomez Tejada, R., Berenstein, T., Gene, R., Citera, G., and Maldonado Cocco, J.A. (1993). Pulmonary hypertension in overlap syndrome. *Arthritis and Rheumatism*, **36** (Suppl.), S276

(abstract).

Lemmer, J.P., Curry, N.H., Mallory, J.H., and Waller, M.V. (1982). Clinical characteristics and course in patients with high titer anti-RNP antibodies. *Journal of Rheumatology*, **9**, 536–42.

LeRoy, E.C. and Medsger, T.A., Jr. (1992). Raynaud's phenomenon: a proposal for classification. *Clinical and Experimental Rheumatology*, **10**, 485–8.

LeRoy, E.C., Mariq, H.R., and Kahaleh, M.B. (1980). Undifferentiated connective tissue syndromes. *Arthritis and Rheumatism*, **23**, 341–3.

Leung, W.H., Wong, K.L., Lau, C.P., Wong, C.K., Cheng, C.H., and Tai, Y.T. (1990). Doppler-echo evaluation of left ventricular diastolic filling in patient with mixed connective tissue disease. *Cardiology*, **77**, 93–100.

Lin, J.D. *et al*. (1992). Pathologic manifestations of the eosinophilia–myalgia syndrome: analysis of 11 cases. *Human Pathology*, **23**, 429–37.

Love, L.A. *et al*. (1991). A new approach to the classification of idiopathic inflammatory myopathy: myositis-specific autoantibodies define useful homogenous patient groups. *Medicine (Baltimore)*, **70**, 360–74.

Love, L.A. *et al*. (1993). Pathological and immunological effects of ingesting L-tryptophan and 1,1'-ethylidenebis(L-tryptophan) in Lewis rats. *Journal of Clinical Investigation*, **91**, 804–11.

Lundberg, I. and Hedfors, E. (1991). Clinical course of patients with anti-RNP antibodies. A prospective study of 32 patients. *Journal of Rheumatology*, **18**, 1511–9.

Lundberg, I., Nennesmo, I., and Hedfors, E. (1992). A clinical, serological, and histopathological study of myositis patients with and without anti-RNP antibodies. *Seminars in Arthritis and Rheumatism*, **22**, 127–38.

Lundberg, I., Nyman, U., Pettersson, I., and Hedfors, E. (1992). Clinical manifestations and anti-(U1)sn RNP antibodies: a prospective study of 29 anti-RNP antibody positive patients. *British Journal of Rheumatology*, **31**, 811–17.

Maddison, P.J. (1991). Mixed connective tissue disease, overlap syndromes, and eosiniphilic fasciitis. *Annals of the Rheumatic Diseases*, **50**, 887–93.

Maddison, P.J., Mogavero, H., and Reichlin, M. (1978). Patterns of clinical disease associated with antibodies to nuclear ribonucleoprotein. *Journal of Rheumatology*, **5**, 407–11.

Magnusson, B. (1993). Collagen diseases in children 0–15 years of age in Sweden 1984–88. *Journal of Rheumatology*, **20** (Suppl.), 61 (abstract).

Marguerie, C. *et al*. (1992). The clinical and immunogenetic features of patients with autoantibodies to the nucleolar antigen PM-Scl. *Medicine (Baltimore)*, **71**, 327–36.

Maricq, H.R. *et al*. (1980). Diagnostic potential of in vivo capillary microscopy in scleroderma and related disorders. *Arthritis and Rheumatism*, **23**, 183–9.

Marshall, J.B. *et al*. (1990). Gastrointestinal manifestations of mixed connective tissue disease. *Gastroenterology*, **98**, 1232–8.

Martin, R.W. *et al*. (1990). The clinical spectrum of the eosinophilia–myalgia syndrome associated with L-tryptophan ingestion. *Annals of Internal Medicine*, **113**, 124–34.

Martin, R.W., Duffy, J., and Lie, J.T. (1991). Eosinophilic fascitis associated with use of L-tryptophan: a case control study and comparison of clinical and histopathologic features. *Mayo Clinic Proceedings*, **66**, 892–8.

Mathews, M.B. and Bernstein, R.M. (1983). Myositis autoantibody inhibits histidyl-tRNA synthetase: a model for autoimmunity. *Nature*, **304**, 177–9.

Mayeno, A.N. and Gleich, G.J. (1994). The eosinophilia–myalgia syndrome: lessons from Germany. *Mayo Clinic Proceedings*, **69**, 702–4.

Mayeno, A.N., Belongia, E.A., Lin, F., Lundy, S.K., and Gleich, G.J. (1992). 3-(Phenylamino)alanine, a novel aniline-derived amino-acid associated with the eosinophilia–myalgia syndrome: a link to the toxic oil syndrome? *Mayo Clinic Proceedings*, **67**, 1134–9.

McHugh, N., James, I., and Maddison, P. (1990). Clinical significance of antibodies to a 68 kDa U1RNP polypeptide in connective tissue disease. *Journal of Rheumatology*, **17**, 1320–8.

Michet, C.J., Doyle, J.A., and Ginsburg, W.W. (1981). Eosinophilic fasciitis. Report of 15 cases. *Mayo Clinic Proceedings*, **56**, 27–34.

Miller, J.J., 3rd. (1992). The fasciitis–morphea complex in children. *American Journal of Diseases of Children*, **146**, 733–6.

Miyata, M. *et al*. (1992). Pulmonary hypertension in MCTD: report of two cases with anticardiolipin antibody. *Clinical Rheumatology*, **11**, 195–201.

Miyawaki, S. and Onodera, H. (1987). Clinical course and prognosis of patients with mixed connective tissue disease. In *Mixed connective tissue disease and anti-nuclear antibodies* (ed. R. Dasukawa and G. Sharp), pp. 331–6. Excerpta Medica, Amsterdam.

Murray-Lyon, I.M., Thompson, R.P.H., Ansell, I.D., and Williams, R. (1970). Scleroderma and primary biliary cirrhosis. *British Medical Journal*, **iii**, 258–9.

Naschitz, J.E. *et al*. (1994). Cancer associated fasciitis panniculitis. *Cancer*, **73**, 231–5.

Nightingale, S.L. (1992). From the Food and Drug Administration. *Journal of the American Medical Association*, **268**, 1828.

Nimelstein, S.H., Brody, S., Mcshane, D., and Holman, H.R. (1980). Mixed connective tissue disease: a subsequent evaluation of the original 25 patients. *Medicine (Baltimore)*, **59**, 239–48.

Nishikai, M. and Reichlin, M. (1980). Heterogeneity of precipitating antibodies in polymyositis and dermatomyositis: characterization of the Jo-1 antibody system. *Arthritis and Rheumatism*, **23**, 881–8.

Nitsche, A., Leiguarda, R.C., Maldonado Cocco, J.A., Lazaro, M.A., and Garcia Morteo, O. (1991). Neurological features in overlap syndrome. *Clinical Rheumatology*, **10**, 5–9.

Norman, D.A. and Fleischmann, R. (1978). Gastrointestinal systemic sclerosis in serologic mixed connective tissue disease. *Arthritis and Rheumatism*, **21**, 811–9.

O'Connell, D.J. and Bennett, R.M. (1977). Mixed connective tissue disease—clinical and radiological aspects of 20 cases. *British Journal of Radiology*, **50**, 620–5.

Oddis, C.V., Medsger, T.A., Jr, and Cooperstein, L.A. (1990). A subluxing arthropathy associated with the anti-Jo 1 antibody in polymyositis/dermatomyositis. *Arthritis and Rheumatism*, **33**, 1640–5.

Oetgen, W.J., Mutter, M.L., Lawless, O.J., and Davia, J.E. (1983). Cardiac abnormalities in mixed connective tissue disease. *Chest*, **83**, 185–8.

Ohtsuka, E., Nonaka, S., Shingu, M., Yasuda, M., and Nobunaga, M. (1992). Sjögren syndrome and mixed connective tissue disease. *Clinical and Experimental Rheumatology*, **10**, 339–44.

Okubo, M., Kurokawa, M., Ohto, H., Nishimaki, T., Nishioka, K., and Kasukawa, R. (1994). Clonotype analysis of peripheral blood T cells and autoantigen-reactive T cells from patients with mixed connective tissue disease. *Journal of Immunology*, **153**, 3784–90.

Pallis, M., Hopkinson, N., and Powell, R. (1991). Nailfold capillary density as a possible indicator of pulmonary capillary loss in systemic lupus erythematosus but not in mixed connective tissue disease. *Journal of Rheumatology*, **18**, 1532–6.

Panush, R.S., Edwards, L., Longley, S., and Webster, E. (1988). 'Rhupus' syndrome. *Archives of Internal Medicine*, **148**, 1633–6.

Pelkonen, P.M. *et al*. (1994). Incidence of systemic connective tissue diseases in children: a nationwide prospective study in Finland. *Journal of Rheumatology*, **21**, 2143–6.

Peller, J.S., Gabor, G.T., Porter, J.M., and Bennett, R.M. (1985). Angiographic findings in mixed connective tissue disease. Correlation with fingernail capillary photomicroscopy and digital photoplethysmography findings. *Arthritis and Rheumatism*, **28**, 768–74.

Piirainen, H.I., Helve, A.T., Tornroth, T., and Pettersson, T.E. (1989). Amyloidosis in mixed connective tissue disease. *Scandinavian Journal of Rheumatology*, **18**, 165–8.

Plotz, P.H., Dalakas, M., Leff, R.L., Love, L.A., Miller, F.W., and Cronin, M.E. (1989). Current concepts in the idiopathic inflammatory myopathies: polymyositis, dermatomyositis, and related disorders. *Annals of Internal Medicine*, **111**, 143–57.

Prakash, U.B.S., Luthra, H.S., and Divertie, M.B. (1985). Intrathoracic manifestations in mixed connective tissue disease. *Mayo Clinic Proceedings*, **60**, 813–21.

Pun, Y.L., Russell, D.M., Taggart, G.J., and Barraclough, D.R. (1991). Pneumatosis intestinalis and pneumoperitoneum complicating mixed connective tissue disease. *British Journal of Rheumatology*, **30**, 146–9.

Query, C.C. and Keene, J.D. (1987). A human autoimmune protein associated with U1 RNA contains a region of homology that is cross-reactive with retroviral p30 ^{gag} antigen. *Cell*, **51**, 211–20.

Rakovec, P., Kenda, M.F., Rozman, B., Zemva, A., and Cibic, B. (1982). Panconductional defect in mixed connective tissue disease. Association with Sjögren's syndrome. *Chest*, **81**, 257–9.

Ranki, A., Kurki, P., Riepponen, S., and Stephansson, E. (1992). Antibodies to retroviral proteins in autoimmune connective tissue disease. Relation to clinical manifestations and ribonucleoprotein autoantibodies. *Arthritis and Rheumatism*, **35**, 1483–91.

Rasmussen, E.K., Ullman, S., Horer-Madsen, M., Sorensen, S.F., and Halberg, P. (1987). Clinical implications of ribonucleoprotein antibody. *Archives of Dermatology*, **123**, 601–5.

Read, C.A., Clauw, D., Weir, C., Da-Silva, A.T., and Katz, P. (1992). Dyspnea and pulmonary function in the L-tryptophan-associated eosinophilia–myalgia syndrome. *Chest*, **101**, 1282–6.

Reichlin, M. and Arnett, F.C. (1984). Multiplicity of antibodies in myositis sera. *Arthritis and Rheumatism*, **27**, 1150–6.

Reichlin, M. and Mattioli, M. (1972). Correlation of a precipitin reaction to an RNA protein antigen and a low prevalence of nephritis in patients with systemic lupus erythermatosus. *New England Journal of Medicine*, **286**, 908–11.

Reichlin, M. *et al*. (1984). Antibodies to a nuclear/nucleolar antigen in patients with polymyositis-overlap syndrome. *Journal of Clinical Immunology*, **4**, 40–4.

Reimer, G. (1990). Autoantibodies against nuclear, nucleolar, and mithochondrial antigens in systemic sclerosis (scleroderma). *Rheumatic Disease Clinics of North America*, **16**, 169–83.

Reimer, G., Huschka, U., Keller, J., Kammerer, R., and Hornstein, O.P. (1983). Immunofluorescence studies in progressive systemic sclerosis (scleroderma) and mixed connective tissue disease. *British Journal of Dermatology*, **109**, 27–36.

Reynolds, T.B., Denison, E.K., Frankl, H.D., Lieberman, F.L., and Peters, R.L. (1971). Primary biliary cirrhosis with scleroderma, Raynaud's phenomena, and telangiectasis. *American Journal of Medicine*, **67**, 302–12.

Satoh, K., Imai, H., Yasuda, T., Wakui, H., Miura, A.B., and Nakamoto, Y. (1994). Sclerodermatous renal crisis in a patient with mixed connective tissue disease. *American Journal of Kidney Diseases*, **24**, 215–8.

Savouret, J.F., Chudwin, D.S., Wara, D.W., Ammann, A.J., Cowan, M.J., and Miller, W.L. (1983). Clinical and laboratory findings in childhood mixed connective tissue disease: presence of antibody to ribonucleoprotein containing the small nuclear ribonucleic acid U1. *Journal of Pediatrics*, **102**, 841–6.

Searles, R.P., Mladinich, E.K., and Messner, R.P. (1978). Isolated trigeminal sensory neruopathy: early manifestation of mixed connective tissue disease. *Neurology*, **28**, 1286–9.

Sharp, G.C. (1987). Diagnostic criteria for classification of MCTD. In *Mixed connective tissue diseases and anti-nuclear antibodies* (ed. R. Kasukawa and G.C. Sharp), pp. 23–32. Excerpta Medica, Amsterdam.

Sharp, G.C. *et al*. (1971). Association of autoantibodies to different nuclear antigens with clinical patterns of rheumatic disease and responsiveness to therapy. *Journal of Clinical Investigation*, **50**, 350–9.

Sharp, G.C., Irvin, W.S., Tan, E.M., Gould, R.G., and Holman, H.R. (1972). Mixed connective tissue disease—an apparently distinct rheumatic disease syndrome associated with a specific antibody to an extractable nuclear antigen (ENA). *American Journal of Medicine*, **52**, 148–59.

Sharp, G.C. *et al*. (1976). Association of antibodies to ribonucleoprotein and Sm antigens with mixed connective-tissue disease, systemic lupus erythematosus and other rheumatic diseases. *New England Journal of Medicine*, **295**, 1149–54.

Shulman, L.E. (1974). Diffuse fasciitis with hypergammaglobulinaemia and eosiniphilia in a new syndrome. *Journal of Rheumatology*, **1** (Suppl.), 46.

Silver, R.M. (1993). Eosinophilia–myalgia syndrome, toxic-oil syndrome, and diffuse fasciitis with eosinophilia. *Current Opinion in Rheumatology*, **5**, 802–8.

Silver, R.M., Sutherland, S.E., Carreira, P., and Heyes, M.P. (1992). Alterations in tryptophan metabolism in the toxic oil syndrome and in the eosinophilia–myalgia syndrome. *Journal of Rheumatology*, **19**, 69–73.

Silver, R.M., *et al*. (1994). A murine model of the eosinophilia-myalgia syndrome induced by 1,1'-ethylidenebis (L-tryptophan). *Journal of Clinical Investigation*, **93**, 1473–80.

Singsen, B.H., Bernstein, B.H., Kornreich, H.K., Koster King, K., Hanson, V., and Tan, E.M. (1977). Mixed connective tissue disease in childhood. A clinical and serologic survey. *Journal of Pediatrics*, **90**, 893–900.

Singsen, B.H., Swanson, V.L., Bernstein, B.H., Heuser, E.T., Hanson, V., and Landing, B.H. (1980). A histologic evaluation of mixed connective tissue disease in childhood. *American Journal of Medicine*, **68**, 710–17.

Springer, M.A., Bock, H.G., Philen, R.M., Hill, R.H., Jr., and Crawford, L.V. (1992). Eosinophilia–myalgia syndrome in a child with phenylketonuria. *Pediatrics*, **90**, 630–3.

Stuart, R.A. and Maddison, P.J. (1991). Atlantoaxial subluxation in a patient with mixed connective tissue disease. *Journal of Rheumatology*, **18**, 1617–20.

Sullivan, W.D. *et al*. (1984). A prospective evaluation emphasizing pulmonary involvement in patients with mixed connective tissue disease. *Medicine (Baltimore)*, **63**, 92–107.

Swygert, L.A., Maes, E.F., Sewell, L.E., Miller, L., Falk, H., and Kilbourne, E.M. (1990). Eosinophilia–myalgia syndrome. Results of national surveillance. *Journal of the American Medical Association*, **264**, 1698–703.

Swygert, L.A., Back, E.E., Auerbach, S.B., Sewell, L.E., and Falk, H. (1993). Eosinophilia–myalgia syndrome: mortality data from the US national surveillance system. *Journal of Rheumatology*, **20**, 1711–17.

Targoff, I.N. (1992). Autoantibodies in polymyositis. *Rheumatic Disease Clinics of North America*, **18**, 455–82.

Targoff, I.N. and Arnett, F.C. (1990). Clinical manifestatioins in patients with antibody to PL-12 antigen (alanyl-tRNA synthetase). *American Journal of Medicine*, **88**, 241–51.

Thiele, D.L. and Krejs, G.J. (1985). Secretory diarrhea in mixed connective tissue disease. *American Journal of Gastroenterology*, **80**, 107–10.

Tiddens, H.A.W.M. *et al*. (1993). Juvenile-onset mixed connective tissue disease: longitudinal follow-up. *Journal of Pediatrics*, **122**, 191–7.

Udoff, E.J., Genant, H.K., Kozin, F., and Ginsberg, M. (1977). Mixed connective tissue disease: the spectrum of radiographic manifestations. *Radiology*, **124**, 613–18.

Udvardy, M., Bodolay, E., Szegedi, G., Harsfalvi, J., Boda, Z., and Rak, K. (1991). Alterations of primary haemostasis in mixed connective tissue disease (MCTD). *Thrombosis Research*, **63**, 281–6

Ueda, Y. *et al*. (1991). Successful treatment of acute right cardiac failure due to pulmonary thromboembolism in mixed connective tissue disease. *Japanese Journal of Medicine*, **30**, 568–72.

Van Den Hoogen, F.H.J. *et al*. (1994). Long-term follow-up of 46 patients with anti-(U1)snRNP antibodies. *British Journal of Rheumatology*, **33**, 1117–20.

Varga, J., Griffin, R., Newman, J.H., and Jimenez, S.A. (1991). Eosinophilic fasciitis is clinically distinguishable from the eosinophilia–myalgia syndrome and is not associated with L-tryptophan use. *Journal of Rheumatology*, **18**, 259–63.

Varga, J., Jimenez, S.A., and Uitto, J. (1993). L-tryptophan and the eosinophilia–myalgia syndrome: current understanding of the etiology and pathogenesis. *Journal of Investigative Dermatology*, **100**, 97S–105S.

Wasicek, C.A., Reichlin, M., Montes, M., and Raghu, G. (1984). Polymyositis and interstitial lung disease in a patient with anti-Jo1 prototype. *American Journal of Medicine*, **76**, 538–44.

Watanabe-Fukunaga, R., Brannan, C.I., Coopeland, N.G., Jenkins, N.A., and Nagata, S. (1992). Lymphoproliferation disorder in mice explained by defects in Fas antigen that mediates apoptosis. *Nature*, **356**, 314–17.

Weiner -Kronish, J.P., Solinger, A.M., Warnock, M.L., Gurg, A., Ordonez, N., and Golden, J.A. (1981). Severe pulmonary involvement in mixed connective tissue disease. *American Review of Respiratory Diseases*, **124**, 499–503.

Weiss, T.D., Nelson, J.S., Woolsey, R.M., Zuckner, J., and Baldasare, A.R. (1978). Transverse myelitis in mixed connective tissue disease. *Arthritis and Rheumatism*, **21**, 982.

Whyte, J., Hough, D., Maddison, P.J., and McHugh, N.J. (1994). The association of primary biliary cirrhosis and systemic sclerosis is not accounted for by cross reactivity between mitochondiral and centromere antigens. *Journal of Autoimmunity*, **7**, 413–24.

5.13.1 Amyloidosis

Gunnar Husby

Definition of amyloid and amyloidosis

The term amyloidosis relates to a heterogeneous group of disorders (rather than a single disease entity) characterized by extracellular deposition of a proteinaceous, fibrillar material—amyloid—in various tissues and organs (Glenner 1980; Husby and Sletten 1986; Husby 1992). The unique amyloid fibril is the principal component of all amyloids irrespective of the clinical expression, tissues, or species involved, or whether they arise spontaneously or are induced in experimental animals.

Amyloid fibrils as seen in the electron microscope are rigid and non-branching with a diameter of 10 to 15 nm, of indefinite length, and consist of polypeptide chains arranged in a twisted b-pleated sheet (Glenner 1980). This specific structure of the fibril proteins determines the tinctorial and optical properties of amyloid, i.e. the affinity for Congo red and the typical green/yellow birefringence seen when amyloid stained with Congo red is viewed in a polarizing microscope (Cooper 1974). The low solubility and relative resistance to proteolytic digestion of the amyloid fibrils contribute to the irreversible and often progressive course of amyloidosis, in many cases leading to death within months or a few years of diagnosis (Husby 1985).

Benditt and Eriksen (1964) observed that in spite of the morphological similarities between amyloids in different clinical settings, amyloid is heterogeneous also with respect to the nature of the amyloid fibrils. Subsequent studies of amyloid extracted from different affected tissues (Pras *et al*. 1968) have revealed that they may be composed of a variety of different protein subunits (Husby and Sletten 1986).

In addition, an extrafibrillar protein called the amyloid-P component or protein **AP**, derived from a normal plasma glycoprotein, termed **SAP**, is invariably present in amyloid, regardless of the type of fibril protein (Pepys 1988). A carbohydrate moiety in the form of glycosaminoglycans and proteoglycans has also been demonstrated in all amyloids deposits so far examined (Kisilevski 1987; Magnus *et al*. 1989; Husby *et al*. 1994b; Stenstad *et al*. 1994).

Amyloid fibril proteins

The chief component of amyloid fibrils is a distinct protein subunit with a relatively small molecular mass—3000 to 30 000 Da in different fibril preparations (Glenner 1980; Husby and Sletten, 1986). Several, apparently non-related proteins can constitute this amyloid fibril subunit in different cases of amyloidosis, a common feature being the ability to assume the b-pleated sheet typical of amyloid (Glenner 1980). A steadily increasing number of such proteins (17 at the present time) have been characterized by their amino acid sequence (Table 1), and in many cases complementary DNA has also been established. The different amyloid proteins are often related to distinct clinical forms of amyloidosis. Indeed, many types of amyloid disease can now be defined and classified by structural analysis of the fibril proteins and/or the genes coding for them (Benditt and Eriksen 1971; Glenner 1980; Husby and Sletten 1986; Husby 1994).

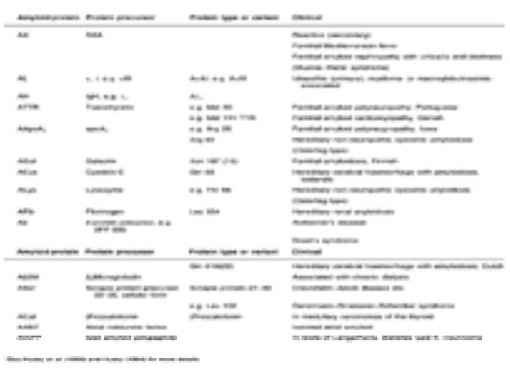


Table 1 Nomenclature and classification of human amyloid and amyloidosis [a]

Certain serum proteins are precursors of the different fibril proteins in the various systemic forms of amyloidosis (Glenner 1980; Husby and Sletten 1986). Two important types of fibril protein related to systemic amyloidosis are the amyloid-L (**AL**) and amyloid-A (**AA**) proteins (Table 1). Protein AL, which consists of homogeneous (monoclonal) immunoglobulin light chains or their aminoterminal fragments is seen in idiopathic and myeloma-associated amyloidosis (AL amyloidosis) (Glenner 1980; Husby 1983). Protein AA (Table 2), which is derived from the serum amyloid-A protein (Husebekk *et al*. 1985), is associated with reactive amyloidosis, familial Mediterranean fever, familial nephropathic amyloidosis with febrile urticaria—the Muckle–Wells syndrome (Muckle 1979; Linke *et al*. 1983)—and spontaneous or experimental amyloidosis in animals (Glenner 1980; Husby and Sletten 1986). These forms of amyloidosis could therefore also be called AA. Serum amyloid A (**SAA**), an apolipoprotein present on high-density lipoprotein in serum, is one of the most sensitive acute-phase proteins described (Marhaug *et al*. 1986). However, its function is largely unknown except for its association with AA amyloidosis and an active role in the metabolism and transport of cholesterol (Meek *et al*. 1994). An extensive review of the histology of SAA/AA proteins and genes and their relation to AA amyloidosis has recently been published (Husby *et al*. 1994a).

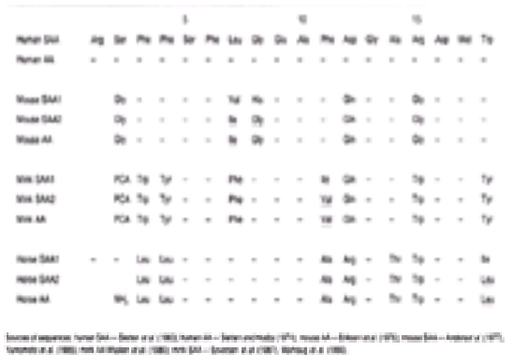


Table 2 Partial aminoterminal amino acid sequence of serum amyloid-A (SAA) and amyloid-A (AA) proteins from man, mouse, mink, and horse, numbered according to human SAA. Alignment is done to maximize homology. Only the amino acid residues that differ from those of human SAA and AA are shown for mouse, mink, and horse. Those amino acid residues that specify amyloidogenic SAA molecules are underlined

Different genetic variants of transthyretin (Table 1), previously termed prealbumin (Husby *et al*. 1990), are associated with the autosomal-dominant familial amyloid polyneuropathies and cardiomyopathies (Nordlie *et al*. 1988; Benson and Wallace 1989). Normal transthyretin makes up systemic senile amyloid (Westermark *et al*. 1986).

Three other serum proteins are also associated with hereditary amyloidosis (Table 1). A variant of apolipoprotein **AI** represents the fibril protein in familial amyloid polyneuropathy from Iowa (Nichols *et al*. 1988), whereas a variant gelsolin, an actin-modulating protein, plays the same role in familial amyloid polyneuropathy of Finnish origin (Maury 1990a). Recently a molecular variant of fibrinogen (Uemichi *et al*. 1994) was found to be associated with hereditary renal amyloidosis, the first form of inherited amyloidosis to be described by Ostertag (1932). A variant of the bacteriolytic enzyme, lysozyme (a product of macrophages and polymorphs normally present in various body secretions), is the amyloid fibril protein found in cases of systemic amyloidosis with predominant renal effects (Pepys *et al*. 1993).

A variant of cystatin C, an enzyme inhibitor, is the fibril protein precursor in inherited cerebral haemorrhage with amyloidosis, Icelandic type (Ghiso *et al*. 1986). The cerebral vessels of members of a Dutch family with a similar clinical picture are affected by amyloid made up by the amyloid b-protein, a 42-amino acid fragment of a variant b-protein precursor (Levy *et al*. 1990). Amyloid b-protein derived from normal b-protein precursor (a cell membrane protein) makes up the cerebral amyloid in Alzheimer's disease and Down's syndrome (Glenner and Murphy 1989).

b_2-Microglobulin, which represents a part of class I histocompatibility molecules, makes up the fibrils in amyloidosis associated with long-term dialysis treatment for renal failure (Gejyo *et al*. 1985). b_2-Microglobulin-derived amyloid affects mainly structures of the locomotor system, like the carpal tunnel, joints, and bone, but systemic distribution is also seen (Kay 1993).

Several hormones are known to form amyloid. The first to be demonstrated was a calcitonin-like molecule in amyloid deposited locally in medullary carcinomas of the thyroid and their metastases (Sletten *et al*. 1976). Atrial natriuretic factor makes up amyloid in isolated atrial amyloidosis of the heart (Johansson *et al*. 1987), and islet amyloid polypeptide plays the same role in amyloid localized to the islets of Langerhans in association with diabetes type II and insulinomas (Westermark *et al*. 1986).

It is thus clear that amyloidosis is heterogeneous not only with respect to clinical expression and protein composition, but also to the functional aspect of the various amyloid precursors.

Pathogenesis

The mechanisms by which the various soluble precursor proteins are converted to insoluble protein aggregates with the unique morphology of amyloid fibrils, and are then deposited in the different target tissues, are not clarified. Also, the predilection of certain amyloid fibrils for particular tissues or organs, often distant from the cells actually producing the fibril protein precursor, remains a mystery.

It is generally accepted that the aetiology and pathogenesis of amyloidosis are multifactorial. Furthermore, the relative importance of the various amyloid-promoting factors may differ in the different types of amyloidosis. Best studied are the pathogenetic mechanisms in the reactive AA type of amyloidosis, as this type is readily induced in experimental animals, and spontaneous AA amyloidosis has been demonstrated in a variety of warm-blooded species, ranging from birds to primates (Husby 1989; Husby 1992).

Amyloidosis may be caused by amyloidogenic precursor protein present in excess amounts, either as a consequence of increased production or decreased clearance. AA amyloid is prevalent in those of the chronic inflammatory diseases that are known to stimulate the acute phase response and thereby production of serum amyloid-A protein (e.g. rheumatoid arthritis), but not in those that are not (e.g. scleroderma). However, AA amyloidosis affects only a minority of patients with rheumatoid arthritis, although most such patients have raised serum amyloid A (Husby 1992). Additional pathogenetic factors are obviously required. Serum amyloid A is polymorphic, which means that more than one gene codes for the protein. There is convincing experimental evidence that serum amyloid-A molecules (Table 2) genetically determined to have particular amyloidogenic structures are important in AA amyloidosis (Meek *et al*. 1986; Husby *et al*. 1994b). The same is probably the case in the inherited amyloidoses related to genetic variants of transthyretin, apo-AI, gelsolin, and b-protein precursor (Table 1).

In most instances, the different amyloid proteins comprise a smaller or larger fragment of their precursors, pointing to an incomplete degradation (Glenner 1980). Defective proteolysis may therefore be implicated in amyloidosis (Fuks and Zucker-Franklin 1985; Husby *et al*. 1994a; Husby *et al*. 1994b).

Amyloid-enhancing factor, a poorly defined material probably consisting of both protein and carbohydrate, is produced by reticuloendothelial cells during persistent inflammation (Hol *et al*. 1986; Kisilevski and Young 1994). This factor precedes the occurrence of amyloid in affected organs in experimental amyloidosis (Kisilevski and Young 1994) and has been shown to shorten the induction time of experimental amyloidosis from weeks to between 24 and 48 h.

Glycosaminoclycans occur in the tissues in close temporal and morphological relationship to amyloid deposition (Snow *et al*. 1987; Stenstad *et al*. 1994) and have been proposed to account for the carbohydrate moiety in amyloid-enhancing factor (Snow *et al*. 1987). The large negative charge of the glycosaminoglycans indicates an effect on the folding of precursor proteins to form fibrils (Snow *et al*. 1987). Glycosaminoglycans may also protect amyloid proteins or their precursors against proteolysis and stabilize the structure of the fibrils in the extracellular milieu. (Husby *et al*. 1994b).

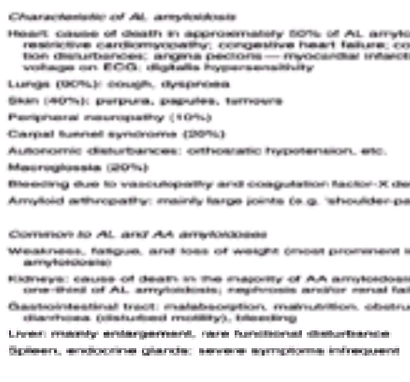
Protein AP, of which SAP is the precursor, is also invariably present in amyloid deposits as an extrafibrillar constituent (Pepys 1988). AP has also been thought to protect amyloid proteins against proteolysis (Husby *et al*. 1994b). A DNA polymorphic site, 5' to the SAP gene, was significantly associated with AA amyloidosis in juvenile rheumatoid arthritis, suggesting an active role in amyloidogenesis (Woo *et al*. 1987). Experiments in our laboratory (Husby *et al*. 1994b) strongly suggest a close interaction between the fibril proteins, glycosaminoglycans, and AP *in vivo*, often mediated by calcium-dependent bindings. Chemical manipulation of this may prove to have therapeutical implications in amyloidosis.

Clinical amyloidosis syndromes

A large number of partly overlapping amyloidosis syndromes have been described, and they have been difficult to classify by their clinical expression, although many attempts at this have been made during the last 150 years (Glenner 1980; Kyle 1982; Husby and Sletten 1986; Pepys 1988). The classification of amyloid and amyloidosis based on the nature of the fibril proteins (Table 1) is therefore of great help for clinical practice as well as scientific work (Husby 1994). Under the following headings, the most important amyloidosis syndromes are described in relation to their protein correlates.

Idiopathic and myeloma-associated AL amyloidosis

The mean age at diagnosis of AL amyloidosis is approximately 60 years. The organ distribution of amyloid is similar in idiopathic and myeloma-associated amyloidoses (Kyle 1982; Kyle *et al*. 1986). The most severe clinical features of AL amyloidosis are caused by accumulation of amyloid in the heart and kidneys (Kyle 1982; Janssen *et al*. 1986; Kyle *et al*. 1986; Stone 1990). Interstitial and vascular amyloidosis of the heart leads to congestive heart failure, conduction disturbances or angina pectoris, sometimes myocardial infarction, and causes about one-half of the deaths associated with this condition (Table 3). Decreased voltage on ECG is common. The patients are hypersensitive to digitalis, which, when given, may cause fatal arrhythmias. The extent of cardiac involvement is the most important predictor for survival in AL amyloidosis (Kyle *et al*. 1986).


Characteristic of AL amyloidosis
Heart: cause of death in approximately 50% of AL amyloidosis; restrictive cardiomyopathy; congestive heart failure; conduction disturbances; angina pectoris — myocardial infarction; low voltage on ECG; digitalis hypersensitivity
Lungs (90%): cough, dyspnoea
Skin (40%): purpura, papules, tumours
Peripheral neuropathy (10%)
Carpal tunnel syndrome (20%)
Autonomic disturbances: orthostatic hypotension, etc.
Macroglossia (20%)
Bleeding due to vasculopathy and coagulation factor-X deficiency
Amyloid arthropathy: mainly large joints (e.g. 'shoulder-pad sign')

Common to AL and AA amyloidoses
Weakness, fatigue, and loss of weight (most prominent in AL amyloidosis)
Kidneys: cause of death in the majority of AA amyloidosis and in one-third of AL amyloidosis; nephrosis and/or renal failure
Gastrointestinal tract: malabsorption, malnutrition, obstruction, diarrhoea (disturbed motility), bleeding
Liver: mostly enlargement, rare functional disturbances
Spleen, endocrine glands: severe symptoms infrequent

Table 3 Clinical features of AL and AA amyloidoses

Glomerular and vascular amyloidosis of the kidneys with consequent nephrosis and/or renal failure causes one-third of the deaths. Pulmonary amyloidosis with cough and dyspnoea is common. Amyloidosis of the skin with purpura, amyloid papules, or tumours is seen in nearly one-half of the patients. Some patients develop peripheral neuropathy (10 per cent) and carpal tunnel syndrome (20 per cent) in addition to autonomic disturbances, for example orthostatic hypotension (Kyle 1982).

Characteristic, but not so common features of AL amyloidosis are macroglossia, seen in 20 per cent of the patients, and amyloid arthropathy affecting mainly large joints. The 'shoulder pad sign' due to amyloid infiltration of the shoulder joint (Fig. 1) is almost pathognomonic. Purpura is a frequent and characteristic finding (Fig. 1). The coagulation factor-X deficiency seen in AL amyloidosis appears to be due to the high affinity of AL-type fibrils for factor X, which may therefore be trapped in the amyloid substance (Furie *et al*. 1981). Together with increased fibrinolysis and amyloid infiltration of the blood vessels, the factor-X deficiency may lead to severe haemorrhages (Kyle 1982).

Fig. 1 (a) Amyloidosis with purpura around the eyelids, lips, and anterior thorax. Macroglossia with inability to close the mouth and amyloid arthropathy with bilateral shoulder pad sign are also present. (Reprinted from the Revised Clinical Collection on the Rheumatic Diseases 1981; used by permission of the American College of Rheumatology). (b) A 14-year-old girl suffering from severe, seropositive juvenile chronic arthritis with onset at 5 years of age, complicated by AA amyloidosis verified by rectal biopsy 6 years later. The clinical picture is dominated by malabsorption and gastrointestinal bleeds with severe weight loss, anaemia, renal failure, hepatosplenomegaly, hormonal disturbances with dwarfism/infantilism, and severe polyarthritis. These manifestations are caused by the underlying arthritic disorder, its treatment with corticosterioids and cytotoxic drugs, and the systemic AA amyloidosis.

Gastrointestinal amyloidosis may be associated with malabsorption, malnutrition, obstruction, diarrhoea (disturbed motility), and bleeding. Amyloidosis of the liver causes enlargement rather than functional disturbances. Amyloid is often present in the spleen, lymph nodes, and endocrine glands but severe symptoms are infrequent.

Localized, so-called tumour-forming, AL has been reported to be present in various organs, of which the respiratory and genitourinary tracts and the skin are most frequent (Husby 1992).

The patients with idiopathic AL amyloidosis regularly have increased numbers of plasma cells in the bone marrow; M components in serum and Bence-Jones proteinuria are also frequent findings, illustrating that this disorder belongs to the immunocyte dyscrasias with the same basic pathogenetic mechanisms as myelomatosis. The major difference between them is that the osteolytic lesions of myelomatosis are not present in idiopathic AL amyloidosis. Another difference is that the k to l ratio, which is approximately 2:1 in myelomatosis, is reversed (1:2) in AL amyloidosis, thus reflecting the greater amyloidogenic potential of l-light chains (Husby and Sletten 1986).

Reactive AA amyloidosis

Reactive AA amyloidosis is mainly associated with long-standing infectious or non-infectious inflammation, and less frequently with cancer, mainly renal cell carcinoma or Hodgkin's disease (Husby 1992). In areas where the incidence of chronic infections like tuberculosis and leprosy has declined, reactive amyloidosis is mostly caused by chronic rheumatic diseases, mainly adult rheumatoid arthritis and juvenile chronic arthritis, and ankylosing spondylitis, with frequencies in living patients of about 3 to 10 per cent, and by occasional cases of Reiter's disease and psoriatic arthropathy (Husby 1985). AA amyloidosis is also seen in some cases of Crohn's disease. Interestingly, this form of amyloidosis is extremely rare in systemic autoimmune rheumatic diseases like systemic lupus erythematosus, dermatomyositis, systemic sclerosis, and Sjögren's syndrome, and these diseases are associated with minimal acute-phase protein responses and hence low concentrations of serum amyloid A.

A marked geographic difference in the prevalence of AA amyloidosis in juvenile chronic arthritis, high (5 to 10 per cent) in European countries and low (0.1 per cent) in the United States, has been attributed to a more frequent occurrence of urinary tract infections caused by *Escherichia coli* in Europe (Filipowicz-Sosnowska *et al*. 1978). *E. coli* endotoxin is a highly potent inducer of the acute phase response as well as of experimental amyloidosis.

Arthritic patients with highly systemic disease are more prone to develop amyloidosis than those with milder disease; however, it is hard to predict individual cases at risk (Husby 1985). HLA studies have failed to disclosed markers for reactive amyloidosis.

Amyloid of the AA type has a tendency to localize in small vessels and parenchymal organs (Table 3). Renal disease (Fig. 2), often with the nephrotic syndrome and/or renal failure, is the chief manifestation and the major cause of death. In a series of 189 patients with rheumatoid arthritis and AA amyloidosis reported by

Wegelius *et al*. (1980) only seven (3.7 per cent) lacked demonstrable clinical signs of renal involvement, and all patients with juvenile chronic arthritis in a series of 51 who developed amyloidosis had proteinuria at the time of diagnosis (Schnitzer and Ansell 1977). Proteinuria in rheumatoid arthritis should always be associated with possible amyloidosis until its cause is eventually disclosed. Haematuria occurs, but less often. Hypertension is uncommon in adult-onset rheumatoid arthritis with amyloidosis, but is found in about half of the patients with amyloidosis associated with juvenile chronic arthritis (Schnitzer and Ansell 1977; Woo 1994). Renal vein thrombosis is frequently found at autopsy.

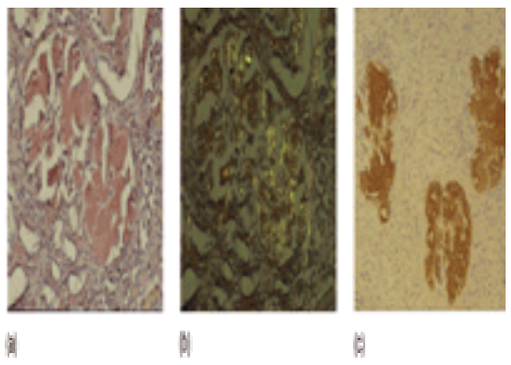


Fig. 2 Renal AA amyloidosis: (a)light microscopy of Congo red-stained glomerular amyloid (original magnification, × 160); (b) polarization microscopy of the same view field revealing the typical green/yellow birefringence of amyloid (× 160); (c) glomerular AA amyloid stained with peroxidase-labelled antiserum to protein AA and examined by light microscopy—no staining was obtained with antisera to other amyloid proteins (× 160).

Infiltration of blood vessels by amyloid, particularly in the gastrointestinal tract, can result in severe bleeding, sometimes life threatening, and malabsorption. AA amyloidosis of the liver is frequent and causes organ enlargement rather than functional disturbance. This type of amyloid may also affect the spleen, endocrine glands, and heart, but without causing severe symptoms. Interestingly, amyloid arthropathy and carpal tunnel syndrome are not features of AA amyloidosis, but are characteristic of both AL and b_2-microglobulin-associated amyloidoses.

No routine laboratory test (except biopsy) can distinguish between arthritic patients with and without AA amyloidosis. The high affinity of protein SAP for amyloid fibrils has recently been utilized in the diagnosis of AS and other systemic amyloidoses for research purposes. Radiolabelled SAP injected intravenously produces scintigraphic images quite accurately demonstrating the distrbution of amyloid *in vivo* (Hawkins 1994). So far, this procedure has been used for research purposes.

Familial Mediterranean fever with AA amyloidosis (see Chapter 5.13.2)

Amyloidosis associated with familial Mediterranean fever is the only form of systemic amyloidosis known to be inherited largely as a recessive trait (Pras 1986). Like other forms of AA amyloidosis, nephropathy is the most important feature and a significant cause of death. Familial Mediterranean fever itself is characterized by attacks of fever, peritonitis, pleuritis, and synovitis with onset during childhood, affecting mainly Sephardic Jews, Anatolian Turks, and Armenians with origin in the Mediterranean area. The prevalence as well as the disease course of amyloidosis in Mediterranean fever are highly heterogeneous in the different ethnic groups (Pras 1986). It appears that the febrile attacks and amyloidosis are inherited independently, and marked heterogeneity with regard to the serum amyloid-A genes has also been observed among patients with this fever syndrome (Sack 1988). A substitution of threonine for phenylalanine at position 69 of amyloid protein A involving all three bases of the codon has been observed in one case of amyloidosis in familial Mediterranean fever (Levin *et al*. 1972), but not confirmed in others.

The causal gene for familial Mediterranean fever has been located to the short arm of chromosome 16 (Pras *et al*. 1992), apparently not related to the human SAA gene family on chromosome 11 (Husby *et al*. 1990), and there is no link between this gene locus and that coding for human SAA present on chromosome 11 (Husby *et al*. 1994b)

Inherited autosomal-dominant amyloidosis

Point mutations causing single amino-acid substitutions in various amyloid protein precursors are associated with a heterogeneous group of familial amyloidoses inherited as autosomal dominants (Table 1) (Benson and Wallace 1989; Glenner and Murphy 1989). The majority of cases are related to different genetic variants of transthyretin. Neuropathy, cardiopathy, nephropathy, and vitreous opacities are the most important clinical problems and occur in a variety of combinations, but also as more or less single clinical entities. Amyloid made up by genetic variants of other proteins (Table 1) may show similar clinical characteristics in addition to causing cerebral amyloid angiopathy and lattice corneal dystrophy. Some examples of inherited, autosomal-dominant amyloidoses, the geographic origin, and protein correlations are listed next (the syndromes are described in more detail by Glenner and Murphy (1989) and Benson (1991)):

1. familial amyloid polyneuropathy related to transthyretin–methionine 30—Portuguese, Swedish, Japanese, and others (Andrade 1952);
2. familial amyloid polyneuropathy, transthyretin–serine 84—Indiana/Swiss, Maryland/German (Benson and Wallace 1989);
3. familial amyloid polyneuropathy, transthyretin–histidine 58—Maryland/German (Nichols *et al*. 1989);
4. familial amyloid cardiomyopathy and neuropathy, transthyretin–alanine 60—Appalachian Indians, (Benson and Wallace 1989);
5. familial amyloid cardiomyopathy (Fig. 3), transthyretin–methionine 111—Danish (Frederiksen *et al*. 1962; Nordlie *et al*. 1988; Nordvåg *et al*. 1992);

Fig. 3 Cardiac amyloidosis. Polarization microscopy of amyloid-laden myocardium stained with Congo red from a patient with transthyretin–methionine 111 familial amyloid cardiomyopathy of Danish origin; massive infiltration of green birefringent amyloid displacing muscle cells is seen (× 250).

6. familial amyloid polyneuropathy, apolipoprotein AI–arginine 26—Iowa (Nichols *et al*. 1988);
7. familial amyloidosis with lattice corneal dystrophy, gelsoline–aspartic acid 187—Finnish (Meretoja 1969; Maury 1990a);
8. hereditary cerebral haemorrhage with amyloidosis, cystatin C–glutamic acid 58—Icelandic (Ghiso *et al*. 1986) and amyloid b-protein–glutamic acid 618—Dutch (Levy *et al*. 1990);
9. the Muckle–Well's syndrome (nephropathy, nerve deafness, and urticaria) related to amyloid protein A (Muckle 1979; Linke *et al*. 1983);
10. hereditary, non-neuropathic systemic amyloidosis with predominant renal involvement (Ostertag 1932) is related to various substitutions in apolipoprotein A1 (Soutar *et al*. 1992), lysozyme (Pepys *et al*. 1993), or the a-chain of fibrinogen (Table 1)—south or north America, the latter being of Irish or Scandinavian descent (Uemichi *et al*. 1994).

Alzheimer's disease

The presenile dementia described by Alzheimer is probably the most common form of amyloidosis (Glenner and Murphy 1989). Amyloid is deposited extracellularly as

Alzheimer's plaques and in the wall of cerebral vessels; even the intraneuronal neurofibrillary tangles are made up by paired helical filaments with the typical b-pleated sheet characteristic of amyloid. The amyloid b-protein in Alzheimer's disease derives from the b-protein precursor whose gene resides on chromosome 21.

The same manifestations of cerebral amyloid composed of the amyloid b-protein is seen in people with Down's syndrome over 40 years of age (Glenner and Murphy 1989). Down's syndrome is associated with trisomy of chromosome 21, and it is conceivable that this leads to an overproduction of b-protein precursor. Altered processing of the amyloid b-protein precursor has been proposed as a pathogenetic factor, and genetic variants of the protein are associated with familial occurrence of Alzheimer's disease. Another risk factor is the type 4 allele of apolipoprotein E (Corder *et al*. 1993). The same type of amyloid is occasionally seen in senile dementia (Glenner and Murphy 1989).

Organ manifestations

Heart

Cardiac amyloidosis manifesting mainly as cardiomyopathy, but also with conduction disturbances and coronary heart disease (Table 3), is a particularly severe manifestation of AL (Kyle 1982; Husby 1983; Janssen *et al*. 1986) and certain forms of inherited amyloidosis (Benson and Wallace 1989). Among the last, the Danish transthyretin–methionine 111 (Nordlie *et al*. 1988) and the Appalachian transthyretin–alanine 60 (Benson and Wallace 1989) are the most important forms (Fig. 3). Cardiac amyloidosis is also a major manifestation of senile systemic amyloidosis related to normal transthyretin (Westermark *et al*. 1990), and senile isolated atrial amyloidosis related to the polypeptide-hormone atrial natriuretic factor (Johansson *et al*. 1987). In general the senile cardiac amyloidoses are benign. AA amyloidosis of the heart localizes mainly to the vasculature without causing major problems.

The respiratory tract

Amyloid may be deposited in any part of the respiratory system, from the nasal cavity to the pulmonary parenchyma and hilar glands. The clinical consequences are most severe in AL amyloidosis (Kyle 1982). Localized (nodular) AL amyloidosis is quite frequent and has a much more favourable prognosis, but may sometimes require surgical removal (Glenner 1980; Husby 1983).

Kidney and urinary tract

The most severe manifestation is renal amyloidosis, the major cause of death in AA amyloidosis (Glenner 1980; Husby 1985; Woo 1994), and frequently also in AL and hereditary amyloidoses related to variants of transthyretin and apo-AI (Glenner 1980; Kyle 1982; Glenner and Murphy 1989). Affected patients present with proteinuria and/or nephrotic syndrome, and frequently develop fatal renal failure. Hypertension and haematuria are less common. Renal biopsy reveals deposition of amyloid in glomeruli, in addition to its peritubular, vascular, and interstitial localization (Fig. 2).

The gastrointestinal tract

The entire gastrointestinal tract is a common target of amyloid deposition, making it an accepted site for diagnosis biopsy (Fig. 4). Intestinal amyloid causes motility disturbances with diarrhoea or constipation, malabsorption, bleeds, and perforation (Table 3) (Glenner 1980; Kyle 1982; Husby 1992). Nutritional disturbance due to macroglossia in AL amyloidosis (Fig. 1) or altered intestinal motility in the inherited neuropathies are quite common causes of death. Fatal gastrointestinal bleeds due to AA amyloidosis has been observed in several patients with rheumatic disease (Husby 1985).

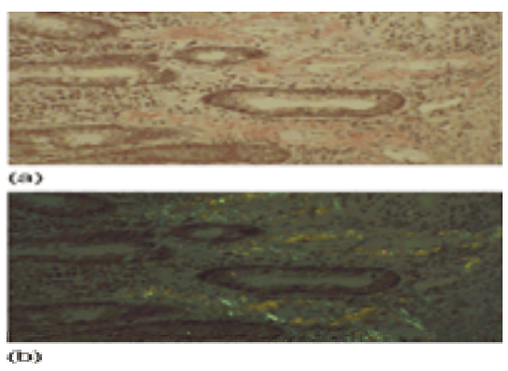


Fig. 4 Light (a) and polarization (b) microscopy of a Congo red-stained section from a rectal biopsy containing vascular and interstitial amyloid in the submucosal layer (× 160). (Photomicrographs kindly provided by Dr Michael Kearney, Tromsø, Norway.)

The endocrine system

Fibrils made up by hormone-related or -derived polypeptides are seen locally in some endocrine tumours. Calcitonin-related amyloid in medullary carcinoma of the thyroid and islet amyloid polypeptide in diabetes type II and insulinomas are best known (Table 1). Endocrine glands are also affected by systemic AA and AL amyloidoses (Husby 1992).

Skin

Amyloid deposits are generally present in subcutaneous fat in both AL and AA, as well as inherited systemic amyloidoses, and thin-needle aspirates of abdominal fat (Fig. 5) have become a useful material for histological diagnosis (Westermark and Stenkvist 1973). In addition, cutaneous amyloid is a common feature of AL amyloidosis in the form of papules, nodules, or purpura (Kyle 1982). The skin is also the site of localized amyloid, e.g. lichen amyloidosus (Black 1976).

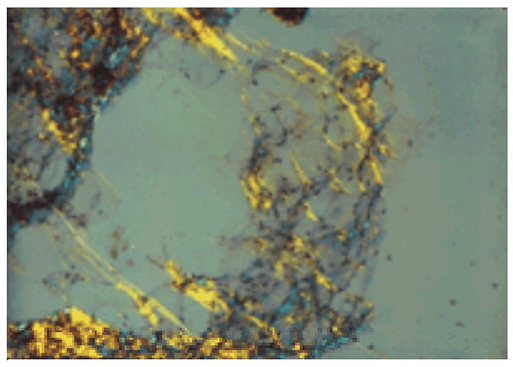


Fig. 5 Polarization microscopy of a subcutaneous, abdominal fat stained with Congo red showing marked deposition of amyloid compatible with the diagnosis of systemic amyloidosis. (Photomicrograph kindly provided by Professor Per Westermark, Linköping, Sweden.)

The locomotor system

The deposition of amyloid in the locomotor system is perhaps most common in b_2-microglobulin-related amyloidosis affecting patients on long-term dialysis for renal failure, who develop carpal tunnel syndrome, arthropathy, and cystic bone lesions, sometimes with pathological fractures, particularly when dialysed for more than 8 years. Ordinary haemodialysis is not able to remove b_2-microglobulin from plasma. Patients with this form of amyloidosis reportedly have abnormally glycosylated b_2-microglobulin capable of inducing inflammation with expression of cytokines and collagenase, which may contribute to destruction of bone and connective tissue (Miyata *et al*. 1994). Although b_2-microglobulin related amyloidosis is undoubtedly of systemic nature, structures of the locomotor system are of marked predilection for amyloid deposition in these patients (Maury 1990b; Kay 1993).

Characteristic of AL amyloidosis, though not so frequent, are the carpal tunnel syndrome (20 per cent), peripheral neuropathy (10 per cent), arthropathy (Fig. 1), or myopathy (less than 5 per cent) (Kyle 1982; Husby 1983). Another infrequent complaint is jaw claudication, which should be considered in the differential diagnosis of temporal arthritis/polymyalgia rheumatica (Gertz *et al*. 1986).

Erosive arthritis occurs in some patients with hereditary amyloidosis, i.e. of the Indiana and Iowa kindreds (Benson and Wallace 1989).

As amyloid arthropathy often has an inflammatory appearance, it may mistakenly be diagnosed as a rheumatic disorder. However, the joint fluid is generally non-inflammatory, sometimes with amyloid-containing synovial debris (Husby 1983).

Localized 'microdeposits' of amyloid in structures of the locomotor system occur with increasing frequency in aged people, mostly in menisci of the knee, joint capsules, and intervertebral discs, sometimes associated with the deposition of calcium pyrophosphate or with osteoarthrosis. The clinical significance of such amyloids, which possibly originate from local tissue proteins, is not clear (Husby and Sletten 1986).

Inclusion body myositis is a subset of chronic polymyositis characterized by muscle fibres with inclusion bodies containing amyloid fibrils. Very interestingly, the same b_2-amyloid protein as that seen in Alzheimer's disease and its precursor have recently been demonstrated to be constituents of these intramuscular amyloid deposits. This shows that b-amyloid can accumulate also outside the central nervous system. (For review see Askanas and Engel 1993)

Diagnosis

The diagnosis of amyloidosis is based on the demonstration of tissue deposits of amyloid. An absolute prerequisite for diagnosis is therefore that the clinician suspects amyloidosis in relevant clinical states. It is a regrettable fact, however, that the diagnosis of amyloidosis is far too often missed during life and is not evident until autopsy. The clinician should realize that AL amyloidosis may occur without an underlying predisposing disorder or in association with monoclonal gammopathies; that rheumatic disorders, long-standing or recurrent infections, and certain malignancies predispose for reactive, AA amyloidosis; and that a family history of amyloidosis may point to inherited disease. Unexplained manifestations that should alert the clinician to consider amyloidosis are: loss of weight (Fig. 1(a) and (b)), fatigue, proteinuria with or without renal failure, restrictive cardiomyopathy, gastrointestinal or respiratory problems, hepatomegaly, cutaneous conditions including purpura, and neuropathy/carpal tunnel syndrome.

It is established that rectal tissue (Fig. 4) which includes the submucosal vessels is a highly representative biopsy in systemic amyloidosis (Husby 1985). There may be bleeding after this procedure, but severe blood loss is infrequent.

In recent years, needle aspirates of abdominal subcutaneous fat (Fig. 5), which are less invasive than rectal biopsy, have increasingly been used for diagnosis (Westermark and Stenkvist 1973). Affected or suspected organs or tissues in the actual case under examination, such as kidneys (Fig. 2), liver, spleen, peripheral nerves, or skin, may be appropriate sites of biopsy but the risk of complications must be considered.

The amyloid carpal tunnel syndrome may be the initial finding in systemic amyloidoses. All tissues removed at surgery for carpal tunnel syndrome should therefore be examined for the presence of amyloid.

Examination of alkaline Congo red-stained tissue sections or aspirated fat smears in a polarizing microscope revealing the apple-green/yellow birefringence characteristic of amyloid deposits is the histochemical method of choice (see Fig. 2,Fig. 3,Fig. 4, and Fig. 5) (Glenner 1980; Kyle 1982; Hawkins 1994). Immunohistochemical methods with specific antisera to classify the various amyloid fibril proteins (Fig. 2) are increasingly used, but are not available for routine diagnostic purposes. Electron microscopy may confirm the diagnosis.

The strong calcium-dependent affinity of protein AP/SAP for amyloid fibrils of any protein type can be utilized diagnostically (Hawkins 1994). Radiolabelled SAP injected intravenously will rapidly and specifically localize to the amyloid deposits to yield high resolution scintigraphic images. This technique is currently at the experimental stage in humans, but is a highly promising diagnostic measure in clinical practice.

Other laboratory tests include analysis of DNA or its protein product to detect genetic variants of proteins known to make up amyloid in the hereditary amyloidoses (Benson and Wallace 1989; Nordvåg *et al*. 1992). Indeed, the use of restriction fragment-length polymorphism combined with *in vitro* amplification of genes using polymerase chain reaction is an extremely sensitive tool for detecting gene carriers and thereby individuals at risk—even before birth (Nichols *et al*. 1989; Nordvåg *et al*. 1992).

Radionuclide imaging using calcium-seeking isotopes, for example $^{99}Tc^{m}$, echocardiography, bone marrow examination, and demonstration of monoclonal immunoglobulins in serum or urine are examples of laboratory methods used in the clinical work-up of patients with amyloidosis (Hazenberg and van Rijswijk 1994).

Therapeutic aspects

In general, amyloidosis is a progressive disease for which there is no cure. As the aetiology and pathogenesis are multifactorial, there are problems with therapy. Prospective, controlled studies of intervention are sparse. The marked prognostic heterogeneity of patients with hereditary amyloidosis makes genetic counselling difficult (Sequeiros and Saraiva 1987). With these problems in mind, the following therapeutic approaches are suggested:

1. reduce the availability of precursor protein to prevent or slow down amyloid formation;
2. attempts at dissolving amyloid deposits *in vivo*;
3. treatment directed towards affected organs.

These are now considered in more detail.

Reduction of amyloid precursor

In reactive AA amyloidosis the synthesis of the precursor serum amyloid A can be reduced by turning down the stimulated hepatic production of acute phase proteins (Husby 1992). This is achieved by effective treatment of the underlying inflammatory or neoplastic disorder, whenever possible. In the face of the poor prognosis of amyloidosis, rather drastic therapeutic intervention may be considered. Cytotoxic drugs have convincingly been shown to improve the prognosis of amyloidosis associated with both adult rheumatoid arthritis (Ahlmen *et al*. 1987; Berglund *et al*. 1993) and juvenile chronic arthritis (Fig. 6) (Woo 1994). Any use of cytotoxic drugs must, however, be weighed against the risk of potentially hazardous adverse effects such as leukaemia. Treatment of familial Mediterranean fever with colchicine prevents the occurrence of febrile attacks as well as AA amyloidosis, and is also effective in the treatment of established amyloidosis in such patients (Zemer *et al*. 1986). A recent review of the literature (Livneh *et al*. 1993) also concluded that colchicine may be added to any therapeutic regimen of AL or AA amyloidosis.

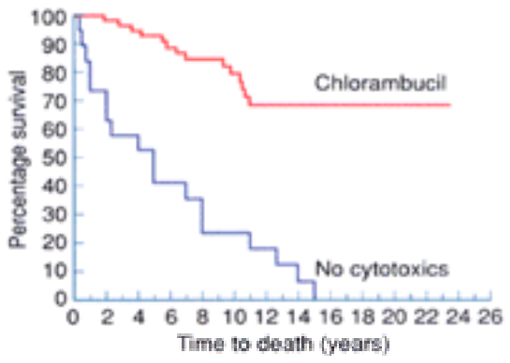


Fig. 6 Survival of juvenile chronic arthritis with amyloidosis with and without treatment using the cytotoxic drug chlorambucil. Although historical controls are used, the markedly improved survival strongly indicates that this can be attributed to treatment. (From Woo (1990) with kind permission from author, editors, and publisher.)

In controlled trials, treatment with the cytotoxic drug melphalan in combination with corticosteroids has been shown (Kyle *et al*. 1990) to improve the symptoms and signs, and possibly also survival, in AL amyloidosis. Studies using historical controls shows that colchicine may exert some additional effect on AL amyloidosis when combined with melphalan and prednisolone (Cohen *et al*. 1987; Kyle *et al*. 1990; Gertz *et al*. 1991). These treatments are thought to work by reducing the number of cells producing the monoclonal immunoglobulin light-chain precursor of amyloid, or to hamper their protein production or secretion. High-dose cytotoxic drugs in combination with stem cell transfusion appears promising (Majolino *et al*. 1993).

A diet where the only source of fat was fish oil high in w-3 polyunsaturated fatty acids has been shown to retard the progression of azocasein-induced AA amyloidoses in mice (Cathcart *et al*. 1987). As fish-oil diets may also suppress the activity of chronic inflammation, a prophylactic or retardant effect on human AA amyloidosis associated with arthritis is an interesting possibility.

Plasmapheresis has been tried in those amyloidoses related to genetic variants of transthyretin, but without documented effect. Of more interest are the recently performed liver transplantations in such patients, which remove the site of production of the variant transthyretin and replace it by a liver harbouring only normal transthyretin genes (Holmgren *et al*. 1993).

Renal transplantation improves the articular complaints and retards the development of amyloid bone cysts in b $_2$-microglobulin-associated amyloidosis (Jadoul *et al*. 1989). A restoration of the normal clearance of b$_2$-microglobulin may explain the therapeutic effect.

***Dissolution of amyloid deposits* in vivo**

Dimethyl sulphoxide partially dissolves amyloid fibrils *in vitro*, and treatment of amyloidotic mice with this compound reduces the amount of amyloid (Isobe and Osserman 1976). It has therefore been tried in the treatment of human AA, AL, and transthyretin-related amyloidoses. A beneficial effect has been reported in occasional patients (Hazenberg and van Rijswijk 1994), but there are no reports of controlled studies. Many patients refuse to take dimethyl sulphoxide because it gives a bad body odour.

Organ-directed therapy

Renal transplantation is increasingly used in amyloid nephropathy, particularly of AA background (Stone 1990). A series of patients had increased survival as well as improvement of quality of life (Hartmann *et al*. 1992). Although the results of renal transplantation are not as good as in non-amyloid nephropathies, more patients should probably be given the chance to receive renal transplant also in AL, transthyretin, and b $_2$-microglobulin-associated amyloidoses (Kay 1993).

The effect of heart transplantation for amyloidosis have been reported in a few patients, most of them with AL type (see Stone 1990); five of six such patients were reportedly alive 21 plus or minus 12 months after transplantation.

Other therapeutic considerations

Patients with cardiac amyloid may have increased sensitivity to digitalis, and calcium-channel blocking agents may aggravate heart failure in such patients. Amyloid appears to bind such agents *in vitro* and possibly also *in vivo* (Stone 1990), and they should be used with caution in amyloidosis patients.

Comprehensive reviews of the clinical and therapeutic aspects of adult (Hazenberg and van Rijswijk 1994) and childhood amyloidoses (Woo 1994) of particular interest in rheumatology have recently been published.

Prognosis

The prognosis of amyloidosis depends on the localization and progress of the tissue deposition of amyloid. Survival time is largely dependent on the time of diagnosis, which varies significantly. It is clear, however, that systemic amyloidosis must be regarded as a severe condition with a high risk of death. AA amyloidosis has been estimated to be the cause of up to 47 per cent of deaths in European patients with juvenile chronic arthritis (Baum and Gutowska 1977). A Finnish report concluded that amyloidosis was the cause of death in 8 per cent of patients with adult rheumatoid arthritis coming to autopsy (Koota *et al*. 1975). On the other hand, the outcome of reactive AA amyloidosis is more favourable than many physicians have previously thought. Reports from Finland and the United States indicate that survival time after diagnosis of amyloidosis associated with rheumatoid arthritis is 4 to 5 years (Wegelius *et al*. 1980; Kyle 1982). In juvenile chronic arthritis with amyloidosis, survival may be even better when treated with cytotoxic drugs for 8 to 9 years (Baum and Gutowska 1977; Schnitzer and Ansell 1977) or longer (Fig. 6) (Woo 1990). There is a large individual variation in survival time for AA amyloidosis, in one report (Tribe *et al*. 1980) ranging from a few weeks up to 20 years.

The prognosis of AL amyloidosis is less favourable. Reported survival rates are less than 2 years for idiopathic type, and 7 months or even less in AL amyloidosis with myeloma (Janssen *et al*. 1986; Kyle 1990). Again, there is a wide range in survival time of individual patients, from weeks up to 5 years or more (Kyle *et al*. 1986). Cardiac amyloid is the most common cause of death and the major determinant of prognosis among patients with AL (Kyle 1990). Urinary excretion of monoclonal l-light chains is associated with inferior survival in AL amyloidosis compared with monoclonal k-chains or no monoclonal protein (Gertz and Kyle 1990)

In the hereditary amyloid polyneuropathies and cardiomyopathies related to variants of transthyretin, there are also wide variations in the course and prognosis among the different affected families as well as between individual gene carriers in the same families. Some patients may die with cachexia before the age of 40 years, whereas others may be in good health at 90 years (Sequeiros and Saraiva 1987).

Chapter References

Ahlmen, M., Ahlmen, J., Svalander, C., and Bucht, H. (1987). Cytotoxic drug treatment of reactive amyloidosis in rheumatoid arthritis with special reference to renal insufficiency. *Clinical Rheumatology*, **6**, 27–38.

Anders, R.F., Natvig, J.B., Sletten, K., Husby, G., and Nordstoga, K. (1977). Amyloid-related serum protein SAA from three animal species: comparison with human SAA. *Journal of Immunology*, **118**, 229–34.

Andrade, C. (1952). A peculiar form of peripheral neuropathy. Familiar atypical generalized amyloidosis with special involvement of the peripheral nerves. *Brain*, **75**, 408–27.

Askanas, V. and Engel, W.K. (1993). New advances in inclusion-body myositis. *Current Opinion in Rheumatology*, **5**, 732–41.

Baum, J. and Gutowska, G. (1977). Death in juvenile rheumatoid arthritis. *Arthritis and Rheumatism*, **20** (Suppl.), 253–5.

Benditt, E.P. and Eriksen, N. (1964). Starch gel electrophoretic analysis of some proteins extracted from amyloid. *Archives of Pathology*, **78**, 325–30.

Benditt, E.P. and Eriksen, N. (1971). Chemical classes of amyloid substance. *American Journal of Pathology*, **65**, 231–52.

Benson, M.D. (1991). Inherited amyloidosis. *Journal of Medical Genetics*, **28**, 73–8.

Benson, M.D. and Wallace, M.R. (1989). Amyloidosis. In *The metabolic basis of inherited disease*, Vol. 2 (6th edn) (ed. C.R. Scriver, A.L. Beaudet, W.S. Sly, and D. Valle), pp. 2439–60. McGraw-Hill, New York.

Berglund, K., Thysell, H., and Keller C. (1993). Results, principles and pitfalls in the management of renal AA-amyloidosis; a 10–21 year followup of 16 patients with rheumatic disease treated with alkylating cytostatics. *Journal of Rheumatology*, **20**, 2051–7.

Black, M.M. (1976). Primary localised amyloidosis of the skin: clinical variants, histochemistry and ultra-structure. In *Amyloidosis* (ed. O. Wegelius and A. Pasternack), pp. 479–511. Academic Press, London.

Cathcart, E.S., Crystal, A.L., Meydani, S.N., and Hayes, K.C. (1987). A fish oil diet retards experimental amyloidosis, modulates lymphocyte function, and decreases macrophage arachidonate metabolism in mice. *Journal of Immunology*, **139**, 1850–4.

Cohen, A.S., Rubinow, A., Anderson, J.J., Skinner, M., Mason, J.H., and Kayne, L.C. (1987). Survival of patients with primary (AL) amyloidosis: colchicine-treated cases from 1976 to 1983 compared with cases seen in previous years (1961 to 1973). *American Journal of Medicine*, **82**, 1182–90.

Cooper, J.H. (1974). Selective amyloid staining as a function of amyloid composition and structure. Histochemical analysis of the alkaline Congo red, standardized toluidine blue, and iodine methods. *Laboratory Investigation*, **31**, 232–8.

Corder, E.H. *et al*. (1993). Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease. *Science*, **261**, 921–3.

Eriksen, N., Ericsson, L.H., Pearsall, N., Lagunoff, D., and Benditt, E.P. (1976). Mouse amyloid protein AA: homology with non-immunoglobulin protein of human and monkey amyloid substance. *Proceedings of the National Academy of Sciences (USA)*, **73**, 964–7.

Filipowicz-Sosnowska, A.M., Rostropowicz-Denisiewicz, K., Rosenthal, C.J., and Baum, J. (1978). The amyloidosis of juvenile rheumatoid arthritis. Comparative studies in Polish and American children. *Arthritis and Rheumatism*, **21**, 699–703.

Frederiksen, T., Gøtzsche, H., Harboe, N., Kiaer, W., and Mellangaard, K. (1962). Primary familial amyloidosis with severe amyloid heart disease. *American Journal of Medicine*, **33**, 328–48.

Fuks, A. and Zucker-Franklin, D. (1985). Impaired Kupffer cell function precedes development of secondary amyloidosis. *Journal of Experimental Medicine*, **161**, 1013–28.

Furie, B., LiAnn Voo, B.S., McAdam, K.P.W.J., and Furie, B.C. (1981). Mechanism of factor X deficiency in systemic amyloidosis. *New England Journal of Medicine*, **304**, 827–30.

Gejyo, F. *et al*. (1985). A new form of amyloid protein associated with chronic hemodialysis was identified as b $_2$-microglobulin. *Biochemical and Biophysical Research Communications*, **129**, 701–6.

Gertz, M.A. and Kyle, R.A. (1990). Prognostic value of urinary protein in primary systemic amyloidosis (AL). *American Journal of Pathology*, **94**, 313–7.

Gertz, M.A., Kyle, R.A., Griffing, W.L., and Hunder, G.G. (1986). Jaw claudication in primary systemic amyloidosis. *Medicine*, **65**, 173–9.

Gertz, M.A., Kyle, R.A., and Greipp, P.R. (1991). Response rates and survival in primary systemic amyloidosis. *Blood*, **77**, 257–62.

Ghiso, J., Jensson, O., and Frangione, B. (1986). Amyloid fibrils in hereditary cerebral hemorrhage with amyloidosis of Icelandic type is a variant of g-trace basic protein (cystatin C). *Proceedings of the National Academy of Sciences (USA)*, **83**, 2974–8.

Glenner, G.G. (1980). Amyloid deposits and amyloidosis. *New England Journal of Medicine*, **302**, 1283–92.

Glenner, G.G. and Murphy, M.A. (1989). Amyloidosis of the nervous system. *Journal of the Neurological Sciences*, **94**, 1–28.

Hartmann, A. *et al*. (1992). Fifteen years' experience with renal transplantation in systemic amyloidosis. *Transplant International*, **5**, 15–8.

Hawkins, P.N. (1994). Diagnosis and monitoring of amyloidosis. In *Reactive amyloidosis and the acute phase response* (ed. G. Husby), *Baillière's Clinical Rheumatology*, **8**, pp. 635–59.

Hazenberg, B.P.C. and van Rijswijk, M.H. (1994). Clinical and therapeutic aspects of AA amyloidosis. In *Reactive amyloidosis and the acute phase response* (ed. G. Husby), *Baillière's Clinical Rheumatology*, **8**, pp. 661–90.

Hol, P.R., Snel, F.W.J.J., Niewold, T.A., and Gruys, E. (1986). Amyloid enhancing factor (AEF) in the pathogenesis of AA-amyloidosis in the hamster. *Virchows Archive B. Cell Pathology*, **52**, 273–81.

Holmgren, G. *et al*. (1993). Clinical improvement and amyloid regression after liver transplantation in hereditary transthyretin amyloidosis. *Lancet*, **341**, 1113–6.

Husby, G. (1983). Immunoglobulin-related (AL) amyloidosis. *Clinical and Experimental Rheumatology*, **1**, 353–8.

Husby, G. (1985). Amyloidosis and rheumatoid arthritis. *Clinical and Experimental Rheumatology*, **3**, 173–80.

Husby, G. (1989). Pathogenesis of AA amyloidosis. In *Acute phase proteins in the acute phase response*, Argenteul Symposium, (ed. M. Pepys), pp. 169–85. Springer-Verlag, London.

Husby, G. (1992). Amyloidosis. *Seminars in Arthritis and Rheumatism*, **22**, 67–82.

Husby, G. (1994). Classification of amyloidosis. In *Reactive amyloidosis and the acute phase response* (ed. G. Husby), *Baillière's Clinical Rheumatology*, **8**, pp. 503–54.

Husby, G. and Sletten, K. (1986). Clinical and chemical classification of amyloidosis 1985. *Scandinavian Journal of Immunology*, **23**, 253–65.

Husby, G. *et al*. (1990). The 1990 guidelines for nomenclature and classification of amyloid and amyloidosis. In *Amyloid and amyloidosis* (ed. J.B. Natvig *et al*.). Kluwer Academic Publishers, Dordrecht.

Husby, G., Marhaug, G., Dowton, B., Sletten, K., and Sipe, J.D. (1994 *a*). Serum amyloid A (SAA): biochemistry, genetics and the pathogenesis of AA amyloidosis. *International Journal of Experimental and Clinical Investigation*, **1**, 119–37.

Husby, G., Stenstad, T., Magnus, J.H., Sletten, K., Nordvåg, B.Y., and Marhaug, G. (1994 *b*). Interaction between circulating amyloid fibril protein precursors and extracellular tissue matrix components in the pathogenesis of systemic amyloidosis. *Clinical Immunology and Immunopathology*, **70**, 2–9.

Husebekk, A., Skogen, B., Husby, G., and Marhaug, G. (1985). Transformation of myloid precursor SAA to protein AA and incorporation in amyloid fibrils in vivo. *Scandinavian Journal of Immunology*, **21**, 283–7.

Isobe, T. and Osserman, E.F. (1976). Effects of dimethyl sulfoxide (DMSO) on Bence-Jones proteins, amyloid fibrils and casein-induced amyloidosis. In *Amyloidosis* (ed. O. Wegelius and A. Pasternack), pp. 247–57. Academic Press, London.

Jadoul, M., Malghem, J., Pirson, Y., Maldague, B., and van Ypersele de Strihou, C. (1989). Effect of renal transplantation on the radiological signs of dialysis amyloid osteoarthropathy. *Clinical Nephrology*, **32**, 194–7.

Janssen, S., van Rijswijk, M.H., Meijer, S., Ruinen, L., and van der Hem, G.K. (1986). Systemic amyloidosis: a clinical survey of 144 cases. *Netherlands Journal of Medicine*, **29**, 376–85.

Johansson, B., Wernstedt, C., and Westermark, P. (1987). Atrial natriuretic peptide deposited as atrial amyloid fibrils. *Biochemical and Biophysical Research Communications*, **148**, 1087–92.

Kay, J. (1993). b $_2$-microglobulin amyloidosis. *Rheumatology Review*, **2**, 22–8.

Kisilevski, R. (1987). From arthritis to Alzheimer's disease: current concepts on the pathogenesis of amyloidosis. *Canadian Journal of Physiology and Pharmacology*, **65**, 1805–15.

Kisilevski, R. and Young, I.D. (1994). Pathogenesis of amyloidosis. In *Reactive amyloidosis and the acute phase response* (ed. G. Husby), *Baillière's Clinical Rheumatology*, **8**, pp. 613–26.

Koota, K., Isomäki, H.A., and Mutru, O. (1975). Death rate and causes of death in patients with rheumatoid arthritis. *Scandinavian Journal of Rheumatology*, **4**, 205–8.

Kyle, R.A. (1982). Amyloidosis. *Clinical Haematology*, **11**, 151–80.

Kyle, R.A. (1990). Primary systemic amyloidosis (AL). In *Amyloid and amyloidosis* (ed. J.B. Natvig *et al*.), pp.147–52. Kluwer Academic Publishers, Dordrecht.

Kyle, R.A., Greipp, P.R., and O'Fallon, W.M. (1986). Primary systemic amyloidosis: multivariate analysis for prognostic factors in 168 cases. *Blood*, **68**, 220–4.

Kyle, R.A., Gertz, M.A., Garton, J.P., and Greipp, P.R. (1990). Primary systemic amyloidosis (AL): a randomized trial of colchicine vs. melphalan and prednisone vs. melphalan, predinisone, and colchicine. In *Amyloid and amyloidosis* (ed. J.B. Natvig *et al*.), pp. 231–4, Kluwer Academic Publishers, Dordrecht.

Levin, M., Franklin, E.C., Frangione, B., and Pras, M. (1972). The amino acid sequence of a major non-immunoglobulin component of some amyloid fibrils. *Journal of Clinical Investigation*, **51**, 2773–6.

Levy, E. *et al*. (1990). Mutation of the Alzheimer's disease amyloid gene in hereditary cerebral hemorrhage, Dutch type. *Science*, **248**, 1124–6.

Linke, R.P., Heilmann, K.L., Nathrath, W.B.J., and Eulitz, M. (1983). Identification of amyloid A proteinin a sporadic Muckle–Wells syndrome. N-terminal amino acid sequence after isolation from formalin-fixed tissue. *Laboratory Investigation*, **48**, 698–704.

Livneh, A., Zemer, D., Langevitz, P., Shemer, J., Sohar, E., and Pras M. (1993). Colchicine in the treatment of AA and AL amyloidosis. *Seminars in Arthritis and Rheumatism*, **23**, 206–14.

Magnus, J.H., Stenstad, T., and Husby, G. (1989). Proteoglycans, glycosaminoglycans and amyloid deposition. In *Reactive amyloidosis and the acute phase response* (ed. G. Husby), *Baillière's Clinical Rheumatology*, **8**, pp. 575–97.

Majolino, I. *et al*. (1993). High-dose therapy and autologous transplantation in amyloidosis-AL. *Haematologica*, **78**, 68–71.

Marhaug, G., Østensen, M., Husby, G., and Husebekk, A. (1986). Serum amyloid A protein, a sensitive inflammation marker. In *Marker proteins in inflammation*, Vol. 3, (ed. J. Bienvenu, J.A. Grim, and P. Laurent), pp. 139–43. Walter de Gryuter, Berlin.

Marhaug, G., Husby, G., Nordstoga, K., and Dowton, S.B. (1990). Mink serum amyloid A protein. Expression and primary structure based on cDNA sequences. *Journal of Biological Chemistry*, **265**, 10049–54.

Maury, C.P.J. (1990*a*). Complete primary structure of amyloid fibril protein in Finnish hereditary amyloidosis: identification of a new type of amyloid protein derived from variant gelsolin. In *Amyloid and amyloidosis* (ed. J.B. Natvig *et al*.). Kluwer Academic Publishers, Dordrecht.

Maury, C.P.J. (1990*b*). b_2-Microglobulin amyloidosis. A systemic amyloid disease affecting primarily synovium and bone in long-term dialysis patients. *Rheumatology International*, **10**, 1–8.

Meek, R.L., Hoffmann, J.S., and Benditt, E.P. (1986). Amyloidogenesis. One serum amyloid A isotype is selectively removed from the circulation. *Journal of Experimental Medicine*, **163**, 499–510.

Meek, R.L., Urieli-Shoval, S., and Benditt, E.P. (1994). Expression of apolipoprotein serum amyloid A mRNA in human atherosclerotic lesions and cultured vascular cells: implications for serum amyloid A function. *Proceedings of the National Academy of Sciences (USA)*, **91**, 3186–90.

Meretoja, J. (1969). Familial systemic paraamyloidosis with lattice dystrophy of the cornea, progressive cranial neuropathy, skin changes and various internal symptoms. A previous unrecognized heritable syndrome. *Annals of Clinical Research*, **1**, 314–24.

Miyata, T. *et al*. (1994). Involvement of b_2-microglobulin modified with advanced glycation end products in the pathogenesis of hemodialysis-associated amyloidosis. *Journal of Clinical Investigation*, **93**, 521–8.

Muckle, T.J. (1979). The 'Muckle–Wells' syndrome. *British Journal of Dermatology*, **100**, 87–92.

Nichols, W.C., Dwulet, F.E., Liepnieks, J., and Benson, M.D. (1988). Variant apolipoprotein A1 as a major constituent of a human hereditary amyloid. *Biochemical and Biophysical Research Communications*, **156**, 762–8.

Nichols, W.C., Padilla, L.-M., and Benson, M.D. (1989). Prenatal detection of a gene for hereditary amyloidosis. *American Journal of Medical Genetics*, **34**, 520–4.

Nordlie, M., Sletten, K., Husby, G., and Ranløv, P.J. (1988). A new prealbumin variant in familian amyloid cardiomyopathy of Danish origin. *Scandinavian Journal of Immunology*, **27**, 119–22.

Nordvåg, B.Y., Husby, G., Ranløv, I., and El-Gewely, M.R. (1992). Molecular diagnosis of the transthyretin (TTR) Met111 mutation in familial amyloid cardiomyopathy of Danish origin. *Human Genetics*, **89**, 459–61.

Orstertag, B. (1932). Demonstration einer eigenartigen familiaren 'Paramyliodose'. *Zentralblat für Allgemeine Pathologie und Pathologische Anatomie (Jena)*, **56**, 253–4.

Pepys, M.B. (1988). Amyloidosis: some recent developments. *Quarterly Journal of Medicine*, **67**, 283–98.

Pepys, M.B. *et al*. (1993). Human lysozyme gene mutations cause hereditary systemic amyloidosis. *Nature*, **362**, 553–7.

Pras, M. (1986). The hereditary amyloidoses. In *Amyloidosis* (ed. J. Marring and M.H. van Rijswijk), pp. 185–93. Martin Nijhoff Publishers, Dortrecht.

Pras, M., Schubert, M., Zucker-Franklin, D., Rimon A., and Franklin, E.C. (1968). The characterization of soluble amyloid prepared in water. *Journal of Clinical Investigation*, **47**, 924–33.

Pras, M. *et al*. (1992). Mapping of a gene causing familial Mediterranean fever to the short arm of chromosome 16. *New England Journal of Medicine*, **326**, 1509–13.

Sack, G.H. Jr. (1988). Serum amyloid A (SAA) gene variation in familial Mediterranean fever. *Molecular Biology and Medicine*, **5**, 61–7.

Schnitzer, T.J. and Ansell, B.M. (1977). Amyloidosis in juvenile chronic polyarthritis. *Arthritis and Rheumatism*, **20**, 245–52.

Sequeiros, J. and Saraiva, M.J.M. (1987). Onset in the seventh decade and lack of symptoms in heterozygotes for the TTR ^{Met30} mutation in hereditary amyloid neuropathy-type I (Portuguese, Andrade). *American Journal of Medical Genetics*, **27**, 345–57.

Sletten, K. and Husby, G. (1974). The complete aminoacid sequence of non–immunoglobulin amyloid fibril protein AS in rheumatoid arthritis. *European Journal of Biochemistry*, **41**, 117–25.

Sletten, K., Westermark, P., and Natvig, J.B. (1976). Characterization of amyloid fibril proteins from medullary carcinoma of the thyroid. *Journal of Experimental Medicine*, **143**, 993–7.

Sletten, K., Marhaug, G., and Husby, G. (1983). The covalent structure of amyloid related serum protein SAA from two patients with inflammatory disease. *Hoppe–Seyler's Zeitschrift für Physiologische Chemie*, **364**, 1039–46.

Snow, A.D., Willmer, J., and Kisilevsky, R. (1987). A close structural relationship between sulfated proteoglycans and AA amyloid fibrils. *Laboratory Investigation*, **57**, 687–98.

Soutar, A.K. *et al*. (1992). Apolipoprotein A1 mutation Arg-60 causes autosomal dominant amyloidosis. *Proceedings of the National Academy of Sciences (USA)*, **89**, 7349–93.

Stenstad, T., Magnus, J.H., and Husby, G. (1994). Characterization of proteoglycans in mouse splenic AA amyloidosis. *Biochemical Journal*, **303**, 663–70.

Stone, M.J. (1990). Amyloidosis: a final common pathway for protein deposition in tissues. *Blood*, **75**, 531–45.

Syversen, V., Sletten, K., Marhaug, G., Husby, G., and Lium, B. (1987). The amino acid sequence of serum amyloid A (SAA) in mink. *Scandinavian Journal of Immunology*, **26**, 763–7.

Tribe, C.R., Bacon, P.A., and Mackenzie, J.C. (1980). Experience with an amyloid clinic. In *Amyloid and amyloidosis* (ed. G.G. Glenner, P.P. Costa, and A. Falcao de Freitas), pp. 179–82. Excerpta Medica, Amsterdam.

Uemichi, T., Liepniks, J.J., and Benson, M.D. (1994). Hereditary amyloidosis with a novel variant fibrinogen. *Journal of Clinical Investigation*, **93**, 731–6.

Waalen, K., Sletten, K., Husby, G., and Nordstoga, K. (1980). The primary structure of amyloid fibril protein AA in endotoxin induced amyloidosis of the mink. *European Journal of Biochemistry*, **104**,

407–12.

Wegelius, O., Wafin, F., Falck, H.M., and Tornroth, T. (1980). Follow-up study of amyloidosis secondary to rheumatid disease. In *Amyloid and amyloidosis* (ed. G.G. Glenner, P.P. Costa, and A. Falcao de Freitas) pp. 183–90, Excerpta Medica, Amsterdam.

Westermark, P. and Stenkvist, B. (1973). A new method for the diagnosis of systemic amyloidosis. *Archives of Internal Medicine*, **132**, 522–3.

Westermark, P., Wernstedt, C., Wilander, E., and Sletten, K. (1986). A novel peptide in the calcitonin gene related peptide family as an amyloid fibril protein in the endocrine pancreas. *Biochemical and Biophysical Research Communication*, **140**, 827–31.

Westermark, P., Sletten, K., Johansson, B., and Cornwell, G.G. (1990). Fibril in senile systemic amyloidosis is derived from normal transthyretin. *Proceedings of the National Academy of Sciences (USA)*, **87**, 2843–5.

Woo, P. (1990). Complications of juvenile arthritis. In *Paediatric rheumatology update* (ed. P. Woo, P.H. White, and B. Ansell), pp. 38–46. Oxford University Press, Oxford.

Woo, P. (1994). Amyloidosis in children. In *Reactive amyloidosis and the acute phase response. Baillière's Clinical Rheumatology*, **6**, 613–26.

Woo, P., O'Brien, Robson, M., and Ansell, B.M. (1987). A genetic marker for systemic amyloidosis in juvenile arthritis. *Lancet*, **ii**, 767–9.

Yamamoto, K.-I. and Migita, S. (1985). Complete primary structures of two major murine serum amyloid A proteins deduced from cDNA sequences. *Proceedings of the National Academy of Sciences (USA)*, **82**, 2915–9.

Zemer, D., Pras, M., Sohar, E., Modan, M., Cabili, S., and Gafni, J. (1986). Colchicine in the prevention and treatment of the amyloidosis of familial Mediterranean fever. *New England Journal of Medicine*, **314**, 1001–5.

5.13.2 Familial Mediterranean fever

Pnina Langevitz and Avi Livneh and Deborah Zemer and Mordechai Pras

History

A genetic disease with a gene frequency as high as that of familial Mediterranean fever in Sephardi Jews must have existed for hundreds or thousands of years. Nevertheless, in the first half of the twentieth century only a few isolated characteristic cases could be traced in medical publications (Janeway and Mosenthal 1908; Alt and Barker 1930). The reason why a disease with such dramatic manifestations has only recently been recognized may be connected with the geographical distribution of its sufferers, who resided mainly in areas where modern medical facilities were not available at the time. This changed in the 1940s and 1950s when the population affected became exposed to advanced medical facilities and research, partly due to migrations.

Siegal (1945) was the first to describe the abdominal attacks as a separate disease entity. Later, many typical cases were included in the heterogeneous case collection termed 'periodic disease'1 by Reimann (1949). French physicians described many typical cases of familial Mediterranean fever in Jewish patients deriving from North Africa (Mamou and Cattan 1952; Siguier *et al*. 1953; Benhamou *et al*. 1954). They were the first to perceive the familial nature and the fatal renal lesion of the disease.

Tel Hashomer Hospital in Israel served in the early 1950s as a referral centre for the new immigrants who came mainly from Mediterranean countries. Enigmatic cases who appeared on the wards with recurrent, short-lived episodes of fever accompanied by peritoneal, pleural, or arthritic inflammation caught the attention of the group led by Professor Harry Heller. In a series of publications (Heller *et al*. 1958; Heller *et al*. 1961a; Heller *et al*. 1961b; Sohar *et al*. 1961; Sohar *et al*. 1967) they defined the clinical features of the disease and its diagnostic criteria, established its genetic nature, mode of transmission, and ethnic distribution, emphasized the role of amyloidosis, and coined the name 'familial Mediterranean fever'.

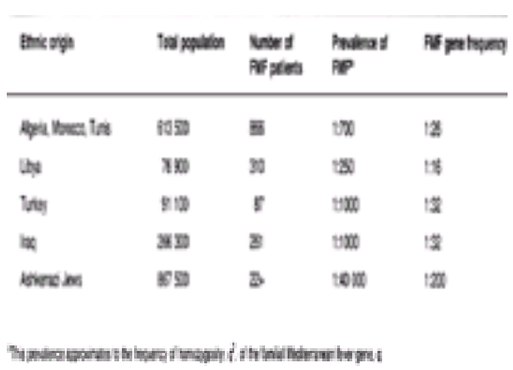
Diagnostic criteria

Despite many attempts at elucidation in the last 35 years the mechanisms leading to the clinical manifestations of the disease are still unknown, as is the underlying basic error of metabolism and the location of the gene. There is no specific diagnostic laboratory test. The diagnostic criteria are based on clinical manifestations.

1. Short attacks of fever recurring at irregular intervals. In most cases the temperature rises to 39–40°C.
2. Painful inflammatory manifestations in the abdomen, chest, joints, or skin, associated with the fever.
3. Nephropathic amyloidosis leading to terminal renal failure early in life.
4. Autosomal-recessive inheritance.
5. Virtual ethnic restriction to Mediterranean stock (Sephardi Jews, Armenians, Anatolian Turks, Arabs, and Ashkenazi Jews).
6. Dramatic response to continuous colchicine treatment in abolishing or reducing febrile attacks.

The inheritance of familial Mediterranean fever

The disease may become manifest as early as during the first year of life, although in most cases onset is later. In two-thirds of our patients the first manifestations appeared during the first decade of life; by the end of the second decade 90 per cent are affected. Only rarely is the onset delayed beyond the age of 40 (Sohar *et al*. 1967). Among our 4000 patients there are only about 100 Ashkenazi Jews and the disease is rare in other Jewish ethnic groups (Yemenite and Iranian). Table 1 shows the prevalence and gene frequency of familial Mediterranean fever, calculated by the number of patients and the size of the total population in each ethnic subgroup.



Ethnic origin	Total population	Number of FMF patients	Prevalence of FMF*	FMF gene frequency
Algeria, Morocco, Tunis	613 500	888	1:700	1:26
Libya	78 900	310	1:250	1:16
Turkey	91 100	87	1:1000	1:32
Iraq	266 300	261	1:1000	1:32
Ashkenazi Jews	867 500	25+	1:40 000	1:200

*The prevalence approximates to the frequency of homozygosity (q^2) of the familial Mediterranean fever gene, q.

Table 1 Prevalence and gene frequency of familial Mediterranean fever (FMF)

The peculiar ethnic restriction and the familial aggregation in familial Mediterranean fever suggests a genetic aetiology. Analysis of 229 of our families led us to conclude that the disease is due to a single, recessive autosomal gene (Sohar *et al*. 1967).

In about 90 affected families, out of several hundreds, the disease occurred in more than one generation. Although in most of these families the inheritance could be considered to be autosomal-recessive, in two the transmission could not be explained by recessive inheritance and must be assumed to be dominant (Yuval *et al*. 1995).

The clinical picture

Attacks

The febrile, painful attacks that are the hallmark of the disease are characterized by marked elevation of body temperature, acute inflammation of the peritoneum, synovia, or pleura, a duration of 12 to 48 h, and complete health between attacks (Sohar *et al*. 1967).

Repeated attacks at irregular intervals and in an unpredictable sequence are typical of the disease: periods of one febrile attack a week can vary with remissions of

weeks, months, or even years, with no apparent cause. During the illness a patient will probably encounter several forms of attacks, but the recurrence of one type over many years is not uncommon.

The most frequent manifestation is the abdominal attack, experienced by 90 per cent of patients; in 68 per cent of these it is the presenting sign. As the most dramatic manifestation, these attacks attract attention and extensive diagnostic efforts (Sohar *et al*. 1967). They are marked by the sudden onset of fever (often with chills) and pain spreading over the entire abdomen from variable points of origin. As the attack gains in intensity, guarding, rebound tenderness, board-like rigidity, distension, and absence of peristalsis appear. Multiple, small fluid levels in the small bowel on radiography combine to suggest an acute abdominal catastrophe. After 6 to 12 h the signs and symptoms recede, and within 24 to 48 h the attack is usually over, leaving the patient as well as before.

Organization of the exudate may result in fibrous adhesions, which in rare cases may give rise to mechanical ileus (Michaeli *et al*. 1966). This may be responsible for chronic subileus in some patients and ascites in others (Zemer *et al*. 1977). It is probably the cause of sterility in some affected women (Ismachovich *et al*. 1973).

The pleural attack has been experienced by 45 per cent of our patients, and in 5 per cent it was the presenting sign (Sohar *et al*. 1967). It assumes the picture of an acute febrile pleuritis, resembling the peritoneal attacks in its abrupt onset, rapid resolution, and unpredictable recurrence. The pleuritic attack may be limited to the chest or may shift to the abdomen. Breathing is painful and breath sounds are diminished on the affected side. There may be radiological evidence of a small exudate in the costophrenic angle, which is difficult to aspirate and which resolves within 48 h. No sequelae of clinical significance have been noted.

Pericarditis is a rare feature of familial Mediterranean fever. We observed clinical attacks of pericarditis in 20 of our patients (A. Livneh, unpublished data). On M-mode echocardiography, pericardial involvement was reported to be more common than in the general population (Dabestani *et al*. 1982). No permanent sequelae have been reported.

The articular attack is the second most common form of attack. It was experienced by 75 per cent of the patients in our series, and was the presenting symptom in 16 per cent of them. Arthritic attacks may recur for years as the only feature of the disease, before other forms appear (Sohar *et al*. 1967).

As a rule, large joints are involved, particularly those of the lower extremities. Arthritic attacks of familial Mediterranean fever may present in two forms: acute, or chronic and protracted. In the acute form the onset is abrupt, fever ranges from 38 to 40°C, and the affected single joint is tender, swollen, and held immobile because of the severe pain. Redness and local heat are frequently less marked than would be expected in so acute a process. The signs and symptoms usually peak in 1 to 2 days and then gradually subside, leaving no residue. The attacks can sometimes be precipitated by minor trauma or effort, such as prolonged walking. Synovial effusion is often demonstrable. The synovial fluid is sterile and varies in appearance from cloudy to purulent, depending upon the acuteness and severity of the synovitis. Resolution of a short attack can occur in as soon as 2 to 3 days, but more commonly takes a week and sometimes nearly a month.

About 5 per cent of patients experience protracted attacks, which persist for more than a month. Usually the hip or knee are involved (Sohar *et al*. 1967; Sneh *et al*. 1977), but episodes in other joints, such as the ankle and, rarely, the temporomandibular or the sternoclavicular, may also assume a protracted course. Rather than recovering after several days, the joint remains markedly swollen and painful, presenting a picture of chronic monoarthritis or in rare cases, chronic oligoarthritis. The affected knee joint, in extreme cases, resembles a fluid bag from which up to 200 ml can be drained (Fig. 1). After several weeks or months, sometimes even after a year or more, the pain subsides spontaneously. During such protracted attacks in a joint, short attacks involving other joints, the abdomen, or chest may occur (Sohar *et al*. 1967).

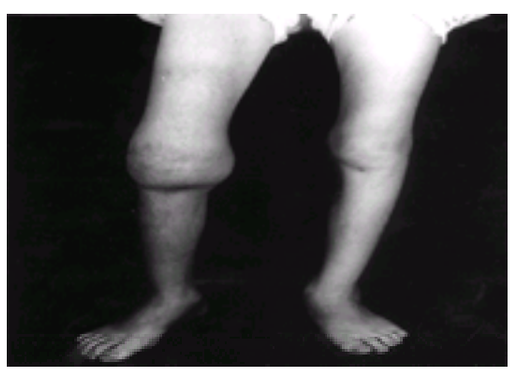


Fig. 1 Protracted arthritis of right knee, which lasted 11 months, in a 15-year-old patient (before colchicine was introduced for treatment).

In some protracted cases, especially in the hips, damage to the joints can be so severe as to cause permanent deformity, which ultimately may require joint replacement (Fig. 2). In 27 of our patients, there was residual incapacity in the affected joint (21 in the hips). Seven hips showed radiologically typical aseptic necrosis of the femoral head, and in 14 sclerosis and narrowing of the joint space was observed. Most of these hips eventually required total prosthetic replacement (Sneh *et al*. 1977). In a summary of 22 total hip replacements performed in 18 patients with familial Mediterranean fever between 1971 and 1985 a relatively high percentage of aseptic loosening of the cemented hip prosthesis was noted. This finding led us to recommend cementless hip prostheses in such patients (Salai *et al*. 1993).

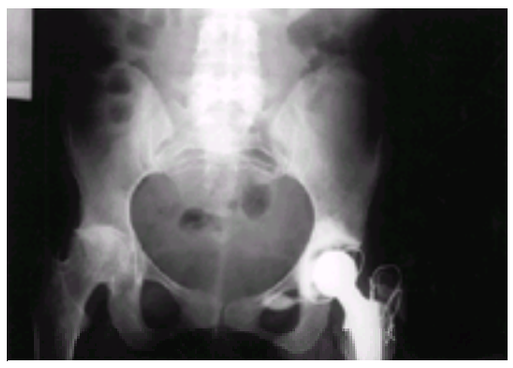


Fig. 2 Results of protracted arthritis of both hips in a 17-year-old girl. State following hip arthroplasty of left hip. Narrowing of the joint space, sclerosis, and aseptic necrosis of the lateral aspect can be seen in the right hip.

Among 160 patients with protracted arthritis we found a small group of 11 in whom the HLA-B27 was negative, who fulfilled the criteria for seronegative spondylarthropathy. Most of these patients responded to therapy with non-steroidal anti-inflammatory drugs, but some of them required the addition of disease-modifying antirheumatic drugs (Langevitz *et al*. 1994a).

Muscle pains occur in about 20 per cent of patients with familial Mediterranean fever. Usually the pain is not severe, appears in the lower extremities after physical exertion, lasts from a few hours to 1 to 2 days, and subsides with rest or non-steroidal anti-inflammatory drugs. In 12 per cent of patients with familial Mediterranean fever a syndrome of protracted febrile myalgia developed, characterized by severe, debilitating myalgia accompanied by fever, abdominal pain, a high erythrocyte sedimentation rate, leucocytosis, and hyperglobulinaemia. In a few patients a mild, short-lasting, vasculitic, non-thrombocytopenic purpura with a deposition of IgA was noted. In patients that were treated by non-steroidal anti-inflammatory drugs the attacks lasted 6 to 8 weeks, but they subsided promptly after a high dose of prednisone (Langevitz *et al*. 1994b). Since colchicine is known to induce neuropathy and myopathy in rare cases, especially in transplanted patients treated with cyclosporin (Yussim *et al*. 1994), it is important to differentiate colchicine-induced myopathy from an attack of protracted febrile myalgia.

Erysipelas-like erythema is one of the most characteristic manifestations of familial Mediterranean fever. It was reported in 11 per cent of affected children, usually combined with arthritis. Rather sharply bordered red patches, hot, tender, and swollen, and 10 to 35 cm^2 in area, appear on the skin of the lower extremities. They are usually located between the knee and ankle, or on the dorsum of the foot or ankle region, and are also accompanied by abrupt elevation of body temperature and last about 24 to 48 h (Sohar *et al*. 1967).

Orchitis (an acute, unilateral, painful swelling and redness of the testis due to inflammation of the tunica vaginalis) has been recognized as a form of attack in children or young adults; it subsides spontaneously after 12 to 24 h, without anatomical residue (Eshel *et al*. 1988).

Elevation of body temperature, sometimes to 40°C for a few hours, occurs frequently, especially in children, as the only expression of an attack (Sohar *et al*. 1967). This phenomenon is often falsely attributed to viral infection, pharyngitis, or tonsillitis.

One-third of our patients complained of pain in the heels or soles of the feet related to mild exertion such as walking or prolonged standing, which subsides after night rest. There are no objective signs of inflammation or elevation of temperature.

Mild splenomegaly of 1 to 4 cm, unrelated to amyloidosis, was found in many patients. In some patients the liver was also palpable. None showed clinical or laboratory malfunction of these organs.

Haematuria, sometimes only microscopic, has been observed in several patients during and between attacks. Mild anaemia with a low serum iron is common. Low levels of haemoglobin (7–10 g per cent) were found in some patients.

Allied conditions

Henoch–Schönlein purpura occurred in over 40 of our patients admitted to our hospital. Most of the patients were children and young adults, and in most of them the disease was characterized by a prolonged and severe course that required steroid therapy in many cases.

Polyarteritis nodosa has been reported in 15 cases of familial Mediterranean fever (Sachs *et al*. 1987; Glikson *et al*. 1989). All 15 were young patients, while polyarteritis nodosa generally occurs in the fifth or six decade of life. Polyarteritis nodosa is a rare disease with an incidence of 5 to 6 per million and the 15 cases recorded in about approximately 10 000 cases of familial Mediterranean fever are more than would be expected by chance (Pras *et al*. 1996).

A relatively high incidence of fibromyalgia (30 per cent) was found in patients with familial Mediterranean fever, especially in those who suffers from back and leg/foot pain (Langevitz *et al*. 1994c).

Various types of glomerulonephritis were reported in few patients with familial Mediterranean fever, including post-streptococcal glomerulonephritis, diffuse mesangial proliferative glomerulonephritis with IgA and IgM deposits and also rapidly progressive glomerulonephritis (Said *et al*. 1992).

Amyloidosis

A genetically determined, AA-type amyloidosis, clinically manifested as nephropathy (Sohar *et al*. 1967; Levine *et al*. 1972) is the fatal lesion in patients with familial Mediterranean fever.

The role of amyloidosis in the natural history of the disease was studied in an untreated group of 470 patients and revealed some typical features. Its clinical presentation occurs at an early age; 90 per cent of the patients who died from amyloidosis were under 40 and six were under 10 (Sohar *et al*. 1967). Subsequent evaluation showed a lower incidence of amyloidosis in some Jewish ethnic groups (Pras *et al*. 1982) and in Armenians (Schwabe and Peters 1974). The onset of the clinical signs of amyloidosis does not correlate with the frequency or intensity of the febrile attacks.

Recent studies showed that pregnancy may have possible deleterious effect on amyloid nephropathy in patients with more advanced renal failure at conception (Livneh *et al*. 1993).

Clinically, the amyloid nephropathy passes through several stages. There is a preclinical stage but since it can only be diagnosed by repeated rectal and renal biopsies or is inadvertently found in an occasional Congo red-stained appendectomy specimen, no attempt was made to determine its duration. Persistent proteinuria in an otherwise healthy patient with familial Mediterranean fever has proved to be a certain indication of renal amyloidosis.

Clinical evidence of extrarenal amyloidosis was scant when patients died from renal failure before chronic dialysis was introduced as a routine treatment for chronic renal failure. Adrenal insufficiency is not apparent despite the severe involvement observed at autopsy (Sohar *et al*. 1967). Some patients show clinical and laboratory evidence of intestinal malabsorbtion (Ravid and Sohar 1974). Following prolongation of life by chronic dialysis and renal transplantation, amyloid deposition in other organs has become more pronounced. In recent years we have observed extrarenal amyloid deposition that interfered with the normal function of certain organs (Table 2). The deposition of amyloid in the small bowel is a particularly grave consequence that caused the death of six patients.

Amyloid cardiomyopathy

Giant hepatomegaly

Amyloid goitre

Addison's disease

Fatal malabsorption

Table 2 Extrarenal amyloidosis in familial Mediterranean fever

The amyloidosis of familial Mediterranean fever is not directly related to the recurrent inflammatory attacks. In some patients it occurs before the appearance of the febrile attacks and in a few it is the only manifestation of the disease (Blum *et al*. 1962). Thirteen per cent of our patients who succumbed to amyloidosis were below 15 years of age, and the youngest died at 5 years after amyloidosis had been manifest for 3 years (Sohar *et al*. 1967; Gafni *et al*. 1968).

Amyloidosis is prevalent mainly in Jewish patients of North African origin. Before the time of colchicine treatment only 11 patients of 418 of North African origin (2.5 per cent) survived and reached 40 years of age without demonstrating clinical signs of amyloidosis (Pras *et al*. 1982). The incidence of amyloidosis in untreated patients is very much higher than in the inflammatory and infectious diseases that predispose to reactive ('secondary') amyloidosis.

Laboratory tests

Laboratory findings are meagre and non-specific. The erythrocyte sedimentation rate is accelerated and acute-phase proteins such as a_2-globulin and fibrinogen are increased, especially during attacks, but also in between (Sohar *et al*. 1967). The special attention accorded to serum amyloid A is due to the fact that its N-terminal fragment is AA protein, which is part of the AA amyloid fibril. Serum amyloid A is raised considerably in amyloidotic patients with familial Mediterranean fever. During attacks, very high concentrations of serum amyloid A are observed, which decrease gradually in the days following an attack. However, even during remissions the

serum amyloid A in patients with familial Mediterranean fever is usually two or three times normal (Knecht *et al*. 1985). Since serum amyloid A is a universal acute-phase protein it is found to be elevated in other inflammatory febrile diseases and cannot therefore be used as a diagnostic measure for familial Mediterranean fever. The absence of LE cells, antinuclear factor, and rheumatoid factor, or the elevation of antistreptolysin-O titres, are relevant in patients with joint involvement.

It has recently been found that the familial Mediterranean fever (*FMF*) gene in our Sephardi families is located on the short arm of chromosome 16 (Pras *et al*. 1992a; Pras *et al*. 1992b). Cloning of the *FMF* gene will elucidate the abnormal biochemical product and the pathogenesis of the febrile inflammatory episodes of this disease.

Treatment (Table 3)

Clinical characteristics	Colchicine	Additional therapies
To control febrile attacks and prevention of amyloidosis	Continuous 1–2 mg daily	
Amyloid nephropathy:		
Normal blood creatinine	2 mg daily	
Abnormal creatinine	1 mg daily	
Acute arthritis	Increase the usual dose by 0.5 mg	NSAIDs for 1–2 weeks
Protracted arthritis	1.5–2 mg	NSAIDs for months, intra-articular corticosteroids
Seronegative spondylarthropathy	1.5–2 mg	NSAIDs, DMARDs
Protracted febrile myalgia	1.5–2 mg	Corticosteroids 1 mg/kg per day

Table 3 Approach to the management of familial Mediterranean fever

Until 1973 therapeutic measures were restricted to alleviating pain. Daily prophylactic treatment with colchicine was suggested by Goldfinger (1972) and assessed by double-blind studies (Zemer *et al*. 1974). The dose required to prevent attacks is not body weight-dependent. Treatment is started with 1 mg colchicine/day, regardless of age or severity of attacks. This dose is increased if necessary to 1.5 to 2 mg, until remission from attacks is achieved. Doses larger than 1 mg must be divided in two. Our experience has shown that if doses of 2 mg/day do not produce remission, further elevation of the dose not improve responsiveness. Omission of a daily dose may be followed promptly by an attack.

Sixty-five per cent of patients enjoy complete remission of attacks if they adhere to their daily dose of colchicine. Partial remission, defined as either a significant decrease in the frequency and severity of all forms of attacks or the remission of one form (usually abdominal) but not of another (usually arthritic), is experienced by an additional 30 per cent. In 5 per cent of treated patients the attack rate remains unchanged. They are maintained on 2 mg/day to prevent amyloidosis.

Our experience showed that continuous prophylactic treatment with colchicine in patients inhibits the development of nephropathic amyloidosis (Zemer *et al*. 1986). Colchicine even reversed the nephrotic syndrome in some patients with amyloidosis (Zemer *et al*. 1992). None of the patients who started treatment without proteinuria has developed amyloidosis during the follow-up period of 21 years, while a control group of non-compliant patients showed the same rate of amyloidosis as would be expected in the natural history of the disease.

Side-effects of colchicine are generally mild. Diarrhoea and nausea are the most common, and usually prove transient and easily controllable.

In 1973, when prophylactic colchicine treatment in familial Mediterranean fever was introduced, we were worried about its possible effects on children, because of its well-known antimitotic action. It is now clear, after hundreds of children have been taking the drug daily for more than 18 years, that in none of the treated patients, including children in their first decade of life, has colchicine caused any deviation from normal in physical examination, routine laboratory tests, linear growth, or sexual development (Zemer *et al*. 1991). The data on the effect of colchicine on fertility in patients with familial Mediterranean fever are controversial (Ehrenfeld *et al*. 1986; Ehrenfeld *et al*. 1987). It is known that a high temperature may cause transient azoospermia. In many of our patients who had suffered from frequent febrile attacks, colchicine treatment improved spermatogenesis by preventing the inflammatory febrile attacks. Fertility is sometimes apparently impaired in women with familial Mediterranean fever, probably due to the induction of early miscarriages or to pelvic adhesions that develop after frequent abdominal attacks. Colchicine, by preventing the abdominal attacks, may lessen the rate of early miscarriages and the development of pelvic adhesions. Four pregnancies producing infants with Down syndrome were reported in colchicine-treated patients (Ravia *et al*. 1991; Rabinovitch *et al*. 1992; and our unpublished data). So, routine amniocentesis to exclude chromosomal aberrations in colchicine-treated patients is recommended.

A daily dose of 1.5 to 2.0 mg colchicine appears to protect the renal graft from amyloidosis (Livneh *et al*. 1992). Although several live kidney transplantations have been successfully performed in some of our patients, cadaver kidney transplant is the treatment of choice in the endstage disease of familial Mediterranean fever.

Chapter References

Alt, H. L. and Barker, H. (1930). Fever of unknown origin. *Journal of the American Medical Association*, **94**, 1457–61.

Benhamou, E., Albou, A., and Griguer, P. (1954). Remarques cliniques biologiques et therapeutiques sur la maladie periodique (a propos de 24 cas personels). *Bulletin et Memoires de la Societe Medical des Hopitaux de Paris*, **70**, 254–8.

Blum, A., Gafni, J., Sohar E., Shibolet, S., and Heller, H. (1962). Amyloidosis as the sole manifestation of familial Mediterranean fever (FMF). *Annals of Internal Medicine*, **57**, 795–9.

Dabestani, A., Noble, L. M., Child, J. S., Krivokapich, M. D., and Schwabe, M. D. (1982). Pericardial disease in familial Mediterranean fever. *Chest*, **81**, 592–5.

Ehrenfeld, M., Levy, M., Margalioth, E. J., and Eliakim, M. (1986). The effects of long-term colchicine therapy on male fertility in patients with familial Mediterranean fever. *Andrology*, **18**, 420–6.

Ehrenfeld, M., Bzezinski, A., Levy, M., and Eliakim, M. (1987). Fertility and obstetric history in patients with familial Mediterranean fever on long-term colchicine therapy. *British Journal of Obstetrics and Gynaecology*, **94**, 1186–91.

Eshel, G., Zemer, D., and Bar-Yochai. A (1988). Acute orchitis in familial Mediterranean fever. *Annals of Internal Medicine*, **109**, 164–5.

Gafni, J., Ravid, M., and Sohar, E. (1968). The role of amyloidosis in familial Mediterranean fever. A population study. *Israel Journal of Medical Sciences*, **4**, 995–9.

Glikson, M., Galun, E., Schlezinger, M., Cohen, D., Haskell, L., Rubinow, A., and Eliakim, M. (1989). Polyarteritis nodosa and familial Mediterranean fever. A report of two cases and review of the literature. *Journal of Rheumatology*, **16**, 536–9.

Goldfinger, S. E. (1972). Colchicine for familial Mediterranean fever. *New England Journal of Medicine*, **287**, 1302.

Heller, H., Sohar, E., and Sherf, L. (1958). Familial Mediterranean fever. *Archives of Internal Medicine*, **102**, 50–71.

Heller, H., Sohar, E., and Gafni, J. (1961*a*). Amyloidosis in familial Mediterranean fever. *Archives of Internal Medicine*, **107**, 539–50.

Heller, H., Sohar, E., and Pras M. (1961*b*) Ethnic distribution and amyloidosis in familial Mediterranean fever (FMF). *Pathologia et Microbiologica*, **24**, 718–32.

Ismachovich, B., Zemer, D., Revach, M., Serr, D. M., and Sohar, E. (1973) The causes of sterility in females with familial Mediterranean fever. *Sterility and Fertility*, **24**, 844–7.

Janeway, T. C. and Mosenthal, H. D. (1908). An unusual paroxysmal syndrome probably allied to recurrent vomiting with a study of the nitrogen metabolism. *Transactions of the Association of American Physicians*, **23**, 504–18.

Knecht, A., De Beer, F. C., and Pras, M. (1985). Serum amyloid A protein in familial Mediterranean fever. *Annals of Internal Medicine*, **102**, 71–2.

Langevitz, P., Livneh, A., Zemer, D., Shemer, J., and Pras M. (1994 *a*). Seronegative spondyloarthropathy (SNSA) in familial Mediterranean fever (FMF). *Arthritis and Rheumatism*, **37**, S203 (Abstr.).

Langevitz, P., Zemer, D., Livneh, A., Shemer, J., and Pras, M. (1994 *b*). Protracted febrile myalgia in patients with familial Mediterranean fever. *Journal of Rheumatology*, **21**, 1708–9.

Langevitz, P. *et al*. (1994*c*). Fibromyalgia in familial Mediterranean fever. *Journal of Rheumatology*, **21**, 1335–7.

Levine, M., Franklin, E. C., Frangione, B. and Pras M. (1972). The amino acid sequence of a major non-immunoglobulin component of some amyloid fibrils. *Journal of Clinical Investigation*, **51**, 2773–6.

Livneh, A., Zemer, D., Siegal, B., Laor, A., Sohar, E., and Pras, M. (1992). Colchicine prevents kidney transplant amyloidosis in familial Mediterranean fever. *Nephron*, **60**, 418–22.

Livneh, A., Cabili, S., Zemer, D., Rabinovitch, O., and Pras, M. (1993). Effect of pregnancy on renal function in amyloidosis of familial Mediterranean fever. *Journal of Rheumatology*, **20**, 1519–23.

Mamou, H. and Cattan R. (1952). La maladie periodique (sur 14 cas personnels dont 8 compliques de nephropathies). *La Semaine des Hopitaux e Paris*, **28**, 1062–70.

Michaeli, D., Pras, M., and Rosen, N. (1966). Intestinal strangulation complicating familial Mediterranean fever (FMF). *British Medical Journal*, **2**, 30–2.

Pras, M., Bronshpigel, N., Zemer, D., and Gafni, J. (1982). Variable incidence of amyloidosis in familial Mediterranean fever among different ethnic groups. *Johns Hopkins Medical Journal*, **150**, 22–6.

Pras, M., Zemer, D., Langevitz, P., and Sohar E. (1992 *a*). Familial Mediterranean fever: a genetic disorder prevalent in Sepharadi Jews. In *Genetic diversity among Jews. Diseases and markers at the DNA level* (ed. B. Bonne-Tamir and A. Adam), pp. 223–7. Oxford University Press.

Pras, M. *et al*. (1992*b*). Mapping of a gene causing familial Mediterranean fever to the short arm of chromosome 16. *New England Journal of Medicine*, **326**, 1509–13.

Pras, M., Langevitz, P., Livneh, A., and Zemer, D. (1996). Vasculitis in familial Mediterranean fever. In *The vasculitides* (ed. B. Ansell, P. A. Bacon, J. T. Lie, and H. Yazici), pp. 412–16. Chapman and Hall, London.

Rabinovitch, O., Zemer, D., Kukia, E., Sohar, E., and Mashiach, S. (1992). Colchicine treatment in conception and pregnancy: two hundred thirty-one pregnancies in patients with familial Mediterranean fever. *American Journal of Reproductive Immunology*, **28**, 245–6.

Ravia, Y., Aviram, A., Marom, M., Chaki, R., Bat-Miriam, M., Barkai, G., and Goldman, B. (1991). Increased rate of meiotic and somatic nondisjunction in FMF patients and their fetuses. *American Journal of Medical Genetics*, **49**, 1245.

Ravid, M. and Sohar, E. (1974). Intestinal malabsorption: first manifestation of amyloidosis in familial Mediterranean fever. *Gastroenterology*, **66**, 446–9.

Reimann, H. A. (1949) Periodic disease. Periodic fever, periodic abdominalgia, cyclic neutropenia, intermittent arthralgia, angioneurotic edema, anaphylactoid purpura and periodic paralysis. *Journal of the American Medical Association*, **141**, 175–82.

Rogers, D. B. *et al*. (1989). Familial Mediterranean fever in Armenians. *American Journal of Medical Genetics*, **34**, 168–72.

Sachs, D., Langevitz, P., Morag B., and Pras, M. (1987). Polyarteritis nodosa in familial Mediterranean fever. *British Journal of Rheumatology*, **26**, 139–41.

Said, R., Hamzeh, Y., Said, S., Tarawneh, M., and Al-Khateeb, M. (1992). Spectrum of renal involvement in familial Mediterranean fever. *Kidney International*, **41**, 414–19.

Salai, M., Langevitz, P., Blankstein, A., Zemer, D., Chechick, A., Pras, M., and Horoszowski, H. (1993). Total hip replacement in familial Mediterranean fever. *Bulletin of the Hospital for Joint Diseases*, **53**, 25–8.

Schwabe, A. D. and Peters, R. S. (1974). Familial Mediterranean fever in Armenians, analysis of 100 cases. *Medicine*, **53**, 453–62.

Siegal, S. (1945). Benign paroxysmal peritonitis. *Annals of Internal Medicine*, **22**, 1–21.

Siguier, F., Zara, M., Funck-Bretano, J. L., and Lagrue, G. (1953). Maladies periodiques a formes degradees evoluant chez plusieurs members d'une meme famille. *Bulletin et Memoires de la Societe Medical des Hopitaux des Paris*, **69**, 679–84.

Sneh, E., Pras, M., Michaeli D., Shain, N., and Gafni J. (1977). Protracted arthritis in familial Mediterranean fever. *Rheumatology and Rehabilitation*, **16**, 102–6.

Sohar, E., Pras, M., Heller, J., and Heller H. (1961). Genetics of familial Mediterranean fever. *Archives of Internal Medicine*, **107**, 529–38.

Sohar, M., Gafni, J., Pras, M., and Heller, H. (1967). Familial Mediterranean fever. A survey of 470 cases and review of literature. *American Journal of Medicine*, **43**, 227–53.

Yussim, A. *et al*. (1994). Gastrointestinal, hepatorenal and neuromuscular toxicity caused by cyclosporine-colchicine interaction in renal transplantation. *Transplantation Proceedings*, **26**, 2825–6.

Yuval, Y., Hemo-Zisser, M., Zemer, D., Sohar, E., and Pras, M. (1995). Dominant Inheritance in two families with familial Mediterranean fever. *American Journal of Medical Genetics*, **57**, 455–7.

Zemer, D. *et al*. (1974). A controlled trial of colchicine in preventing attacks of familial Mediterranean fever. *New England Journal of Medicine*, **291**, 932–4.

Zemer, D., Cabili, S., Revach, M., and Shain, N. (1977). Constrictive pericarditis in familial Mediterranean fever. *Israel Journal of Medical Sciences*, **13**, 55–8.

Zemer, D., Pras, M., Sohar, E., Modan, B., Cabili, S., and Gafni, J. (1986). Colchicine in the prevention and treatment of the amyloidosis of familial Mediterranean fever. *New England Journal of Medicine*, **314**, 1001–5.

Zemer, D., Livneh, A., Danon Y. L., Pras M., and Sohar E. (1991). Long-term colchicine treatment in children with familial Mediterranean fever. *Arthritis and Rheumatism*, **34**, 973–7.

Zemer, D., Livneh, A., and Langevitz, P. (1992). Reversal of the nephrotic syndrome by colchicine in amyloidosis of familial Mediterranean fever. *Annals of Internal Medicine*, **116**, 426.

5.13.3 Panniculitis

Jeffrey P. Callen

Panniculitis refers to inflammation within the subcutaneous fat (Patterson 1983). Panniculitis is probably a dynamic process that progresses through inflammation with neutrophils to lymphocytes to histiocytes and ends with fibrosis (Thiers 1988). The panniculitis can become granulomatous when in the histiocytic phase. The exact nature of the infiltrate perhaps depends upon when the biopsy is taken in relation to the age of the lesion being sampled. The panniculitides have been divided into four categories based on histopathological criteria: (i) septal panniculitis, (ii) lobular panniculitis, (iii) mixed with septal and lobular components, and (iv) panniculitis with vasculitis. Table 1 presents one of the currently accepted classifications for the panniculitides. Frequently the panniculitides are associated with systemic disease. Often the separation of one syndrome from another is possible only after a period of observation.

Table 1 Classification of the panniculitides

Clinical features

The prototypic septal panniculitis is erythema nodosum (Soderstrom 1982). A relatively common process, erythema nodosum is usually acute and self-limited. Erythema nodosum occurs most commonly in young adult women, but any age or sex can be affected. The typical clinical presentation is the sudden onset of one or more tender, erythematous nodules on the anterior tibial surface (Fig. 1). The nodules are deep and are better palpated than visualized. As the lesions age they may soften and develop an ecchymotic appearance. Over a 4- to 6-week period, the lesions eventually heal without scar formation. Ulceration of the primary process is extremely rare. While other symptoms may be present, they are usually those of an associated condition.

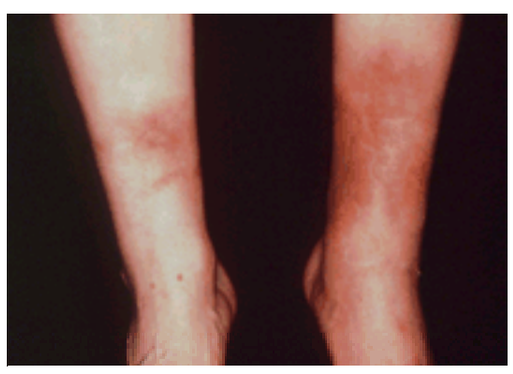


Fig. 1 Erythema nodosum—multiple tender subcutaneous erythematous nodules.

Although erythema nodosum is usually an acute process, many patients with chronic or recurrent disease have been described. Terms such as chronic erythema nodosum, erythema nodosum migrans, subacute nodular migratory panniculitis (Vilanova's disease), or septal granulomatous panniculitis have been used to characterize these patients (Prestes *et al*. 1990). The patient with erythema nodosum migrans tends to be a middle-aged (mean 45 to 50 years) woman. The disease is often present for several years, and is most common on the legs. As with acute erythema nodosum, accompanying symptoms may be present, but are usually a result of the associated condition.

Causative or associated conditions are present in about 50 per cent of patients with acute, recurrent, or chronic variants of erythema nodosum. The associated conditions can be broken into three broad categories: infectious diseases, therapeutic agents, or systemic diseases (usually inflammatory). Some of the known associations are listed in Table 2.

Infections:	
Streptococcal pharyngitis	
Tuberculosis	
Valley fever (coccidioidomycosis)	
Blastomycosis	
Histoplasmosis	
Psittacosis	
Yersinia colitis	
Salmonella gastroenteritis	
Cat scratch fever	
Leprosy	
Drugs:	
Antibiotics — penicillins, sulphonamides	
Birth control pills	
Systemic processes or diseases:	
Pregnancy	
Sarcoidosis	
Inflammatory bowel disease	
Collagen-vascular disorders — dermatomyositis, lupus erythematosus, scleroderma	
Malignancy (rare)	
Sweet's syndrome	
Idiopathic	

Table 2 Some aetiological causes of or associations with erythema nodosum

The infectious agents associated with erythema nodosum primarily tend to affect the respiratory or gastrointestinal tracts and are most often bacterial or fungal in origin. The most common drugs linked with the disease are antibiotics and oral contraceptives. Pregnancy, particularly in its second trimester, is a known association; and erythema nodosum will recur with subsequent pregnancies or with oral contraceptive use. A specific variant of sarcoidosis is associated with erythema nodosum, known as Löfgren's syndrome. This is an acute, often self-resolving variant in which erythema nodosum occurs in association with asymptomatic bilateral hilar lymphadenopathy, arthritis, and anterior uveitis. Crohn's disease (granulomatous colitis and regional enteritis) and ulcerative colitis have been associated with erythema nodosum. Patients with these inflammatory bowel diseases develop erythema nodosum that parallels the activity of the bowel disease. Panniculitis can occur with the collagen-vascular diseases (Winkelmann 1983), but it may not be best to classify the process as erythema nodosum (see below).

Weber-Christian disease

This is characterized by multiple recurrent subcutaneous nodules (Fig. 2), with accompanying fever (Panush *et al*. 1985). Histopathologically, the disease is characterized by a lobular panniculitis with an early neutrophilic infiltrate, fat degeneration, foamy histiocytes, and giant-cell formation. Eventually fibrosis occurs, and this, in addition to the destruction of fat, results in the clinical finding of an atrophic scar. Other clinical features that commonly occur are arthralgias and myalgias. Some patients also have recurrent abdominal pain. In addition to the skin lesions, any area of the body containing fat can be affected by Weber-Christian disease. Several cases of mesenteric panniculitis have been reported, as has involvement of the heart, lungs, liver, and/or kidneys (Lemley *et al*. 1991). The disease is chronic, but can result in death in 10 to 15 per cent of cases. Some of the patients with Weber-Christian disease have had multiple surgical procedures because of an acute inflammatory lesion in the presence of fever.

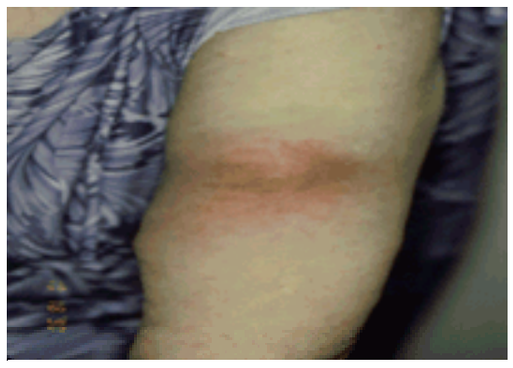


Fig. 2 Weber-Christian disease—this patient developed multiple recurrent subcutaneous lesions with accompanying fever and arthritis.

The laboratory abnormalities associated with Weber-Christian disease include an elevated sedimentation rate, anaemia, leucopenia or leucocytosis, depression of complement components, and evidence of circulating immune complexes. There have been several reports of a $_1$-antitrypsin deficiency in patients with Weber-Christian disease (Breit *et al*. 1983), however, the meaning of this finding is not clear. In addition, patients who have lupus erythematosus or pancreatic disease have been diagnosed with this form of panniculitis; however, their accompanying laboratory abnormalities would be those associated with the primary process.

Panniculitis associated with a$_1$-antitrypsin deficiency

Several groups of patients with a lobular or septal panniculitis have been ound to have a deficiency of a $_1$-antitrypsin (Smith *et al*. 1987). In a study of 96 patients with panniculitis, Smith *et al*. (1989) found 15 patients with a$_1$-antitrypsin deficiency. There were differences in the clinical and histopathological manifestations, but very little difference in the associated conditions present in either group. Specifically, the group with a $_1$-antitrypsin deficiency was more likely to have ulceration and drainage, and correspondingly had greater amounts of fat necrosis and destruction of elastic tissue. Geller and Su (1994) have suggested that the finding of splaying of neutrophils between collagen bundles is highly suggestive of this form of panniculitis. Furthermore, Smith *et al*. (1987) believed that induction of lesions by trauma was more likely in the enzyme deficient group. The recognition of these patients may be important on several grounds. First, debridement of the lesion should be avoided; second, in patients felt to have factitial panniculitis, a $_1$-antitrypsin deficiency should be considered; third, these patient should be evaluated for pulmonary disease and should be counselled to avoid smoking; and fourth, therapy with a $_1$-proteinase inhibitor concentrate may be helpful. A recent report documented the effectiveness of doxycycline and postulated that its effects were mediated through anticollagenase properties (Humbert *et al*. 1991).

Pancreatic panniculitis

Some patients with pancreatic diseases develop subcutaneous fat necrosis (lobular panniculitis) (Fig. 3), with accompanying polyarthritis and osseous intramedullary fat necrosis (Wilson *et al*. 1983). A variety of changes have been implicated in the development of this process including pancreatitis, pancreatic carcinoma (acinar cell), pancreatitis secondary to cholelithiasis, post-traumatic lesions, pancreatic ischaemia, pancreatic pseudocyst (Zimmerman-Gorska *et al*. 1986), and a pancreatic difusum (a congenital pancreatic abnormality) (Huber and Asaad 1986). It is not clear whether the elevated lipase in the circulation is primarily involved in the pathogenesis of the fat necrosis, or whether its presence follows the fat necrosis (Simkin *et al*. 1983). Histopathologically there is extensive fat necrosis with a basophilic alteration of lipocytes. Ghost cells with absent nuclei are also common. Treatment centres on non-specific measures for control of the panniculitis (see below) and reduction of the pancreatic inflammation or removal of a pancreatic tumour.

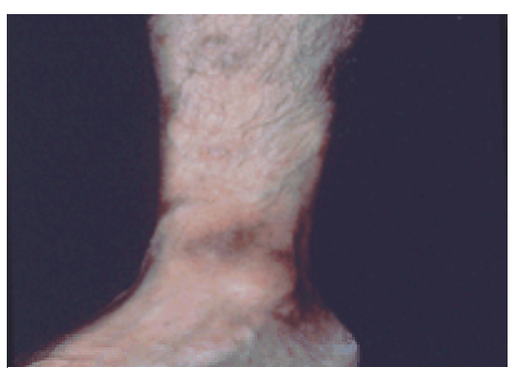


Fig. 3 Pancreatic panniculitis. Tender, erythematous subcutaneous nodules in a patient with pancreatitis (by courtesy of Robert Schosser, MD of Lexington, KY, USA).

Calcifying panniculitis of renal failure

Patients with renal failure often have abnormal calcium–phosphorus metabolism. Rarely, these patients develop acute, erythematous, tender indurated nodules as a manifestation of calciphylaxis (Lugo-Somolinos *et al*. 1990; Ivker *et al*. 1995). The panniculitic lesions can progress to necrosis and ulceration (Fig. 4). Laboratory evaluation reveals a normal calcium level and hyperphosphataemia with an elevated calcium–phosphorus product. This condition must be differentiated from metastatic calcification and pancreatic panniculitis. The prognosis is poor, and treatment is aimed at a correction of the calcium–phosphorus imbalance. Parathyroidectomy may at times be helpful.

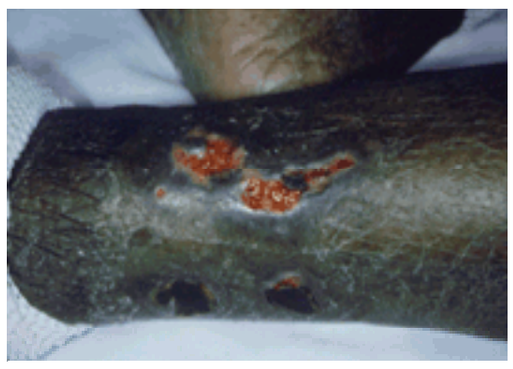


Fig. 4 Calcifying panniculitis (calcifylaxis) in this patient with renal failure. The reaction has resulted in necrosis and ulceration of the nodular lesions on the lower extremity.

Post-steroid panniculitis

Panniculitis following withdrawal of corticosteroid therapy is a rare entity that seems to be limited to children (Roenigk *et al*. 1964). Patients reported have been treated with corticosteroids for a wide array of problems including leukaemia, nephrotic syndrome, rheumatic carditis, and encephalopathy. Interestingly, the panniculitis may clear upon readministration of the corticosteroids. The pathogenesis of this rare complication is not understood.

Lipoatrophic panniculitis

Several conditions have been described in children that often result in lipoatrophy following the inflammatory reaction. There exists a spectrum which perhaps includes Rothman-Makai syndrome (lipogranulomatosis subcutanea), lipoatrophic panniculitis, lipophagic panniculitis of childhood (Winkelmann *et al*. 1989), and localized lipoatrophy (atrophic connective tissue disease panniculitis) (Peters and Winkelmann 1980). These children tend to have multiple erythematous lesions, most commonly on the extremities, which resolve with subcutaneous atrophy (Fig. 5) (Roth *et al*. 1989). The patients often are febrile. They may have associated 'autoimmune' phenomena such as juvenile chronic arthritis, Hashimoto's thyroiditis, or diabetes mellitus (Billings *et al*. 1987). There is no known effective therapy, but some patients have responded to oral corticosteroids, oral antimalarials, or oral dapsone.

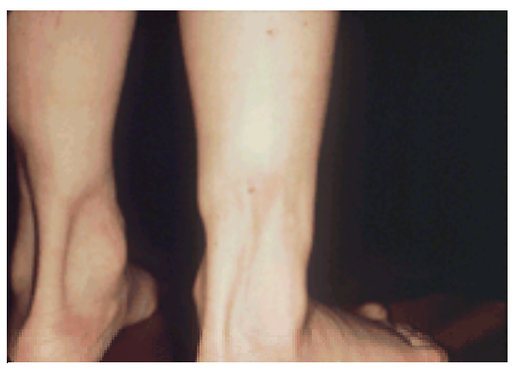


Fig. 5 Lipoatrophy following an active panniculitis of unknown cause in this child.

Histiocytic cytophagic panniculitis

Histiocytic cytophagic panniculitis was described by Crotty and Winkelmann (1981) as a chronic histiocytic disease of the subcutaneous fat, with accompanying inflammatory panniculitis, fever, serositis, and 'reticuloendotheliomegaly'. This is an extremely rare entity, having been reported in a small number of patients. The process may be a primary skin disorder of unknown cause, but has also been linked to neoplastic processes such as lymphoma and malignant histiocytosis (Barron *et al*. 1985). Histopathologically, the fat contains both T cells and histiocytes (Perniciaro *et al*. 1994). Haemorrhagic complications, perhaps due to thrombocytopenia, have occurred in half of the patients (Crotty and Winkelmann 1981). Alegre and Winkelmann (1989) have suggested that aggressive therapy with cytotoxic agents be used early in this process.

Factitial panniculitis

Factitial panniculitis from external trauma or from the injection of foreign substances is not uncommon, and should be considered in any patient with panniculitis and unusual clinical or histopathological features (Fig. 6). In traumatic lesions an organizing haematoma is often demonstrated histopathologically; whereas with the injection of foreign material, refractile bodies, or a 'Swiss cheese' effect are encountered. Occasionally spectroscopic and/or chromatographic techniques are necessary to identify the causative injected material.

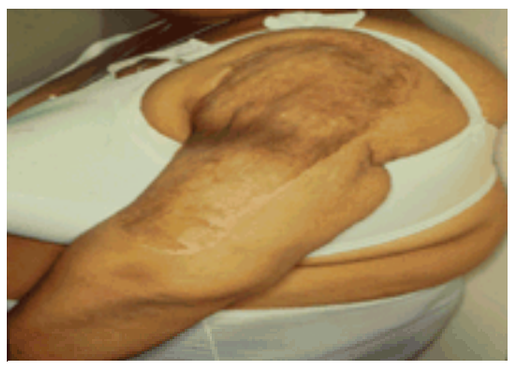


Fig. 6 Factitial panniculitis—this patient developed multiple lesions, which were surgically drained or excised; only after careful inspection of several specimens by polarized microscopy was refractile material found.

Lupus erythematosus panniculitis (lupus profundus)

This is a rare manifestation of chronic cutaneous lupus erythematosus occurring in less than 3 per cent of cases of systemic lupus erythematosus, and less than 1 per cent of cases with cutaneous lupus erythematosus (Izumi 1985). The lesions are tender, red-blue, subcutaneous nodules (Fig. 7) that may eventually ulcerate or atrophy. Calcification may occur in late lesions (Fig. 8) The lesions tend to occur on the face, upper arms, and/or buttocks. The lesions may underlie a typical lesion of discoid lupus erythematosus. Trauma may initiate or worsen these lesions. Facial lesions are common, and in a recent report, lupus panniculitis accounted for periparotid swelling and was confused clinically with a possible neoplastic process (White *et al*. 1993). Histopathological changes include a panniculitis which is both lobular and septal. The overlying epidermis and dermis often demonstrate changes of lupus erythematosus, in which case the diagnosis of lupus erythematosus panniculitis can be histopathologically confirmed. Although one-half of the patients have four or more of the criteria for systemic lupus erythematosus, the activity of the panniculitic lesion does not seem to follow the course of the systemic disease (Tuffanelli 1982).

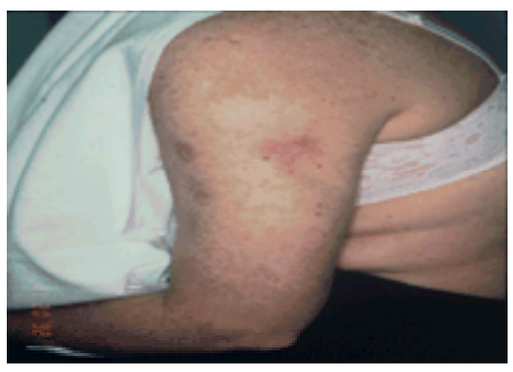


Fig. 7 Lupus erythematosus panniculitis. Erythematous to violaceous nodules with resultant atrophy of the subcutaneous tissue.

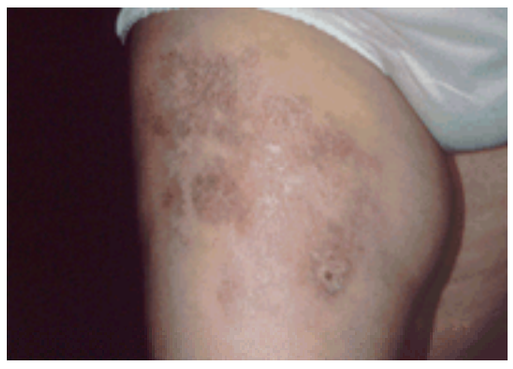


Fig. 8 Calcified subcutaneous nodules of lupus erythematosus.

Sclerosing panniculitis

Jorizzo *et al*. (1991) coined the term sclerosing panniculitis to describe a group of patients with well-circumscribed, indurated, inflammatory plaques of the lower extremity (Fig. 9). These lesions most frequently occur in women and are often accompanied by signs of venous insufficiency. There may be a history of prior thrombophlebitis (Alegre *et al*. 1988). Histopathologically this disorder is characterized by fat necrosis, sclerosis, and a lobular panniculitis. Fat microcysts with foci of membranous fat necrosis are also commonly observed in the later stages of the disease. Sclerosing panniculitis is probably a manifestation of venous insufficiency and thus therapy should include support stockings, elevation, and rest. Measures for the prevention of phlebitis are also warranted. Low-dose aspirin or other non-steroidal anti-inflammatory drugs may be helpful. Intralesional injection of triamcinolone acetonide may also be helpful.

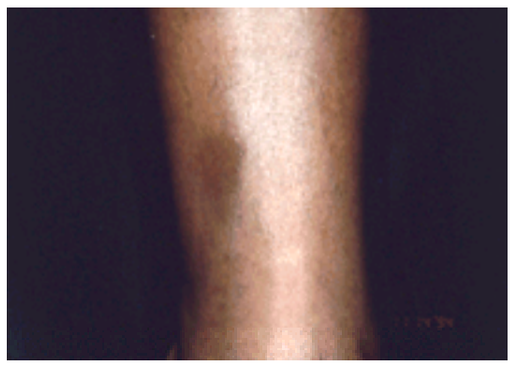


Fig. 9 Sclerosing panniculitis. Tender, subcutaneous nodule in a patient with peripheral oedema.

Panniculitis with vasculitis

Nodular vasculitis, polyarteritis nodosa (both cutaneous and systemic varieties), and small vessel vasculitis may involve subcutaneous vessels and result in

inflammatory or ischaemic changes in the subcutaneous fat. Erythema induratum is a form of nodular vasculitis thought to be due to tuberculosis.

Evaluation

The evaluation of the patient with panniculitis should include a careful history, physical examination, and a deep incisional biopsy. Findings from this examination will usually allow appropriate classification. Patients with erythema nodosum should be screened for the possibility of infection of the upper respiratory tract, and a throat swab for a rapid streptococcal screen, a skin test with PPD (purified protein derivative), and a chest radiograph should be obtained. Inflammatory bowel disease or infectious enteritis are usually symptomatic and thus it is rarely necessary to perform endoscopic or radiographic procedures in these patients. In patients in whom Weber-Christian disease or pancreatic panniculitis are being considered, tests for enzymatic abnormalities such as amylase, lipase, or a $_1$-antitrypsin deficiency should be ordered. In all patients the possibility of a coexistent collagen vascular disease should be considered.

Differential diagnosis

The differential diagnosis of erythema nodosum involves distinguishing it from (i) other forms of panniculitis (all of which may present with tender subcutaneous nodules), (ii) insect bites, (iii) thrombophlebitis, or (iv) cellulitis (Black 1988). Insect bites may be tender, red, infiltrated lesions. The distribution is dependent upon the type of bite, and in general the patient lacks associated symptomatology. Often there is a puncture site within the lesion. Cellulitis rarely manifests as subcutaneous nodules. The lesion of cellulitis may be tender and erythematous, but often there is an accompanying fever and a lymphangitic streak ascending from the primary area of involvement. When the diagnosis is in question, a wedge biopsy or deep punch biopsy which includes subcutaneous fat will be helpful.

The differential diagnosis of Weber-Christian disease includes other lobular panniculitides, erythema nodosum, and lupus erythematosus panniculitis (lupus profundus). Erythema nodosum can be separated on clinical and histopathological grounds. Lupus profundus occurs in conjunction with typical skin lesions of lupus erythematosus or in patients with systemic lupus erythematosus. This latter group may be difficult to distinguish from Weber-Christian disease, particularly in view of the reports of Weber-Christian disease in some patients with systemic lupus erythematosus. Perhaps the diagnosis of Weber-Christian disease should be made only in the absence of associated disorders such as lupus erythematosus, scleroderma, dermatomyositis, malignancy, pancreatic disease, sarcoidosis, or factitial panniculitis (Diesk and Panush 1981).

Treatment

The treatment of erythema nodosum and other panniculitides first involves assessment of a causative disease and its treatment. In the absence of a treatable disorder, therapy is symptomatic (Callen 1985). Acute erythema nodosum is often self-limited, thus non-toxic therapies are advised. Bed rest and leg elevation are very helpful in controlling symptomatology. In patients who need to continue to be ambulatory, support stockings or tights may be helpful (Lehman 1980). Aspirin or other non-steroidal anti-inflammatory agents may be helpful. My experience with aspirin has not produced results prior to toxicity and therefore my choice is to use oral indomethacin at a dosage of 25 to 75 mg per day (Barr and Robinson 1981).

In patients with chronic erythema nodosum or frequent recurrences, oral potassium iodide at a dosage of 300 to 900 mg per day has been useful in open clinical trials. My experience with 15 patients with erythema nodosum treated with oral potassium iodide has yielded control of the disease in 10 patients (unpublished observation). Furthermore, in some of the responding patients, when the drug is stopped or the dosage is lowered, the disease has relapsed, only to respond again with reinstitution of therapy. Other therapies which may be considered include oral corticosteroids, colchicine (Lupton *et al*. 1988), or an immunosuppressive agent. Systemic corticosteroids which are almost always effective, are often complicated by iatrogenic Cushing's syndrome. Colchicine has not been effective in any of the patients that I have treated. Azathioprine and cyclosporin have been used successfully in some patents. One adjunctive therapy that can be helpful is intralesionally injected triamcinolone acetonide (3 to 5 mg/ml).

There is no specific therapy for Weber-Christian disease. Reports have centred on the use of anti-inflammatory agents including aspirin, non-steroidal anti-inflammatory drugs, oral corticosteroids, antimalarials, and immunosuppressives including cyclophosphamide and cyclosporin (Usuki *et al*. 1988). In addition, colchicine, dapsone, and potassium iodide may be effective in individual cases.

Patients with lupus erythematosus panniculitis may respond to antimalarials or to intralesional injections of triamcinolone. Corticosteroids and/or immunosuppressives are rarely necessary for lupus erythematosus panniculitis.

Conclusion

Panniculitides form a wide array of syndromes that perhaps can be separated on the basis of clinical features, associated disorders, and/or histopathological features. Unfortunately, several of the syndromes have overlapping features, the pathogenesis is not understood for most, and therapeutic options are similar for all. The differential diagnosis often depends on an adequate specimen for histopathological investigation. Thus a deep, fusiform, incisional biopsy or a wedge biopsy, often including fascia, should be performed. Furthermore, a biopsy of the youngest lesion should be taken. Sections should be serially cut to identify the pattern of the panniculitis: lobular, septal, vasculitis, or mixed. The patient should be tested for underlying processes including ingestants, infections, malignancy, and autoimmune disorders. Treatment, in the absence of an underlying disease, is aimed at control of the inflammatory reaction with agents such as non-steroidal anti-inflammatory drugs, corticosteroids, dapsone, potassium iodide, antimalarials, colchicine, or immunosuppressive agents.

Chapter References

Alegre, V.A. and Winkelmann, R.K. (1989). Histiocytic cytophagic panniculitis. *Journal of the American Academy of Dermatology*, **20**, 177–85.

Alegre, V.A., Winkelmann, R.K., and Aliaga, A. (1988). Lipomembranous changes in chronic panniculitis. *Journal of the American Academy of Dermatology*, **131**, 241–5.

Barr, W.G. and Robinson, J.A. (1981). Chronic erythema nodosum treated with indomethacin. *Annals of Internal Medicine*, **95**, 661.

Barron, D.R., Davis, P.R., Pomeranz, J.R., Hines, J.D., and Park, C.H. (1985). Cytophagic histiocytic panniculitis. *Cancer*, **29**, 2538–42.

Billings, J.K., Milgramm, S.S., Gupta, A.K., Headington, J.T., and Rasmussen, J.E. (1987). Lipoatrophic panniculitis: a possible autoimmune inflammatory disease of fat. *Archives of Dermatology*, **123**, 1662–6.

Black, M.M. (1988). Panniculitis: problems with diagnosis. *Australasian Journal of Dermatology*, **29**, 79–84.

Breit, S.N., Clark, P., Robinson, J.P., Luckhurst, E., Daekins, R.L., and Penny, R. (1983). Familial occurrence of a $_1$-antitrypsin deficiency and Weber-Christian disease. *Archives of Dermatology*, **119**, 198–202.

Callen, J.P. (1985). Erythema nodosum. In *Current therapy in dermatology*, (ed. T.T. Provost and E.R. Farmer), pp. 159–60. Mosby, St Louis.

Crotty, C. and Winkelmann, R.K. (1981). Cytophagic histiocytic panniculitis with fever, cytopenia, liver failure, and terminal hemorrhagic diathesis. *Journal of the American Academy of Dermatology*, **4**, 181–94.

Diesk, A. and Pannush, R.S. (1981). Factitial Weber-Christian disease: a case report. *Journal of Rheumatology*, **8**, 129–32.

Geller, J.D. and Su, W.P.D. (1994). A subtle clue to the histopathologic diagnosis of early a $_1$-antitrypsin deficiency panniculitis. *Journal of the American Academy of Dermatology*, **31**, 241–5.

Huber, R.M. and Asaad, D.M. (1986). Panniculitis associated with a pancreas difusum. *Journal of the American Academy of Dermatology*, **14**, 331–4.

Humbert, P., Faivre, B., Gibey, R., and Agache, P. (1991). Use of anti-collagenase properties of doxicycline in treatment of a $_1$-antitrypsin deficiency panniculitis. *Acta Dermatologica and Venereology*, **71**, 194–8.

Ivker, R.A., Woosley, J., and Briggaman, R.A. (1995). Calciphylaxis in three patients with end-stage renal disease. *Archives of Dermatology*, **131**, 63–8.

Izumi, A.K. (1985). Lupus erythematosus panniculitis. *Clinics in Dermatology*, **3**, 69–78.

Jorizzo, J.L., White, W.L., Zanolli, M.D., Greer, K.E., Solomon, A.R., and Jetton, R.L. (1991). Sclerosing panniculitis: a clinicopathologic assessment. *Archives of Dermatology*, **127**, 544–8.

Lehman, D.W. (1980). Control of chronic erythema nodosum with naproxen. *Cutis*, **26**, 66–7.

Lemley, D.E. *et al*. (1991). Cardiac manifestations of Weber-Christian Disease: report and review of the literature. *Journal of Rheumatology*, **18**, 756–60.

Lugo-Somolinos, A., Sánchez, J.L., Méndez-Coll, J., and Joglar, F. (1990). Calcifying panniculitis associated with polycystic kidney disease and chronic renal failure. *Journal of the American Academy of Dermatology*, **22**, 743–7.

Lupton, G.P., Slagel, G.A., and Distelmeier, M.R. (1988). Treatment of nodular vasculitis with colchicine. *Journal of the Association of Military Dermatology*, **14**, 24–6.

Panush, R.S., Vonker, R.A., Diesk, A., Langley, S., and Caldwell, J.R. (1985). Weber-Christian disease. *Medicine*, **64**, 181–90.

Patterson, J.W. (1983). Panniculitis: a guide to differential diagnosis. *Journal of the Association of Military Dermatology*, **9**, 36–9.

Perniciaro, C., Winkelmann, R.K., and Ehrhardt, D.R. (1994). Fatal systemic cytophagic histiocytic panniculitis: a histopathological and immunohistochemical study of multiple organ sites. *Journal of the American Academy of Dermatology*, **31**, 901–5.

Peters, M.S. and Winkelamnn, R.K. (1980). Localized lipoatrophy (atrophic connective tissue disease panniculitis). *Archives of Dermatology*, **116**, 1363–8.

Prestes, C.A., Winkelmann, R.K., and Su, W.P.D. (1990). Septal granulomatous panniculitis: comparison of the pathology of erythema nodosum migrans (migratory panniculitis) and chronic erythema nodosum. *Journal of the American Academy of Dermatology*, **22**, 477–83.

Roenigk, H.H., Haserick, J.R., and Arundell, F.D. (1964). Post steroid panniculitis. Report of a case and review of the literature. *Archives of Dermatology*, **90**, 387–91.

Roth, D.E., Schickler, K., and Callen, J.P. (1989). Annular atrophic connective tissue panniculitis of the ankles. *Journal of the American Academy of Dermatology*, **21**, 1152–6.

Simkin, P.A., Brunzell, J.D., Wisner, D., Fiechter, J.R., Cardin, J.S., and Wilkins, R. F. (1983). Free fatty acids in the pancreatic arthritis syndrome. *Arthritis and Rheumatism*, **26**, 127–32.

Smith, K.C., Pittelkow, M.R., and Su, W.P.D. (1987). Panniculitis associated with severe a-1-antitrypsin deficiency. *Archives of Dermatology*, **123**, 1655–61.

Smith, K.C., Su, W.P.D., Pittelkow, M.R., and Winkelmann, R.K. (1989). Clinical and pathologic correlations with 96 patents with panniculitis, including 15 patients with deficient levels of a_1-antitrypsin. *Journal of the American Academy of Dermatology*, **21**, 1192–6.

Soderstrom, R.M. (1982). Erythema nodosum: an update. *Current Concepts in Skin Diseases*, Spring, pp. 3–7.

Thiers, B.H. (1988). Panniculitis. In *Common problems in dermatology* (ed. K.E. Greer), pp. 267–71. Yearbook Medical Publishers, Inc., New York.

Tuffanelli, D.L. (1982). Lupus erythematosus (panniculitis) profundus. A classic revisited commentary and report of 22 cases. *Hawaii Medical Journal*, **41**, 394–7.

Usuki, K., Kitamura, K., Urabe, A., and Takakin, F. (1988). Successful treatment of Weber–Christian disease by cyclosporin A. *American Journal of Medicine*, **85**, 276–8.

White, W.L., Sherertz, E.F., Berg, D., and Clark, R.E. (1993). Periparotid lupus erythematosus panniculitis. *Archives of Pathology and Laboratory Medicine*, **117**, 535–9.

Wilson, H.A., Askar, A.D., Neiderhizer, D.H., Johnson, M., Andrew, B.S., and Hoskins, L.C. (1983). Panniculitis with atrophy and subcutaneous fat necrosis. *Arthritis and Rheumatism*, **26**, 121–6.

Winkelmann, R.K. (1983). Panniculitis in connective tissue diseases. *Archives of Dermatology*, **119**, 336–44.

Winklemann, R.K., McEvoy, M.T., and Peters, M.S. (1989). Lipophagic panniculitis of childhood. *Journal of the American Academy of Dermatology*, **21**, 971–8.

Zimmerman-Gorska, I., Urbaniak, M., and Karwowski, A. (1986). Coexistence of arthritis, subcutaneous fat necrosis, and pseudocyst of pancreas. *Rheumatology International*, **6**, 45–58.

5.13.4 Neutrophilic dermatoses

Jeffrey P. Callen

The neutrophilic dermatoses are a group of non-infectious disorders characterized by the presence of an angiocentric, vessel-based, primary neutrophilic inflammatory cell infiltrate (Table 1) (Jorizzo 1988). These disorders can be further divided into those which lead to destruction of vessel walls (vasculitis) and those that do not destroy the vessels. The disorders to be discussed in this chapter will be those in which the vessel wall is not destroyed. For a full discussion of Behçet's disease see Chapter 5.11.7, the vasculitides see Chapter 5.11.1,Chapter 5.11.2,Chapter 5.11.3,Chapter 5.11.4,Chapter 5.11.5,Chapter 5.11.6,Chapter 5.11.7 and Chapter 5.11.8, and familial Mediterranean fever see Chapter 5.13.2. The remaining disorders are linked by the presence of similar associated processes, massive cutaneous neutrophilic infiltrates, occasionally overlapping clinical features, and similar approaches to therapy.

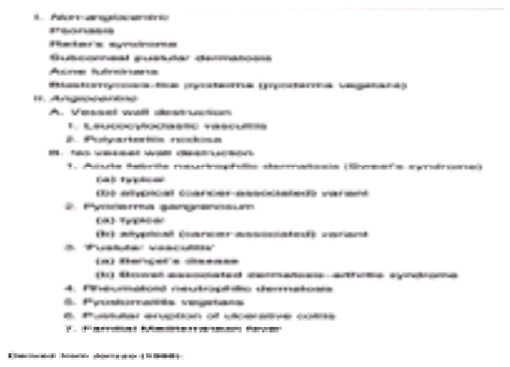


Table 1 Non-infectious neutrophilic dermatoses

Sweet's syndrome (acute febrile neutrophilic dermatosis)

In 1964, R.D. Sweet described a group of patients with one or more attacks of painful, erythematous plaques, accompanied by fever, arthralgias, and leucocytosis (Sweet 1964). The histopathological correlate was a massive neutrophilic infiltration of the dermis in the absence of vessel wall destruction or demonstrable infection. Sweet termed the process acute febrile neutrophilic dermatosis, but it has become known as Sweet's syndrome. Since Sweet's original description, more than 500 cases have been reported (von den Driesch *et al*. 1994).

This syndrome is more frequent in women (female:male, 3.7:1) between the ages of 30 and 70 years (mean age 52.6 years). Children with Sweet's syndrome have also been reported (Boatman *et al*. 1994). The disease may be preceded by symptoms suggestive of an upper respiratory tract infection. The skin lesions are felt to be distinctive, but may simulate several other processes. The characteristic lesion is a well defined, erythematous plaque with a mamillated surface, which may give the clinical impression of microvesiculation (Fig. 1). There is rarely any accompanying epidermal change or ulceration, and the lesions usually heal without scar formation. Pustules may stud the surface or may be a major feature of the process. A relatively common clinical variant is a tender, erythematous nodule which clinically resembles erythema nodosum (Cohen *et al*. 1992). Genital lesions have been reported, but are rare (Banet *et al*. 1994). The lesions occur in crops, and may be initiated by a variety of traumatic injuries (pathergy) such as a needle stick, wound debridement, or burn (Fig. 2). The lesions are accompanied by fever and malaise in most patients, and myalgias and/or arthralgias in about half. Headache, nausea, vomiting, diarrhoea, and/or conjunctivitis may also occur in some patients. Untreated lesions resolve over 6 to 8 weeks, however, many patients continue to produce new lesions chronically or recurrently.

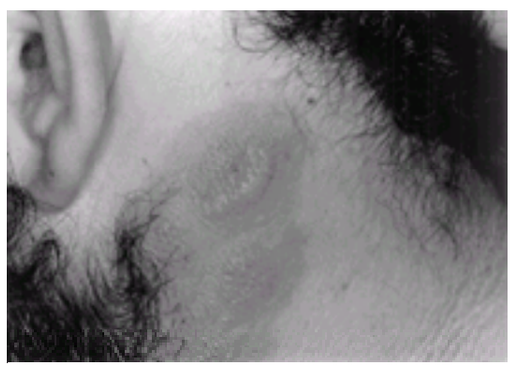


Fig. 1 Sweet's syndrome. This patient developed erythematous plaques with a central mamillated 'microvesicular' surface.

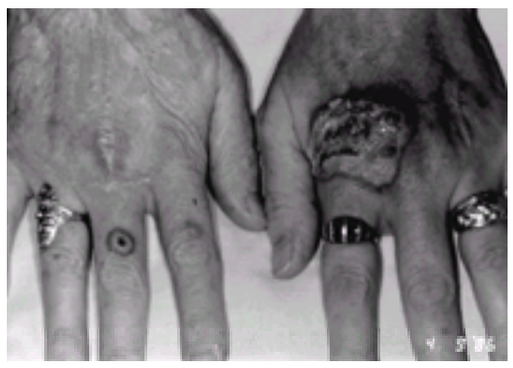


Fig. 2 This patient's pustular lesion on an erythematous base was initiated after a kitchen accident in which boiling water splattered on her hand.

The laboratory findings include a leucocytosis which is composed of mature neutrophils. White blood cell counts generally range from 10 to 20 000 cells/mm ^{3}. The remainder of the blood count is within normal limits, except in patients who have leukaemia-associated Sweet's syndrome. The erythrocyte sedimentation rate is

frequently elevated. On rare occasions a patient may have proteinuria. Histopathologically there is a dense dermal infiltrate composed of mature polymorphonuclear leucocytes. The infiltrate may be more pronounced in perivascular areas, and leucocytoclasia is frequent, but the vessel walls are spared. Oedema in the papillary dermis may be intense coinciding with the microvesicular lesions observed clinically. Immunofluorescence microscopy has been negative in a small number of cases in which it has been reported.

Sweet's syndrome has been reported with a variety of diseases. Von den Driesch *et al*. (1994) has suggested that Sweet's syndrome can be subdivided into four groups: (i) classic or idiopathic, (ii) parainflammatory, (iii) paraneoplastic, and (iv) pregnancy associated (Cohen *et al*. 1993). While idiopathic cases are the most frequent, paraneoplastic Sweet's syndrome is the most frequently identified association and myelogenous leukaemia or preleukaemia account for most of the paraneoplastic conditions (Cooper *et al*. 1983; Cohen *et al*. 1988) Sweet's syndrome is not clinically or histopathologically different among the four groups; however, in the presence of leukaemia, the patients more frequently tend to be anaemic or thrombocytopenic and may have haemorrhagic lesions (Fig. 3). Other associated processes in paraneoplastic Sweet's syndrome include benign monoclonal gammopathy, lymphoma, myelodysplastic disorders, and various solid tumours (including breast, stomach, genitourinary, and colon most commonly) (Cohen *et al*. 1993). Parainflammatory Sweet's syndrome has also been reported in conjunction with lupus erythematosus, rheumatoid arthritis (Trentham *et al*. 1976; Harary 1983), Sjögren's syndrome (Prystowsky *et al*. 1978), inflammatory bowel disease (Crohn's disease and ulcerative colitis), Behçet's disease, thyroiditis, various infections (including HIV, hepatitis, mycobacteria, cytomegalovirus, salmonella), and drug hypersensitivity (Johnson and Grimwood 1994; Piette *et al*. 1994). It is not known how frequently paraneoplastic Sweet's syndrome occurs, but most authorities quote a figure ranging from 10 to 20 per cent, with about 40 to 50 per cent of these patients having a haematological malignancy.

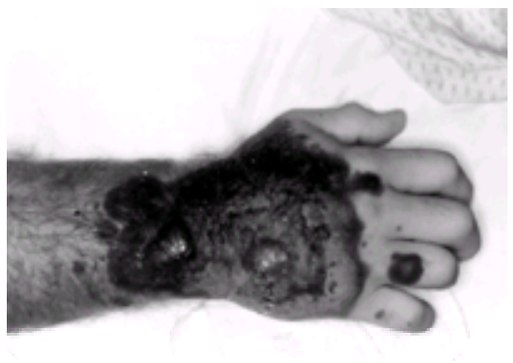


Fig. 3 Sweet's syndrome in a patient with acute myelogenous leukaemia. The purpura is due to bleeding into the lesions associated with thrombocytopenia.

The pathogenesis of Sweet's syndrome is not known. Tests for circulating immune complexes, tissue-bound immunoglobulins, or complement have generally been negative. Kemmett and Hunter (1990) reported perinuclear antineutrophil cytoplasmic antibodies (**p-ANCA**) in 6 patients with Sweet's syndrome, but believed this to be an epiphenomenon. Von den Driesch (1994) was unable to demonstrate p- or c-ANCA in any of his 10 patients who were tested. Studies of neutrophil function have not shown a consistent abnormality. Furthermore, abnormalities of T cells and proinflammatory cytokines such as g-interferon or interleukin 8 have not been reproducibly reported.

Sweet's syndrome is usually an acute, steroid-responsive, self-limited disease. In general, a 2-week tapering course of oral prednisone (40 to 60 mg/day) is effective. One or more exacerbations requiring brief reinstitution of corticosteroids are common. From Sweet's initial report and the many later ones, it appears that the process can follow a chronic course, and the use of steroid-sparing agents should be considered. In reports of individual or small groups of patients, dapsone, potassium iodide, indomethacin, doxycycline, clofazimine, colchicine, metronidazole, isotretinoin, methotrexate, chlorambucil, cyclosporin, and pulse dosage of methylprednisolone have been successfully used (Hoffman 1977; Horio *et al*. 1980; Subhisa *et al*. 1983; Aram 1984; Case *et al*. 1989; Banet *et al*. 1994; von den Driesch 1994)

Pyoderma gangrenosum

Pyoderma gangrenosum is an uncommon, ulcerative, cutaneous condition with distinctive clinical characteristics (Callen 1990). Frequently there is an associated systemic disease. The diagnosis is made by exclusion of other processes that may cause cutaneous ulcers. Like patients with Sweet's syndrome, patients with pyoderma gangrenosum are often pathergic, with lesions sometimes developing after minor trauma.

The ulcerations of classical pyoderma gangrenosum are frequently clinically characteristic. The border is well defined with a deep erythematous to violaceous colour (Fig. 4). The lesion extends peripherally and often the border overhangs the ulceration (undermined) as the inflammatory process spreads within the dermis, only secondarily causing necrosis of the epidermis. The lesions may be single, or may occur in crops, often beginning as a discrete pustule with a surrounding inflammatory erythema. The lesions may occur on any surface, but are most common on the legs. Pain is a prominent feature and is sometimes so severe that narcotics are required for symptomatic relief. As the lesion heals, scar formation occurs and the resulting scar is often described as cribiform (Fig. 5).

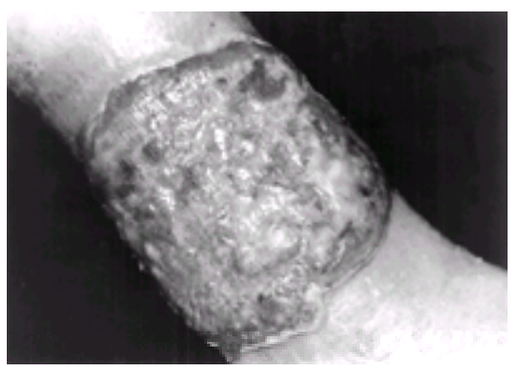


Fig. 4 Pyoderma gangrenosum. Typical large ulceration with an undermined violaceous border.

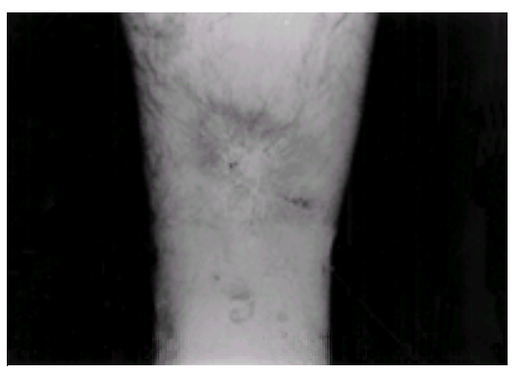


Fig. 5 Healed lesion of pyoderma gangrenosum.

Several variants of pyoderma gangrenosum have been described. The pustular eruption of ulcerative colitis was first reported by O'Loughlin and Perry (1978). In this process the patient is acutely ill with fever and develops multiple sterile pustules (Fig. 6) (Fenske *et al*. 1983; Callen and Woo 1985). The lesions may regress without scarring, or some may progress into a typical lesion of pyoderma gangrenosum. Biopsy of the early lesion reveals sheets of mature polymorphonuclear leucocytes.

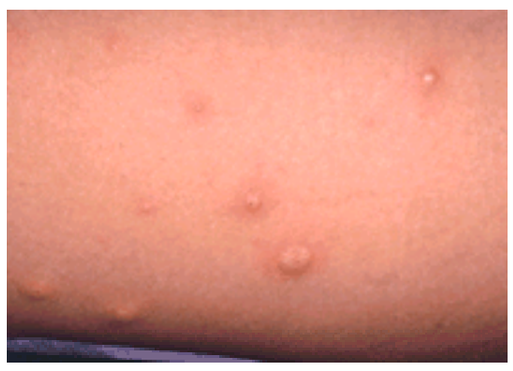


Fig. 6 Vesiculopustular eruption of ulcerative colitis.

Peristomal pyoderma gangrenosum is a recently recognized variant that occurs in patients with ulcerative colitis or Crohn's disease who have had abdominal surgery and have an ileostomy or colostomy (Keltz *et al*. 1993). The ulceration (Fig. 7) may occur as an early or late phenomenon. Perhaps irritation from the ileostomy or colostomy appliance is involved in the induction of this process (pathergy). These ulcerations must be differentiated from infections, dermatitis, or extension of the underlying bowel disease (Crohn's only).

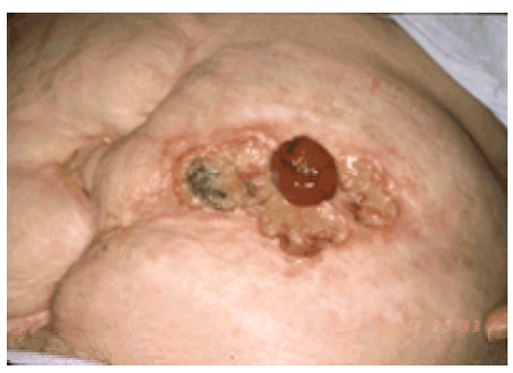


Fig. 7 Peristomal pyoderma gangrenosum in a patient with an ileostomy after bowel surgery for ulcerative colitis.

Vulvar pyoderma gangrenosum is another recently recognized variant (McCalmont *et al*. 1991). Except for its location, the ulceration is otherwise typical of pyoderma gangrenosum (Fig. 8). This variant should be differentiated from Behçet's disease.

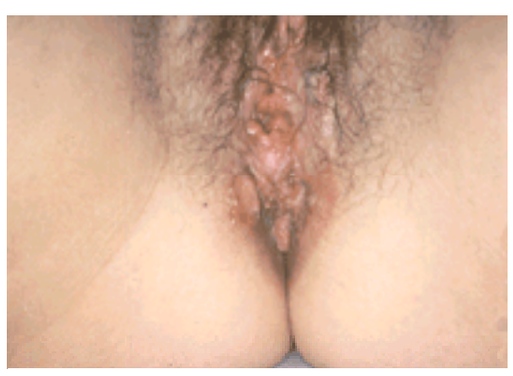


Fig. 8 Vulvar pyoderma gangrenosum.

Another variant is pyostomatitis vegetans. This process is one in which chronic, pustular, eventually vegetative erosions develop on the mucous membranes (Fig. 9), most notably in the oral cavity (Van Hale *et al*. 1985). Most of these patients have had inflammatory bowel disease, and some have had ulcerative skin lesions similar to pyoderma gangrenosum (Storwick *et al*. 1994).

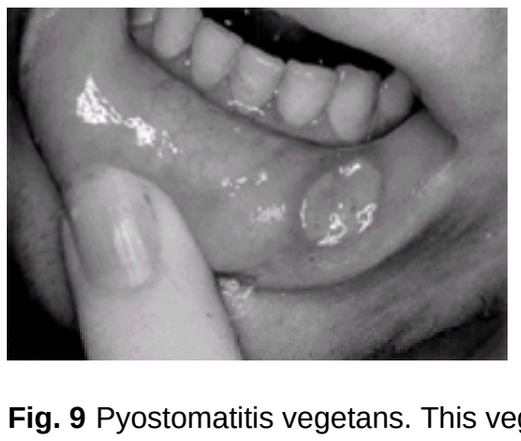


Fig. 9 Pyostomatitis vegetans. This vegetative ulceration occurred in a patient with ulcerative colitis.

A condition known as malignant pyoderma is distinguished from pyoderma gangrenosum by three features: (i) lesions predominantly on the head and neck (atypical for pyoderma gangrenosum), (ii) lack of associated systemic diseases, and (iii) the absence of undermined borders and surrounding erythema (Perry *et al*. 1968). The distinctiveness of this variant has been questioned (Wernikoff *et al*. 1987; Newman and Frank 1993).

Lastly, there is a variant known as atypical or bullous pyoderma gangrenosum. In this the ulceration is more superficial, there is often a bullous, blue-grey margin (Fig. 10), and the upper extremities and face are more commonly affected (Perry and Winkelmann 1972; Romano and Safai 1979). This variant has been reported with

haematological disease, specifically preleukaemic conditions such as myeloid metaplasia (Callen *et al*. 1976), or acute myelogenous leukaemia. At times the separation of atypical pyoderma gangrenosum from leukaemia-associated Sweet's syndrome is difficult.

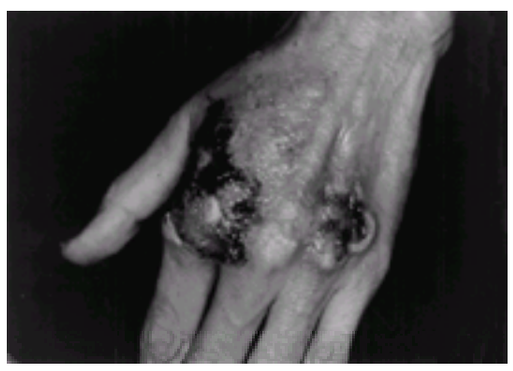


Fig. 10 Atypical pyoderma gangrenosum in a patient with a preleukaemic state. The lesion is a shallow ulcer with a bullous, blue-grey border.

The histopathological features of pyoderma gangrenosum are not specific, but are useful in ruling out other causes of cutaneous ulceration. There is controversy over what is the initial histopathological change, with some classifying the process as a neutrophilic dermatosis (Jorizzo 1988), and others believing that the initial changes involve lymphocytic infiltrate, endothelial cell swelling, and fibrinoid necrosis of the vessel wall (a lymphocytic vasculitis) (Su *et al*. 1986). Regardless, the lesion does not involve a leucocytoclastic vasculitis, nor is granuloma formation compatible with a diagnosis of pyoderma gangrenosum.

P>The aetiology and pathogenesis of pyoderma gangrenosum are not understood, but a variety of abnormalities of the immune system have been described. Although associated conditions are common (Table 2), perhaps a quarter to a half of the patients have 'idiopathic' disease. The most common associated conditions are inflammatory bowel disease, arthritis, paraproteinaemia, and haematological malignancy.

Table 2 Diseases associated with pyoderma gangrenosum

Initial reports of pyoderma gangrenosum emphasized the association with ulcerative colitis (Perry 1959). Eventually, it has become recognized that regional enteritis and Crohn's disease (granulomatous colitis) are found with pyoderma gangrenosum as often as ulcerative colitis (Powell *et al*. 1985; Prystowsky *et al*. 1989). In the most recent accounts, inflammatory bowel disease has constituted about 15 to 20 per cent of the associated phenomena. In addition, early reports stressed the relationship of pyoderma gangrenosum to the activity of the bowel disease, and even suggested that some patients' pyoderma gangrenosum lesions may benefit from surgical resection of the inflamed bowel (ulcerative colitis only). Callen *et al*. (1989) have reported pyoderma gangrenosum in association with inactive terminal ileitis, and Talansky *et al*. (1983) also have reported lack of effect of bowel resection in some patients with pyoderma gangrenosum and ulcerative colitis.

Arthritis is a frequent finding with pyoderma gangrenosum. In some of the later reports, arthritis has been the most frequent associated condition (Powell *et al*. 1985; Prystowsky *et al*. 1989). Five of the nine patients with arthritis reported by *Prystowsky et al*. (1989) had inflammatory bowel disease associated arthritis. In general, the arthritis associated with pyoderma gangrenosum is a symmetrical polyarthritis which may be seronegative or seropositive. Spondylitis, although it may occur in conjunction with inflammatory bowel disease associated with pyoderma gangrenosum, has not been reported independently with pyoderma gangrenosum.

A variety of malignancies have been reported with pyoderma gangrenosum, most commonly myelogenous leukaemia or preleukaemia. This association may be more common with the atypical variants of pyoderma gangrenosum. Although a variety of solid tumours have been reported, their presence is probably coincidental.

Paraproteinaemia, in general a benign variety, has been reported with pyoderma gangrenosum. With newer techniques of protein separation, it appears that 15 per cent of patients with pyoderma gangrenosum may have a benign monoclonal gammopathy, most often of IgA variety (Murray 1983; Prystowsky *et al*. 1989). The development of myeloma in these patients with pyoderma gangrenosum appears to be very rare.

The diagnostic evaluation (Table 3) of the patient presumed to have pyoderma gangrenosum has two objectives: (i) to rule out other causes of cutaneous ulceration, and (ii) to determine whether there is a treatable, systemic, associated disorder. There is not a specific laboratory or histopathological test that is diagnostic. Moreover, some of the associated disease processes may be clinically silent.

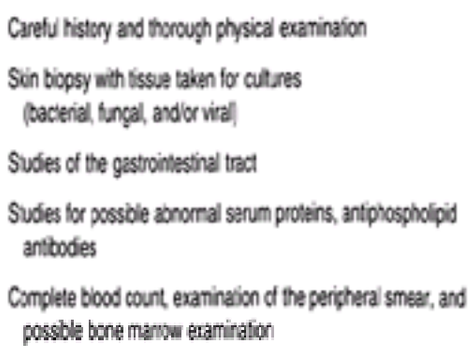


Table 3 Diagnostic evaluation of the patient with pyoderma gangrenosum

The differential diagnoses of ulcerative cutaneous lesions include: (i) infectious diseases, (ii) halogenodermas, (iii) vasculitis, (iv) insect bites, (v) venous or arterial insufficiency (including occlusive disease associated with antiphospolipid antibodies), and (vi) factitial ulcerations. Cultures should be taken from both exudate and tissues.

To test for the presence of an associated disorder, another series of examinations should be made. A thorough historical evaluation and examination of the gastrointestinal tract should be undertaken in conjunction with a gastroenterologist. Radiographic procedures may include an upper gastrointestinal series and a barium enema. Flexible sigmoidoscopy or colonoscopy, or both, may also be done, with appropriate biopsies being taken. A complete blood count, careful evaluation

of the peripheral smear, and possibly a bone marrow aspirate or biopsy will help rule out the presence of an associated haematological malignant process. Serum protein electrophoresis, serum immunodiffusion studies, and possibly, serum and urine immunoelectrophoresis will help to eliminate a diagnosis of an associated monoclonal gammopathy or myeloma. Multiple reports of pyoderma gangrenosum-like leg ulcers in patients with antiphospholipid antibodies have appeared and tests such as VDRL, anticardiolipin antibody, and partial thromboplastin time are now standard in the evaluation of a patient with pyoderma gangrenosum (Babe *et al*. 1992).

There is not a specific, uniformly effective treatment for pyoderma gangrenosum. Although systemic treatment may affect the underlying disease process in some patients with chronic ulcerative colitis (Mir-Madjlessi 1985), it sometimes becomes necessary to consider colectomy. Some patients' skin lesions will respond to bowel resection, but there are patients in whom total colectomy, including removal of the rectosigmoid colon, does not lead to a remission (Talansky *et al*. 1983). In one report a patient developed pyoderma gangrenosum 10 years after total colectomy (Cox *et al*. 1986). In mild cases, local measures, such as dressings, elevation, rest, topical agents, or intralesional injections may be sufficient to control the disease process. Compresses, wet to dry dressings, or the newer bio-occlusive semipermeable dressings may be useful. Cleansing or therapy with antibacterial agents such as hydrogen peroxide or benzoyl peroxide have been reported to be beneficial in an occasional patient. Hyperbaric oxygen has also been reported, in a small number of cases, to be effective. Superpotent topical corticosteroids and intralesional injections of corticosteroids (Gardner and Archer 1972) may be beneficial in some patients. Care must be taken to avoid introducing an infection and to limit the potential systemic effects of corticosteroids that arise from injecting large doses intralesionally. In general, the periphery of the lesions are injected, but the ulcer base may also be injected. Other topical approaches include the use of sodium cromoglycate, nitrogen mustard, and 5-aminosalicyclic acid (DeCock and Thorne 1980; Tsele *et al*. 1992).

In patients who do not respond to topical or local therapies, or whose severe, rapid course warrants the use of a systemic agent, sulphonamides, sulphones, or corticosteroids have been the most commonly used agents. Perry (1959) reported that oral sulphasalazine is effective in patients both with and without inflammatory bowel disease. Dapsone in doses of up to 400 mg per day has often been used as a monotherapy, or as an adjunctive steroid-sparing agent. Usually the drug is administered in lower doses (100 to 150 mg per day), and the usual precautions and pretherapy evaluation are necessary. The mechanism of action of the sulphanomides and sulphones in this process is not understood, but effects on the polymorphonuclear leucocyte may be a factor. Another antileprosy agent, clofazimine, has also been reported to be successful in some patients with pyoderma gangrenosum. Finally, several other antibiotics have been used successfully in individual cases; these include minocycline and rifampicin.

Systemic corticosteroids have been used extensively in patients with pyoderma gangrenosum and its variants and are generally believed to be very effective. Usually, large doses (40 to 120 mg per day) are necessary in order to induce a remission of the disease. These doses, used over the long term, will frequently result in steroid-related side-effects. In the studies by *Holt et al*. (1980), 6 of their 12 patients treated with corticosteroids developed serious steroid complications, and 4 of these 6 died as a result of the therapy. To avoid the complications of long-term steroid use, Johnson and Lazarus (1982), and subsequently others, have used pulse therapy with 1 g of methylprednisolone given intravenously each day for a period of 5 days. Maintenance of the remission was accomplished with oral corticosteroids every other day. Prystowsky *et al*. (1989) have reported the experience of Lazarus with a further eight patients. They found that remissions occurred in five of the patients, and that they were usually able to remove oral corticosteroid therapy and often lower the dose of other therapies. Pulse therapy is not without side-effects, which include sudden death. In the hands of Prystowsky *et al*., and my experience with three patients, this therapy has primarily resulted in transient hyperglycaemia; however, it should only be used with great caution and proper monitoring.

Immunosuppressive agents have been suggested for use in patients who fail to respond to other therapies, particularly systemic corticosteroids, or who develop steroid-related side-effects. Individual reports using oral azathioprine, cyclophosphamide (Newell and Malkinson 1983), chlorambucil (Burruss *et al*. 1994), cyclosporin (Matis *et al*. 1992), tacrolimus (FK 506) (Abu-Elmagd *et al*. 1993), or methotrexate (Teitel 1993) have suggested that, at least in some patients, these agents may be successful. Intravenous pulses of cyclophosphamide (Zonana-Nacach *et al*. 1994) or immunoglobulin (Gupta *et al*. 1995) have also been successful in individual patients. The mode of action of the immunosuppressive agents is not understood.

Rheumatoid neutrophilic dermatitis

Ackerman (1978) described a neutrophilic dermatosis in patients with rheumatoid arthritis. This is apparently a rare manifestation of rheumatoid arthritis, and has only been reported in a small number of patients (Scherbenske *et al*. 1989; Sanchez and Cruz 1990; Lowe *et al*. 1992). The patients are described as having symmetric, erythematous nodules and plaques on the extensor surfaces of the joints (Fig. 11). There is an apparent predilection for the dorsa of the hands and arms. It is not clear whether this condition is clinically or histopathologically distinct, specifically, whether it can be differentiated from Sweet's syndrome. An effective therapy has not been described.

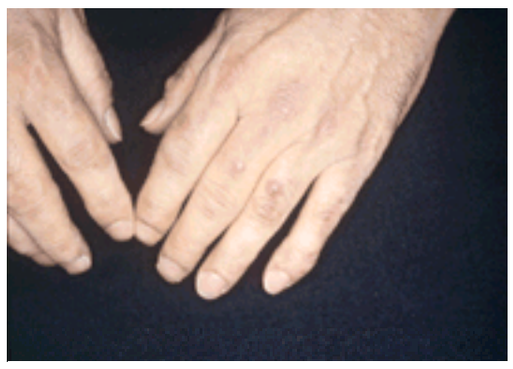


Fig. 11 Rheumatoid neutrophilic dermatosis. This patient developed multiple vesiculopustular plaques on the dorsum of his hands.

Bowel-associated dermatosis–arthritis syndrome

Patients who had undergone bowel bypass surgery for morbid obesity have occasionally developed scattered pustular lesions (Fig. 12) and arthritis. This became known as the bowel bypass syndrome, and was felt to be an immune-complex disease caused by bacterial overgrowth in the blind loop. Treatment with antibiotics was often effective in clearing the cutaneous lesions and improving the joint symptoms. Later, Jorizzo *et al*. (1983) coined the term 'bowel-bypass syndrome without bowel bypass' or 'bowel-associated dermatosis–arthritis syndrome'. They reported on four patients with this syndrome of whom two had blind loops due to Billroth II procedures, one had ulcerative colitis, and one had Crohn's disease. Dicken (1984) reported two similar patients who had had Roux-en-Y procedures with resultant blind loops. Clinically, these patients present with a widespread eruption characterized by pustules on an erythematous or necrotic base. The lesions may be few in number, or may be extensive. Ulceration is rare. The appearance of the lesions is often accompanied by fever, arthralgias or a true inflammatory arthropathy, and myalgias. The arthritis accompanying this process is generally symmetrical, non-deforming, and most frequently involves the small joints such as the wrists, ankles, metacarpophalangeal, proximal interphalangeal, and metatarsophalangeal joints. Histopathologically, the disease resembles Sweet's syndrome. In fact, only after the report by Jorizzo *et al*. (1983) did we recognize that a patient reported by our group (Bechtel and Callen 1981) probably would have been more correctly diagnosed as having the bowel-associated dermatosis–arthritis syndrome (Fig. 13) rather than Sweet's syndrome, because of his prior Billroth II procedure. This disease is presumed to be due to immune complexes (Jorizzo *et al*. 1984) and while anti-inflammatory therapy is at times helpful, antibiotics frequently control the process, or bowel surgery (to remove 'blind' loops) will reverse it.

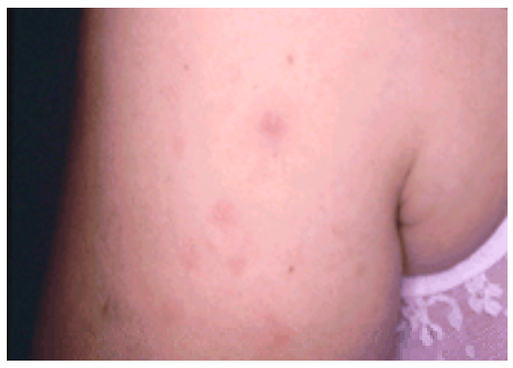


Fig. 12 Scattered pustules on an erythematous base in a patient with a bowel bypass for morbid obesity.

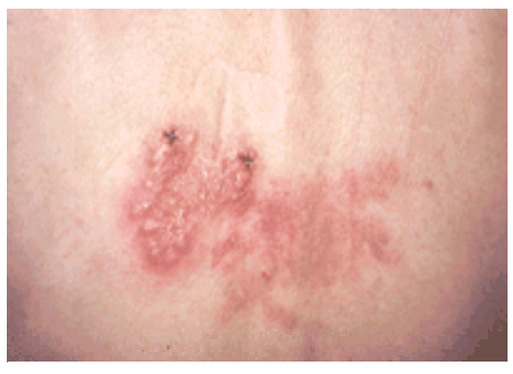


Fig. 13 Vesiculopustular eruption in a patient with previous ulcer surgery (Billroth II) and a blind loop—the so-called 'bowel bypass syndrome without a bowel bypass'.

Chapter References

Abu-Elmagud, K. *et al*. (1993). Resolution of severe pyoderma gangrenosum in a patient with streaking leukocyte factor disease after treatment with tacrolimus (FK 506). *Annals of Internal Medicine*, **119**, 595–8.

Ackerman, A. (1978). *Histologic diagnosis of inflammatory skin diseases: a method by pattern analysis*, pp. 449–50. Lea & Febiger, Philadelphia.

Aram, H. (1984). Acute febrile neutrophilic dermatosis (Sweet's syndrome): response to dapsone. *Archives of Dermatology*, **120**, 245–7.

Babe, K.S. Jr., Gross, A., Leyva, W.H., and King, L.E. Jr. (1992). Pyoderma gangrenosum associated with antiphospholipid antibodies. *International Journal of Dermatology*, **31**, 588–90.

Banet, D.E., McClave, S.A., and Callen, J.P. (1994). Oral metronidazole, an effective treatment for Sweet's syndrome in a patient with associated inflammatory bowel disease. *Journal of Rheumatology*, **21**, 1766–8

Bechtel, M.A. and Callen JP (1981). Acute febrile neutrophilic dermatosis (Sweet's syndrome). *Archives of Dermatology*, **117**, 664–6.

Boatman, B.W., Taylor, R.C., Klein, L.E., and Cohen, B.A. (1994). Sweet's syndrome in children. *Southern Medical Journal*, **87**, 193–6.

Burruss, J.B., Farmer, E., and Callen, J.P. (1994). Chlorambucil is an effective corticosteroid-sparing agent for recalcitrant pyoderma gangrenosum. *Arthritis and Rheumatism*, **37**, S275 (Abstract).

Callen, J.P. (1990). Pyoderma gangrenosum and related disorders. *Dermatologic Clinics*, **7**, 1249–59.

Callen J.P. and Woo T.Y. (1985). Vesiculopustular eruption in a patient with ulcerative colitis. *Archives of Dermatology*, **121**, 339–44.

Callen J.P., Dubin H.V., and Gherke, C.F. (1976). Pyoderma gangrenosum and agnogenic myeloid metaplasia. *Archives of Dermatology*, **113**, 1585–6.

Callen, J.P., Case, J.D., and Sager, D. (1989). Chlorambucil—an effective, corticosteroid-sparing therapy for pyoderma gangrenosum. *Journal of the American Academy of Dermatology*, **21**, 515–9.

Case, J.D., Smith, S.Z., and Callen, J.P. (1989). The use of pulse methylprednisolone and chlorambucil in the treatment of Sweet's syndrome. *Cutis*, **44**, 125–9.

Cohen, P.R. (1993). Pregnancy-associated Sweet's syndrome: world literature review. *Obstetrical and Gynecological Survey*, **48**, 584–7.

Cohen, P.R., Talpaz, M., and Kuzrock, R. (1988). Malignancy-associated Sweet's syndrome: review of the world literature. *Journal of Clinical Oncology*, **6**, 1887–97.

Cohen, P.R., Holder W.R., and Rapini R.P. (1992). Concurrent Sweet's syndrome and erythema nodosum: a report, world literature review and mechanism of pathogenesis. *Journal of Rheumatology*, **19**, 814–20.

Cohen, P.R., Holder, W.R., Tucker S.B., Kono, S., and Kurzrock, R. (1993). Sweet syndrome in patients with solid tumors. *Cancer*, **72**, 2723–31.

Cooper, P.H., Innes, D.J. Jr., and Greer, K.E. (1983). Acute febrile neutrophilic dermatosis (Sweet's syndrome) and myeloproliferative disorders. *Cancer*, **51**, 1518–26.

Cox, N.H., Peebles-Brown, A., and Mackie, R.M.(1986). Pyoderma gangrenosum occurring 10 years after protocolectomy for ulcerative colitis. *British Journal of Hospital Medicine*, **36**, 363–5.

DeCock, K.M. and Thorne, M.G. (1980). Treatment of pyoderma gangrenosum with sodium cromoglycate. *British Journal of Dermatology*, **102**, 231–3.

Dicken, C.H. (1984). Bowel-associated dermatosis–arthritis syndrome: bowel bypass syndrome without bowel bypass. *Mayo Clinic Proceedings*, **59**, 43–6.

Fenske, N.A. *et al*. (1983). Vesiculopustular eruption of ulcerative colitis. *Archives of Dermatology*, **119**, 664–9.

Gupta, A.K., Shear, N.H., and Sauder, D.N. (1995). Efficacy of human intravenous immune globulin in pyoderma gangrenosum. *Journal of the American Academy of Dermatology*, **32**, 140–2.

Harary, A.M. (1983). Sweet's syndrome associated with rheumatoid arthritis. *Archives of Internal Medicine*, **143**, 1993–5.

Hoffman, G.S. (1977). Treatment of Sweet's syndrome (acute febrile neutrophilic dermatosis) with indomethacin. *Journal of Rheumatology*, **4**, 201–6.

Holt, P.J.A. *et al*. (1980). Pyoderma gangrenosum. *Medicine*, **59**, 114–33.

Horio, T., Immamura, S., Danno, K., Furnkawa, F., and Ofuji, S. (1980). Treatment of acute febrile neutrophilic dermatosis (Sweet's syndrome) with potassium iodide. *Dermatologica*, **160**, 341–7.

Johnson, M.L. and Grimwood, R.E. (1994). Leukocyte colony-stimulating factors: a review of associated neutrophilic dermatoses and vasculitides. *Archives of Dermatology*, **130**, 77–81.

Johnson, R.B. and Lazarus, G.S. (1982). Pulse therapy. *Archives of Dermatology*, **118**, 76–84.

Jorizzo, J.L. (1988). Neutrophilic dermatoses: Sweet's syndrome and pyoderma gangrenosum. In *Inflammation: basic principles and clinical correlates*, (ed. J.I. Gallin, I.M. Goldstein, and R. Syndermoin), pp. 785–802. Raven Press, New York.

Jorizzo, J.L. *et al*. (1983). Bowel bypass syndrome without bowel bypass: bowel-associated dermatosis–arthritis syndrome. *Archives of Internal Medicine*, **143**, 457–61.

Jorizzo, J.L. *et al*. (1984). Bowel-associated dermatosis–arthritis syndrome: immune complex-mediated vessel damage and increased neutrophil migration. *Archives of Internal Medicine*, **144**, 738–40.

Keltz, M., Lebwohl, M., and Bishop, S. (1993). Peristomal pyoderma gangrenosum. *Journal of the American Academy of Dermatology*, **27**, 360–4.

Kemmett, D. and Hunter, J.A.A.(1990). Sweet's syndrome: a clinicopathologic review of twenty-nine cases. *Journal of the American Academy of Dermatology*, **23**, 503–7.

Lowe, L., Kornfeld, B., Clayman, J., and Golitz, L.E. (1992). Rheumatoid neutrophilic dermatosis. *Journal of Cutaneous Pathology*, **19**, 48–53.

Matis, W.L., Ellis, C.N., Griffiths, C.E.M., and Lazarus, G.S. (1992). Treatment of pyoderma gangrenosum with cyclosporine. *Archives of Dermatology*, **128**, 1060–4.

McCalmont, C.S., Leshin, B., White, W.L., Greiss, F.C. Jr., and Jorizzo, J.L. (1991). Vulvar pyoderma gangrenosum. *International Journal of Gynecology and Obstetrics*, **35**, 175–8.

Mir-Madjlessi, S.H., Taylor, J.S., and Farmer, R.G.(1985). Clinical course and evolution of erythema nodosum and pyoderma gangrenosum in chronic ulcerative colitis: a study of 42 patients. *American Journal of Gastroenterology*, **80**, 615–20.

Murray, J.C. (1983) Pyoderma gangrenosum and IgA gammopathy. *Cutis*, **32**, 503–7.

Newell, L.M. and Malkinson, F.D. (1983). Pyoderma gangrenosum: response to cyclophosphamide therapy. *Archives of Dermatology*, **119**, 477–86.

Newman, W.P. and Frank, H.J. (1993). Pyoderma gangrenosum of the orbit. *Eye*, **7**, 89–94.

O'Loughlin, S. and Perry, H.O. (1978). A diffuse pustular eruption associated with ulcerative colitis. *Archives of Dermatology*, **114**, 1061–4.

Perry, H.O. (1959). Pyoderma gangrenosum. *Southern Medical Journal*, **62**, 899–908.

Perry, H.O. and Winkelmann, R.K. (1962). Bullous pyoderma gangrenosum and leukemia. *Archives of Dermatology*, **106**, 901–5.

Perry, H.O. *et al*. (1968). Malignant pyoderma. *Archives of Dermatology*, **98**, 561–74.

Piette, W.W., Trapp, J.F., O'Donnell, M.J., Argenyi, Z., Talbot, E.A., and Burns, C.P. (1994). Acute neutrophilic dermatosis with myeloblastic infiltrate in a leukemia patient receiving all-*trans*-retinoic acid therapy. *Journal of the American Academy of Dermatology*, **30**, 293–7.

Powell, F.C. *et al*. (1985). Pyoderma gangrenosum: a review of 86 patients. *Quarterly Journal of Medicine*, **55**, 173–86.

Prystowsky, J.H., Kahn, S.N., and Lazarus, G.S. (1989). Present status of pyoderma gangrenosum. *Archives of Dermatology*, **125**, 57–64.

Prystowsky, S.D., Fye, K.H., Goette, K.D., and Daniels, T.E. (1978). Acute febrile neutrophilic dermatosis (Sweet's syndrome) associated with Sjögren's syndrome. *Archives of Dermatology*, **114**, 1234–5.

Romano, J. and Safai, B. (1979). Pyoderma gangrenosum and myeloproliferative disorders. *Archives of Internal Medicine*, **139**, 932–4.

Sanchez, J.L. and Cruz, A.(1990). Rheumatoid neutrophilic dermatosis. *Journal of the American Academy of Dermatology*, **22**, 922–5.

Scherbenske, J.M., Benson, P.M., and Lupton, G.P. (1989). Rheumatoid neutrophilic dermatosis. *Archives of Dermatology*, **125**, 1105–8.

Storwick, G.S., Prihoda, M.B., Fulton, R.J., and Wood, W.S. (1994). Pyodermatitis–pyostomatitis vegetans: a specific marker for inflammatory bowel disease. *Journal of the American Academy of Dermatology*, **31**, 336–41.

Su, W.P.D. *et al*. (1986). Histopathologic and immunopathologic study of pyoderma gangrenosum. *Journal of Cutaneous Pathology*, **13**, 323–30.

Subhisa, S. *et al*. (1983). Colchicine in the treatment of acute febrile neutrophilic dermatosis (Sweet's syndrome). *British Journal of Dermatology*, **108**, 99–01.

Sweet, R.D. (1964). An acute febrile neutrophilic dermatosis. *British Journal of Dermatology*, **76**, 349–56.

Talansky, A.L. *et al*. (1983). Does intestinal resection heal the pyoderma gangrenosum of inflammmatory bowel disease? *Journal of Clinical Gastroenterology*, **108**, 580–1.

Teitel, A.D. (1994). Treatment of pyoderma gangrenosum with methotrexate. *Arthritis and Rheumatism*, **36**, S163 (Abstract).

Trentham, D.E., Masi, A.T., and Bale, G.F. (1976). Arthritis with an inflammatory dermatoses resembling Sweet's syndrome: report of a unique case and review of the literature on arthritis associated with the inflammatory dermatoses. *American Journal of Medicine*, **61**, 424–32.

Tsele, E., Yu, R.C.H., and Chu, A.C. (1992). Pyoderma gangrenosum—response to topical nitrogen mustard. *Clinical and Experimental Dermatology*, **17**, 437–40.

Van Hale, H.M. *et al*. (1985). Pyostomatitis vegetans: a reactive mucosal marker for inflammatory disease of the gut. *Archives of Dermatology*, **121**, 94–8.

von den Driesch, P. (1994). Sweet's syndrome (acute febrile neutrophilic dermatosis). *Journal of the American Academy of Dermatology*, **31**, 535–56.

von den Driesch, P. *et al*. (1994). Sweet's syndrome: therapy with cyclosporin A. *Clinical and Experimental Dermatology*, **19**, 274–7.

Wernikoff, S. *et al*. (1987). Malignant pyoderma or pyoderma gangrenosum of the head and neck? *Archives of Dermatology*, **123**, 371–5.

Zonana-Nacach, A., Jiménez-Balderas, J., Martínez-Osuna, P., and Mintz, G. (1994). Intravenous cyclophosphamide pulses in the treatment of pyoderma gangrenosum associated with rheumatoid arthritis: report of 2 cases and review of the literature. *Journal of Rheumatology*, **21**, 1352–6.

5.13.5 Sarcoidosis

Barry Bresnihan

Sarcoidosis is a multisystem disease of unknown aetiology characterized by the presence of multiple, non-caseating granulomas in involved tissues. Symptoms and signs depend on the severity of disease in the affected organ systems. Tissues that may be involved include joints, bones and muscles.

Aetiology

Environmental factors are important in the aetiology of sarcoidosis. Firstly, the acute form of sarcoidosis presents more frequently during spring months (Poukkula *et al*. 1986), suggesting that an environmental agent with a higher prevalence during the spring may be causative. A number of localized outbreaks of sarcoidosis have suggested the possibility of an infectious agent (Veale and FitzGerald 1990). Others have suggested a relation between sarcoidosis and occupation (Edmondstone 1988).

Genetic factors also appear to be important. There is an increased incidence of sarcoidosis within families. The British Thoracic and Tuberculosis Association (1973) reported 121 patients among 59 families. In one series of 114 patients there was a 9.6 per cent frequency of sarcoidosis in the siblings of index cases (Brennan *et al*. 1984).

A number of studies have examined associations between HLA alleles and sarcoidosis, but no linkages with disease susceptibility have yet been identified (Mehra and Bovornkitti 1988). However, some investigators have suggested associations between HLA-B8 and sarcoid arthropathy (Brewerton *et al*. 1977), and more recently an association between HLA-B8, -DR3 and acute sarcoidosis with arthritis has been demonstrated (Kremer 1986).

Pathology

The characteristic histological feature of sarcoidosis found in all affected tissues is a well-defined, round or oval granuloma composed of compact, radially arranged epithelioid cells, a few multinucleate giant cells and a rim of lymphocytes (Fig. 1). Caseation is absent. A small area of fibrinoid necrosis may be present. Sarcoid granuloma may be divided into early, intermediate, and late stages. The epithelioid cells are derived from circulating monocytes. Giant cells are formed from fusion of epithelioid cells. Their size varies between 150 and 300 µm. Most giant cells are of foreign-body or Langhans type and may contain inclusion bodies, which represent metabolic end-products. The lymphocytes are derived from peripheral blood and consist predominantly of CD4+, T-helper cells (Hunninghake and Crystal 1981). There is evidence to support the suggestion that the accumulation of CD4+ lymphocytes may result from increased production of interleukin 2 (**IL-2**) (Semenzato *et al*. 1993). Soluble IL-2 receptor has been demonstrated in active sarcoidosis (Ina *et al*. 1992). Increased expression of the IL-2 receptor on mononuclear cells and alveolar macrophages correlated positively with serum concentrations of angiotensin-converting enzyme (Pforte *et al*. 1993). Studies of T-cell receptors have suggested that T cells accumulate in sarcoid lesions in response to persistent stimulation by a specific but unidentified antigen (Du Bois *et al*. 1992; Forrester *et al*. 1993).

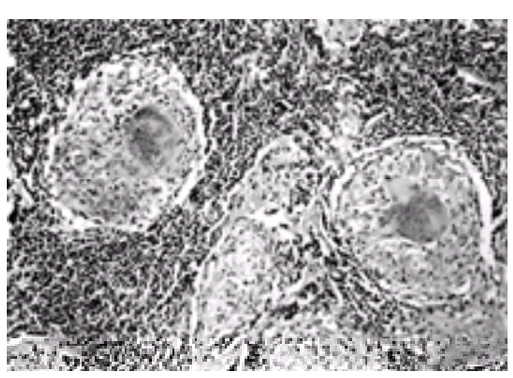


Fig. 1 Sarcoid granuloma. Two, well-formed, non-caseating granulomas are demonstrated in lymph-node tissue. Central multinucleated giant cells are present, surrounded by epithelioid cells and some lymphocytes. (By courtesy of Dr Mary McCabe, St. Vincent's Hospital, Dublin.)

A number of recent studies have suggested a role for several cytokines, including tumour necrosis factor-a, IL-1b, and IL-6, in disease pathogenesis and granuloma formation (Shakoor and Hamblin 1992; Pueringer *et al*. 1993; Steffen *et al*. 1993).

Epidemiology

Sarcoidosis occurs worldwide, but prevalence, clinical manifestations and outcome vary widely in different areas (Tierstein and Lesser 1983). Sarcoidosis is recognized more often in developed than in underdeveloped countries, in Western rather than in Far Eastern countries, and in Northern rather than Southern Europe. Moreover, the prevalence within a given region may vary. For example, Sweden and Denmark have prevalence rates of approx. 60 per 100 000 population, while Finland has a prevalence of 7.5 per 100 000. The overall prevalence rate in the United Kingdom has been estimated as 20 per 100 000, and in Ireland as 30 per 100 000. The highest reported prevalence in the United Kingdom (of up to 200 per 100 000) was observed in Irish immigrant women of child-bearing age living in London. In the United States, sarcoidosis is 10 to 17 times more prevalent among black than white individuals.

General clinical features

Sarcoidosis may involve almost any tissue in the body. Approximately half of patients with sarcoidosis will present with pulmonary symptoms such as dyspnoea, cough or chest pain (Sharma 1984a). A further 20 per cent approximately will present with an abnormal routine chest radiograph and minimal or no symptoms. Approximately 1 in 4 patients will present with non-specific constitutional symptoms including fever, weight loss and anorexia. Less than 10 per cent of patients will present with features confined to extrapulmonary systems such as the musculoskeletal system.

Acute sarcoidosis

Acute or subacute sarcoidosis may present as an illness of explosive onset, usually characterized by fever, erythema nodosum, and hilar lymphadenopathy. The association of hilar lymphadenopathy and erythema nodosum is often referred to as Lofgren syndrome (Lofgren 1953). Erythema nodosum appears to be more common among Swedish, Irish, Puerto Rican, and Mexican women of child-bearing age. This form of sarcoidosis has a high prevalence of spontaneous remission and a good prognosis. The chest radiograph clears within 1 year of presentation in more than 60 per cent of patients with acute sarcoidosis.

Chronic sarcoidosis

Chronic sarcoidosis is less common than acute sarcoidosis and has a subtle onset with an insidious progressive and highly variable clinical course. Chest radiographs demonstrate extensive parenchymal infiltration in more than 90 per cent of patients with chronic sarcoidosis. Pulmonary function tests confirm a restrictive lung defect. Other manifestations are listed in Table 1.

Lung: parenchymal disease
Skin: lupus pernio; skin plaques or nodules
Eye: uveitis; conjunctivitis; keratoconjunctivitis sicca
Lymphatic system: peripheral lymphadenopathy; splenomegaly
Bone marrow infiltration
Liver: hepatic granulomata; portal hypertension; hepatic failure
Kidney: nephrocalcinosis; renal calculi; granulomatous infiltration; glomerular disease; renal arteritis
Heart: cardiomyopathy; conduction abnormalities
Nervous system: cranial and peripheral neuropathy; papilloedema; intracerebral lesions; meningitis; seizures; spinal cord involvement; psychiatric disorders
Endocrine: pituitary; hypothalamus; thyroid; parathyroid; adrenal
Reproductive organs: ovaries; testes
Gastrointestinal tract and pancreas
Salivary and lacrimal glands
Nose, tonsils, larynx

Table 1 Clinical manifestations in chronic sarcoidosis

Laboratory investigations

Routine investigations

Most reports are that anaemia is uncommon in sarcoidosis (Sharma 1984b), occurring in approx. 5 per cent of patients. Haemolytic anaemia has been reported, but is rare. Leucopenia occurs in approximately one-third of patients, sometimes in the absence of splenomegaly and may reflect granulomatous infiltration of the bone marrow. Neutrophilia and polycythaemia are rare. There is eosinophilia in approximately one-quarter of patients. Thrombocytopenia has long been recognized as a relatively common complication of sarcoidosis. An elevated erythrocyte sedimentation rate occurs during acute disease episodes, particularly in association with erythema nodosum.

Reports of hypercalcaemia in sarcoidosis vary widely between 2 and 63 per cent. The reasons for such wide variation are unclear. The frequency tends to be higher in North American series. In a worldwide review of more than 3 000 patients, James *et al*. (1976) reported a frequency of 11 per cent. The serum calcium fluctuates during the course of the disease and may be raised during acute or subacute phases.

Mild elevation of serum bilirubin and alkaline phosphatase are common (Sharma 1984c). Severe jaundice and disturbance of liver function is uncommon, even in the presence of extensive liver involvement.

Urinary hydroxyproline concentrations are considerably increased in acute sarcoidosis, returning to normal as disease manifestations resolve. Hydroxyproline excretion is normal in chronic sarcoidosis. Hypercalcinuria occurs more frequently than hypercalcaemia, and has been reported in 49 per cent of one series where only 13 per cent had hypercalcaemia. Significant proteinuria has been reported in approximately one-third of patients.

Angiotensin-converting enzyme

Angiotensin-converting enzyme catalyses the conversion of angiotensin I to vasoactive angiotensin II and inactivates bradykinin. Increased concentrations of serum angiotensin-converting enzyme were first observed in 83 per cent of patients with active sarcoidosis. The incidence has varied in subsequent series. In one international study of almost 2000 patients, 57 per cent had elevated serum concentrations of this enzyme (Studdy and James 1983). Angiotensin-converting enzyme is produced mainly by the epithelioid cells in the granulomas and enzyme activity reflects the total granulomatous 'load' in the body. Serial measurements of this enzyme are useful when monitoring the course of the disease.

Kveim test

The intracutaneous injection of a saline suspension of previously validated human sarcoid spleen or lymph node will give rise to a non-caseating granulomatous reaction after 2 to 6 weeks in approximately three-quarters of patients with sarcoidosis (Kataria *et al*. 1980) The Kveim test is generally regarded as specific for sarcoid, but has a number of disadvantages, including variability of results with different antigen sources, the 2- to 6-week delay and the inability to treat with corticosteroids during the wait, the need for punch biopsy, and variability in histological interpretation. Newer diagnostic procedures yield diagnostic material more rapidly.

Diagnostic imaging

The chest radiograph may be normal, or demonstrate hilar lymphadenopathy with or without pulmonary infiltration, pulmonary infiltration without adenopathy or advanced and irreversible fibrosis. Computed tomography may demonstrate adenopathy or early granulomatous infiltration that is undetectable on plain radiographs.

Beaumont *et al*. (1982) described [^{67}Ga]citrate scanning of the lungs as a very sensitive method for detecting granuloma formation in sarcoidosis. Several studies have demonstrated that ^{67}Ga scanning of the lungs reflects disease activity and response to therapy (Lawrence *et al*. 1983; Baughmann *et al*. 1984). Lung uptake of ^{67}Ga is greatest in patients with acute disease and high serum concentrations of angiotensin-converting enzyme. Uptake of ^{67}Ga does not correlate with quantitative pulmonary function nor with the clinical course and outcome. It is uncertain why ^{67}Ga accumulates at sites of granulomatous inflammation.

Histology

Sarcoidosis may resemble other diseases such as lymphoma and tuberculosis. Many institutions have a policy of obtaining histological confirmation of the diagnosis in all cases. Transbronchial lung biopsy has become the most widely used biopsy technique and is both highly sensitive and selective (Tierstein 1983). Open lung biopsy and mediastinal lymph-node biopsy are also highly sensitive and selective options, but, being more invasive, have some disadvantages. Other tissues likely to yield positive histological findings include scalene node, liver, skin, muscle, conjunctiva, lacrimal and minor salivary glands, and spleen.

Bronchoalveolar lavage

Bronchoalveolar lavage provides samples of the alveolar secretions for analysis. Lavage fluids may be studied for cellular content and inflammatory mediators. Analysis of bronchoalveolar fluids is a useful aid in the differential diagnosis of some interstitial lung diseases (Hunninghake *et al*. 1979). Patients with acute pulmonary sarcoidosis generally have high lymphocyte counts and high T helper-/T suppressor-cell ratios in bronchoalveolar fluids. Serial analysis of these lavage fluids may provide valuable information on the response to therapy.

Musculoskeletal manifestations

Joint disease

Distinctive patterns of arthropathy are associated with both acute and chronic sarcoidosis (James *et al*. 1976). In acute sarcoidosis, transient flitting arthralgias may precede the emergence of fever, erythema nodosum (Fig. 2), and hilar adenopathy (Fig. 3) (Gumpel *et al*. 1967; Siltzbach and Duberstein 1968; Spilberg *et al*. 1969). The joints most frequently presenting with arthritis after the development of erythema nodosum are the ankles and knees, although others may be involved occasionally. Approximately 65 per cent of patients presenting with acute sarcoidosis will have articular features. Prominent signs of inflammation are present, with marked erythema and tenderness of involved joints causing pain and limited motion. The arthritis is usually symmetrical and migratory, and persists for periods of between 3 to 4 days and 3 to 4 months. It may be difficult on occasion to distinguish features of erythema nodosum from articular and periarticular inflammation of the ankle joints. Ultrasonographic examination may be helpful in this context. In one study, joint effusions were demonstrated in 6 of 24 consecutive patients examined and tenosynovitis was identified in 8 (Kellner *et al*. 1992). Joint effusions in other joints are usually not detectable. When an effusion is present it is usually only mildly inflammatory, with less than 1000 leucocytes/mm^3, predominantly lymphocytes and large mononuclear cells. Occasionally an inflammatory effusion is aspirated, with leucocyte counts of greater than 40 000/mm^3, predominantly neutrophils. Needle biopsy specimens of synovial membrane in acute sarcoid arthropathy show mild, non-specific synovitis consistent with the clinical impression that much of the inflammation is periarticular. After recovery, acute sarcoidosis only occasionally recurs.

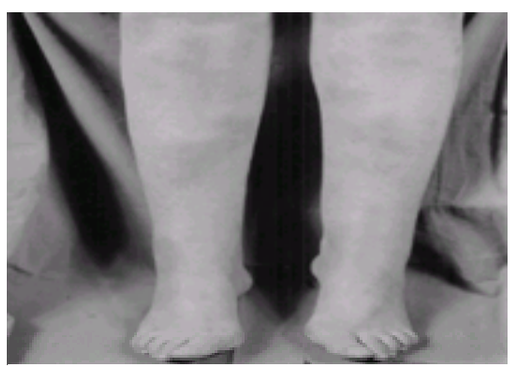


Fig. 2 Erythema nodosum is characteristic of acute sarcoidosis. Tender, red swellings appear on the shins, thighs, and upper limbs. Resolving erythema nodosum resembles painful bruises. Erythema nodosum usually resolves fully within 3 to 6 weeks.

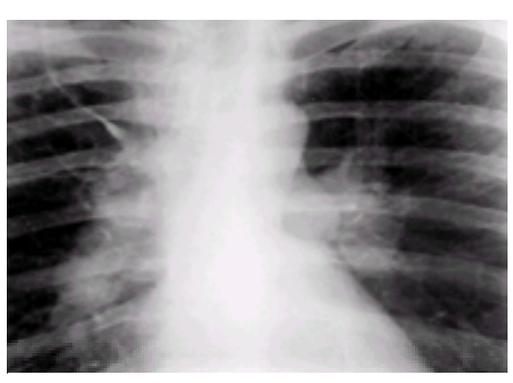


Fig. 3 Bilateral hilar lymphadenopathy due to sarcoidosis may be discovered on routine chest radiographs in asymptomatic individuals. It also frequently presents in acute sarcoidosis with fever, erythema nodosum, and arthritis.

Chronic arthritis is relatively uncommon in sarcoidosis (Gumpel *et al*. 1967; Siltzbach and Duberstein 1968; Spilberg *et al*. 1969; Grigor and Hughes 1976). It usually involves several joints and may appear either early or late in the course of the disease. Chronic monoarticular arthritis is rare. Large joints such as knees and ankles, elbows and wrists are most frequently involved. Hips and shoulders may also be affected. Involvement of spinal and temporomandibular joints has also been reported. Chronic arthritis involving the small joints of hands and feet is rare, except when secondary to osseous disease involving the phalanges (see below). Chronic polyarthritis is more frequently observed in women than men with sarcoidosis. A history of erythema nodosum is unusual in patients with chronic arthritis. Acute exacerbations may occur over a period of years, especially during periods of generalized disease activity. Chronic sarcoid arthritis is characterized by synovial thickening and effusions. Histological examination of synovial membrane may reveal typical non-caseating granulomas (Fig. 4). However, in some patients with chronic arthritis the histological features in synovial membrane are non-specific (Palmer and Schumacher 1984). Chronic arthritis may progress to joint deformity and destruction.

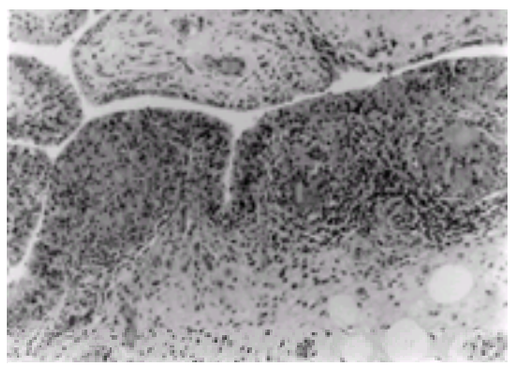


Fig. 4 Chronic arthritis in sarcoidosis showing thickening of the synovial lining layer, mononuclear cell infiltration, and non-caseating granulomas.

Bone disease

Bone involvement occurs in approx. 5 per cent of all patients with sarcoidosis. Bone cysts containing characteristic granulomas are most frequently observed in the phalanges of the hands and feet (Gumpel *et al*. 1967; James *et al*. 1976). Bone cysts are occasionally observed in long bones, pelvis, vertebrae, and the bones of the skull. Bone lesions are most frequently found in patients with persistent disease and are particularly frequent in those with lupus pernio (Fig. 5) and other chronic skin

lesions (Fig. 6). However, sarcoidosis limited to osseous disease of fingers and toes has been described (Schriber and Firooznia 1975). Bone lesions may be asymptomatic; frequently they are identified fortuitously after routine radiographic examination. Bone scanning techniques may be more sensitive than routine radiography in detecting early osseous sarcoidosis. Pain may be absent, even when obvious bone swelling is present.

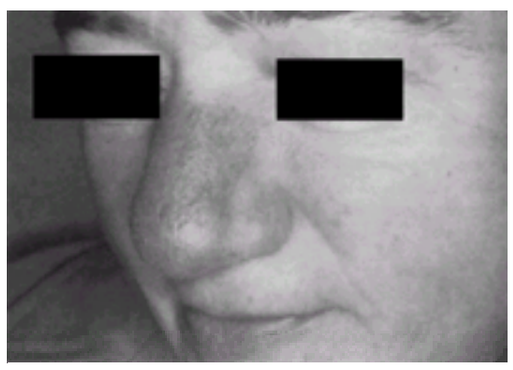


Fig. 5 Lupus pernio, presenting as a bluish-red or violaceous swelling of the nose extending on to the cheek, is a characteristic of chronic sarcoidosis and is frequently associated with bone lesions.

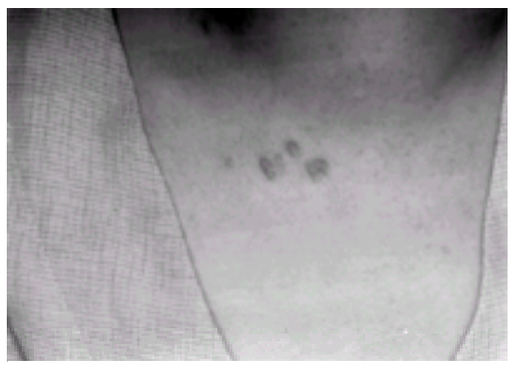


Fig. 6 Chronic cutaneous sarcoidosis, which may accompany bone involvement.

The most characteristic presentation of osseous sarcoidosis is of swollen, sausage-like digits resulting from cyst formation in the phalanges (Fig. 7). Single or multiple digits may be involved. Radiographic examination will demonstrate lytic lesions of various sizes ranging from minute cortical defects to intraosseous cysts causing diffusely expanded phalanges or cysts in phalangeal or metatarsal heads (Fig. 8 and Fig. 9). Less frequent radiographic changes include thickening of cortical bone with a fine reticular alteration of the trabecular pattern, acrosclerosis of distal phalanges or, rarely, gross bone destruction. Osseous sarcoidosis may cause secondary articular changes in the joints of the hands and feet.

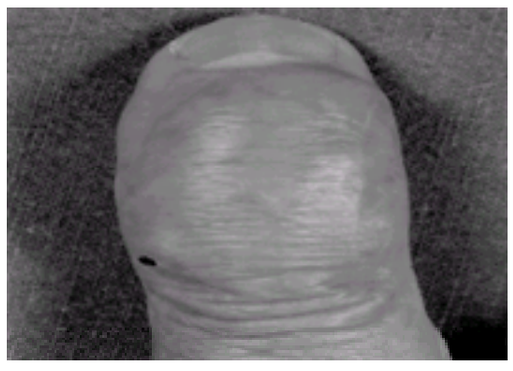


Fig. 7 Sarcoid dactylitis caused by granuloma formation in phalanges and surrounding soft tissue.

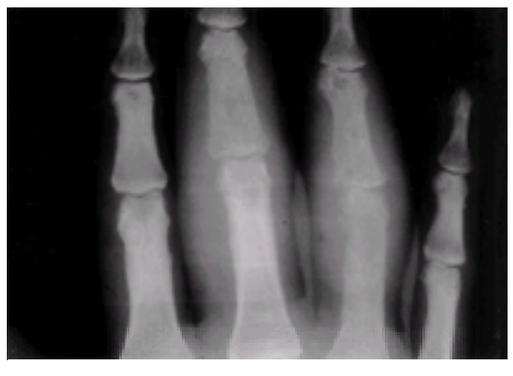


Fig. 8 Bone sarcoidosis presenting as dactylitis of the third and fourth fingers. The radiological changes are relatively early. Note the lacy, reticular pattern in the third middle and the fourth middle and proximal phalanges, associated with prominent soft-tissue swelling.

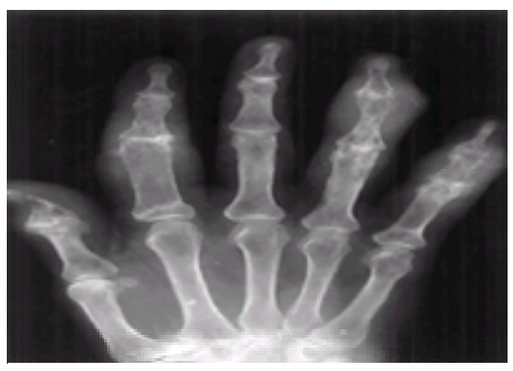


Fig. 9 Bone sarcoidosis involving four digits sparing the middle finger. Soft-tissue swelling, bone cysts and joint-space narrowing are prominent. There is expansion of

the second proximal phalanx.

Vertebral sarcoidosis may present with back pain or with neurological impairment resulting from cord compression. Radiographic examination of the spine will demonstrate extensive lytic or sclerotic changes in one or more vertebrae at any level. Pathological fracture has been reported as a complication of long-bone involvement. Skull involvement may appear as multiple lytic defects associated with overlying soft tissue swellings containing granulomas.

Muscle disease

Sarcoidosis of skeletal muscle is usually asymptomatic. In the early acute stages, especially in the presence of erythema nodosum, asymptomatic granulomatous muscle involvement is common, with a prevalence rate ranging between 50 and 80 per cent. Random muscle biopsy may, therefore, be a useful diagnostic procedure in patients presenting with erythema nodosum. Investigators have emphasized the need for generous sampling and examination of multiple tissue sections.

Muscle sarcoidosis has two clinical types: nodular and myopathic (Hinterbuchner and Hinterbuchner 1964). The nodular type often involves extremities and causes solitary or multiple nodules. The myopathic type involves muscles symmetrically and diffusely without causing an intramuscular mass. Slowly progressive myalgia, weakness and atrophy usually occur. Symptomatic muscular sarcoidosis may accompany the acute disease onset and is characterized by fever and severe pain and tenderness, usually involving the proximal muscles of the upper and lower limb girdles (Douglas *et al*. 1973). The histological appearances in acute symptomatic muscular sarcoidosis are identical to those in asymptomatic patients and consist of characteristic granulomatous lesions located between apparently normal muscle fibres. The electromyographic features are those of idiopathic polymyositis. A number of reports have described patients with isolated sarcoid muscle disease without apparent involvement of other organ systems. It is possible that these patients have granulomas in other tissues without their causing clinical manifestations. In patients who appear to have isolated muscle sarcoidosis, systematic clinical, radiological, and histological evaluation will usually reveal evidence of multisystem disease.

^{67}Ga scintigraphy is a useful diagnostic method for demonstrating myopathic disease, whereas abnormalities detected by magnetic resonance imaging may define the nodular type of muscular sarcoidosis (Otake 1994). Computed tomography, ultrasonography, and angiography were less useful in identifying sarcoid muscle disease.

Childhood sarcoidosis

Sarcoidosis is rare in childhood. Chronic arthritis and tenosynovitis have been described (Rosenberg *et al*. 1983). Arthritis usually affects knees and ankles, and is characterized by boggy thickening of synovia or tendon sheaths and effusion, causing minimal pain and limitation (Fig. 10). Arthritis usually occurs in children with onset of sarcoidosis before the age of 5 years, and is frequently associated with ocular and cutaneous manifestations. Sarcoid arthritis in children usually follows an indolent clinical course, with minimal or absent constitutional symptoms.

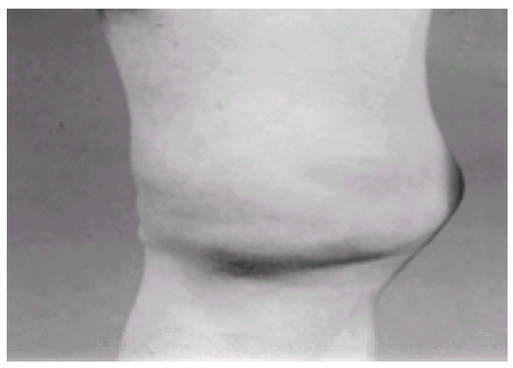


Fig. 10 Chronic sarcoid arthritis with boggy synovial thickening of the knee presenting in a 5-year-old boy who also had uveitis and elevated serum angiotensin-converting enzyme.

Treatment

Acute sarcoidosis with erythema nodosum and hilar lymphadenopathy is usually a self-limiting illness. The arthralgias and arthritis are transient. Symptoms are usually relieved by rest and non-steroidal anti-inflammatory drugs. Occasionally, a short course of corticosteroid therapy may be justified to relieve very acute symptoms. Exacerbations of chronic arthritis are usually associated with active multisystem disease, so that therapeutic decisions usually depend on the severity of the non-articular features. Thus, most patients with chronic arthritis will require corticosteroids for control of systemic disease.

In osseous sarcoidosis, therapy is not indicated for asymptomatic patients. Corticosteroid agents are usually prescribed for symptomatic osseous sarcoidosis, and healing of lytic and destructive lesions has been documented. Similarly, no therapy is required for asymptomatic muscle disease. Corticosteroids are usually recommended for symptomatic patients, although some uncertainties remain about the efficacy of corticosteroid therapy, particularly in chronic sarcoid myopathy.

Chapter References

Baughmann, R. P., Fernandez, M., Bosken, C. H., Mantil, J., and Hurtubise, P. (1984). Comparison of gallium-67 scanning bronchoalveolar lavage and serum angiotensin-converting enzyme levels in pulmonary sarcoidosis: predicting response to therapy. *American Review of Respiratory Diseases*, **129**, 676–81.

Beaumont, D., Herry, S. Y., Sapene, M., Bourguet, P., Larzul, J. J., and De Labarthe, B. (1982). Gallium 67 in the evaluation of sarcoidosis: correlation with serum angiotensin converting enzyme and bronchoalveolar lavage. *Thorax*, **37**, 11–18.

Brennan, J. J., Crean, P., and FitzGerald, M. X. (1984). High prevalence of familial sarcoidosis in an Irish population. *Thorax*, **39**, 14–18.

Brewerton, D. *et al*. (1977). HL-A antigens in sarcoidosis. *Clinical and Experimental Immunology*, **27**, 277–9

British Thoracic and Tuberculosis Association (1973). Familial association in sarcoidosis. *Tubercle*, **54**, 87–98.

Douglas, A. C., Macleod, J. G., and Matthews, J. D. (1973). Symptomatic sarcoidosis of skeletal muscle. *Journal of Neurology, Neurosurgery and Psychiatry*, **36**, 1034–40.

Du Bois, R. M., Kirby, M., Balbi, B., Saltini, C., and Crystal R. D.(1992). Lymphocytes that accumulate in the lung in sarcoidosis have evidence of recent stimulation of the T-cell antigen receptor. *American Review of Respiratory Diseases*, **145**, 1205–11.

Edmondstone, W. M. (1988). Sarcoidosis in nurses: is there an association? *Thorax*, **43**, 342–3.

Forrester, J. M., Newman, L. S., Wang, Y., King, T. E., Jr, and Kotzin, B. L. (1993). Clonal expansion of lung V1+ T cells in pulmonary sarcoidosis. *Journal of Clinical Investigation*, **91**, 292–300.

Grigor, R. R. and Hughes, G. R. V. (1976). Chronic sarcoid arthritis. *British Medical Journal*, **2**, 1044.

Gumpel, J. M., Johns, C. J., and Shulman, L. E. (1967). The joint disease of sarcoidosis. *Annals of the Rheumatic Diseases*, **26**, 194–205.

Hinterbuchner, C. N. and Hinterbuchner, L. P. (1964). Myopathic syndrome in muscular sarcoidosis. *Brain*, **87**, 355–66.

Hunninghake, G. W. and Crystal, R. G. (1981). Pulmonary sarcoidosis. A disorder mediated by excess helper T lymphocyte activity at sites of disease activity. *New England Journal of Medicine*, **305**,

429–34

Hunninghake, G., Gadek, J. E., Kawanami, O., Ferrans, V. J., and Crystal, R. (1979). Inflammatory and immune processes in the human lung in health and disease; evaluation of bronchoalveolar lavage. *American Journal of Pathology*, **97**, 149–206.

Ina, Y., Takada, K., Sato, T., Yamamoto, M., Nodura, M., and Morishita, M. (1992). Soluble interleukin 2 receptors in patients with sarcoidosis: possible origin. *Chest*. **102**, 1128–33.

James, D. G. *et al*. (1976). A worldwide review of sarcoidosis. *Annals of the New York Academy of Science*, **278**, 321–34.

Kataria, Y. *et al*. (1980). Kveim antigen Cr-1: its sensitivity and specificity in sarcoidosis, a comparative study. In *Sarcoidosis and other granulomatous diseases* (ed. W. J. Williams and B. H. Davies), p. 660. Alpha Omega, Cardiff.

Kelner, H., Spathling S., and Herzer, P. (1992). Ultrasound findings in Lofgren's syndrome: is ankle swelling caused by arthritis, tenosynovitis or periarthritis? *Journal of Rheumatology*, **19**, 38–41.

Kremer, J. M. (1986). Histologic findings in siblings with acute sarcoid arthritis; association with the HLA B8, DR3 phenotype. *Journal of Rheumatology*, **13**, 593–7.

Lawrence, E. C., Teague, R. B., Gottlieb, M. S., Jhingran, S. G., and Lieberman, J. (1983). Serial changes in markers of disease activity with corticosteroid treatment in sarcoidosis. *American Journal of Medicine*, **74**, 747–56.

Lofgren, S. (1953). Primary pulmonary sarcoidosis. I. Early signs and symptoms. *Acta Medica Scandinavica*, **145**, 424–31.

Mehra, N. K. and Bovornkitti, S. (1988). HLA and sarcoidosis. *Sarcoidosis*, **5**, 87–9.

Otake, S. (1994). Sarcoidosis involving skeletal muscle: imaging findings and relative value of imaging procedures. *American Journal of Roentgenology*, **162**, 369–75.

Palmer, D. G. and Schumacher, H. R. (1984). Synovitis with non-specific histologic changes in synovium in chronic sarcoidosis. *Annals of the Rheumatic Diseases*, **43**, 778–82.

Pforte, A. *et al*. (1993). Concomitant modulation of serum-soluble interleukin-2 receptor and alveolar macrophage interleukin-2 receptor in sarcoidosis. *American Review of Respiratory Diseases*, **147**, 717–22.

Poukkula, A., Huhti, T., Lilja, M., and Saloheimo, M. (1986). Incidence and clinical picture of sarcoidosis in a circumscribed geographical area. *British Journal of Diseases of the Chest*, 80, 138–47.

Pueringer, R. J., Shwartz, D. A., Dayton, S. C., Gilbert, S. R., and Hunninghake, G. W. (1993). The relationship between alveolar macrophage TNF, interleukin-1 and PGE_2 release, alveolitis and disease severity in sarcoidosis. *Chest*, **103**, 832–8.

Rosenberg, A. M., Yee, E. H., and MacKenzie, J. W. (1983). Arthritis in childhood sarcoidosis. *Journal of Rheumatology*, **10**, 987–90.

Schriber, R. A. and Firooznia, H. (1975). Extensive phalangeal cystic lesions. Sarcoidosis limited to the hands and feet? *Arthritis and Rheumatism*, **18**, 123–8.

Semenzato, G., Zambello, P., Trentin, L., and Agostini, C. (1993). Cellular immunity in sarcoidosis and hypersensitivity pneumonitis: recent advances. *Chest*, **103**, 138–43.

Shakoor, Z. and Hamblin, A. S. (1992). Increased CD11/CD18 expression on peripheral blood leucocytes of patients with sarcoidosis. *Clinical and Experimental Immunology*, **90**, 99–105.

Sharma, O. P. (1984*a*). Clinical Features. In *Sarcoidosis—clinical management*, pp. 22–4, Butterworths, London.

Sharma, O. P. (1984*b*). Laboratory investigations. In *Sarcoidosis—clinical management*, pp. 143–51. Butterworths, London.

Sharma, O. P. (1984*c*). Sarcoidosis of the liver. In *Sarcoidosis—clinical management*, pp. 80–3. Butterworths, London.

Siltzbach, L. E. and Duberstein, J. L. (1968). Arthritis in sarcoidosis. *Clinical Orthopaedics and Related Research*, 57, 31–49.

Spilberg, I., Siltzbach, L. E, and McEwen, C. (1969). The arthritis of sarcoidosis. *Arthritis and Rheumatism*, **12**, 126–37.

Steffen, M. *et al*. (1993). Increased secretion of tumor necrosis factor-a, interleukin-1b and interleukin-6 by alveolar macrophages from patients with sarcoidosis. *Journal of Allergy and Clinical Immunology*, **91**, 939–49.

Studdy, P. R. and James D. G. (1983). The sensitivity and specificity of serum angiotensin-converting enzyme in sarcoidosis and other diseases. In (Proceedings) *Ninth international conference on sarcoidosis and other granulomatous disorders* (ed. J. Cretien, J. Marsac, and J. C. Saltiel), pp. 679–81. Pergamon, Paris.

Teirstein, A. S. (1983). Fibre-optic bronchoscopy in the diagnosis of sarcoidosis. In *Sarcoidosis and other granulomatous diseases of the lung* (ed. B. L. Fanburg), p 323. Marcel Decker, New York.

Teirstein, A. S. and Lesser, M. (1983). World distribution and epidemiology of sarcoidosis. In *Sarcoidosis and other granulomatous diseases of the lung* (ed. B. L. Fanburg), pp. 101–34. Marcel Decker, New York.

Veale, D. and FitzGerald, O. (1990). Acute sarcoid arthropathy—an infective cause? *British Journal of Rheumatology*, **29**, 158–9.

5.13.6 The chronic, infantile, neurological, cutaneous, and articular syndrome (CINCA)

Anne-Marie Prieur

With increasing knowledge in paediatric rheumatology, new inflammatory entities have been more accurately defined. Closer scrutiny of particular sites of joint involvement and long-term follow-up now distinguish certain syndromes from systemic-onset juvenile chronic arthritis. The heterogeneity of so-called systemic-onset juvenile arthritis occurring during the first year of life was stressed several years ago by Prieur and Griscelli (1983). The children in question all had high fever, skin rash, and polyarthritis, but careful analysis of the symptoms allowed us to distinguish various entities. Among these, a group of patients displayed a chronic syndrome that also involved the central nervous system, including the eye and other sensory organs. The autonomy of this syndrome, which in Europe is designated as chronic, infantile, neurological, cutaneous, and articular syndrome (**CINCA**), is now accepted, although its cause remains a mystery.

Clinical presentation

Background

In definitive descriptions of CINCA, the first symptoms occur in early infancy and are often present at birth. The course of the disease is chronic, with intermittent flares associated with fever, and enlargement of lymph nodes and spleen. There have been individual case reports of a disease with these characteristics (Campbell and Clifton 1950; Lorber 1973; Ansell *et al*. 1975, Lampert *et al*. 1975; Fajardo *et al*. 1982), mostly preceding our suggestion that this could be a specific syndrome (Prieur and Griscelli 1981), and others have since confirmed its autonomy (Hassink and Goldsmith 1983; Yarom *et al*. 1985; Kaufman and Lovell 1986). A comprehensive description of the disease is given in Prieur *et al*. (1987). The syndrome is now more easily identified and more fully documented (Torbiak *et al*. 1989).

We have recognized five new cases, in particular a family in which father and daughter, and probably grandmother and great-grandfather, had the disease. The CINCA syndrome tends to be sporadic, but three families where more than one member is affected are known (see also Campbell and Clifton 1950; Ansell *et al*. 1975).

Perinatal events

The frequency of perinatal events is an important feature of this syndrome. These have been identified in half of the neonates in whom details of the perinatal period are known. After generally uneventful pregnancies (except for minor viral infections in a few cases), half of the babies are premature. The average birth weight is 2600 g (range 1700 to 4700 g) after a 38-week pregnancy (range 33 to 42 weeks). Five instances of anomaly in the umbilical cord have been found (umphaloceles or umphalitis). In one case, a histological study of the placenta showed thickened vessel walls with thrombosis and microcalcifications, and infiltration of the umbilical cord by polymorphonuclear cells. There was no evidence of infection in any of the infants, despite the occurrence of respiratory distress in three and icterus in five.

Signs and symptoms

In all cases, the association of skin rash, joint involvement, and central nervous manifestations is present.

Rash

The rash is present at birth in three-quarters of the children or occurs within the first 6 months after birth in the others. In a few patients, it has been noticed later or information on this period of life is lacking. The clinical expression is similar in all patients, although variable in intensity. It resembles the rash of Still's disease but is more intense, often having the appearance of a non-pruritic urticaria (Fig. 1). The rash is exacerbated during flare-ups of other symptoms of the disease. Skin biopsy shows a normal epidermis, mild inflammation in the dermis, and perivascular aggregation of polymorphonuclear cells with some eosinophils. No immunoglobulin or complement deposits have been found by immunofluorescence.

Fig. 1 The skin rash and patella overgrowth with knee deformity. (Reproduced from Prieur *et al*. 1987, with permission.)

Central nervous and sensory anomalies

These are important manifestations of this syndrome. They are usually discovered at the onset of symptoms that could be related to a chronic meningeal irritation, such as headaches, vomiting, seizures, or transitory episodes of hemiplegia. A chronic meningitis has been identified in 23 of 25 patients investigated; these patients had a mild leucocytosis with polymorphs (and sometimes eosinophils), and/or increased protein concentrations in the cerebrospinal fluid. Anomalies in the cerebrospinal fluid may vary at follow-up. Extensive attempts to demonstrate chronic infection, including electron microscopy of cells in the cerebrospinal fluid, have been negative. Clinical follow-up of the central nervous involvement often shows an increase in the meningeal irritation, with headaches and sometimes seizures. Spasticity of the legs is not uncommon, but to date no persistent neurological defect has been observed. A low IQ has been noticed in several patients.

The skull has an increased cranial volume and there is delay in the closure of the anterior fontanelle. In some patients, there are calcifications of the falx and dura (Fig. 2). Computed tomography (**CT**) often reveals a mild ventricular dilatation (Fig. 3).

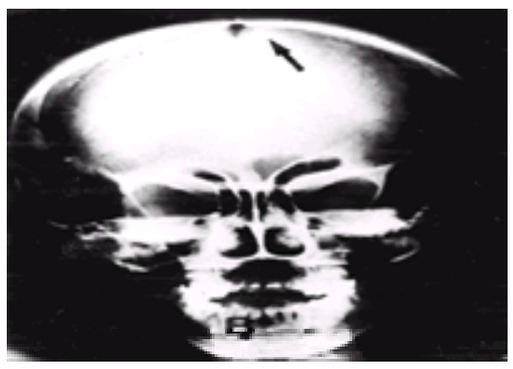


Fig. 2 Calcifications on the falx and dura in a 12-year-old patient. (Reproduced from Prieur *et al*. 1987, with permission.)

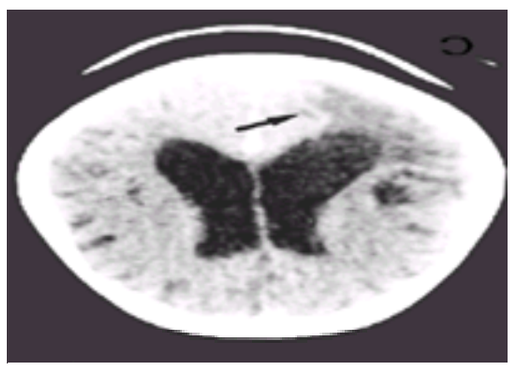


Fig. 3 CT scan: enlargement of ventricles with an area of cerebral change of probable vascular origin. (Reproduced from Prieur *et al*. 1987, with permission.)

Sensory anomalies are progressive. Varying degrees of perceptive deafness are present in older children, often requiring the use of a hearing aid. Eye involvement can be severe: optic atrophy, papillitis, conjunctivitis, keratitis, uveitis, and chorioretinitis have been observed and can lead to a progressive visual defect (Fig. 4). Hoarseness is frequent.

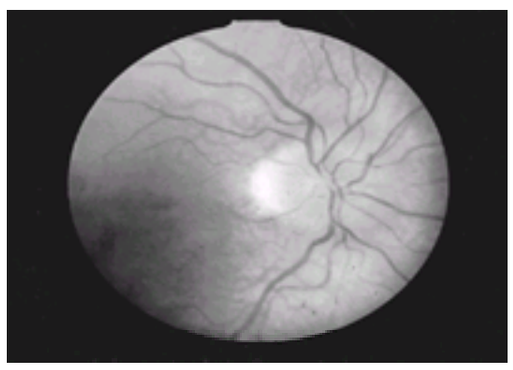


Fig. 4 Papillitis.

Joint involvement

The severity of joint involvement varies among patients. The knees are most frequently affected (see Fig. 1), then elbows, wrists, ankles, and small joints of the hand. In some patients, the joint symptoms are mild, and manifest as transitory inflammation during flare-ups. In this group, the first articular manifestations are generally observed after the age of 2 years. In other patients, the arthropathy may present as early as 10 days of age (Hassink and Goldsmith 1983). The main finding is an overgrowth of the epiphyseal plate of the bone, resulting in hard bony enlargement. The synovial fluid exudate may show a non-specific inflammatory reaction. Progressive contractures and limitation of motion due to patellar and/or epiphyseal overgrowth occur (see Fig. 1).

The radiographic anomalies of affected joints are progressively more characteristic with increasing age in half of the patients. The first finding is swelling of periarticular soft tissues, sometimes with early periosteal reaction of the diaphysis (Kaufman and Lovell 1986). The most characteristic modifications are in the metaphysis and epiphysis of long bones, giving an irregular ossification with abnormal trabeculae. These anomalies may involve growth cartilage plates (Fig. 5); one biopsy showed an irregular metachromasia and completely disorganized columns of cartilaginous cells (Fig. 6). Surprisingly, no inflammatory cells were found in the abnormal cartilage (Fig. 6).

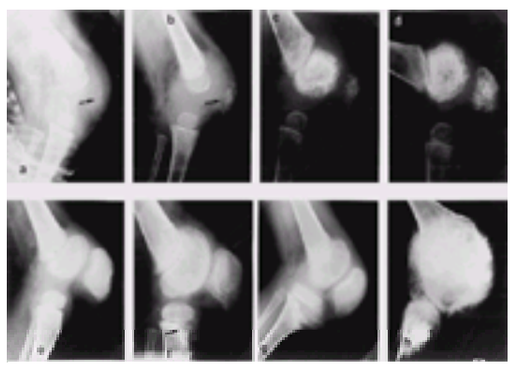


Fig. 5 Follow-up of radiographic anomalies in the knee at various ages: (a) 1 year; (b) 1 year and 6 months; (c) 2 years and 6 months; (d) 3 years and 9 months; (e) 5 years; (f) 6 years (arrow: late modification of tibial growth cartilage); (g) 8 years; (h) 33 years (another patient). (Reproduced from Prieur *et al*. 1987, with permission.)

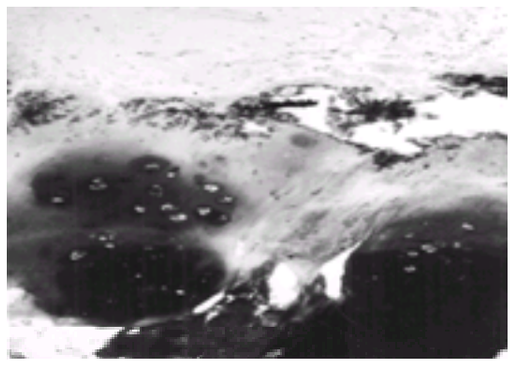


Fig. 6 Biopsy of growth-cartilage overgrowth: irregular metachromasia and disappearance of the normal columns of cartilaginous cells. Note the absence of inflammatory cells in this abnormal cartilaginous tissue. (By courtesy of Giulia Coumot-Witmer.)

Morphological changes

These patients have morphological changes in common. A progressive growth retardation with height below the third percentile is very frequent, but no defect in growth hormone has been found. The heads of these patients have a similar appearance, with enlargement (often with frontal bossing) and blond hair making these unrelated children look like siblings (Fig. 7(a and b)). The adults also resemble each other, as shown in Fig. 7(c and d). A saddle-back nose is frequent and the patient shown in Fig. 7(c) looks surprisingly like one of the patients in the first published descriptions in 1950 (Campbell and Clifton 1950). There is also clubbing of the fingers and toes (Fig. 8). Hands and feet appear short and thick. In some patients, the palms and soles have a wrinkled appearance.

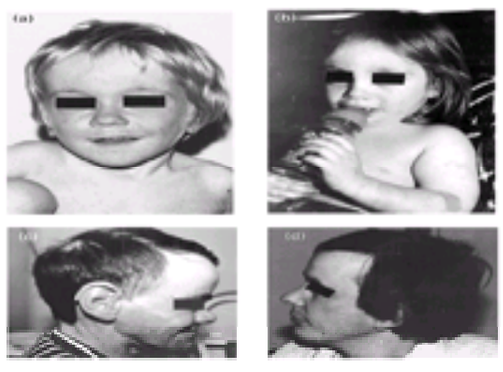


Fig. 7 Morphological aspects of two unrelated children (a) and (b) and two unrelated adults (c) and (d). Patients (b) and (d) are daughter and father.

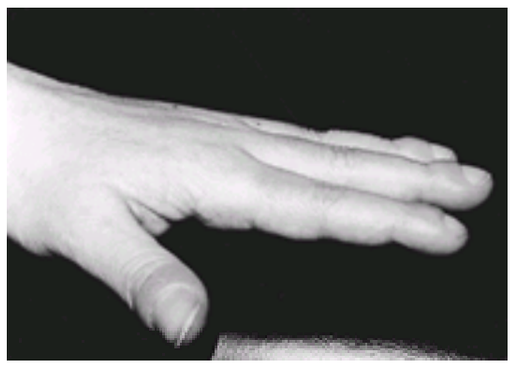


Fig. 8 Clubbing of finger in a 20-year-old girl.

Adult features

These children may reach adulthood, although eight deaths have been reported in children with this syndrome. The causes of death in five of the children were: secondary amyloidosis (Prieur *et al*. 1987); sepsis (Fajardo *et al*. 1982); leukaemia, possibly related to chlorambucil therapy (Prieur and Griscelli 1981); and subacute necrotizing leucoencephalopathy after head injury (Lampert 1986). An adult died of gangrene of the foot (Warin 1977).

Follow-up

Long-term follow-up reveals a slow worsening of most symptoms despite various therapeutic trials. Non-steroidal anti-inflammatory drugs induce pain relief. Steroids are partially effective for fever and pain, but not for skin lesions or joint disease. Immunosuppressive drugs have no dramatic effect; disease-modifying drugs are ineffective. Physiotherapy improves the functional status in patients with severe arthropathy.

Pathogenesis

The pathophysiology of this syndrome is unknown. There are indications of a non-specific inflammation: hypochromic anaemia, leucocytosis with a predominance of polymorphonuclear neutrophils and eosinophils, high platelet counts, and an increase in erythrocyte sedimentation rate and other acute-phase reactants. A polyclonal stimulation of immunoglobulin synthesis has been found in all patients. Surprisingly, no circulating immune complexes have been detected; no autoantibodies or immunodeficiency have been found. The microscopic architecture of the lymph nodes is preserved except for features of chronic inflammation, with an infiltration of the subcortical T-cell region by polymorphonuclear cells and often eosinophils. Extensive investigations for an infectious agent have been unsuccessful. Evidence of a retroviral infection was sought in leucocytes and cells from cerebrospinal fluid in a recently diagnosed patient, but the outcome was negative (C. Rouzioux, unpublished data).

This disorder is an unremitting inflammatory process that begins at birth. Most of the organs or tissues involved, except cartilage, show chronic inflammation. Cartilage may be the target organ in this syndrome, with the clustering of growth retardation, anomalies of ossification of skull and epiphyses, saddle-back nose, hoarseness suggesting laryngeal localization, and progressive deafness. Indeed, some indications of such a mechanism were suggested by the presence of a toxic effect of the serum from these patients on normal human cartilage cells in culture (Prieur *et al*. 1988). However, the origin of this disease remains unknown, even though one may speculate on an initiating event that takes place before birth. It represents an original and new disorder that must be distinguished from previously described rheumatic diseases of childhood.

Chapter References

Ansell, B.M., Bywaters, E.G.L., and Elderkin, F.M. (1975). Familial arthropathy with rash, uveitis and mental retardation. *Proceedings of the Royal Society of Medicine*, **68**, 584–5.

Campbell, A.M.G. and Clifton F. (1950). Adult toxoplasmosis in one family. *Brain*, **73**, 281–90.

Fajardo, J.E., Geller, T.J., Koenig, H.M., and Kleine, M.L. (1982). Chronic meningitis, polyarthritis, lymphadenitis and pulmonary hemosiderosis. *Journal of Pediatrics*, **101**, 738–40.

Hassink, S.G. and Goldsmith T.J. (1983). Noeonatal onset multisystem inflammatory disease. *Arthritis and Rheumatism*, **26**, 668–73.

Kaufman, R.A. and Lovell, D.J. (1986). Infantile-onset multisystem inflammatory disease: radiologic findings. *Radiology*, **160**, 741–6.

Lampert, F. (1986). Infantile multisystem inflammatory disease: another case of a new syndrome. *European Journal of Paediatrics*, **144**, 593–6.

Lampert, F. *et al*. (1975). Infantile chronic relapsing inflammation of the brain, skin and joints. (Letter.) *Lancet*, **i**, 1250–2.

Lorber, J. (1973). Syndrome for diagnosis: dwarfing; persistently open fontanelle; recurrent meningitis; recurrent subdural effusions with temporary alternate-sided hemiplegia; high tone deafness; visual defect with pseudopapilloedema; slowing intellectual development; recurrent acute polyarthritis; erythema marginatum, splenomegaly and iron resistant hypochromic anemia. *Proceedings of the Royal Society of Medicine*, **66**, 1070–1.

Prieur, A.M. and Griscelli, C. (1981). Arthropathy with rash, chronic meningitis, eye lesions and mental retardation. *Journal of Pediatrics*, **99**, 79–83.

Prieur, A.M. and Griscelli, C. (1983). Aspect nosologique des formes systémiques de l'arthrite juvénile début très précoce: a propos de dix-sept observations. *Annales de Pediatrie* (*Paris*), **30**, 565–9.

Prieur, A.M. *et al*. (1987). A chronic, infantile, neurological, cutaneous and articular (CINCA) syndrome: a specific entity analysed in 30 patients. *Scandinavian Journal of Rheumatology*, **66**, 57–68.

Prieur, A.M., Cournot-Witmer, G., Plachot, J.J., Griscelli, C., and Corvol, M.T. (1988). *In vitro* study of growth cartilage in children with the infantile onset neurologic, cutaneous and articular syndrome. Inhibitory effect of patient's sera on cultured normal human chondrocyte metabolism. *Arthritis and Rheumatism*, **31**, 578.

Torbiak, R.P., Dent, P.B., and Colkshott, W.P. (1989). NOMID–a neonatal syndrome of multisystem inflammation. *Skeletal Radiology*, **18**, 359–64.

Warin, R.P. (1977). Familial vasculitis (life-long morbiliform urticaria, neurological changes and arthritis). *British Journal of Dermatology*, **97**, 30–1.

Yarom, A., Rennebohm, R.M., and Levinson, J.E. (1985). Infantile multisystem inflammatory disease: a specific syndrome? *Journal of Pediatrics*, **106**, 390–6.

5.13.7 Multicentric reticulohistiocytosis

P. J. Maddison

Multicentric reticulohistiocytosis is a rare systemic disease that is recognized clinically by the combination of typical papular and nodular skin lesions and a severe and destructive polyarthritis, although virtually any organ system of the body can be involved. The term 'multicentric reticulohistiocytosis' was introduced by Goltz and Layman (1954) to distinguish the disorder from solitary cutaneous nodules, termed reticulohistiocytomas, which have identical histological appearances but are not associated with systemic disease. There are a number of synonyms, which include lipoid dermatoarthritis, reticulohistiocytosis, giant-cell reticulohistiocytosis, and normocholesterolaemic xanthomatosis. At one time it was thought to be a lipid-storage disease but no consistent abnormality of serum or intracellular lipids has been identified. It is thought generally that the histiocytes and giant cells that characterize the lesions contain a non-specific accumulation of glycoprotein and lipids, and that this disease is a histiocytic granulomatous reaction to an unknown stimulus (Campbell and Edwards 1991). However, in one case, type VI collagen inclusions usually found in lymphohistiocytic neoplasms were demonstrated, supporting the concept of a proliferative rather than an inflammatory background for this condition (Fortier-Beaulieu *et al*. 1993).

Clinical features

Multicentric reticulohistiocytosis is primarily a disorder of adults, being slightly more common in women (Barrow and Holubar 1969), with a mean age of onset in the fifth decade. About 85 per cent of reported cases are caucasoid, but it can also occur in other ethnic groups (Lesher and Allen 1984). Typical cases have been reported in adolescence (Raphael *et al*. 1989), but in most instances the childhood form differs from that in adults in having a familial component, a marked tendency to involve the eye (glaucoma, uveitis, and cataracts), and a lack of giant cells in tissue sections, the last feature being considered an essential feature of multicentric reticulohistiocytosis (Zayid and Farraj 1973).

Skin nodules and a destructive polyarthritis are the most common features (Barrow and Holubar 1969; Chevrant-Breton 1977; Rapini, 1993). Arthritis is frequently the presenting feature. It has a similar distribution to rheumatoid arthritis, including the temporomandibular joints and atlantoaxial involvement (Gold *et al*. 1975), but also affects the distal interphalangeal joints. Commonly involved sites are the hands (75 per cent), knees (70 per cent), shoulders (65 per cent), wrists (65 per cent), hips (60 per cent), ankles (60 per cent), elbows (60 per cent), feet (60 per cent), and spine (50 per cent) Often it has a persistent course with progressive joint destruction, sometimes leading to the picture of arthritis mutilans with 'opera glass' deformities of the hand developing in approximately 20 per cent (Barrow and Holubar 1969). Radiographs of the joints show symmetrical erosions resembling rheumatoid arthritis. However, osteoporosis is not marked (Maki *et al*. 1995) and there is often prominent involvement of the distal interphalangeal joints, and the spread of erosions from the periphery of the joint produces apparent widening of the joint space (Fig. 1). Analysis of joint fluid shows low to moderate cell counts, mostly with a preponderance of mononuclear cells but sometimes of neutrophils. Synovial biopsy shows the presence of lipid-laden giant cells and histiocytes, as found in the involved skin.

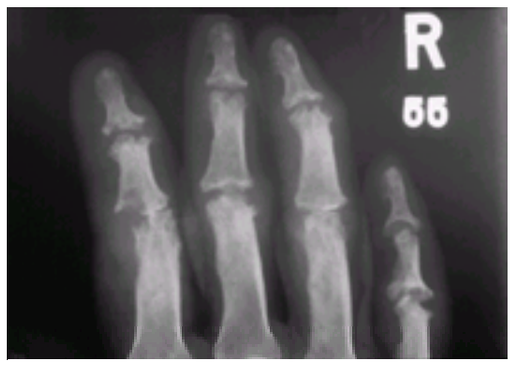


Fig. 1 Radiograph of the hand in multicentric reticulohistiocytosis.

Mucocutaneous involvement is apparent in approximately 90 per cent of patients at the time of presentation but may occur in the months following the development of arthritis. Occasionally it is the first feature. Skin lesions consist of numerous, usually non-pruritic, skin-coloured, yellowish or red-brown nodules, ranging from a few millimetres to several centimetres in diameter. Occasionally, however, pruritus is a feature, and can be severe and precede the overt lesions. The nodules mostly involve the hands in a periungual distribution, together with nodules scattered on the fingers and on the face, where lesions develop on the ears, at the corners of the mouth, and the side of the nose adjacent to the alae (Fig. 2 and Fig. 3). Extensive involvement of the face produces a leonine facies. Lesions can also occur on the forehead, chest, back, and over the olecranon process and knees. Xanthelasma develops in approximately 25 per cent. The oral mucous membranes are involved in approximately half, with, for example, numerous flesh-coloured papules on the buccal mucosa, sides of the tongue, surface of the palate, and gingiva. Occasionally the pharynx and larynx are affected (Katz and Anderson 1988). To date there have been no reports of vaginal or perianal involvement.

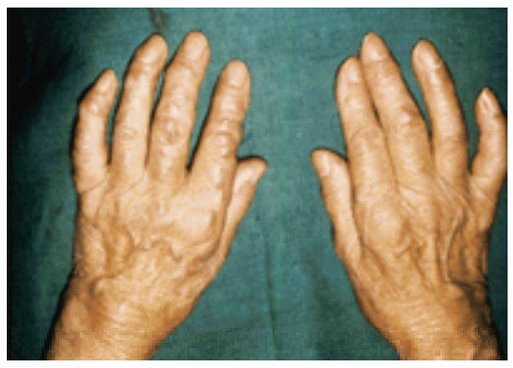


Fig. 2 Multicentric reticulohistiocytosis: nodules involving the hands.

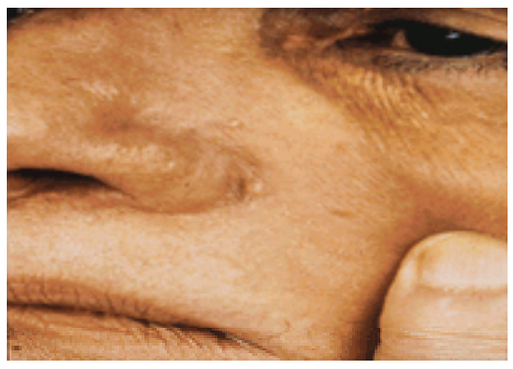


Fig. 3 Multicentric reticulohistiocytosis: involvement of the face.

Although multicentric reticulohistiocytosis usually affects the skin and joints, a large variety of systemic features can occur, reflecting the potential involvement of most tissues and organs of the body. There are also reported associations with various types of malignancy. Constitutional manifestations occur in about one-third of patients, and include fever and weight loss; cases have been reported with pericarditis, pulmonary involvement, and myositis (Anderson *et al*. 1968; Fast 1976; Lesher and Allen 1984; Widman *et al*. 1988; Gao *et al*. 1990). There have also been case reports of lymph node enlargement, splenomegaly, marrow infiltration resulting in pancytopenia, salivary gland involvement, thyroid infiltration leading to hypothyroidism, and carpal tunnel syndrome (Warin *et al*. 1957; Orkin *et al*. 1964; Ehrlich *et al*. 1972; Furey *et al*. 1983; Finelli *et al*. 1986; Rene *et al*. 1990).

Laboratory abnormalities occur, but are inconsistent, and include a raised erythrocyte sedimentation rate, abnormal liver function tests, and elevated creatine phosphokinase. Paraproteinaemia of g heavy-chain type has been encountered but serological tests for such as rheumatoid factor are otherwise usually normal. Gallium scintiscans have been used to evaluate multicentric reticulohistiocytosis but the abnormalities are non-specific (Widman *et al*. 1988).

Approximately 25 per cent of patients with multicentric reticulohistiocytosis have been reported to have malignant disease (Catterall and White 1978). Associations have mostly been with internal carcinomas of the colon, breast, bronchus, cervix, ovary, and stomach. Cases have also been associated with sarcomas and lymphomas. Patients with malignant melanoma and pleural mesothelioma have also been reported (Coupe *et al*. 1987; Gibson *et al*. 1995). In one review of the literature (Nunnick *et al*. 1985), the onset of both diseases occurred in several patients within 1 year. The onset of multicentric reticulohistiocytosis in adulthood necessitates, therefore, a thorough clinical examination, including a chest radiograph, to look for signs of malignant disease. It is not necessary, however, to embark on invasive procedures to look for occult neoplasms. It should be borne in mind that there may be a bias towards reporting cases with underlying malignancy, especially previously unreported tumours, and some of the associations may be purely coincidental.

Pathology

Histologically, the nodular infiltrates, which have a granular or ground-glass appearance, consist of multinucleate (up to 20 nuclei) giant cells of the foreign-body type and histiocytes of the monocyte–macrophage lineage (Heathcote *et al*. 1985) admixed with small numbers of CD4+ lymphocytes and B cells. Histochemically, the cytoplasm of the giant cells contains diastase-resistant, periodic acid–Schiff-positive material thought to be glycoprotein, neutral fats, phospholipids, and iron (Barrow and Holubar 1969). Ultrastructural studies show the presence of pleomorphic cytoplasmic inclusions (Degas *et al*. 1975; Caputo *et al*. 1981). The giant cells and mononuclear histiocytes express the surface markers of the monocyte–macrophage lineage (Salisbury *et al*. 1990; Zegler *et al*. 1994), and cytokines and enzyme products of activated macrophages have been detected (Lotti *et al*. 1988; Zagala *et al*. 1988). Negative staining with markers such as CD1 and the absence of Birbeck granules helps to confirm that the cells are not derived from Langerhans cells, in contrast to histiocytosis X (Salisbury *et al*. 1990).

Diagnosis

The diagnosis is confirmed by skin or synovial biopsy. Solitary reticulohistiocytomas have identical histological findings but there is no evidence of systemic disease (Zegler *et al*. 1994).

Diseases that may possibly be confused with multicentric reticulohistiocytosis on the basis of clinical or histological findings are:

1. rheumatoid or psoriatic arthritis;
2. sarcoid dactylitis—associated with discrete tubercles that lack foam cells or giant cells typical of multicentric reticulohistiocytosis;
3. xanthomas;
4. fibroxanthoma;
5. histiocytosis X—usually presents in childhood and proliferating cells are epidermotropic Langerhans cells rather than of true monocyte–macrophage lineage;
6. generalized eruptive histiocytoma—not associated with arthritis and lesions do not exhibit the typical multinucleate giant cells;
7. tendon sheath giant-cell tumours—solitary, well-circumscribed nodules in the hands;
8. Farber's disease—lipogranulomatosis, usually fatal in infancy.

Treatment and course

The disease often runs a waxing and waning course and sometimes it stabilizes. It is difficult to predict the course in the individual case but as a rule the disease 'burns out' after 5 to 8 years, with regression of the cutaneous nodules (Lyell and Carr 1959; Albert *et al*. 1960), leaving the patient with severe joint deformities. Occasionally, systemic involvement can be severe enough to be life threatening (Barrow and Holubar 1969) but otherwise the prognosis is dominated by whether or not there is an associated malignancy.

Various drug treatments have been proposed to suppress the condition, including corticosteroids, adrenocorticotrophic hormone, salicylates, antimalarials, and cytotoxic agents. There are no controlled trials, as the condition is so rare, and therefore reports of efficacy should be interpreted with caution (Table 1).

Drug	Comments
Prednisolone	
Taylor (1977)	Single case history with improvement in pruritis and slow resolution of skin lesions
Zoale *et al*. (1980)	High-dose prednisolone effective in one patient
Cytotoxic drugs	
Hanauer (1973)	Remission on cyclophosphamide (one case)
Ehrlich *et al*. (1972)	Response in one patient to combination of azathioprine and prednisolone
Krey *et al*. (1974)	No response to azathioprine (one patient)
Gold *et al*. (1975)	Cyclophosphamide improved skin and joint disease (one case)
Brandt *et al*. (1982)	Improvement on cyclophosphamide with reduction in erosions on hand radiograph (one case); previous improvement of skin lesions with topical nitrogen mustard
Hanser *et al*. (1983)	Good response to chlorambucil before patient died from breast carcinoma
Doherty *et al*. (1984)	Good response first to cyclophosphamide, then to chlorambucil (one case)
Coupe *et al*. (1987)	Skin and joint manifestations responded well to cyclophosphamide until the patient died from associated mesothelioma
Grinberg *et al*. (1988)	Five patients treated with cyclophosphamide, one with chlorambucil, all improved, five went into remission, one relapsed after treatment stopped
Kurtz *et al*. (1988)	Response to cyclophosphamide until the patient died of recurrent melanoma
Other agents	
Davies *et al*. (1968)	Improvement in two patients on combination of hydroxychloroquine and prednisolone
Gold *et al*. (1977)	Improvement on rifampicin plus ethambutol in a patient with coexistent tuberculosis
Doherty *et al*. (1984)	No response to penicillamine (one case)
Kaur (1988)	No response to methotrexate (one case)
Gaumetou *et al*. (1991)	Good response to methotrexate resulting in prolonged remission (one case)

Table 1 Drug treatments for multicentric reticulohistiocytosis

Corticosteroids are of limited use. Skin infiltration does not usually respond, although there may be relief of troublesome pruritus (Taylor 1977). Joint symptoms

sometimes improve (Davies *et al*. 1968).

There is a consensus that alkylating agents are the treatment of choice and there are several reports of improvement in skin, joint, and systemic disease with cyclophosphamide in particular (Ginsburg *et al*. 1989). Patients with long-lasting remission after therapy with cyclophosphamide or chlorambucil have been reported (Hanauer 1972; Brandt *et al*. 1982; Coupe *et al*. 1987; Ginsburg *et al*. 1989; Kenik *et al*. 1990), but some do relapse when treatment is stopped (Ginsburg *et al*. 1989). Methotrexate is also reportedly successful (Gourmelen *et al*. 1991), but not in all cases (Giam 1988).

Chapter References

Albert, J., Bruce, W., Allen, A. C., and Blank, M. (1960). Lipoid dermatoarthritis: reticulohistiocytoma of the skin and joints. *American Journal of Medicine*, **28**, 661–7.

Anderson, T. E., Carr, A. J., Chapman, R. J., Downie, A. W., and Maclean, G. D. (1968). Myositis and myotonia in a case of multicentric reticulohistiocytosis. *British Journal of Dermatology*, **80**, 39–45.

Barrow, M. V. and Holubar, K. (1969). Multicentric reticulohistiocytosis: a review of 33 patients. *Medicine (Baltimore)*, **48**, 287–305.

Brandt, F., Lipman, M., Taylor, J. R., and Halprin, K. M. (1982). Topical nitrogen mustard therapy in multicentric reticulohistiocytosis. *Journal of the American Academy of Dermatology*, **6**, 260–2.

Campbell, D. A., and Edwards, N. L. (1991). Multicentric reticulohistiocytosis: systemic macrophage disorder. *Baillières Clinical Rheumatology*, **5**, 301–19.

Caputo, R., Alessi, E., and Berti, E. (1981). Collagen phagocytosis in multicentric reticulohistiocytosis. *Journal of Investigative Dermatology*, **76**, 342–6.

Catterall, M. and White, J. (1978). Multicentric reticulohistiocytosis and malignant disease. *British Journal of Dermatology*, **98**, 221–4.

Chevrant-Breton, J. (1977). La reticulo-histiocytose multicentrique: revue de la litterature recente (depuis 1969). *Annales de Dermatologie et de Venereologie (Paris)*, **104**, 745–9.

Coupe, M. O., Whittaker, S. J., and Thatcher, N. (1987). Multicentric reticulohistiocytosis. *British Journal of Dermatology*, **116**, 245–7.

Davies, M. E., Roenigk, H. H., Hawk, W. A., and O'Duffy, J. D. (1968). Multicentric reticulohistiocytosis: report of a case with histochemical studies. *Archives of Dermatology*, **97**, 543–7.

Degas, R., Civatte, J., Bonvalet, D., and Bouhin, F. (1975). Reticulo-histiocytoma gigantocellulaire. *Annales de Dermatologie et de Syphiligraphie (Paris)*, **102**, 315–17.

Doherty, M., Martin, M. F., and Dieppe, P. A. (1984). Multicentric reticulohistiocytosis associated with primary biliary cirrhosis: successful treatment with cytotoxic agents. *Arthritis and Rheumatism*, **27**, 844–8.

Ehrlich, G., Young, I., Nosheny, S., and Katz, W. (1972). Multicentric reticulohistiocytosis (lipoid dermatoarthritis): a multisystem disorder. *American Journal of Medicine*, **52**, 830–40.

Fast, A. (1976). Cardiopulmonary complications in multicentric reticulohistiocytosis. *Archives of Dermatology*, **112**,1139–41.

Finelli, L. G., Tenner, L. K., and Ratz, J. L. (1986). A case of multicentric reticulohistiocytosis with thyroid involvement. *Journal of the American Academy of Dermatology*, **15**, 1097–1100.

Fortier-Beaulieu, M., Thomine, E., Boullie, M. C., Le Loet, X., Lauret, P., and Hemet, J. (1993). *American Journal of Dermatopathology*, **15**, 587–9.

Furey, N., DiMauro, J., Eng, A., and Shaw, J. (1983). Multicentric reticulohistiocytosis with salivary gland involvement and pericardial effusion. *Journal of the American Academy of Dermatology*, **8**, 679–84.

Gao, I. K., Goronzy, J. J., and Weyand, C. M. (1990). Ten-year follow-up of a patient with multicentric reticulohistiocytosis associated with myopathy. *Scandinavian Journal of Rheumatology*, **19**, 437–41.

Giam, Y. C. (1988). Multicentric reticulohistiocytosis. *Annals of the Academy of Medicine of Singapore*, **17**, 548–50.

Gibson, G., Cassidy, M., O'Connell, P., and Murphy, G. M. (1995). Multicentric reticulohistiocytosis associated with recurrence of malignant melanoma. *Journal of the American Academy of Dermatology*, **32**, 134–6.

Ginsburg, W. W., O'Duffy, J. D., and Morris, J. L. (1989). Multicentric reticulohistiocytosis: response to alkylating agents in six patients. *Annals of Internal Medicine*, **111**, 384–8.

Gold, R. H., Metzger, A. L., and Mirra, J. M. (1975). Multicentric reticulohistiocytosis (lipoid dermato-arthritis): an erosive polyarthritis and distinctive clinical, roentgenographic and pathological features. *American Journal of Roentgenology, Radiation Therapy and Nuclear Medicine*, **124**, 610–24.

Gold, K. D., Sharp, J. T., Estrada, R. G., Duffy, J., and Person, D. A. (1977). Relationship between multicentric reticulohistiocytosis and tuberculosis. *Journal of the American Medical Association*, **237**, 2213–14.

Goltz, R. W. and Layman, C. W. (1954). Multicentric reticulohistiocytosis of the skin and synovia. *Archives of Dermatology and Syphilology*, **69**, 717–22.

Gourmelen, O. *et al*. (1991). Methotrexate treatment for multicentric reticulohistiocytosis. *Journal of Rheumatology*, **18**, 627–8.

Hanauer, L. B. (1972). Reticulohistiocytosis: remission after cyclophosphamide therapy. *Arthritis and Rheumatism*, **15**, 636–40.

Heathcote, J. G., Guenther, L. C., and Wallace, A. C. (1985). Multicentric reticulohistiocytosis: A report of a case and a review of the pathology. *Pathology*, **17**, 601–8.

Heenan, P. J., Quirk, C. J., and Spagnolo, D. V. (1983). Multicentric reticulohistiocytosis: a light and electron microscopic study. *Australasian Journal of Dermatology*, **24**, 122–6.

Katz, R. W. and Anderson, K. F. (1988). Multicentric reticulohistiocytosis. *Oral Surgery, Oral Medicine, Oral Pathology*, **65**, 721–5.

Kenik, J.G., Fok, F., Huerter, C.J., Hurley, J.A., and Stanoshek, J.F. (1990). Multicentric reticulohistiocytosis in a patient with malignant melanoma: a response to cyclophosphamide and a unique cutaneous feature. *Arthritis and Rheumatism*, **33**, 1047–51.

Krey, P. R., Comerford, F. R., and Cohen, A. S. (1974). Multicentric reticulohistiocytosis: fine structural analysis of the synovium and synovial fluid cells. *Arthritis and Rheumatism*, **17**, 615–33.

Lesher, J. L. and Allen, B. S. (1984). Multicentric reticulohistiocytosis. *Journal of the American Academy of Dermatology*, **11**, 713–23.

Lotti, T., Santucci, M., Casigliani, R., Fabbri, P., Bondi, R., and Panconesi, E. (1988). Multicentric reticulohistiocytosis: report of three cases with evaluation of tissue proteinase activity. *American Journal of Dermatopathology*, **10**, 497–504.

Lyell, A. and Carr, A. (1959). Lipoid dermato-arthritis (reticulohistiocytosis). *British Journal of Dermatology*, **71**, 12–21.

Maki, D. D., Caperton, E. M., and Griffiths, H. J. (1995). Radiologic case study. Multicentric reticulohistiocytosis. *Orthopedics*, **18**, 77–81.

Nunnick, J. C., Krusinski, P. A., and Yates, J. W. (1985). Multicentric reticulohistiocytosis and cancer: a case report and review of the literature. *Medical Pediatric Oncology*, **13**, 273–9.

Orkin, M., Goltz, R. W., Good, R. A., Michael, A., and Fisher, I. (1964). A study of multicentric reticulohistiocytosis. *Archives of Dermatology*, **89**, 640–54.

Raphael, S. A., Cowdery, S. L., Faerber, E. N., Lischner, H. W., Shumacher, R., and Tourtellotte, C. D. (1989). Multicentric reticulohistiocytosis in a child. *Journal of Paediatrics*, **114**, 266–9.

Rapini, R. P. (1993). Multicentric reticulohistiocytosis. *Clinics in Dermatology*, **11**, 107–11.

Rene, J., Starz, T., and Miller, E. B. (1990). Multicentric reticulohistiocytosis and Sjögren's syndrome. *Arthritis and Rheumatism*, **33**, 1870–1.

Salisbury, J., Hall, P. A., Williams, H. C., Mangi, M. H., and Mufti, G. J. (1990). Multicentric reticulohistiocytosis: detailed immunophenotyping confirms macrophage origin. *American Journal of Surgical Pathology*, **14**, 687–93.

Taylor, D. R. (1977). Multicentric reticulohistiocytosis. *Archives of Dermatology*, **113**, 320–2.

Warin, R. P., Evans, C. D., Hewitt, M., Taylor, A. L., Price, C. H., and Middlemiss, J. H. (1957). Reticulohistiocytosis (lipoid dermato-arthritis). *British Medical Journal*, **1**, 1387–91.

Widman, D., Swayne, L. C., and Rozan, S. (1988). Multicentric reticulohistiocytosis: Assessment of pulmonary disease by gallium-67 scintigraphy. *Journal of Rheumatology*, **15**, 132–5.

Zagala, A., Guyot, A., Benson, J. C., and Phelip, X. (1988). Multicentric reticulohistiocytosis: a case with enhanced interleukin-1 prostaglandin E_2 and interleukin-2 secretion. *Journal of Rheumatology*, **15**, 136–8.

Zayid, I. and Farraj, F. (1973). Familial histiocytic dermatitis: a new syndrome. *American Journal of Medicine*, **54**, 793–800.

Zeale, P. J., Miner, D., Honig, S., Waxman, M., and Bartfield, H. (1985). Multicentric reticulohistiocytosis: a cause of dysphagia with response to corticosteroids. *Arthritis and Rheumatism*, **28**, 231–4.

Zegler, B., Soyer, H. P., Misch, K., Orchard, G., and Wilson-Jones, E. (1994). Reticulohistiocytoma and multicentric reticulohistiocytosis. Histopathologic and immunophenotypic distinct entities. *American Journal of Dermatopathology*, **16**, 577–84.

5.13.8 Hyperlipidaemias

Keng Hong Leong, J. Reckless and Neil John McHugh

Introduction

The relationship of lipid disorders to rheumatic diseases is multifaceted. For instance several abnormalities in the lipid and lipoprotein profile (or dyslipidaemia) have been documented in various forms of chronic arthritis. The dyslipidaemia may depend on the type or activity of the joint disease or may be influenced by other factors such as treatment or systemic complications of the disease. Whenever the dyslipidaemia is present, its severity will be influenced by diet and lifestyle factors, and also by the underlying genetic predisposition. Renal involvement in systemic lupus erythematosus is an example where the disease, treatment of the disease, or complications of the disease such as nephrotic syndrome may each have a separate influence on lipid metabolism. The dyslipidaemia may potentially influence disease outcome itself by accelerating atheroma, although the relative contribution of dyslipidaemia to the increased mortality from vascular disease observed in several rheumatic disorders is uncertain.

Musculoskeletal symptoms may be a manifestation of an underlying hyperlipidaemia and will be described in more detail. There is some evidence that potentially toxic lipid particles, such as oxidized low density lipoprotein (**LDL**), may accumulate within joints (Winyard *et al*. 1993), as well as having more recognized proinflammatory effects on other tissues. Oxidized LDL may be a target for potentially pathogenic autoantibodies that cross-react with phospholipids (Vaarala *et al*. 1993). Conversely, dietary modification of lipid intake may have beneficial effects on inflammatory disease. Dietary supplementation with polyunsaturated fatty acids of the omega 3 and omega 6 series may be effective therapy for certain rheumatic diseases as it is in animal models of inflammatory arthritis. This chapter will cover some of the above relationships between lipids and rheumatic disease and give an outline of appropriate management where relevant.

Lipid metabolism

A comprehensive review of lipid metabolism would not be appropriate but a brief reminder of certain aspects relevant to rheumatology is given here.

Lipid composition

Lipoproteins are complex particles that transport lipids such as triglycerides and cholesterol through the plasma. The composition of a typical lipoprotein particle is shown in Fig. 1. Each particle is composed of a non-polar core containing variable proportions of hydrophobic lipids (triglycerides and cholesterol ester) surrounded by a polar coat of phospholipids and free cholesterol, within which there are various apolipoproteins. This outer coat allows the particle to be soluble in aqueous plasma. The apolipoproteins have both structural roles in the lipoproteins and functional roles as enzyme activators or as ligands with high affinity for specific tissue-cell receptors. There are five main types of lipoprotein, defined by their relative density, that differ in their composition of lipid and the type of apolipoprotein (Table 1). Apolipoprotein A (**apoA**) is the major apolipoprotein in high density lipoprotein (**HDL**) and apolipoprotein B (**apoB**) is the major structural apolipoprotein of the other lipoproteins, chylomicrons, very low dense lipoprotein (**VLDL**), intermediate density lipoprotein (**IDL**), and LDL. The full apoB (apoB100) is found in VLDL, IDL, and LDL, having structural and functional components, but a truncated form, apoB48, having only the structural role, is present in chylomicrons made by gut cells. The main cholesterol-carrying lipoproteins are LDL and HDL while the main triglyceride-carrying lipoproteins are chylomicrons and VLDL. As VLDL is catabolized, remnants are formed known as IDL which are normally rapidly removed from the liver or converted to LDL.

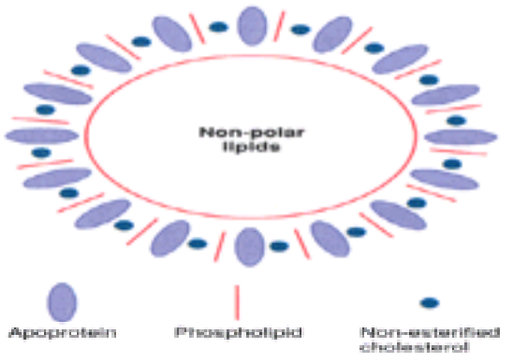


Fig. 1 Composition of typical lipoprotein particle.

Lipoprotein	Major lipids	Major lipoproteins
Chylomicrons and remnarts	Dietary triglycerides	AI, AII, B48, CI, CII, CIII, E
VLDL	Endogenous triglycerides	B48, CI, CII, CIII, E
IDL	Cholesteryl esters, triglycerides	B100, CIII, E
LDL	Cholesteryl esters	B100
HDL	Cholesteryl esters	AI, AII

Table 1 The major classes of lipoproteins

Lipid subclasses

The major lipoproteins are heterogeneous and subclasses can be detected by methods such as precipitation or ultracentrifugation. HDL can be separated into two main fractions—HDL2 and HDL3. HDL3 accepts cholesterol from cholesterol-replete tissues, converts to larger HDL2, which in turn delivers cholesterol to the liver or other lipoproteins. LDL exists in three main subclasses—LDL1, LDL2, and LDL3—which can be separated by cumulative flotation ultracentrifugtion. The normal LDL pattern is predominantly LDL2 with some LDL1 and less LDL3, but under certain circumstances this pattern can shift towards smaller particles with a predominance of small, dense LDL3 which is cholesterol depleted and relatively triglyceride rich. LDL3 is less readily recognized by normal LDL receptors and is more readily recognized (and much more likely to be taken up) by macrophages in the arterial subintimal space which form the foam cells of the fatty streak of early atheroma (Campos *et al*. 1992). VLDL can be separated into three main fractions—VLDL1, VLDL2, and VLDL3. The larger, more buoyant, triglyceride-rich VLDL1 particles, formed especially when there is increased hepatic triglyceride synthesis, are more likely to be precursors for LDL3, while small VLDL particles favour generation of the larger (more normal) LDL subclasses. Changes in the proportion of lipoprotein subclasses will not be apparent on basic lipid screening, yet this has an important bearing on the risk of atherosclerosis and may be influenced by metabolic and systemic alterations that accompany chronic rheumatic disorders. For example triglyceride lipases that have important roles in determining lipoprotein composition are influenced by cytokines that are upregulated in inflammatory conditions (as discussed below).

Exogenous pathway of lipid transport

Chylomicrons carry dietary triglycerides and cholesterol from the intestine to adipose tissue and skeletal muscle. Lipoprotein lipase liberates free fatty acid and monoglycerides which cross the endothelium to the peripheral tissue for further metabolism. The chylomicron remnant is removed by the liver.

Endogenous pathway of lipid transport

Triglycerides are synthesized in the liver and incorporated into VLDL particles. VLDL is a substrate for lipoprotein lipase in peripheral tissue where triglycerides are hydrolysed. VLDL becomes IDL which is either taken back up by the liver or loses further triglyceride and apolipoproteins apart from apoB and becomes cholesterol-rich LDL. Both the liver and extrahepatic tissues that utilize cholesterol recognize LDL via receptor-mediated pathways. Otherwise a variable amount of LDL is scavenged by the reticuloendothelial system and particularly by macrophages. HDL can accept cholesterol from cholesterol-replete tissues. Cholesterol is esterified on the surface of HDL by the enzyme lecithin:cholesterol acyltransferase to become a HDL core cholesterol ester. Subsequently, cholesterol ester in HDL can be exchanged with VLDL and LDL through the action of cholesterol ester transfer protein, or HDL cholesterol can be returned to the liver.

Lipoprotein (a)

ApoA is a complex protein with a protease domain and multiple repeats, called kringles, which have considerable homology with plasminogen. It is linked by a single disulphide bridge to the apoB of a LDL molecule to give lipoprotein (a). The apo (a) gene is highly polymorphic and lipoprotein (a) levels are mainly genetically determined. Interest in lipoprotein (a) stems from its reported association with the presence, severity, and progression of coronary heart disease (Terres *et al*. 1995). However, the risk of atherosclerosis may be most when high levels of lipoprotein (a) are associated with concomitant elevation in LDL–cholesterol levels (Maher and Brown 1995). High lipoprotein (a) levels are linked to premature macrovascular disease, but may not be as good a predictor of atheroma in the long term. Recent work in transgenic mice suggests that lipoprotein (a) acts to inhibit plasminogen activation, leading to reduced activation of transforming growth factor-b which has several adverse effects on vessel wall function (Grainger *et al*. 1994).

Lipids and inflammation

Essential fatty acids and inflammatory mediators

In addition to being core components of complex lipids, fatty acids are either incorporated into cell membranes or serve as precursors for biologically active metabolites, such as prostaglandins and leukotrienes. Essential fatty acids such as linoleic acid cannot be synthesized by mammals and are required in the diet. Polyunsaturated fatty acids contain at least two double bonds in their carbon backbone. Arachidonic acid is the commonest dietary polyunsaturated fatty acids and is a precursor for the potent proinflammatory 2-series prostanoids, via the action of cyclo-oxygenase, and for the 5-series leukotrienes, via the action of 5-lipoxygenase. Metabolites of arachidonic acid are also intimately involved in the early events controlling coagulation and vascular homeostasis. Platelet-derived thromboxane-A_2 is a powerful vasoconstrictor and causes platelet aggregation, whereas endothelial-derived prostacyclin, or PGI_2, causes vasodilatation and opposes platelet aggregation.

Omega-3 fatty acids are essential polyunsaturated fatty acids found in fish oils and include eicosapentaenoic acid. They are precursors of the 3-series prostanoids and the 5-series leukotrienes which competitively inhibit formation of proinflammatory eicosanoids derived from arachidonic acid. Another natural source of essential fatty acids is evening primrose oil which contains predominately gammalinoleic acid. This is a precursor of prostaglandin E_1 and of 15-OH-dihomo-gammalinoleic acid. Prostaglandin E_1 has anti-inflammatory effects and 15-OH-dihomo-gammalinoleic acid inhibits both 5- and 12-lipoxygenases, which generate proinflammatory eicosanoids.

Lipid peroxidation and antioxidants

Polyunsaturated fatty acids are the main target for lipoperoxidative injury by oxygen free radicals. In particular, polyunsaturated fatty acids and their attached phospholipid in plasma LDL can be oxidized by endothelial cells and macrophages. Oxidized LDL may exert several proinflammatory effects by virtue of its chemotactic, cytotoxic, and, possibly, immunogenic properties. LDL is modified by close contact with endothelial cells and becomes a much more ready substrate for scavenger uptake by subintimal macrophages. Hence, endothelially minimally-modified LDL is taken up by receptors on scavenger cells which become lipid-laden foam cells, representing an early event in atheroma formation. LDL3 in particular is more likely to be oxidized, to be immunogenic, to be poorly recognized by the LDL receptor, and more readily taken up by the scavenger pathway. Therefore, the generation of oxidized LDL may be an important factor in atherosclerosis, in addition to contributing to proinflammatory mechanisms in rheumatological disorders.

Certain micronutrients, such as selenium and vitamins A and E, act as antioxidants by preventing the accumulation of oxygen free radicals (Fox 1996). Levels of antioxidants may also have relevance to the susceptibility of LDL to oxidation. As the fat-soluble, antioxidant vitamins are carried on lipoproteins such as LDL, the antioxidant levels may be reflected in the delay time before oxidation propagation occurs in LDL.

Types of hyperlipidaemia

Various patterns of elevated lipoproteins have been recognized (Table 2). However, most patterns can be caused by several different genetic disorders or they may be secondary to other metabolic disturbances. Many hyperlipidaemias are the result of environmental factors exacerbating an underlying monogenic or, more usually,

polygenic hyperlipidaemic tendency or background. Nonetheless it is worth considering a framework to classify the primary inherited hyperlipidaemias, as shown in Table 3, together with what is known about their genetic basis. The pattern of secondary hyperlipidaemia associated with several rheumatic disorders will be discussed.

Lipoprotein pattern	Lipoprotein	Lipid
Type 1	Chylomicrons	Triglycerides
Type 2a	LDL	Cholesterol
Type 2b	LDL and some VLDL	Cholesterol and triglycerides
Type 3	Chylomicron remnants and IDL	Triglycerides and cholesterol
Type 4	VLDL	Triglycerides
Type 5	VLDL and chylomicrons	Triglycerides (and some cholesterol)

Table 2 Patterns of dyslipidaemia

Genetic disorder	Biochemical defect	Lipoproteins elevated (pattern)	Clinical findings
Lipoprotein lipase deficiency	Lipoprotein lipase deficiency	Chylomicrons (1)	Pancreatitis, eruptive xanthomata
Apoprotein CII deficiency	Apoprotein CII deficiency	Chylomicrons and VLDL (1 or 5)	Pancreatitis
Familial type 4 hyperlipoproteinaemia	Abnormal apoprotein E	Chylomicron remnants and IDL (3)	Palmar and tuberous xanthomata, premature atherosclerosis
Familial hypercholesterolaemia	LDL receptor deficiency	LDL (2A)	Tendon xanthomata, premature atherosclerosis
Familial hypertriglyceridaemia	Unknown	VLDL (4)	Eruptive xanthomata
Familial combined hyperlipidaemia	Unknown	LDL and VLDL (2a, 2b, 4)	

Table 3 Primary hyperlipidaemias

It is worth mentioning here that several drugs used in the treatment of hyperlipidaemia are known to give rise to rheumatic complaints, although this is extremely rare. Myopathy has been described in patients treated with statins, fibrates, or nicotinic acid (Zuckner 1990; Goldman *et al*. 1989; Ross Peirce *et al*. 1990). Rhabdomyolysis is also possible, although rare (Hino *et al*. 1996). However, early reports included patients on cyclosporin following heart transplantation, where cyclosporin impaired renal clearance of statins. Recently, the Committee on the Safety of Medicines have calculated that the statin or the fibrate use risk of serious myositis is less than one per 100 000 patient years exposure. Use of statin and fibrate together may increase the risk and milder cases of myositis may be missed. There are case reports of drug-induced lupus caused by lovostatin and clofibrate (Ahmad 1991; Howard and Brown 1973).

Musculoskeletal manifestations of primary hyperlipidaemia

Musculoskeletal symptoms or signs may be the first manifestation of a hyperlipidaemia. A transient polyarthritis is especially common in patients with familial hypercholesterolaemia. Larger rather than smaller peripheral joints tend to be affected. In 1968, Khadchadurian studied 14 families from whom 18 homozygotes with type 2 hyperlipoproteinaemia were identified (Khadchadurian 1968). Ten of these patients experienced a migratory polyarthritis resembling rheumatic fever. A monoarthritis or oligoarthritis may also occur in heterozygotes, commonly affecting the knee or occasionally the first metatarsophalangeal joint. Oligoarthritis and periarticular hyperaesthesia have been described in patients with type 4 hyperlipidaemia (Goldman *et al*. 1972; Buckingham *et al*. 1975). Articular manifestations have not been reported for type 1 or type 3 hyperlipidaemia.

Recurrent tendinitis is a common problem in patients with familial hypercholesterolaemia. Tendinous, tuberous, or periosteal xanthomata may also be present. Common sites for tendon xanthomata are the tendo-Achillis, patellar tendons and extensor tendons of the hands and feet. Sometimes, tenosynovitis may occur without obvious xanthomata. Tendinitis may occur during the early weeks of statin use, usually in patients with very high cholesterol levels. These episodes may result from the mobilization of deposited tissue cholesterol as blood levels fall with treatment, rather analogous to the acute gout that accompanies urate mobilization in the initial weeks of allopurinol treatment. Palmar xanthomata are typically seen in type III, remnant lipidaemia (familial dysbetalipoproteinaemia). Type I, V, and severe IV hyperlipoproteinaemias may be associated with eruptive xanthomatas over the knees, buttocks, shoulders, and back.

In addition to diet and change of lifestyle, for primary hypercholesterolaemia requiring drug therapy, statins (3-hydroxy-3-methylglutaryl coenzyme A, HMG Co-A, reductase inhibitors) are the agents of choice, with or without resins (bile acid sequestrants) where tolerated. For mixed lipidaemia fibrate drugs (fenofibrate, ciprofibrate, bezafibrate, or gemfibrozil) are first choice.

In a recent study of 80 patients with hyperlipidaemia, tendon xanthomatas, particularly of the tendo-Achillis, were found in about 50 per cent of patients with either adult familial hypercholesterolaemia or with mixed hyperlipidaemia (increased cholesterol and triglycerides), often with an associated tendinitis (Klemp *et al*. 1993). Oligoarthritis was seen only in patients with mixed hyperlipidaemia. The manifestations improved with lipid-lowering therapy.

Most cases of acute arthropathy associated with hyperlipidaemia resolve spontaneously. The synovial fluid is usually non-inflammatory. Positively birefringent suspected lipid crystals have been aspirated from a bursitis in a patient with hypercholesterolaemia (Schumaker and Michaels 1989). Subcutaneous cholesterol nodules mimicking rheumatoid nodules and gouty tophi have been described in patients with an apparently normal lipid profile (Szachnowski and Bridges 1994).

Lipids in joints

Lipid microspherules have been found in the joints of patients with rheumatoid arthritis. The mean levels of apoAI, apoB, and cholesterol levels were significantly higher in the synovial fluid of 12 patients with untreated rheumatoid arthritis than in eight patients with degenerative joint disease (Ananth *et al*. 1993). A likely explanation is the increased permeability for these lipoprotein constituents across inflamed synovial membranes. Lazarevic *et al*. did not find concomitant systemic lipid abnormalities nor the presence of antilipoprotein antibodies in such patients (Lazarevic *et al*. 1993a).

Whether synovial lipoproteins are involved in the pathogenesis of rheumatoid arthritis is unclear. Winyard *et al*. have reported positive immunostaining for oxidized LDL in synovial membranes from six patients with rheumatoid arthritis (Winyard *et al*. 1993). The staining was confined to foamy macrophages in the region of synovial blood vessels. The authors point out that inflamed synovial tissue has all the appropriate requirements for enhanced lipid peroxidation, such as macrophage enrichment, extensive iron deposits, and reduced concentrations of antioxidants such as vitamin E. Also, peroxidation products, such as 4-hydroxynonenal, have been detected in synovial fluid (Selley *et al*. 1992). Fairburn *et al*. found reduced levels of the lipid soluble antioxidant a-tocopherol, which terminates the process of lipid peroxidation, suggesting depleted levels may point to increased oxidative activity (Fairburn *et al*. 1992).

The pattern of dyslipidaemia in rheumatic disorders

Low levels of HDL have been found in several chronic inflammatory disorders and seem related to disease activity. Ilowite *et al*. identified such a pattern in a set of paediatric systemic lupus erythematosus patients studied longitudinally (Ilowite *et al*. 1988). The mean HDL–cholesterol and apoA1 were markedly lower in untreated

and clinically active systemic lupus erythematosus patients compared to controls. Levels of HDL–cholesterol and apoA1 returned to normal as the systemic lupus erythematosus became less active. In our own patients in Bath, low levels of HDL were found with systemic lupus erythematosus (Leong *et al*. 1995) and psoriatic arthritis (Jones *et al*. 1994) which was most noticeable in those patients with active disease. In rheumatoid arthritis, low levels of HDL–cholesterol have been found, which return to normal with treatment of the arthritis (Lazarevic *et al*. 1992; Svenson *et al*. 1987a; Svenson *et al*. 1987b).

Low levels of HDL are normally associated with high triglycerides epidemiologically. However, there are conflicting results for levels of other lipoprotein components apart from HDL in rheumatic disorders, which may be explained by confounding influences such as corticosteroid treatment, renal disease, and possibly altered catabolic states associated with chronic inflammation. Ilowite *et al*. (Ilowite *et al*. 1988) found an increase in total triglycerides and VLDL–cholesterol in children with systemic lupus erythematosus. Our systemic lupus erythematosus patients selected for absence of renal disease or corticosteroid treatment had lower levels of total cholesterol, triglycerides, and total LDL (although LDL-3 was increased) but VLDL was unchanged (Leong *et al*. 1995). A similar pattern was seen in patients with psoriatic arthritis where only LDL-3 was increased (Jones *et al*. 1994). Svenson *et al*. studied 48 patients with untreated rheumatoid arthritis and 21 with seronegative spondylarthropathy. Compared to healthy controls, the patients with active rheumatic disease had lower levels of VLDL–cholesterol and VLDL–triglyceride, as well as lower HDL–triglyceride and HDL–cholesterol (Svenson *et al*. 1987a). Lazarevic *et al*. also found low levels of LDL–cholesterol in patients with active rheumatoid arthritis or psoriatic arthritis (Lazarevic *et al*. 1992).

In individuals without rheumatic disorders, the usual effect of endogenous steroids (Cushing's syndrome) or of exogenous steroid therapy is predominantly to increase serum cholesterol, due to increased LDL–cholesterol and often VLDL–cholesterol. There is often increased triglycerides but this is less obvious unless there are additional reasons for hypertriglyceridaemia, such as diabetes (Havel *et al*. 1980) or underlying polygenic mixed hyperlipidaemia. HDL–cholesterol may rise with prednisolone treatment (Zimmerman *et al*. 1984).

The use of corticosteroids or nephrotic syndrome contributes to a type 2b hyperlipidaemia (raised cholesterol with moderately raised triglycerides). Ettinger *et al*. compared 46 systemic lupus erythematosus patients with 30 controls and found that systemic lupus erythematosus patients on corticosteroids had higher levels of total cholesterol, total triglyceride, and LDL–cholesterol (Ettinger *et al*. 1987; Ettinger and Hazzard 1988). Those who were not on steroids had levels similar to controls, apart from a lower HDL–cholesterol. Similar results were found in a study of 100 Singaporean Chinese patients with systemic lupus erythematosus (Leong *et al*. 1994). In the latter study, the type 2b pattern was associated with renal involvement and the use of corticosteroids. MacGregor *et al*. found raised concentrations of triglyceride and apoB in systemic lupus erythematosus patients treated with corticosteroids (MacGregor *et al*. 1992).

Evidence for increased mortality from atherosclerosis in rheumatic disorders

Rheumatoid arthritis is associated with an increased mortality. Mortality is higher in those with more severe disease (Wolfe *et al*. 1994). Prior *et al*. studied 489 consecutive patients with classical rheumatoid arthritis followed-up for a mean of 11.2 years and found a three-fold increase in overall mortality compared with age and gender-specific rates in the general population (Prior *et al*. 1984). Cardiovascular deaths were 2.5 times higher than expected values compared to age and gender-matched population figures. Rasker and Cosh followed a cohort of 100 patients for 18 years (Rasker and Cosh 1981). Of the 43 patients who died, 16 deaths were related to rheumatoid arthritis and 27 were unrelated. In the latter group, 14 were cardiac deaths and 8 were cerebrovascular events. However, results from other studies recording causes of death in 2262 rheumatoid arthritis patients from 13 centres, suggests that cardiovascular disease accounts for 40 per cent of deaths in rheumatoid arthritis which was no different from the general population (Pincus and Callahan 1986).

Urowitz *et al*. recognized a bimodal pattern of mortality in systemic lupus erythematosus with early deaths due to active disease or infections and late events related to atherosclerotic complications (Urowitz *et al*. 1976). Further analysis of 665 patients from the same centre largely confirmed the earlier findings, with about a five-fold overall excess death rate in systemic lupus erythematosus (Abu-Shakra *et al*. 1995). The authors concluded that atherosclerosis is a major cause of death and morbidity in patients with systemic lupus erythematosus. It is not clear whether the late cardiovascular disease is related primarily to corticosteroid treatment or to other manifestations of the disease. Petri *et al*. studied prospectively the risk factors for coronary artery disease in 229 systemic lupus erythematosus patients (Petri *et al*. 1992). Nineteen patients (8.3 per cent) had coronary artery disease which was associated with age at systemic lupus erythematosus diagnosis, duration of corticosteroid use, requirement for antihypertensive treatment, maximum cholesterol level, and obesity.

Many studies, including autopsy studies, have shown premature coronary atherosclerosis and myocardial infarction in relatively young patients with systemic lupus erythematosus. Bulkley and Roberts compared necroscopy findings of 36 systemic lupus erythematosus patients on corticosteroids with findings in patients not on corticosteroids (Buckley and Roberts 1975). Subepicardial and myocardial fat was increased in all the former group. In 42 per cent of the 18 patients who received corticosteroids for more than 1 year, the lumen of at least one of the three major coronary arteries was narrowed by more than 50 per cent by atherosclerotic plaques. In contrast, none of the 17 patients who received corticosteroids for less than 1 year, and who were on average 5 years older, had such findings. Systemic hypertension was twice as common in patients who received corticosteroids for longer than 1 year.

A number of other chronic rheumatic disorders are associated with a shortened life expectancy (Callahan and Pincus 1995). Patients with renal involvement and scleroderma, vasculitis, or systemic lupus erythematosus generally have the poorest prognosis. However, conditions thought to have a better outlook, such as ankylosing spondylitis, may also have an excess death rate (Radford *et al*. 1977; Rigby 1991). Current data suggests that there is a tendency for premature death in patients with rheumatic diseases. One significant contributor is the presence of cardiovascular or cerebrovascular disease. Longer-term, prospective studies with appropriate population controls will help clarify risk factors. As our ability to manage patients improves due to our increasing knowledge, survival of patients will be prolonged, as has already happened with systemic lupus erythematosus. Awareness of potential problems from premature atherosclerosis will become increasingly important.

Mechanisms for dyslipidaemia

Several mechanisms may account for the dyslipidaemia associated with chronic inflammatory disorders. Lipoprotein lipase is a membrane-bound enzyme responsible for liberation of fatty acids from VLDL resulting in formation of LDL. Reduced mass or inhibition of activity of lipoprotein lipase is an attractive hypothesis as it would account for reduced clearance of VLDL–triglyceride and an associated rise in plasma triglycerides. Normal or slightly low LDL–cholesterol would be expected as a result of lower precursor input into LDL—and LDL concentrations in plasma would tend to fall. HDL levels would also tend to fall because there would be a slower reduction in VLDL size, and therefore less VLDL surface coat given up to provide HDL.

Lipoprotein lipase activity is inhibited by several regulatory cytokines such as tumour necrosis factor-a (Fried and Zechner 1989), interleukin-1, and g-interferon (Querfeld *et al*. 1990), that are upregulated in inflammation. Furthermore, 24-h infusion of recombinant human tumour necrosis factor-a is associated with a decrease in serum cholesterol and HDL levels (Spriggs *et al*. 1988). However, tumour necrosis factor-a and other cytokines may also stimulate hepatic lipid secretion and increase *de novo* fatty acid synthesis (Feingold and Grunfeld 1992). Lipoprotein lipase is also an insulin-dependent enzyme and its synthesis in adipose and muscle is increased by insulin treatment. Whether the dyslipidaemia in active rheumatic disease may be due to a form of insulin resistance is not known. Hypertriglyceridaemia, low HDL, small dense LDL3, and insulin resistance (known as the atherogenic lipid profile) is often accompanied by glucose intolerance or diabetes, hypertension, hyperuricaemia, or gout, and procoagulant changes with high PA1–1 and fibrinogen.

Alternatively, an increased production of acute phase proteins by the liver in inflammation may occur at the expense of lipoprotein production. Also, acute phase reactants may interfere with HDL metabolism. Kumon *et al*. showed that in patients with rheumatic disease, total HDL cholesterol, apoA-I, and apoA-II were lower than in normal subjects and were inversely correlated with plasma concentrations of serum amyloid A protein (Kumon *et al*. 1993). It was suggested that serum amyloid A protein may displace apoA-I and apoA-II on HDL particles, and especially on HDL3.

Corticosteroids cause increased hepatic production of VLDL resulting in elevated plasma cholesterol, plasma triglyceride, HDL-C, VLDL-C, and LDL-C. Proteinuria leads to increased generalized hepatic protein synthesis, and hence lipoprotein synthesis which may also be driven by urinary loss of apoA of HDL. Conversely, control of hyperlipidaemia may retard the development of glomerulosclerosis (Kasiske *et al*. 1988).

Antibody-mediated mechanisms may also be important. Lazarevic *et al*. found antilipoprotein antibodies against VLDL and LDL in one-third of patients with active rheumatoid arthritis, but not in patients with osteoarthritis, psoriatic arthritis, or healthy blood donors (Lazarevic *et al*. 1993b). Autoantibodies against LDL have been demonstrated in diabetes mellitus. More recently anti-apoA1 antibodies have been described in systemic lupus erythematosus (Merrill *et al*. 1995). Rarely, antibodies against apoC-II may be associated with severe hypertriglyceridaemia, due to myeloma and clonal production of the antibody.

Rheumatic disorders

Systemic lupus erythematosus

Mechanisms accounting for vasculopathy

Factors to account for the increase in vascular disease in systemic lupus erythematosus include corticosteroid treatment, renal disease, dyslipidaemia, an underlying coagulopathy, vasculitis, and, possibly, antibody or immune-complex mediated injury. In support of the latter, Kabakov *et al*. demonstrated that in the presence of lupus sera, lipid accumulation in cultured smooth muscle cells is increased up to six fold compared to cells cultured with normal human sera from healthy donors (Kabakov *et al*. 1992). Incubation of the smooth muscle cells with circulating immune complexes isolated from lupus sera caused a three to four-fold increase in intracellular cholesterol level. The atherogenic effect was thought to be due to LDL-containing immune complexes.

Lipoprotein (a)

There are conflicting findings concerning the levels of lipoprotein (a) in systemic lupus erythematosus and the risk of thrombosis. Takegoshi reported a case of an 18-year-old systemic lupus erythematosus patient who had severe elevations of lipoprotein (a) associated with myocardial infarction and cerebral infarction (Takegoshi *et al*. 1990). Elevated levels of lipoprotein (a) levels have been reported in a few studies of patients with systemic lupus erythematosus (Matsuda *et al*. 1994; Borba *et al*. 1994) but not in association with thrombosis, which is in accordance with our own unpublished results. However, one recent study found increased lipoprotein (a) levels in systemic lupus erythematosus patients with myocardial and cerebral infarction (Kawai *et al*. 1995).

Lipids and antiphospholipid antibodies

The interplay between lipid abnormalities and antiphospholipid antibodies may be complex. A plasma cofactor which forms part of the antigenic complex recognized by antiphospholipid antibodies has been identified as b $_2$-glycoprotein I or apoH. Whether lipoproteins other than b $_2$-glycoprotein I are involved in the pathogenesis of the antiphospholipid syndrome is unknown. There may be an additive risk of thrombotic clinical events if both antiphospholipid antibodies and hyperlipidaemia are present together (Garrido *et al*. 1994). Lahita *et al*. found an association between anticardiolipin antibodies and low levels of HDL and apoAI, although cholesterol levels were also low in these patients (Lahita *et al*. 1993).

Antibodies to oxidized LDL have been described in patients with systemic lupus erythematosus (Vaarala *et al*. 1993). Monoclonal antibodies to cardiolipin have been derived from a mouse model of the antiphospholipid syndrome that have reactivity against b $_2$-glycoprotein and cross-react with oxidized LDL (Mizutani *et al*. 1995). In a prospective study of middle-aged dyslipidaemic men, there was a correlation between anticardiolipin levels and antibodies to oxidized LDL and the presence of both had an additive risk for myocardial infarction (Vaarala *et al*. 1995).

Rheumatoid arthritis

Most studies have shown that patients with rheumatoid arthritis have a lipid profile that increases the risk of atherosclerosis, especially when samples are studied during active phases of the disease (Lazarevic *et al*. 1992; Svenson 1987a; Svenson 1987b; Rantapää-Dahlqvist *et al*. 1991). The most consistent abnormality is a reduction in HDL. Magaro *et al*. reported lower levels of apoA1 compared with controls, although apoB was also reduced (Magaro *et al*. 1991). Lipoprotein (a) levels may also be increased (Rantapää-Dahlqvist *et al*. 1991). Therefore, dyslipidaemia associated with active rheumatoid arthritis needs to be considered as a contributing factor to the excess death rate observed in mortality studies.

Psoriatic arthritis

Patients with psoriatic arthritis have a similar pattern of dyslipidaemia to that occurring in rheumatoid arthritis (Lazarevic *et al*. 1992). Jones *et al*. found an atherogenic profile with low total HDL and HDL3, and elevated LDL3 and lipoprotein (a) levels (Jones *et al*. 1994). Elevated LDL3 may be particulary important as this small, dense subclass of LDL is much more likely to be taken up by macrophages in the arterial subintimal space, possibly explaining its association with coronary artery disease (Campos *et al*. 1992).

There may also be an imbalance in fatty acids and antioxidants either contributing to disease pathogenesis or as a secondary metabolic consequence of chronic inflammation. Azzini *et al*. found abnormalities in red blood cell fatty acid composition with increased total fatty acids and decreased w-6 polyunsaturated fatty acids (Azzini *et al*. 1995). Reduced levels of the antioxidant selenium and increased plasma concentrations of copper were also found (Azzini *et al*. 1995).

Systemic sclerosis

Vascular impairment secondary to reperfusion injury and the formation of free radicals are postulated as disease mechanisms in scleroderma. Bruckdorfer *et al*. reported that LDL in patients with systemic sclerosis are more susceptible to oxidation (Bruckdorfer *et al*. 1995). In addition to its toxic effects on endothelium, oxidized LDL may potentially contribute to the increased matrix synthesis seen in systemic sclerosis by upregulating certain adhesion molecules and growth factors such as platelet-derived growth factor.

Gout

There is a clear association between gout and type 4 hyperlipidaemia. Decreased HDL–cholesterol and increased VLDL levels have been reported (Ulreich *et al*. 1985; Matsubara *et al*. 1989). The association is unlikely to be causal as patients with gout often have other predisposing factors for hyperlipidaemia, such as obesity, increased alcohol consumption, and altered nutritional habit. Furthermore, in the disorder of insulin resistance the syndrome complex includes hyperuricaemia with or without gout, hypertension, mixed dyslipidaemia, glucose intolerance or diabetes mellitus, central obesity, hyperfibrinogenaemia, high PAI-1, and associated coronary heart disease and peripheral vascular disease.

Hyperlipidaemia may also be more directly associated with decreased renal excretion of uric acid. Renal excretion of urate is lower in hyperuricaemic–hyperlipidaemic patients than in hyperuricaemic–normolipidaemic patients. Tinahones *et al*. reported increased VLDL levels, diminished fractional excretion of uric acid, and increased apoCIII/CII ratios in patients with hyperuricaemia and hypertriglyceridaemia (Tinahones *et al*. 1995). ApoCIII inhibits the hydrolysis of triglycerides by lipoprotein and hepatic lipase, whereas apoCII is an activator of lipoprotein lipase. An altered apoCIII/CII ratio would lengthen the time VLDL remains in the plasma. The mechanism is unknown whereby lipids are directly or indirectly associated with altered renal handling of urate.

Management

The use of essential fatty acids in the treatment of rheumatic diseases

The use of essential fatty acids to treat inflammation has a supportive in vivo and biochemical basis, although the clinical efficacy and cost-effectiveness of such treatment in man remains uncertain. In animal models, marine oils have been effective in suppressing inflammation. Eicosapentaenoic acid supplementation prevented renal disease from developing in NZB/NZW mice (Prickett *et al*. 1981). A fish oil diet had a beneficial effect on the severity of collagen-induced arthritis (Cathcart and Gonnerman 1991) and on lupus-like features in MRL-*lpr* mice (Robinson *et al*. 1986).

Prostaglandins are synthesized from essential fatty acids (see above). The addition of omega 3 or omega 6 polyunsaturated fatty acids to the diet may shift the balance of prostaglandin metabolism towards substances that have less of a proinflammatory action. The two most extensively studied therapies have been fish oils and evening primrose oil. The former contains high levels of polyunsaturated fatty acids of the omega-3 series (such as eicosapentaenoic acid and docosahexaenoic acid) and the latter predominantly gammalinoleic acid.

P>Several studies in rheumatoid arthritis have shown some beneficial effects of diets containing essential fatty acids in fish oils, such as a reduced requirement for non-steroidal anti-inflammatory drugs (Lau *et al*. 1993). In a study of 66 patients randomized to fish oil or corn oil, Kremer *et al*. reported a decreased number of

tender joints, morning stiffness, and global arthritis activity and pain in the fish oil group (Kremer *et al*. 1995). Geusens *et al*. studied 90 patients on various dosage regimes of fish versus olive oil and found an improvement in the patients' global evaluation and physicians' assessment of pain in those patients taking the higher dose of fish oil (2.6 g/day) (Geusens *et al*. 1994). It is worth mentioning that olive oil may also have some active beneficial effects as it contains mainly unsaturated fats. Gammalinoleic acid at a dose of 1.4 g/day was better than placebo in improving joint tenderness and swelling in patients with rheumatoid arthritis (Leventhal *et al*. 1993).

There is some evidence that fish oils may be beneficial in other conditions such as psoriasis and systemic lupus erythematosus. Bittiner reported decreased pruritus, erythema, and scaling of skin psoriasis in 28 patients treated with fish oil compared to a placebo (Bittiner *et al*. 1988). A number of studies have reported an improvement in skin lesions with eicosapentaenoic acid. However, a preparation containing eicosapentaenoic acid, gammalinoleic acid, and decosahexaenoic acid (Efamol) had no significant beneficial effect in patients with psoriatic arthritis despite a fall in leukotriene B4 (Veale *et al*. 1994). It is possible that an insufficient dose of icosapentaenoic acid was used in the latter study. In a study of 27 patients with active systemic lupus erythematosus, symptoms were favourably affected by an eicosapentaenoic acid enriched diet in 14 patients compared to placebo (Walton *et al*. 1991). Another benefit of a fish oil diet is an alteration of the thromboxane/prostacyclin ratio towards substantially less thrombotic thromboxane and a substantially more antithrombotic prostacyclin.

It is important to note that the amount of essential fatty acids needed to produce clinical effects are often in excess of the dose available in commercial preparation. Also, it may be difficult to achieve such doses by ingestion of natural food substances. Fahrer *et al*. found that 7.5 g of fish oil, the therapeutic dose in many studies, is the equivalent of 700 g of fish (Fahrer *et al*. 1991) and carries a 70 calorie load. This means that four to six meals of fish per week are required in order to induce a potentially useful anti-inflammatory effect.

Longer-term studies with larger numbers of patients are required to help clarify the therapeutic role of essential fatty acids. It is possible that essential fatty acid supplementation may be a reasonable treatment alternative to non-steroidal anti-inflammatory agents, especially when the latter are contraindicated. Fish oils can improve moderate to marked hypertriglyceridaemia in some individuals, but do not alter LDL-cholesterol much—a dose of 10 g of fish oil (as Maxepa) being needed, equivalent to 90 calories daily.

Rationale for lipid-lowering treatment

Rheumatic diseases may well carry an increased risk of morbidity and mortality from premature atherosclerosis, although more evidence is needed. Abnormalities in the LDL subfraction are clearly present in patients on long-term corticosteroid treatment (Ettinger *et al*. 1987; Ettinger and Hazzard 1988; MacGregor *et al*. 1992). Many studies have shown the relationship between hypercholesterolaemia and the risk of coronary artery disease. The Framingham study showed that the risk of myocardial infarction was associated with higher levels of LDL-cholesterol and inversely associated with HDL levels (Castelli *et al*. 1986), as did the study of 362 262 American men screened from the Multiple Risk Factor Intervention Trial (Neaton *et al*. 1992). Therefore the dyslipidaemia present in chronic rheumatic diseases would appear to increase the risk of atheroma.

There is good evidence that the treatment of lipid abnormalities influences outcome, with reduced coronary heart disease deaths and events. In older studies, this was shown for cholesterol lowering with cholestyramine (Lipid Research Clinics Program 1984) or with gemfibrozil where triglyceride was also lowered and HDL cholesterol raised (Frick *et al*. 1987). More recently, the more potent statins have shown reduction in coronary heart disease deaths and events, in coronary artery surgery requirements, and in total mortality, both in primary (West of Scotland Coronary Prevention Study) (Shepherd *et al*. 1995) and in secondary (Scandinavian Simvastatin Survival Study) (Scandinavian Simvastatin Survival Study 1994) prevention studies. Many studies (e.g. Blankenhorn *et al*. 1987) have shown slowing of progression, and some regression, of atheroma in patients actively treated for hyperlipidaemia after coronary artery graft surgery.

Approach to treatment

Risk factors for hyperlipidaemia need to be assessed for the individual patient. A basic fasting lipid profile may be appropriate on any patient with a chronic inflammatory rheumatic disease, although subtle changes in lipid subfraction composition may be missed. Other associated risk factors, such as hypertension, diabetes, and smoking, are important and need to be fully discussed with the patient and managed.

Dietary intervention plays a major role in the management of hyperlipidaemia. While dietary change may produce only modest change in lipid levels, dietary modification may be all that is necessary in moderate hyperlipidaemia. Target levels for treatment of dyslipidaemia have been set in national and international guidelines (National Cholesterol Education Program 1994 ; Pyorala *et al*. 1994). In systemic lupus erythematosus, benefit of dietary change at 6 months has been shown in patients receiving corticosteroids (Hearth-Holmes *et al*. 1995). In adolescents with systemic lupus erythematosus, dietary modification with fish oil supplementation improved the lipid profile (Ilowite *et al*. 1995). However, a significant number of systemic lupus erythematosus patients might require further pharmacological therapy for persistent dyslipidaemia, with statins, fibrates, or other agents.

The role of hydroxychloroquine

Hydroxychloroquine is often used in the treatment of mild systemic lupus erythematosus and rheumatoid arthritis. Besides its immunomodulatory properties, hydroxychloroquine may have a protective role against hyperlipidaemia induced by corticosteroids. In a case–control study, patients with systemic lupus erythematosus on hydroxychloroquine had 35 to 45 per cent lower levels of total triglyceride, VLDL-triglyceride, LDL-triglyceride, and apoCIII than patients not on treatment (Hodis *et al*. 1993). Wallace *et al*. found that hydroxychloroquine was strongly associated with low levels of cholesterol, triglycerides, and LDL-cholesterol regardless of concomitant corticosteroid use (Wallace *et al*. 1990). In a cohort longitudinal study involving 264 patients with systemic lupus erythematosus (Petri *et al*. 1994) using a regression model for steroid use, a change in prednisolone dose of 10 mg was associated with a change in cholesterol level of 7.5 mg (+/– 1.46 SD) and a weight gain of 2.5 kg (+/– 0.6 SD). On the other hand, hydroxychloroquine at 200 mg or 400 mg/day were both associated with lower serum cholesterol. However, no prospective randomized study of hydroxychloroquine on lipoprotein and lipid levels has been carried out.

Conclusions

There is sufficient evidence to suggest that lipid metabolism is altered in chronic rheumatic diseases in a manner which may promote accelerated atherosclerosis. Further long-term studies are needed to determine the risk more precisely and that contributed by treatments such as corticosteroids. Dietary intervention would seem a sensible first treatment step, and lipid lowering agents may be required, appropriate target levels being cholesterol less than 5.0 mmol/l, triglyceride less than 2 mmol/l, and LDL–cholesterol less than 3.5 mmol/l. A prudent, low-fat diet, partly supplemented by mono- and polyunsaturated fatty acids, with high fibre should be encouraged. Whether dietary supplementaion with marine oils has a place in reducing requirements for anti-inflammatory agents is uncertain, although the dose of fat supplement may be unacceptable.

Lipid peroxidation products such as oxidized LDL may play an important role in promoting endothelial and synovial injury in addition to atheroma and, together with the possible protective role of antioxidants, warrant further study. Where patients have risk factors for coronary heart disease in addition to those from the rheumatic disease itself, for example in patients with systemic lupus erythematosus and hypertension, attention to all possible risk factors may be necessary to help prevent coronary heart disease. Treatment of risk factors is essential in patients who have manifested macrovascular disease, where there is adequate further life expectancy.

Chapter References

Abu-Shakra, M., Urowitz, M.B., Gladman, D.D., and Gough, J. (1995). Mortality studies in systemic lupus erythematosus. Results from a single centre. 1. Causes of death. *Journal of Rheumatology*, **22**, 1259–64.

Ahmad, S. (1991). Lovastatin-induced lupus erythematosus. *Archives of Internal Medicine*, **151**, 1667–8.

Ananth, L., Prete, P.E., and Kashyap, M.L. (1993). Apolipoproteins A-I and B and cholesterol in synovial fluid of patients with rheumatoid arthritis. *Clinical and Experimental Metabolism*, **42**, 803–6.

Azzini, M., Girelli, D., Olivieri, O., Guarini, P., Stanzial, A.M., Frigo, A., Milanino, R., Bambara, L.M., and Corrocher, R. (1995). Fatty acids and antioxidant micronutrients in psoriatic arthritis. *Journal of Rheumatology*, **22**, 103–8.

Bittiner, S.B., Tucker, W.F.G., Cartwright, I., and Bleehen, S.S. (1988). A double-blind, randomised, placebo-controlled trial of fish oil in psoriasis. *Lancet*, **i**, 378–80.

Blankenhorn, D.H., Nessim, S.A., and Johnson, R.L. (1987). Beneficial effects of combined colestipol-niacin therapy on coronary atherosclerosis and coronary venous bypass grafts. *Journal of the*

American Medical Association, **257**, 3233.

Borba, E.F., Santos, R.D., Bonfa, E., Vinagre, C.G., Pileggi, F.J.C., Cossermelli, W., and Maranhao, R.C. (1994). Lipoprotein(a) levels in systemic lupus erythematosus. *Journal of Rheumatology*, **21**, 220–3.

Bruckdorfer, K.R., Hillary, J.B., Bunce, T., Vancheeswaran, R., and Black, C.M. (1995). Increased susceptibility to oxidation of low-density lipoproteins isolated from patients with systemic sclerosis. *Arthritis and Rheumatism*, **38**, 1060–7.

Buckingham, R.B., Bole, G.G., and Bassett, D.R. (1975). Polyarthritis associated with Type IV lipoproteinemia. *Archives of Internal Medicine*, **135**, 286–90.

Buckley, B.H. and Roberts, W.C. (1975). The heart in systemic lupus erythematosus and the changes induced in it by corticosteroid therapy—a study of 36 necrosopy patients. *American Journal of Medicine*, **58**, 243–63.

Callahan, L.F. and Pincus, T. (1995). Mortality in rheumatic diseases. *Arthritis Care and Research*, **8**, 229–41.

Campos, H., Genest, Jr J.J., Blijlevan, S.E., McNamara, J.R., Jenner, J.L., Ordovas, J.M., Wilson, P.N.F., and Schaefer, E.J. (1992). LDL particle size and coronary artery disease. *Arteriosclerosis and Thrombosis*, **12**, 187–95.

Castelli, W.P., Garrison, R.J., and Wilson, P.W.F. (1986). Incidence of coronary heart disease and lipoprotein cholesterol levels. *Journal of the American Medical Association*, **256**, 2835.

Cathcart, E.S. and Gonnerman, W.A. (1991). Fish oil fatty acids and experimental arthritis. *Rheumatic Disease Clinics of North America*, **17**, 235–42.

Ettinger, W.H. and Hazzard, W.R. (1988). Elevated apolipoprotein-B levels in corticosteroid-treated patients with systemic lupus erythematosus. *Journal of Clinical Endocrinology and Metabolism*, **67**, 425–8.

Ettinger, W.H., Goldberg, A.P., and Apple-Bowden, D., Hazzard, W.R. (1987). Dyslipoproteinemia in SLE—effect of corticosteroids. *American Journal of Medicine*, **83**, 503–8.

Fahrer, H., Hoeflin, F., Lauterberg, B.H., Peheim, E., Levy, A., and Vischer, T.L. (1991). Diet and fatty acids: can fish substitute for fish oil? *Clinical and Experimental Rheumatology*, **9**, 403–6.

Fairburn, K, Grootveld, M., Ward, R.J., Abiuka, C., Kus, M., Williams, R.B., Winyard, P.G., and Blake, D.R. (1992). Alpha-tocopherol, lipids and lipoproteins in knee-joint synovial fluid and serum from patients with inflammatory joint disease. *Clinical Science*, **83**, 657–64.

Feingold, K.R. and Grunfeld, C. (1992). Role of cytokines in inducing hyperlipidaemia. *Diabetes*, **41** (Suppl. 2), 97–101.

Fox, R.I. (1996). Fifth international symposium on Sjögren's syndrome. *Arthritis and Rheumatism*, **39**, 195–6.

Frick, M.H., Elo, O., and Haapa, K. (1987). Helsinki heart study: primary prevention trial with gemfibrozil in middle-aged men with dyslipidaemia. *New England Journal of Medicine*, **317**, 1237–45.

Fried, S.K. and Zechner, R. (1989). Cachetin/tumor necrosis factor decreases human adipose tissue lipoprotein lipase mRNA levels, synthesis, and activity. *Journal of Lipid Research*, **30**, 1917–23.

Garrido, J.A., Peromingo, J.A.D., Sesma, P., and Pia, G. (1994). More about the link between thrombosis and atherosclerosis in autoimmune diseases: Triglycerides and risk for thrombosis in patients with antiphospholipid antibodies. *Journal of Rheumatology*, **21**, 2394.

Geusens, P., Wouters, C., Nijs, J., Jiang, Y., and Dequeker, J. (1994). Long-term effect of omega-3 fatty acid supplementation in active rheumatoid arthritis. A 12-month, double-blind, controlled study. *Arthritis and Rheumatism*, **37**, 824–9.

Goldman, J.A., Glueck, C.J., and Abrams, N.R. (1972). Musculoskeletal disorders associated with Type IV hyperlipoproteinemia. *Lancet*, **2**, 449–52.

Goldman, J.A., Fishman, A.B., and Lee, J.E. (1989). The role of cholesterol-lowering agents in drug-induced rhabdomyolysis and polymyositis. *Arthritis and Rheumatism*, **32**, 358–9.

Grainger, D.J., Kemp, P.R., Liu, A.C., Lawn, R.M., and Metcalfe, J.C. (1994). Activation of transforming growth factor-b is inhibited in transgenic apolipoprotein (a) mice. *Nature*, **370**, 460–2.

Havel, R.J., Goldstein, J.L., and Brown, M.S. (1980). *Metabolic control and disease*, pp. 393–494. WB Saunders, Philadelphia.

Hearth-Holmes, M., Baethge, B.A., Broadwell, L., and Wolf, R.E. (1995). Dietary treatment of hyperlipidemia in patients with systemic lupus erythematosus. *Journal of Rheumatology*, **22**, 450–4.

Hino, I., Akama, H., Furuya, T., Ueda, H., Taniguchi, A., Hara, M., and Kashiwazaki, S. (1996). Pravastatin-induced rhabdomyolysis in a patient with mixed connective tissue disease. *Arthritis and Rheumatism*, **39**, 1259–61.

Hodis, H.N., Quismorio, F.P. Jr, Wickham, E., and Blankenhorn, D.H. (1993). The lipid, lipoprotein, and apolipoprotein effects of hydroxychloroquine in patients with systemic lupus erythematosus. *Journal of Rheumatology*, **20**, 661–5.

Howard, E.J. and Brown, S.M. (1973). Clofibrate-induced antinuclear factor and lupus-like syndrome. *Journal of the American Medical Association*, **226**, 1358–9.

Ilowite, N.T., Samuel, P., Ginzler, E., and Jacobson, M.S. (1988). Dyslipoproteinemia in pediatric systemic lupus erythematosus. *Arthritis and Rheumatism*, **31**, 859–63.

Ilowite, N.T., Copperman, N., Leicht, T., Kwong, T., and Jacobson, M.S. (1995). Effects of dietary modification and fish oil supplementation on dyslipoproteinemia in pediatric systemic lupus erythematosus. *Journal of Rheumatology*, **22**, 1347–51.

Jones, S.M., Lloyd, J.L., Barnes, J., Stirling, C.A., Harris, C.P.D., Reckless, J.P.D., and McHugh, N.J. (1994). Lipoproteins and their subfractions in psoriatic arthritis. *Arthritis and Rheumatism*, **37**, S206.

Kabakov, A.E., Tertov, V.V., Saenko, V.A., Poverenny, A.M., and Orekhov, A.N. (1992). The atherogenic effect of lupus sera: SLE-derived immune complexes stimulate the accumulation of cholesterol in cultured smooth muscle cells from human aorta. *Clinical Immunology and Immunopathology*, **63**, 214–20.

Kasiske, B.L., O'Donnell, M.D., Cleary, M.P., and Keane, W.F. (1988). Treatment of hyperlipidemia reduces glomerular injury in obese Zucker rats. *Kidney International*, **33**, 667–72.

Kawai, S., Mizushima, Y., and Kaburaki, J. (1995). Increased serum lipoprotein(a) levels in systemic lupus erythematosus with myocardial and cerebral infarctions. *Journal of Rheumatology*, **22**, 1210–11.

Khadchadurian, A.K. (1968). Migratory polyarthritis in familial hypercholesterolaemia (type II hyperlipoproteinemia). *Arthritis and Rheumatism*, **11**, 385–93.

Klemp, P., Halland, A.M., Majoos, F.L., and Steyn, K. (1993). Musculoskeletal manifestations in hyperlipidaemia: a controlled study. *Annals of the Rheumatic Diseases*, **52**, 44–8.

Kremer, J.M., Lawrence, D.A., Petrillo, G.F., Litts, L.L., Mullaly, P.M., Rynes, R.I., Stocker, R.I., Parhami, N., Greenstein, N.S., Fuchs, B.R., Mathur, A., Robinson, D.R., Sperling, R.I., and Bigaouette, J. (1995). Effects of high-dose fish oil on rheumatoid arthritis after stopping nonsteroidal antiinflammatory drugs. Clinical and immune correlates. *Arthritis and Rheumatism*, **38**, 1107–14.

Kumon, Y., Suchiro, T., Ikedo, Y., Yoshida, K., Hashimoto, K., and Ohno, F. (1993). Influence of serum amyloid A protein on HDL in chronic inflammatory disease. *Clinical Biochemistry*, **26**, 505–11.

Lahita, R.G., Rivkin, E., Cavanagh, I., and Romano, P. (1993). Low levels of total cholesterol, high-density lipoprotein, and apolipoprotein A1 in association with anticardiolipin antibodies in patients with systemic lupus erythematosus. *Arthritis and Rheumatism*, **36**, 1566–74.

Lau, C.S., Morley, K.D., and Belch, J.J.F. (1993). Effects of fish oil supplementation on non-steroidal anti-inflammatory drug requirement in patients with mild rheumatoid arthritis—a double-blind placebo controlled study. *British Journal of Rheumatology*, **32**, 982–9.

Lazarevic, M.B., Vitic, J., Mladenovic, V., Myones, B.L., Skosey, J.L., and Swedler, W.I. (1992). Dyslipoproteinemia in the course of active rheumatoid arthritis. *Seminars in Arthritis and Rheumatism*, **22**, 172–80.

Lazarevic, M.B., Skosey, J.L., and Vitic, J. (1993 *a*). Cholesterol crystals in synovial and bursal fluid. *Seminars in Arthritis and Rheumatism*, **23**, 99–103.

Lazarevic, M.B., Vitic, J., and Myones, B.L. (1993 *b*). Anti-lipoprotein antibodies in rheumatoid arthritis. *Seminars in Arthritis and Rheumatism*, **22**, 385–91.

Leong, K.H., Koh, E.T., Feng, P.H., and Boey, M.L. (1994). Lipid profiles in patients with systemic lupus erythematosus. *Journal of Rheumatology*, **21**, 1264–7.

Leong, K.H., Stirling, C.A., Lloyd, J., Reckless, J., Maddison, P.J., and McHugh, N.J. (1995). SLE patients have abnormal lipid profiles compared to age and gender matched controls (abstract).

Lupus, **4**, 149.

Leventhal, L.J., Boyce, E.G., and Zurier, R.B. (1993). Treatment of rheumatoid arthritis with gammalinoleic acid. *Archives of Internal Medicine*, **119**, 867–73.

Lipid Research Clinics Program (1984). The Lipid Research Clinics coronary primary prevention trial results. I. Reduction in incidence of coronary heart disease. *Journal of the American Medical Association*, **251**, 351–64.

MacGregor, A.J., Dhillon, V.B., Binder, A., Forte, C.A., Knight, B.C., Betteridge, D.J., and Isenberg, D.A. (1992). Fasting lipids and anticardiolipin antibodies as risk factors for vascular disease in systemic lupus erythematosus. *Annals of the Rheumatic Diseases*, **51**, 152–5.

Magaro, M., Altomonte, L., Zoli, A., Mirone, L., and Ruffini, M.P. (1991). Serum lipid pattern and apolipoprotein A-I and B100 in active rheumatoid arthritis. *Zeitschrift für Rheumaforschung*, **50**, 168–70.

Maher, V.M.G. and Brown, B.G. (1995). Lipoprotein (a) and coronary heart disease. *Current Opinions in Lipidology*, **6**, 229–35.

Matsubara, K., Matsuzawa, Y., Jiao, S., Takama, T., Masaharu, K., and Tami, S. (1989). Relationship between hypertriglyceridemia and uric acid production in primary gout. *Metabolism*, **38**, 689–701.

Matsuda, J., Gotoh, M., Gohchi, K., Saitoh, N., and Tsukamoto, M. (1994). Serum lipoprotein(a) level is increased in patients with systemic lupus erythematosus irrespective of positivity of antiphospholipid antibodies. *Thrombosis Research*, **73**, 83–4.

Merrill, J.T., Rivkin, E., Shen, C., and Lahita, R.G. (1995). Selection of a gene for apolipoprotein A1 using autoantibodies from a patient with systemic lupus erythematosus. *Arthritis and Rheumatism*, **38**, 1655–9.

Mizutani, H., Kurata, Y., Kosugi, S., Shiraga, M., Kashiwagi, H., Tomiyama, Y., Kanakura, Y., Good, R.A., and Matsuzawa, Y. (1995). Monoclonal anticardiolipin autoantibodies established from the (New Zealand White×BXSB)F_1 mouse model of antiphospholipid syndrome cross-react with oxidised low-density lipoprotein. *Arthritis and Rheumatism*, **38**, 1382–8.

National Cholesterol Education Program (1994). Second report of the expert panel on detection, evaluation, and treatment of high blood cholesterol in adults (adult treatment panel II). *Circulation*, **89**, 1333–445.

Neaton, J.D., Blackburn, H., Jacobs, D., *et al*. (1992). Serum cholesterol level and mortality findings for men screened in the multiple risk intervention trial. *Archives of Internal Medicine*, **152**, 1490–500.

Petri, M., Perez-Gutthann, S., Spence, D., and Hochberg, M.C. (1992). Risk factors for coronary artery disease in patients with systemic lupus erythematosus. *American Journal of Medicine*, **93**, 513–19.

Petri, M., Lakatta, C., Magder, L., and Goldman, D. (1994). Effect of prednisone and hydroxychloroquine on coronary artery disease risk factors in systemic lupus erythematosus: A longitudinal data analysis. *American Journal of Medicine*, **96**, 254–9.

Pincus, T. and Callahan, L.F. (1986). Taking mortality in rheumatoid arthritis seriously—predictive markers, socioeconomic status and comorbidity. *Journal of Rheumatology*, **13**, 841–5.

Prickett, J.D., Robinson, D.R., and Steinberg, A.D. (1981). Dietary enrichment with the polyunsaturated fatty acid eicosapentaenoic acid prevents proteinuria and prolongs survival in NZB×NZW F_1 mice. *Journal of Clinical Investigation*, **68**, 556–9.

Prior, P., Symmons, D.P.M., Scott, D.L., Brown, R., and Hawkins, C.F. (1984). Cause of death in rheumatoid arthritis. *British Journal of Rheumatology*, **23**, 92–9.

Pyorala, K., de Backer, G., Graham, I., Poole-Wilson, P., and Wood, D. (1994). Prevention of coronary heart disease in clinical practice: recommendations of the task force of the European Society of Cardiology, European Atherosclerosis Society and European Society of Hypertension. *Atherosclerosis*, **110**, 121–61.

Querfeld, U., Ong, J.M., Prehn, J., Carty, J., Saffari, B., Jordan, S.C., and Kern, P.A. (1990). Effects of cytokines on the production of lipoprotein lipase in cultured human macrophages. *Journal of Lipid Research*, **31**, 1379–86.

Radford, E.P., Doll, R., and Smith, P.G. (1977). Mortality among patients with ankylosing spondylitis not given x-ray therapy. *New England Journal of Medicine*, **297**, 572–6.

Rantapää-Dahlqvist, S., Wällberg-Jonsson, S., and Dahlén, G. (1991). Lipoprotein (a), lipids and lipoproteins in patients with rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **30**, 366–8.

Rasker, J.J. and Cosh, J.A. (1981). Cause and age of death in a prospective study of 100 patients with rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **40**, 115–20.

Rigby, A.S. (1991). Review of UK data on the rheumatic diseases—5: ankylosing spondylitis. *British Journal of Rheumatology*, **30**, 50–3.

Robinson, D.R., Prickett, J.D., Makoul, G.T., Steinberg, A.D., and Colvin, R.B. (1986). Dietary fish oil reduces progression of established renal disease in (NZB×NZW)F1 mice and delays renal disease in BXSB and MRL/I strains. *Arthritis and Rheumatism*, **29**, 539–46.

Ross Peirce, L., Wysowski, D.K., and Gross, T.P. (1990). Myopathy and rhabdomyolysis associated with lovastatin gemfibrozil combination therapy. *Journal of the American Medical Association*, **264**, 71–5.

Scandinavian Simvastin Survival Study Group (1994). Randomised trial of cholesterol lowering in 4444 patients with coronary heart disease: results of the Scandinavian Simvaastin Survival Study (4S). *Lancet*, **344**, 1383–9.

Schumaker, H.R. and Michaels, R. (1989). Recurrent tendinitis and Achilles tendon nodule with positively birefringent crystals in a patient with hyperlipoproteinaemia *Journal of Rheumatology*, **16**, 1387–9.

Selley, M.L., Bourne, D.J., and Bartlett, M.R. (1992). Occurrence of (E)-4-hydroxy-2-nonenal in plasma and synovial fluid of patients with rheumatoid arthritis and osteoarthritis. *Annals of the Rheumatic Diseases*, **51**, 481–4.

Shepherd, J., Cobbe, S.M., Ford, J., Isles, C.G., Lorimer, A.R., MacFarlane, P.W., McKillop, J.H., and Packard, C.J. (1995). Prevention of coronary heart disease with pravastatin in men with hypercholesterolaemia. West of Scotland Coronary Prevention Study Group. *New England Journal of Medicine*, **333**, 1301–7.

Spriggs, D.R., Sherman, M.L., Michie, H., Arthur, K.A., Imamura, K., Wilmore, D., Frei, E., and Kufe, D.W. (1988). Recombinant human tumor necrosis factor administered as a 24-hour intravenous infusion. *Journal of the National Cancer Institute*, **80**, 1039–44.

Svenson, K.L.G., Lithell, H., Hällgren, R., and Vessby, B. (1987 *a*). Serum lipoprotein in active rheumatoid arthritis and other chronic inflammatory arthritides. I. Relativity to inflammatory activity. *Archives of Internal Medicine*, **147**, 1912–16.

Svenson, K.L.G., Lithell, H., Hällgren, R., and Vessby, B. (1987 *b*). Serum lipoprotein in active rheumatoid arthritis and other chronic inflammatory arthritides. II. Effects of anti-inflammatory and disease-modifying drug treatment. *Archives of Internal Medicine*, **147**, 1917–20.

Szachnowski, P. and Bridges, A.J. (1994). Subcutaneous cholesterol nodules. *Journal of Rheumatology*, **21**, 2391–2.

Takegoshi. T., Haba, T., Hirai, J., Saga, T., Kitoh, C., and Mabuchi, H. (1990). A case of hyperLp(a)aemia associated with SLE, suffering from myocardial infarction and cerebral infarction. *Japanese Journal of Medicine*, **29**, 77–84.

Terres, W., Tatsis, E., Pfalzer, B., Beil, U., Beisiegel, U., and Hamm, C.W. (1995). Rapid angiographic progression of coronary heart disease in patients with elevated lipoprotein (a). *Circulation*, **91**, 948–50.

Tinahones, F.J., Collantes, E., C-Soriguer, F.J., González-Ruiz, A., Pineda, M., Anón, J., and Sánchez Guijo, P. (1995). Increased VLDL levels and dimished renal excretion of uric acid in hyperuricaemic-hypertriglycerideaemic patients. *British Journal of Rheumatology*, **34**, 920–4.

Ulreich, A., Korner, G.M., Pfeiffer, K.P., Sedlmay, R., and Rainer, F. (1985). Serum lipids and lipoproteins in patients with primary gout. *Rheumatol International*, **5**, 73–7.

Urowitz, M.B., Bookman, A.A.M., Koehler, B.E., Gordon, D.A., Smythe, H.A., and Ogryzlo, M.A. (1976). The bimodal mortality pattern of systemic lupus erythematosus. *American Journal of Medicine*, **60**, 221–5.

Vaarala, O., Alfthan, G., Jauhiainen, M., Leirisalo-Repo, M., Aho, K., and Palosuo, T. (1993). Crossreaction between antibodies to oxidised low-density lipoprotein and to cardiolipin in systemic lupus erythematosus. *Lancet*, **341**, 923–5.

Vaarala, O., Manttari, M., Manninen, V., Tenkanen, L., Puurunen, M., Aho, K., and Palosuo, T. (1995). Anti-cardiolipin antibodies and risk of myocardial infarction in a prospective cohort of

middle-aged men. *Circulation*, **91**, 23–7.

Veale, D.J., Torley, H.I., Richards, I.M., O'Dowd, A., Fitzsimmons, C., Belch, J.J.E., and Sturrock, R.D. (1994). A double-blind placebo controlled trial of efamol marine on skin and joint symptoms of psoriatic arthritis. *British Journal of Rheumatology*, **33**, 954–8.

Wallace, D.J., Metzger, A.L., Stecher, V.J., Turnbull, B.A., and Kern, P.A. (1990). Cholesterol-lowering effect of hydroxychloroquine in patients with rheumatic disease: reversal of deleterious effects of steroids on lipids. *American Journal of Medicine*, **89**, 322–6.

Walton, A.J.E., Snaith, M.L., Locniskar, M., Cumberland, A.G., Morrow, W.J.W., and Isenberg, D.A. (1991). Dietary fish oil and the severity of symptoms in patients with systemic lupus erythematosus. *Annals of the Rheumatic Diseases*, **50**, 463–6.

Winyard, P.G., Tatzber, F., Esterbauer, H., Kus, M.L., Blake, D.R., and Morris, C.J. (1993). Presence of foam cells containing oxidised low density lipoprotein in the synovial membrane from patients with rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **52**, 677–80.

Wolfe, F., Mitchell, D.M., and Sibley, J.T. (1994). The mortality of rheumatoid arthritis. *Arthritis and Rheumatism*, **37**, 481–94.

Zimmerman, J., Fainaru, M., and Eisenberg, S. (1984). The effects of prednisone therapy on plasma lipoproteins and apolipoproteins. A prospective study. *Metabolism*, **33**, 521–6.

Zuckner, J. (1990). Drug-induced myopathies. *Seminars in Arthritis and Rheumatism*, **19**, 259–68.

5.14 Soft-tissue rheumatism

Brian Hazleman

Introduction

Soft-tissue injuries, although commonly overlooked in the planning and provision of health care, are of major and increasing importance. Fortunately there has been an increase in interest in these lesions concomitant with an increase in our understanding of the associated disorders.

When rheumatology first developed as a medical subspeciality its basis was in the articular diseases, a concept that was soon found inadequate as it became apparent that other synovial structures, such as bursal and tendon sheaths, were equally affected by the same conditions. Later, the term 'abarticular rheumatism' was used to distinguish the extra-articular rheumatic complaints from the better-defined articular rheumatism. Because joints are hard and bony, and abarticular rheumatism by definition does not affect joints, the all-embracing term for this group of complaints, now in international usage, is soft-tissue rheumatism. The heterogeneity of soft-tissue rheumatism poses considerable problems in arriving at sensible treatment protocols and systems of management.

Lesions of tendons and their sheaths, fascias, bursae, joint capsules, and the tenoperiosteal junction (the enthesis) cause much illness and loss of productivity. They constitute a significant proportion of the workload of general medical practices and of hospital accident, orthopaedic, and rheumatology departments. As biopsy and surgery are rarely employed in their diagnosis and treatment, histological data are scanty and their pathological backgrounds poorly understood. While there are adequate anatomical and clinical features to allow identification of individual conditions, diagnosis is often imprecise and management remains largely empirical.

Any or all of these lesions may occur in association with overt systemic disease, as, for example, in inflammatory arthritis or infection, but a large proportion occur in its absence. In these circumstances local causes, such as chronic, repetitive, low-grade trauma, excessive and unaccustomed use, either at work or at play, may be responsible. These factors may also cause partial interruption of the blood supply, resulting in incomplete attempts at healing and degeneration, which render the extra-articular structures more vulnerable in the middle-aged and elderly people in whom these lesions predominate.

One feature common to all soft-tissue syndromes is a tendency to spontaneous remission. Many of the lesions take only weeks to improve and few persist with significant symptoms beyond 6 months. Therefore, a clear diagnosis allows the doctor to reassure the patient that arthritis is not present and that the prognosis is good. Failure to resolve is often due to further injury. Few of the lesions require complete rest, but most respond to selective rest of the region involved.

An accurate diagnosis of soft-tissue rheumatism can be made by taking a careful history, considering possible trauma or overuse, and carrying out a systematic examination. For most local lesions, further investigation is unnecessary and radiographs that reveal the expected, age-related degenerative changes can often confuse. It is usually apparent if there is a more generalized condition that requires appropriate investigation.

Classification

Disorders can be classified by clinicopathological process (tendinitis), by anatomical region (shoulder pain), or by aetiology (repetitive strain injury). There is no universally accepted classification. The following is practical and problem-orientated, and divides the conditions into (i) generalized and (ii) localized. The generalized forms may be further divided into those associated or not with inflammation (Table 1). Those that are localized may be regional, involving a few sclerotomes, or affecting a specific site (Table 1).

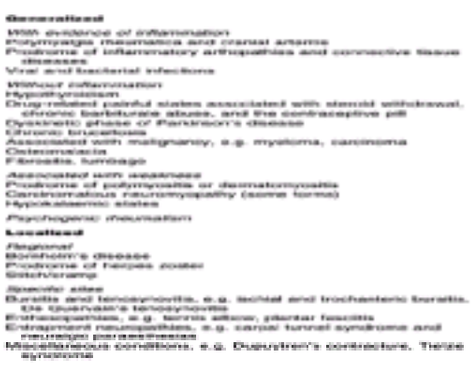


Table 1 Classification of soft-tissue rheumatism

Fibromyalgia or the 'fibrositis syndrome' is a particular variant of soft-tissue rheumatism. It is characterized by widespread aching and multiple tender sites. The tender sites are central to the definition and our understanding of this condition.

Generalized soft-tissue lesions

These may result from underlying disease, and most of the primary conditions can be diagnosed by careful clinical and laboratory assessment. These conditions will

not be discussed here, and polymyalgia rheumatica has been considered elsewhere (Chapter 5.11.5).

The diagnosis of 'psychogenic rheumatism' must be made with caution and after the exclusion of other diseases. Chronic pain may produce psychological overtones. There are several features that will suggest a psychological illness, since overt psychiatric symptoms are not usually present in rheumatic disorders but are common in patients who are unhappy at work or at home. Suggestive features include written lists of symptoms, inconsistent or negative physical findings on repeated examinations, and inappropriate concern with serious future disability.

Localized soft-tissue lesions

The major structures involved and associated lesions are listed in Table 2. Examination should permit accurate localization of the anatomical structure involved. The conditions frequently encountered in clinical practice are listed in Table 3.

Structure	Lesion
Tendons, tendon sheaths	Rupture, degeneration tendinitis, peritendinitis, tenosynovitis, ganglia
Tenoperiosteal junction	Enthesopathies, apophysitis
Bursae	Bursitis—acute or chronic
Fascias	Fasciitis, Dupuytren's contracture
Ligaments	Sprain, strain, tear

Table 2 Localized soft-tissue lesions

Site	Lesion
Shoulder	Rotator-cuff lesions Capsulitis of glenohumeral joint
Elbow	Lateral epicondylitis Medial epicondylitis Olecranon bursitis
Wrist and hand	Carpal tunnel syndrome Dupuytren's contracture De Quervain's tenosynovitis
Trunk	Non-specific neck and low back pain Trapezius muscle spasm 'Costochondritis'
Hip	Ischial and trochanteric bursitis Meralgia paresthetica
Knee	Bursitis Ligamentous sprains
Ankle and foot	Achilles tendinitis/peritendinitis Plantar fasciitis

Table 3 Common soft-tissue lesions and sites

Economic effects and burden on health-care services

Soft-tissue rheumatism is the most commonly encountered rheumatic cause of sickness absences from work, accounting for 44 per cent of medically certified rheumatic spells and 6 per cent of all incapacity spells. As is known from their natural history, however, the amount of time lost from work is comparatively less impressive. Thus the average duration of an incapacity spell for soft-issue rheumatism (21 days), when aggregated, amounts to 2.5 per cent of rheumatic days or 3.5 per cent of all days lost from work. These figures include absences on industrial injury benefit as well as sickness benefit, the former making up about one-fifth of the total. There is an age association in days lost. Traumatic disorders show no very marked trend, but the non-traumatic complaints are associated with an increase in days lost that is progressive with age; the rate is at least six times greater in the oldest as compared with the youngest quinquennium. This pattern reflects the fact that the average duration of incapacity spells grows longer with age, those occurring in the oldest group being more than twice as long as those in younger people. The overall pattern of spells is generally similar in most of the diagnostic categories, but muscle tendon and fascial lesions are unusual for the undue length of some of their incapacity spells.

In the United Kingdom alone the resultant loss of working days from soft-tissue lesions is likely to cost the country almost a billion pounds a year in lost productivity, apart from the value of social security payments, lost tax revenue, and health and social services applied to this problem. Behind these huge sums is a great deal of pain and misery, and much disruption to family life and work arrangements. The economic effects, loss of time from work, and the load that such patients present to both general practitioners and hospital doctors, are clearly immense. Greater attention needs to be focused on these diseases.

Epidemiology

Studies on the incidence of soft-tissue injuries are sadly inadequate as epidemiological studies have hitherto only infrequently been performed. Indeed, Dixon (1979) has described soft-tissue rheumatism as 'the great outback of rheumatology, a vast frontier land, ill-defined and little explained, its features poorly categorised and far from internationally agreed'. The fact that any or all of these soft-tissue lesions may be associated with overt systemic disease makes the classification and interpretation of data difficult. The lack of universally acceptable, defined criteria for these injuries has been a major obstacle to conducting epidemiological studies.

The absence of specific diagnostic tests for the soft-tissue syndromes highlights the importance of developing diagnostic criteria, based on combinations of the predominant clinical features, that are sensitive enough to exclude those without the syndrome. Most of these syndromes will be covered by the inclusion criteria of pain at a specific site or sites, local tenderness, limitation of movement, and pain on specific resisted movement or movements. There is a need for rigorous diagnostic criteria that take account of the lack of a 'gold standard' reference test, the continuous distribution pattern of the symptoms and signs, the variation in disease expression at differnt times, and the inter- and intraobserver variation in eliciting the specific clinical signs. Diagnostic criteria are required for defining occurrence, measuring outcome, and assessing the effect of intervention, for example in recruitment to clinical trials.

A population survey by the Swedish National Central Bureau of Statistics indicates a prevalence of all forms of soft-tissue rheumatism of 1.6 per cent in men and 3.6 per cent in women (Bjelle *et al*. 1990). Chronic complaints attributed to non-articular rheumatism as well as miscellaneous back disorders were estimated to occur in about 3 per cent of the adult (18–70 years) population of the United States in the 1976 National Health Interview Survey. In comparison, about 3 per cent of symptoms were attributable to rheumatoid arthritis and 8 per cent to degenerative joint disease. Together, soft-tissue lesions comprise one-third of all rheumatic diseases seen by family physicians and 3 to 4 per cent of all ambulatory medical-care visits (Wood *et al*. 1979).

There is considerable confusion in diagnostic terminology and a wide variation in reported incidence between countries, probably reflecting the contrasting, as well as legal, attitudes to these conditions. In general the increased incidence of these disorders appears to be due to new technology leading to advanced automation and mechanization. Occupational disorders are of interest because, theoretically at least, it should be possible to devise methods of working that would prevent their development. Ergonomic and behavioural studies have much to offer. Certainly sufferers are at present at considerable hazard of repeated attendance, as indicated by the reported rates of recrudescence for 'beat' knee (12 per cent), 'beat' elbow (7 per cent), and tenosynovitis of the wrist (6 per cent) (Wood *et al*. 1979). It would be of interest to learn why only some individuals exposed to particular methods of working develop soft-tissue injuries. Automation and mechanization have led to the need for repetitive movements that may produce repetitive strain injuries and associated overuse syndromes; this aspect will be discussed later.

Crude differences in age patterns may be discerned between traumatic and non-traumatic forms of soft-tissue rheumatism. The frequency of spells of strains and

sprains tends to reduce after the thirties, whereas non-traumatic conditions show a less marked decline and this does not appear until after the forties. Rates in women tend in general to be less than in men.

Of localized problems, the painful shoulder ranks highest in frequency and consequently has been better documented than most other soft-tissue lesions. Regional shoulder complaints rank fifth amongst the regional rheumatic diseases as a cause of incapacity or of visiting a physician. Data from general medical practices in the United Kingdom (Department of Health and Social Security 1986) suggest that approx. 1 in 170 of the adult population will present to their general practitioner with a new episode of shoulder pain each year. This contrasts with 1 in 30 for back pain.

Elderly people frequently have considerable functional impairment: 20 per cent of the elderly hospital population had shoulder disorders; less than one-fifth of these had sought medical attention (Chard and Hazleman 1987). The association between shoulder pain and occupation was highlighted in a study of the medical outpatient departments of six heavy industries by Bjelle *et al*. (1980). Approximately one-third of the outpatient visits were for non-traumatic musculoskeletal complaints, of which 30 to 50 per cent were neck–shoulder complaints. In the United States, the Health and Nutrition Examination Survey 1971–1975 studied the adult population aged 25 to 74 years. It found that 4.2 per cent had suffered from shoulder complaints and that shoulder abnormalities observed by a physician were present in 1.2 per cent; almost 40 per cent had associated neck complaints. In over 80 per cent the complaints had occurred during the year immediately before the study (Miller 1973). Health-care and insurance data support the influence of age and sex in relation to shoulder complaints amongst workers in production industries. The influence of age may partly be due to the frequent finding of degenerative changes of the rotator cuff that make this tissue highly vulnerable to trauma after the fifth decade. Also, women run a higher risk than men of developing shoulder complaints when working in production industries. Further studies are required to determine whether the age, sex, and occupation pattern seen with shoulder lesions is equally applicable to other soft-tissue injuries.

Tennis elbow or lateral epicondylitis affects 1 to 3 per cent of the population. It occurs mostly between 40 and 60 years of age, and usually affects the dominant arm. Most sufferers experience a recurrence of symptoms within 18 months. While some 40 to 50 per cent of tennis players suffer with tennis elbow, it is more frequent and severe in older players. Less than 5 per cent of cases in total are related to tennis and it is found more often in non-athletes. Labelle *et al*. (1992) report that an average of 62 days is lost from work per patient in industrial workers.

Defining the outcome

There is a very low rate of referral of soft-tissue problems to hospital and thus outcomes from hospital series do not reflect the outcomes from the syndrome as a whole. There is also considerable variation between different hospitals in the case severity of new attendees. This is partly a reflection of waiting-list time, and partly of the interests of general practitioners and hospital clinicians. Also, many conditions are self-limiting: half of patients presenting to general practitioners with soft-tissue back pain are better within a week and 90 per cent by a month.

Pathogenesis

Perhaps one of the biggest hindrances to progress in understanding soft-tissue rheumatism has been a negative attitude to the affected tissues, which are still thought to be inert, homogeneous structures. Their role in joint mechanics and pathology is often considered to be passive and secondary to that of other joint structures such as bone, cartilage, and synovium. These misconceptions are belied by recent studies demonstrating that these tissues are indeed metabolically active, interesting, and worthy of investigation.

As outlined above, many local causes of soft-tissue rheumatism are related to chronic, repetitive, low-grade trauma and excessive and unaccustomed use (both at work and at play), which both may partially interrupt the local blood supply. Incomplete attempts at healing and degeneration will render soft tissues more vulnerable in the older people in whom such lesions are more common. Since the vascular supply to adult tendon is poor, healing of these lesions is slow. Poor tendon repair and degeneration would appear to explain the chronicity of tendon lesions.

Tendon and ligaments: composition and function

Tendons are clusters of parallel collagen fibrils interspersed with a few fibrocytes; the collagen is almost all type I. Larger tendons consist of multiple fascicles separated by loose connective tissue containing nerve fibres and blood vessels. Ligaments may be of a similar structure, but usually exhibit less regular orientation of collagen and contain elastin. They function passively, being structurally superior to tendons for constant tension. Tendons, in contrast, must provide flexibility and resistance to tension, with concentration of force to an attachment area. Loading a tendon causes a brief alignment of fibres, then linear extension to the limit of elasticity. According to Hooke's law this is followed by failure of individual fibres and then rupture. The area of weakness should be the attachment to bone; but the loading is distributed evenly by a mechanism called 'fibre weave'. Therefore, tendons often fail at a distance from the attachment or enthesis.

The fan-shaped point of attachment of tendon to bone (the enthesis) is an area of great stress. It differs histologically from the rest of the tendon. The enthesis is traditionally made up of four zones: tendon, fibrocartilage, mineralized fibrocartilage, and bone. In the zone of mineralized fibrocartilage there is an increase in smooth endoplasmic reticulum, lysosomes, and intracellular lipid, and a more prominent Golgi apparatus.

Blood supply

Arterioles pass between the fascicles in the tendon. These freely communicate and are accompanied by venules and lymphatic vessels. Extensions of vessels from the muscles pass through the tendon and anastomose with those from the periosteum. Studies with radioactive isotopes have demonstrated that the enthesis has a particularly rich blood supply. It is believed that the blood supply to tendon is relatively abundant compared to the metabolic demands. However, there is reduced blood flow when tendons are under tension.

Nerve supply

This is sensory, there being no definite evidence of vasomotor control. Close to the attachment to muscles are found specialized afferent receptors. The nerve fibres from the tendon mainly follow the branch of the motor nerve to the muscle but some fibres form small branches that pass directly into nearby peripheral nerves.

Tensile strength

The tensile strength of normal tendon is greatly in excess of ordinary demands. This strength requires an intact blood supply, even partial interruption leads to a reduction. Normal tendon with a cross-sectional area of 1 cm^2 can support 600 to 1000 kg, which is similar to half that of steel.

Tendon lesions

Rupture

Severe trauma directly to a tendon will result in rupture. Marked strain on a tendon may also do this, although avulsion fracture at site of attachment is more likely. Analyses of spontaneously ruptured tendons have shown degenerative changes that include alterations in the size and orientation of collagen fibres, with increased deposition of proteoglycans between the fibres. The tenocytes have enlarged vacuoles, sometimes containing lipids, and cell necrosis is sometimes found. In some cases calcium is deposited (Kannus and Jozsa 1991).

Similar changes are found in human tendons removed at operation for degeneration of the rotator cuff and lateral epicondylitis (Chard *et al*. 1994).

Tenosynovitis

Inflammation of the synovial lining of the sheath through which a tendon moves may be the result of an inflammatory arthritic condition but is more commonly caused by trauma. It is often the result of unaccustomed exercise or repetitive work action. Tenderness, swelling, and crepitus on palpating the moving tendon are characteristic. When severe and chronic the tendon sheath becomes thickened, especially over bony prominences. Secondary thickening of the tendon distally may occur, producing snapping on movement of the thickened area through the inflamed sheath, for example trigger finger. Sites of tenosynovitis include the thumb (De

Quervain's tenosynovitis), finger flexors, flexor carpi radialis, common peroneal sheath, and tibialis posterior tendon. Treatment includes rest, splinting, local steroid injections, and sometimes surgical release.

Tendinitis

This is a poorly defined group of conditions where pain is attributed to strain or injury to tendons and their attachments to bone. In many cases the lesion is believed to be tenoperiosteal, for example lateral humeral epicondylitis (tennis elbow) and jumper's knee, and thus comes under the term enthesopathy. The most important tendon lesions in terms of frequency and severity are those affecting the shoulder, and these produce the majority of cases of shoulder pain.

Pathological studies (Fig. 1)

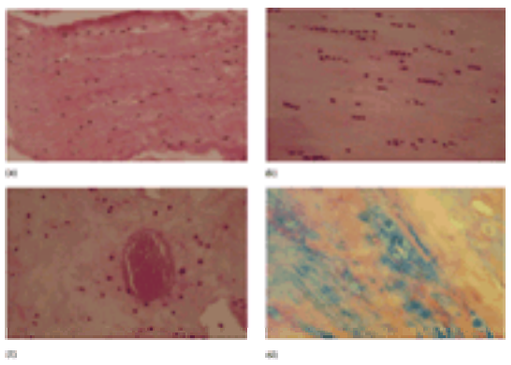


Fig. 1 Common histological findings in the human supraspinatus tendon: (a) tendon from a 19-year-old cadaver showing normal fibre structure with elongated fibroblasts between parallel fibre bundles (H and E); (b) ruptured tendon from a 65-year-old patient with tendinitis, showing region of fibrocartilage with rows of rounded, chondrocyte-like cells, which may be a pathological change or a normal adaptation to compressive forces acting on the tendon (H and E); (c) ruptured tendon from a 55-year-old patient with tendinitis, showing severe degenerative changes with no distinct fibrillar structure, scattered round cells, and a calcific deposit in the matrix (H and E); (d) specimen from a 76-year-old showing deposition of glycosaminoglycan between collagen fibrils and associated with rounded cells (Alcian blue/periodic acid–Schiff).

In vivo work has shown that tendon healing after damage probably involves the reaction of surrounding tissue and adhesion, without participation of the tendon tissue itself. Other studies have revealed that tendons are capable of responding to injury, although it does appear that most actively dividing cells are derived from the superficial part of the tendon.

The majority of cells obtained from tendon that are capable of replication *in vitro* are derived from the superficial layer (epithelium). The growth characteristics of the cell lines have been established by investigating DNA synthesis using thymidine incorporation in response to stimulation by fetal calf serum. No significant reduction in growth response with increasing age was found. Initial light-microscopic analyses of normal rotator-cuff tendons (Chard *et al*. 1989) suggest a general thickening of blood vessels, fewer tendon fibroblasts, increased mucopolysaccharide, and increased calcification with increasing age. An increase in cells resembling chondrocytes and hyperplasia of arteriolar intimal are seen (Yamanaka and Fukuda 1991).

Attention has been drawn to the unifying concept of the enthesopathies which suggests that local ischaemia is the common denominator. The enthesis is always at risk because the working muscle takes up most of the blood at the expense of the tendon and insertion. Other contributory factors include overstressing, muscular hypertonus, and excessive cooling. Endogenous factors such as impaired vascularization, metabolic disorders, endocrine disorders, trophic disorders, toxic damage, and even psychological factors may also influence damage to the enthesis. Tendons without sheaths are better supplied with blood than those with, and dynamic blood-flow studies *in vivo* in these regions may yield valuable information. This section has focused on the tendon and its pathology. Lesions of bursae, fascias, and muscle will be discussed later in the chapter.

Collagen is the basic framework of all soft tissues. Once a tear occurs within the collagen bundles, the defect is replaced by haphazard, loose connective tissue formed in the blood clot that initially fills the torn area. Thus the intrinsic structural strength may be reduced significantly, leading to impaired power, mobility, and skill, and further tears.

Tendon cells are metabolically responsive and capable of repair; they maintain the matrix composition by a balance between anabolic and catabolic processes (Abrahamsson 1991). Biochemical and cell biological studies have shown that diseased tendons have an increased concentration of dermatan and chondroitin sulphates, and a threefold increase in hyaluronan (Riley *et al*. 1994), indicating that a change in proteoglycan synthesis has occurred. There is also an increase in cell numbers, a reduced collagen content, possibly caused by an increase in collagen degradation, and in the majority an increased content of collagen type III. These changes are consistent with inflammation and a fibroproliferative response, presumably in an attempt to repair the tendon defect, although it is not known if this process is primary or secondary to the tendon rupture.

A highly specialized 'fibrocartilage' develops in regions of tendons exposed to compression (Evans *et al*. 1990), which differ biochemically and structurally from tension-bearing regions and have characteristics somewhere between those of classic tendon regions and articular cartilage (Vogel *et al*. 1993). Although type I collagen remains the principal matrix component, the cellular activity in these compressed regions includes the synthesis of both type II collagen and aggrecan.

Growth factors have anabolic effects on tissues, increasing matrix synthesis and reducing matrix breakdown. Cytokines such as interleukin 1 and tumour necrosis factor-a often have the opposite effect, decreasing matrix synthesis and up-regulating the proteinases that promote matrix breakdown. Recent studies have shown that tendons *in vitro* respond to interleukin 1 by altering the synthesis of matrix components and proteinases such as matrix metalloproteinases (Dalton *et al*. 1995). The production of excessive enzyme activity can exceed that of local inhibitors (such as tissue inhibitor of metalloproteinases) and so lead to collagen breakdown.

Recent interest has centred on the role of prostaglandin synthesis and release in response to soft-tissue injury due either to trauma or overuse. Prostaglandins may act synergistically with other inflammatory mediators such as histamine, serotonin, and bradykinin to potentiate both swelling and pain. Muscle injuries are associated with bleeding to a greater or lesser degree, and interstitial haematomas produce marked pain and loss of function. Muscle regeneration is a slow process. Strains affect muscles. With minor strains only fibres are damaged, with the muscle sheath left intact. With more severe strains there is partial or complete rupture of fibres and sheath. A sprain is an overstretch injury of a ligament. It may affect only a few fibres or lead to rupture.

Generalized soft-tissue conditions

Repetitive strain syndrome

Occupational repetition strain injuries have become, particularly in recent years, a significant source of disability at work. There is considerable confusion in diagnostic terminology and a wide variation in reported incidence between countries, perhaps reflecting the contrasting medical and more particularly, legal, attitudes to these conditions. Flexor tenosynovitis, rotator-cuff tendinitis, and lateral epicondylitis have all been described as 'occupationally' induced clinical syndromes; they have high rates of relapse—6 per cent for tenosynovitis of the wrist, for example. Repetitive strain injury has, however, more diffuse clinical features and no clearly defined pathological basis. Two conditions are currently recognized as work-related industrial injuries: PDA4—cramp of the hand or forearm, in people with an occupation entailing prolonged periods of handwriting, typing, or other repetitive movements of the finger, hand, or arm; and PDA8—traumatic inflammation of the tendons of the hand or forearm or the associated tendon sheaths in any occupation entailing manual labour or frequent or repeated movements of the hand or wrist. In general, as mentioned above, the increased incidence of these disorders appears to be due to new technology with advanced automation and mechanization. This has led to a need for rapid, repetitive movements, often of just the arms or even the hands and wrists alone. Although actual physical workloads may be lighter, it is this increased rate of work concentrated locally on an individual's musculoskeletal system that results in repetitive (repetition) strain injuries, and associated overuse syndromes. The influence of occupation on how we use or overuse parts of our body is striking. This suggests that were alternative ways of accomplishing various

activities to be developed, a part of the problem could be controlled. Ergonomic and behavioural studies would seem to have much to offer. Definitions of the association between work and rheumatic complaints have been suggested (World Health Organization 1988). They emphasize the multifactorial nature of work-related musculoskeletal complaints in industrial workers. The load from work may not only be physical; stress factors also are important in some workplaces. Predisposing factors such as age and anthropometric features must also be considered. Repetitive strain syndrome can be defined as follows:

1. a chronic pain syndrome;
2. affecting one or both neck–arm regions;
3. occurs in activities requiring controlled posture, often of a repetitive nature;
4. psychological factors contribute to the syndrome.

The clinical features consist of:

1. chronic pain in neck, chest wall, arm, and hand;
2. inability to achieve previous work performance or carry out full leisure activities;
3. variable hand–arm–forearm swelling with poor grip strength and taut proximal muscles;
4. often mild algodystrophy;
5. poor sleep pattern, often with mood changes.

Terminology

Arm pain associated with work, commonly known as repetitive strain injury, is currently sweeping the developed world. The most recent 'epidemic' began in Australia in the 1980s and has now 'spread' to the United States and United Kingdom. The term repetitive strain injury is perhaps misleading in that it implies that an actual 'injury' has been caused by repetitive movement. This has so far been difficult to establish, given the diffuse clinical features and absence of pathology in most cases. Also, there is no certain pathogenesis for repetitive strain injury that would justify the use of the term 'strain' and many sufferers develop their symptoms as a result of static rather than repetitive or dynamic muscle load (Littlejohn 1986).

Although an association between certain occupations and the incidence of repetitive strain injury has been described, there have been few studies designed to assess the connection properly. In Australia there was a sudden, dramatic increase in reporting symptoms that lasted around 18 months and was followed by an equally rapid reduction to baseline. The rise in cases was mirrored by increasing numbers of media reports on repetitive strain injury. This can be seen in Hocking's study carried out by Telecom Australia (Hocking 1987). Annual reports of repetitive strain injury increased from 109 in 1981 to over 1700 in 1985. Telephonists were most affected and telegraphists least. This order is the inverse of the keystroke rates, which are a few hundred an hour for a telephonist and over 12 000 an hour for a telegraphist. Cleland (1987) has described repetitive strain injury as a model of social iatrogenesis—the treatments, advice, and expectations provided by a wide range of people have provided an environment that has focused attention on discomfort about, and apprehension of, the potentially damaging effects of pain in the workplace. Through this process it is likely that a minor discomfort will be transformed into a protracted, painful, and disabling condition, precluding effective work and degrading quality of life.

The rapid decline in repetitive strain injury in Australia was due to several different factors. Work practices were changed, appropriate work breaks were instituted, and advice given about comfortable work stations. The term repetitive strain injury was also disposed of; if no specific diagnosis could be made, the condition was designated a regional pain syndrome.

Clinical features

Pain is the principal feature; this may be associated with inability to perform routine work or household activities. There may have been an underlying ache in the arm, shoulder or neck for some weeks before the onset of the pain syndrome, or it may follow a specific soft-tissue injury. Usually the symptoms begin around the wrist or forearm or elbow, and within days or weeks spread into the upper arm, shoulder, and neck. They may then involve the opposite side. Up to 20 per cent of patients develop more generalized pain involving the low back, buttock, and leg regions.

Generalized fatigue is common, and pain may fluctuate with activity, emotional stress or temperature change. In addition, patients report disturbance of their usual sleep pattern, with frequent waking and a feeling of tiredness in the morning. Examination shows an altered pain threshold.

Vasomotor changes in the forearm may be apparent and in the minority all the features of an algodystrophy syndrome may occur. The patient describes dysaesthesias in the hand and poor grip strength. There is a sensation of swelling but usually no objective evidence of this. Tender points in such as the first web space in the hand are also present. There is no evidence of synovitis, tenosynovitis, or neurological abnormality.

The syndrome is therefore one of regional pain, but the usual course is of prolonged pain for several months or years, irrespective of treatment or outcome of compensation. Once the regional pain syndrome has developed an exaggeration of tenderness when the patient is assessed for medicolegal purposes may be noted. No racial, social, or professional group has been spared, but the self-employed are not usually affected.

Several psychological factors come into play in occupations requiring rapid, repetitive movements. Indeed a significant body of opinion follows the concept of repetitive strain injury as a largely 'hysterical' or psychogenic illness (Awerbach 1985). A more accurate assessment would seem to be that physical, ergonomic, environmental, and psychological factors are involved in its onset and development.

Treatment

Management is similar to that of other chronic pain syndromes. Prevention and early diagnosis are important. If the full syndrome has developed the patient must be reassured that there is no tissue damage or injury, and that the symptoms are best thought of as a form of' overuse strain' that will recover with suitable modifications of activities.

It is necessary to treat any triggering ergonomic factor, either by modifying work activity or using a cervical collar. The patient should be encouraged to carry out regular exercise that involves stretching of tight, tender regions such as neck and shoulder. Sleep disturbance may require treatment with a hypnotic or low-dose tricyclic antidepressant.

If the syndrome has arisen in connection with paid employment, all the parties involved, including management, trades unions and third-party insurers, should be informed. An early return to work should be encouraged, but modification of work activity may be necessary when the pain is severe. Prolonged rest does not help, nor does the use of anti-inflammatory drugs or extensive physical treatments. Chronic pain is treated by gradually increasing activity under supervision.

Many patients with this problem are involved in disputes over compensation, litigation, or disability assessment. It has been suggested that the occurrence of chronic pain syndromes may be encouraged in those countries where compensation schemes include any symptoms that occur at work. This happened in Australia during the 1980s and is thought to have contributed to the pain syndrome 'epidemic' in that country. Prolonged medicolegal processes counteract any appropriate treatment programme (Littlejohn 1989).

The term repetitive strain injury implies a cause without defining the lesion. Ireland (1988) has emphasized that this condition predominantly affects women employed in low-paid, monotonous, unprestigious jobs. The employees placed the blame on the employment and thus on the employer. However, a small number of patients with this syndrome do present with genuine pain and disability and do not respond to current management. If no specific musculoskeletal diagnosis can be made, then one should look at social and psychological factors that might promote the pain. These should be addressed and the patient's pain put in its proper context. Occupational issues concerning a constrained and rigid posture, the work cycle, time spent in any one activity, work satisfaction, and social intervention at work are important.

The advice 'if it hurts, stop' runs contrary to all accepted principles of the behavioural management of chronic pain and may lead to chronicity. The involvement of a

clinical psychologist in the team, and an active occupational rehabilitation programme, helps most people to return to work.

Hand and arm problems of musicians

Musicians are prone to a variety of problems in the upper limb that produce significant disability; these include overuse syndromes, entrapment neuropathies, and focal motor dystonias. The diagnosis can be difficult as symptoms may be mild and only occur on playing. Episodes may be triggered by changes in repertoire, technique, and instrument, or by increases in daily playing time. The weight of a wind instrument such as the clarinet or oboe frequently leads to pain in the muscles of the first web space. Percussionists have a low prevalence of regional pain syndrome, presumably because their playing is more intermittent.

In reviewing several series of musculoskeletal injuries in musicians, Hoppmann and Patrone (1989) reported that of 179 injured musicians, 62 per cent had overuse syndrome, 18 per cent had nerve-entrapment thoracic-outlet syndrome, and 10 per cent had problems of motor control. Some of the musculoskeletal problems seen in musicians are listed in Table 4. Diagnosis and management are helped by a detailed knowledge of the instruments used and the specific dynamics of music making. Musculoskeletal problems are common in musicians of all ages and levels of skill. Recent research has included attempts to better define overuse, focal dystonia, and the role of hypermobility.

Overuse injury:
Specific tenosynovitis
Non-specific diffuse forearm pain
Entrapment neuropathies
Focal dystonia
Thoracic-outlet syndrome
Hypermobility syndrome
Osteoarthritis

Table 4 Musculoskeletal problems of musicians

Treatment must take into account the specific injury and the instrument played. Table 5 lists some of the treatments that are generally applicable. Greater consideration is being given to the ergonomics of music making and technique is being assessed as a risk factor.

Rest:
Complete or incomplete rest periods into practice
Technique:
Correct problem; noticing posture, muscle tone and movements that aggravate injury
Physical or occupational therapy:
Splints, exercise, adaptive devices
Relaxation techniques:
May include electromyographic feedback
Non-steroidal anti-inflammatory drugs or local steroid injections
Surgery:
Nerve-entrapment release

Table 5 Treatment of musculoskeletal injuries in musicians

Musculoskeletal problems in dancers

Overuse injuries are common, and many joints are stressed by ranges of motion exceeding those that anatomy readily permits. The most common problems are back pain, and damage to the ankle and foot. Some 70 to 80 per cent of professional dancers complain of back pain, and the repetitive movements of dance may lead to overuse injuries such as tendinitis, neuritis, and stress fractures (Ramel and Moritz 1994). Pre-existing hypermobility may predispose to injury, and in the case of stress fractures, osteoporosis secondary to eating disorders and amenorrhoea may play a part. In one ballet company 104 dancers sustained 309 injuries over a 3-year period (Garrick and Requa 1993).

Fibromyalgia

The two cardinal features of fibromyalgia are generalized chronic pain and diffuse tenderness at discrete anatomical sites. Fibromyalgia has been found in 2 to 5 per cent of the population, predominantly affecting women in their forties and fifties, and is recognized as the second most common disorder seen in North American rheumatological practice. In some patients the syndrome complicates the illness associated with established rheumatic diseases.

Fibromyalgia is a recognizable syndrome characterized by chronic, diffuse pain, an absence of inflammatory or structural musculoskeletal abnormalities, and a range of symptoms that include fatigue, and sleep and mood disturbances. Physical examination and laboratory testing are unrevealing, except for the presence of pain on palpation of characteristic soft-tissue sites, the tender points.

Despite the recognition of fibromyalgia by the World Health Organization in 1992, it remains a controversial condition and its existence as a distinct entity remains uncertain (Cohen and Quintner 1993). However, the concept of fibromyalgia is a useful one, allowing many investigations to be avoided and appropriate advice on treatment to be given.

Fibromyalgia may overlap with symptoms of, and the patient further impaired by, anxiety and depression; patients with fibromyalgia have high scores on anxiety and depression questionnaires. The term fibromyalgia syndrome does not imply causation and merely describes the most common symptoms. Many patients with 'chronic fatigue syndrome' fulfil the criteria for fibromyalgia and represent one end of a spectrum of presentation. Evidence for triggering viral infections is lacking in the majority of patients, and, unlike fibromyalgia, most 'postviral' syndromes are self-limiting and are not associated with tender sites. Common presenting symptoms are listed in Table 6.

Fatiguability: Often extreme, following minimal exertion Predominantly axial, often aggravated by stress and cold Diffuse and unresponsive to analgesics/NSAIDs (all severely limit daily activities)
Objective swelling of extremities
Paraesthesias, dysaesthesia of hands, feet
Waking unrefreshed
Poor concentration, depressed
Headache, diffuse abdominal pain
Altered bowel habit
Urinary frequency, urgency

NSAIDs, non-steroidal anti-inflammatory drugs.

Table 6 Common symptoms of fibromyalgia

The only tests required to exclude alternative diagnoses that may present with widespread pain, weakness or fatigue are listed in Table 7.

Differential diagnosis	Investigations
Hypothyroidism	Full blood count and ESR
Systemic lupus erythematosus	Calcium, creatine kinase
Inflammatory myopathy	Thyroid function
Hyperparathyroidism	Antinuclear factor
Osteomalacia	Plasma proteins
Parkinsonism	
Myeloma, carcinomatosis	
Polymyalgia rheumatica	

ESR, erythrocyte sedimentation rate.

Table 7 Differential diagnosis and investigation of fibromyalgia

The tender sites are central to the definition and understanding of this condition. The neck and back show many changes, but the other tender sites reveal no changes sufficient to account for the marked tenderness, and certainly no inflammation. The same sites become tender in regional pain syndromes. The points (Fig. 2) are unknown to the patient, and often quite far from the region of referred pain; so that the pattern cannot be exaggerated for psychological reasons. The tenderness can be quantified, whereas pressure over other sites shows a normal pain threshold.

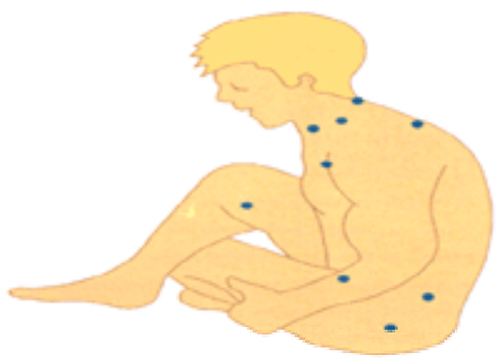


Fig. 2 The 'tender' points.

However, it should be emphasized that the very existence of this condition is a matter of considerable debate. Doubts have been expressed about the validity of both the symptoms reported by patients and the signs observed; some deny the existence of the condition because no objective pathological findings are demonstrable. However, in 1990 the American College of Rheumatology published criteria for the classification (Table 8). According to this set of criteria (American College of Rheumatology 1990), fibromyalgia is a syndrome of widespread pain, by definition affecting both sides of the body and the upper and lower segments. Its symptoms may also include sleep disturbance, fatigue, and stiffness. The most important feature is, however, the 'tender point' count, first described by Smythe (1989). Smythe has suggested that the pain experienced by patients with fibromyalgia is referred pain from mechanical stresses in the lower neck and low back. Studies by Lewis and Kellgren in the 1930s (Kellgren 1939) demonstrated that irritation of superficial and deep structures in the spine can produced both referred pain and also referred tenderness. Maigne (1972) demonstrated that irritation of posterior branches of spinal nerves could produce referred pain, referred tenderness, skin-rolling tenderness, and reactive hyperaemia in the anatomical distribution of the involved nerves. However, these observations have not been confirmed in patients with fibromyalgia, and the diffuse nature of pain in fibromyalgia suggests that spinal factors alone are not responsible for the clinical syndrome.

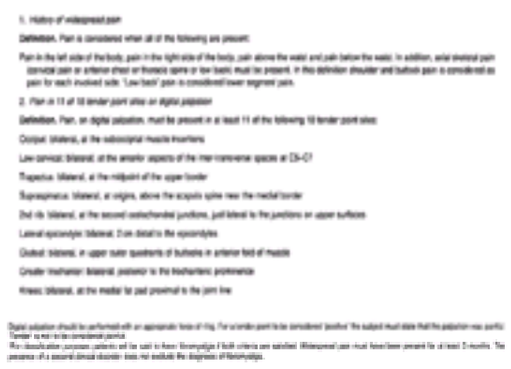


Table 8 American College of Rheumatology criteria for fibromyalgia

In drawing up the criteria for the diagnosis of fibromyalgia syndrome there has been much emphasis on the presence of tender points. It might therefore be asked whether it is possible to distinguish between fibromyalgia and a myofascial trigger-point pain syndrome simply by finding tender points in the former and trigger points in the latter. Myofascial trigger-point pain syndromes usually develop after trauma to the affected muscle or muscles, and the pain is alleviated by deactivating the hyperactive nerve endings at trigger points by one or other of a variety of methods, including the insertion of dry needles. Janet Trowell has written extensively on the subject and shown that each muscle in the body has its own specific pattern of trigger-point pain referral (Travell and Simons 1983).

Rheumatologists in general are not in the habit of looking for trigger points and those that have made a special study of fibromyalgia have never adequately addressed themselves to the question of whether some or perhaps all of the tender points in this condition could be trigger points. It would seem that most rheumatologists simply examine muscles for tender points and do so by applying pressure to them with outstretched fingers rather than by rolling the fingers firmly across them, and because of this tend to overlook the presence of any trigger point containing palpable bands.

It could be argued that tender points and trigger points must be one and the same, as the nociceptive hyperactivity that must be responsible for making the points so tender must also make the nerve endings at these sites at least potentially capable of triggering pain. Recently, it has been suggested that in patients with fibromyalgia, trigger points and tender points many represent the same abnormality. There are no longitudinal studies of myofascial pain to determine if a subset of patients evolves into a more characteristic fibromyalgia syndrome.

Aggravating factors for fibromyalgia include exhaustion, lack of fitness, disturbed sleep (which may affect the way pain signals are modulated or transmitted), and anxiety. The following are commonly associated: tension headaches, paraesthesiae, subjective swelling of joints, and primary dysmenorrhoea.

In 1975, Smythe and Moldofsky recorded electromyograms and electroencephalograms in 10 patients suffering from so-called 'fibrositis' (Moldofsky *et al*. 1975). The electroencephalographic changes showed that each patient had a disturbance of stage IV, non-rapid eye movement (**REM**) sleep. This disturbance is brought about by the rapid, 8 to 10 cycles/s a-rhythm, normally found in REM sleep, intruding into the usual slow, 1 to 2 cycles/s d-rhythm of stage IV deep sleep. Sleep disturbances have subsequently been reported in 60 to 90 per cent of patients with fibromyalgia, but are also recognized as not specific, seen in many other conditions, and in as many as 15 per cent of healthy individuals (Scheuler *et al*. 1988).

A further study in healthy normal volunteers showed that the same a-intrusion into the d-wave pattern of deep sleep could be produced experimentally by disturbing the volunteers' sleep by hand contact or a buzzer (Scheuler *et al*. 1988). When this intrusion occurred the sleepers developed general musculoskeletal pain and tenderness similar to that seen in 'fibrositis'. After a few nights of normal sleep the pain disappeared. These studies led to increasing recognition of fibrositis and the concept of fibrositis as we understand it today.

Some patients have no difficulty in getting off to sleep, consider that they sleep soundly, and yet awake with a general feeling of tiredness, fatigue, and general stiffness. Others experience this feeling after a light, restless sleep. It is not known whether changes in sleep physiology in fibromyalgia are the primary disturbance, with the muscle pains occurring secondary to these, or whether similar electroencephalographic changes develop as a secondary event in those whose sleep is disturbed as a result of pain. There is still insufficient evidence to link fatigue with specific sleep disturbances.

Psychological disturbances are seen only in the minority of people with fibromyalgia. However, emotional upsets and stress seem capable of bringing on the symptoms. Case–control studies on patients with fibromyalgia and matched controls with rheumatic diseases have shown a higher prevalence in fibromyalgia of depression, sexual and physical abuse, and eating disorders preceding the onset of the disease. Psychological stresses are frequent before its onset in adolescents.

There have been studies suggesting a deficiency of the neurotransmitter serotonin in fibromyalgia, with lower concentrations of its metabolites in cerebrospinal fluid. However, if serotonin deficiency was central to the pathophysiology of fibromyalgia, we would expect improvement on treatment with serotonergic drugs such as fluoxetine, which does not appear to be the case. Endorphins and substance P are other neurotransmitters that have been extensively studied in fibromyalgia. Endorphins were found to be normal, whilst substance P was increased in cerebrospinal fluid; however, healthy controls were used, which prevents firm conclusions (Russell *et al*. 1994).

Treatment

Whether we label sufferers as having fibromyalgia or not, there is no doubt that many have diffuse pain, sleep disturbances, fatigue, and tender points. There is no specific treatment and the prognosis is poor, but individual patients may be helped, by an adequate explanation of the condition, to learn to live better with it and so avoid further unnecessary investigations and drug treatments.

At present, tricyclic antidepressants and aerobic exercises are the treatments that have been most extensively studied. Both have a moderate degree of benefit. Cognitive behavioural approaches and multidisciplinary treatment programmes have also been used in an attempt to help patients gain control over their symptoms, but both are time-consuming and expensive and require fuller study. Aspects of management are summarized in Table 9 and of treatment in Table 10.

Educate patient

Educate patient's family

Keep investigations to minimum and stop ineffective drug therapy

Use interventions to improve sleep disturbance and improve aerobic fitness

Table 9 Principal strategies for management of fibromyalgia

Low-dose amitriptyline:
- Initially give 25 mg at night

Graded aerobic exercise regimen:
- Individualized for patient
- Encourage frequent but small amounts
- Encourage continuation despite pain
- Set increasing weekly targets

Coping strategies:
- Behavioural therapy
- Yoga

Table 10 Treatment of fibromyalgia

As many patients have suffered disappointment, blows to self-confidence and esteem, together with pain and exhaustion, they often find it hard to believe they can recover; they therefore require much encouragement from doctor and therapist. The absence of inflammation in soft tissue is now accepted. This explains the lack of efficacy of non-steroidal anti-inflammatory drugs and corticosteroids. Patients with fibromyalgia are aerobically unfit. There is no evidence to suggest that they have any primary muscle abnormality nor do they suffer from defects in energy metabolism. Aerobic fitness training unfortunately only produces slight benefit.

Tricyclic antidepressants in small doses of 10 to 30 mg are often helpful in improving the quality of sleep, in reducing morning stiffness, and in alleviating pain. However, overall the prognosis is poor and the condition tends to take a protracted course (Felson and Goldenberg 1986).

Conclusions

The opponents of the concept of fibromyalgia consider it a 'non-disease' and suggest that the label is likely to create a population of 'worried well', with adverse social and psychological consequences. The Australian experience with the diagnosis of repetitive strain injury warns us of the possibility of negative disease labelling. The converse argument is that the term fibromyalgia provides the patient with a structure for understanding the condition. Many patients have had numerous diagnoses and investigations, and the average duration of symptoms before the diagnosis of fibromyalgia is made is 5 years. Once it is made and explained, anxiety and frustration often disappear.

Opponents of the fibromyalgia label argue that by labelling a non-disease doctors are legitimizing patients' sickness behaviour. However, the converse is that the term fibromyalgia, put in its proper perspective, provides a health-care professional with the basis to recommend activity rather than inactivity, work modification rather than

work termination, and coping strategies.

The clinical syndrome that we now call fibromyalgia has been present for a long time. The facts that it is poorly understood and there are no objective physical and laboratory abnormalities do not mean that the patient is not ill or in distress. Physical and emotional pain may be disabling. We must guide patients away from sickness behaviour; we must also focus on the complicated psychosocial and biological factors that distinguish individuals who cope well with their symptoms from those who remain disabled in chronic pain (Goldenberg 1995). Physicians have taken an entity that existed for centuries, given it a new name, and created a major health issue by elevating it in importance and suggesting it deserves disability coverage (Carette 1995).

Hypermobility

The term 'hypermobility syndrome' is used to describe a common disorder in which seemingly otherwise healthy individuals present with articular and/or spinal symptoms for which no explanation is forthcoming, other than their joint hypermobility. The degree of hypermobility relates to the degree of ligamentous laxity. These patients are susceptible to torn muscles and ligaments, and are also predisposed to traumatic and degenerative changes of joints, and prolapsed discs. Dislocation of joints is more likely to occur. Persistent low back pain in the absence of an identifiable anatomical lesion in hypermobile patients is called 'the loose back syndrome'.

Normal individuals have a wide range of joint mobility, apart from the influence of age, sex, and race. Hypermobility diminishes steadily throughout childhood and more slowly during adulthood. Females show a greater range of joint movement than males. Generalized ligamentous laxity, the prerequisite of joint hypermobility, is seen in about 10 per cent of healthy individuals; the majority suffer no ill effects. The diagnosis of joint hypermobility depends on the ability to perform a series of passive joint manoeuvres. The Beighton scale (Beighton *et al*. 1989) is the measurement of choice (Table 11); the maximum score is 9.

Scoring 1 point on each side
Passive dorsiflexion of the fifth metacarpophalangeal joint to 90°
Apposition of the thumb to the flexor aspect of the forearm
Hyperextension of the elbow beyond 10°
Hyperextension of the knee beyond 10°
Scoring 1 point
Forward trunk flexion placing hands flat on floor with knees extended

Table 11 The 9-point Beighton scoring system for joint hypermobility

The diagnosis is commonly missed, for clinicians are trained to look for loss of range of joint and spinal motion and so fail to recognize an increase in range. Hypermobility syndrome has to be distinguished from the less benign heritable disorders of connective tissues such as the Marfan and Ehlers–Danlos syndromes and osteogenesis imperfecta, with which it shares a number of common features. This distinction is not always easy. Echocardiography, ophthalmic examination, and genetic studies may be required to define the phenotype in individual families.

Tissues that rely for their structural integrity on the tensile strength of collagen may be affected. The skin may be soft and develop striations. An association between mitral-valve prolapse, aortic incompetence, and hypermobility has been reported. There is an increased incidence of abdominal hernia, and rectal and uterine prolapse. Bone fragility may also be present and stress fractures can occur.

Articular features

The clinical effects of hypermobility depend on its degree and are irrespective of its cause. Patients present with a wide variety of traumatic and overuse lesions including joint or tendon-sheath synovitis, friction lesions at insertions of tendons or ligaments, rotator-cuff lesions, or back pain. Others suffer the effects of joint instability, such as flat feet and recurrent dislocation or subluxation. Arthralgia can occur and tends to improve over the years as the joints lose some of their hypermobility. There is also presumptive evidence to suggest that premature osteoarthritis may be a direct consequence of joint hypermobility.

There is a female preponderance of 85 per cent in this syndrome, which reflects the greater laxity of their joints. All patients have a varied pattern of locomotor disorders and it has been suggested that people with hypermobility do not present with different rheumatological problems from other patients; but that they have them in greater variety.

Hypermobility in healthy individuals should not be considered as just a liability. It seems to be an important selection factor enabling them to compete successfully in athletics, music, and ballet. Since inherent joint laxity enables the dancer to achieve impressive ranges of joint movement without effort, hypermobile individuals tend to be selected for ballet schools, upon which their connective tissue abnormality ceases to be an asset and becomes a liability.

Patients with hypermobility syndrome present with an unusual collection of seemingly unconnected locomotor symptoms. Unless the clinical examination is thorough and looks for increased range of joint and spinal movement, the syndrome will be overlooked. It is a common condition and patients are often pleased that someone can explain their symptoms.

Aetiology

Studies of families have provided evidence for a dominant mode of inheritance. Within individual families, females are more frequently and more severely affected than males. Females tend to present with arthralgia and mitral-valve prolapse, while males tend to develop dislocations and back pain (Child 1986). The mitral-valve prolapse is usually mild and has no haemodynamic significance.

Skin biopsies show abnormalities in the architecture of the collagen bundles. The normal gradation between the coarser deep and the fine superficial bundles is lost, and a more uniform deposition of the fine bundles prevails throughout the dermis. At the biochemical level there are raised ratios of collagen type III:types II + I, indicating a significant imbalance in the two major collagen types present.

Treatment

Management (Table 12) must first include an explanation of the symptoms. Patients need to know they have a condition that doctors recognize and that they do not have Marfan syndrome with its less favourable prognosis.

1. Discuss condition and reassure that pain is not imaginary
2. Inform that although children may inherit joint hypermobility they will not necessarily have symptoms
3. Patients with mitral-valve prolapse require antibiotic prophylaxis
4. Anti-inflammatory drugs no more effective than analgesics
5. Stretching exercises aim to restore movement into hypermobile range
   Conventional physiotherapy disappointing
6. Patients should try to avoid activities that provoke or aggravate symptoms

Table 12 Management of hypermobility

Symptoms arising in unstable, weight-bearing joints may be relieved by an appropriate exercise routine, in an attempt to allow muscle to compensate for ligamentous laxity. Many of the complications of hypermobility (osteoarthrosis, ligamentous or muscle tears) are treated along conventional lines. However, the arthralgia and the 'loose back syndrome' can be difficult to treat, apart from noting aggravating factors and trying to avoid them. These pains respond poorly to non-steroidal anti-inflammatory drugs and conventional physiotherapy.

Localized soft-tissue lesions

The painful shoulder

More than 90 per cent of lesions causing the painful shoulder result from extracapsular soft-tissue conditions (Table 13). The glenohumeral joint is a multiaxial joint that permits the greatest freedom of movement of any joint in the body, allowing placement of the hand for optimal function, although this is at the expense of stability. Ligamentous support is important in maintaining the stability of the shoulder joints, and muscles act as prime movers at the shoulders as well as providing some dynamic stability at the glenohumeral joint.

Rotator-cuff lesions

Tendinitis:
- Supraspinatus
- Acute calcific supraspinatus
- Infraspinatus and subscapularis

Rupture:
- Partial
- Complete

Bicipital syndromes

Tenosynovitis of long head

Rupture of long head

Subacromial bursitis

Usually secondary to adjacent pathology

Frozen shoulder (adhesive capsulitis)

Shoulder–hand syndrome and referred pain

Table 13 The painful shoulder

Shoulder pain may be seen in association with several medical conditions and may be referred from cervical, thoracic, or abdominal sources. The mechanism of injury may also help in making a diagnosis. A fall on to an outstretched arm can give rise to glenohumeral instability in the younger person or a rotator-cuff tear in the elderly person. A fall on to the point of the shoulder may result in injury to the rotator cuff or acromioclavicular joint. Throwing injuries tend to stress the capsule and ligaments of the glenohumeral joint, and can also give rise to rotator-cuff or bicipital tendinitis.

Complaints of shoulder pain are frequently related to occupation. Bjelle *et al*. (1979) have documented this association: they monitored the medical departments of six heavy industries and found that approximately one-third of visits were for non-traumatic musculoskeletal complaints, of which 30 to 50 per cent were neck–shoulder complaints.

Hagberg (1981a) and Hagberg (1981b) studied the relation between task demand and shoulder pain, and assessed endurance and fatigue. It was found that if the arm is held abducted and in forward flexion at a right angle, the supraspinatus and upper trapezius demonstrate fatigue within 5 min. With a variable workload and repetitive forceful shoulder flexion, the lower trapezius fatigues and there is associated discomfort in the lower neck for as long as 24 h after exertion. Workers were also filmed to gain objective estimates of workload. These physiological insights suggest task modifications that might lead to more comfortable working patterns. Job rotation, shorter exposure times, and job sharing may all lessen shoulder discomfort. These studies suggest that shoulder or neck disorders in industrial workers are multifactorial in origin and that pre-existing disease is important among the causative factors. Workers with chronic shoulder pain were significantly older and women were at higher risk than men of developing shoulder complaints.

A pain history is essential as the location and type of pain varies between conditions. Pain referred from the cervical spine is often maximal over the suprascapular region, with associated paraesthesias or pain in the upper limb. Acromioclavicular and sternoclavicular pain is usually well localized in the involved joint. Pain from the rotator cuff is usually felt at the outer aspect of the upper arm. Capsulitis tends to give rise to an intense aching deep in the shoulder. Night pain tends to be of two main types, either a sharp pain associated with movement indicative of a rotator-cuff tendinitis or acromioclavicular pathology, or pain of a deep, constant aching more suggestive of capsulitis or a rotator-cuff tear.

Examination should include close inspection for deformity, muscle wasting, and abnormalities of scapulohumeral movement. Palpation should assess the presence of tenderness, swelling, and instability as well as trigger points. Active and passive ranges of motion of both shoulders should be assessed in the planes of abduction, forward flexion, and external rotation, both with the arm by the side and at 90° abduction (Table 14).

Lesion	Painful arc	Pain increased by:
Supraspinatus tendinitis, calcific deposit or incomplete tear	Yes	Resisted abduction
Infraspinatus tendinitis	Yes	Resisted external rotation
Acromioclavicular joint disease	Yes; pain begins later in abduction (not below 90°) and increases as full elevation is reached	Local palpation; resisted adduction
Subscapularis	No	Resisted internal rotation
Bicipital tendinitis	No	Resisted flexion and supination of the elbow Tender bicipital groove

Table 14 Clinical features of some extracapsular shoulder lesions

Periarticular conditions affecting the shoulder can be loosely grouped into those with and without capsulitis. If there is no capsular involvement, then passive joint movement is largely unaffected whereas active movement may be limited by pain or weakness. In capsulitis there is a generalized restriction of movement (Dalton 1989).

Rotator-cuff disorders

As outlined above, these range from mild tendinitis following an episode of glenohumeral instability in the young patient to a complete tear in an older patient. The cuff is subjected to stresses when the arm is in the raised position, and impingement can occur as the supraspinatus tendon is compressed between the humeral head and the overlying acromion, coracoacromial ligament, and the inferior border of the acromioclavicular joint. As the rotator cuff becomes inflamed and thin its function as a depressor of the humeral head is compromised, and migration of the humeral head due to the unopposed action of the deltoid gives rise to further impingement. In the degenerative cuff this can result in a cuff arthropathy with degenerative changes in the subacromial and glenohumeral joints (Neer 1983).

Rotator-cuff tendinitis

While any of the tendons of the rotator cuff may be affected by tendinitis, it most commonly affects the supraspinatus portion of the cuff close to its insertion to the humeral head. Overuse, with resultant wear and tear and relative avascularity, may be important in inducing degeneration of the tendon.

In the young adult, rotator-cuff tendinitis usually presents acutely after an activity such as throwing. In middle-aged individuals the onset is more gradual, reflecting the underlying chronic change in the tendon, with pain aggravated by movement into abduction on elevation. Pain at night occurs when lying on the affected side. Active movements may be restricted by pain, but passive range is usually maintained and in the more chronic cases a secondary capsulitis may further restrict the shoulder movement.

Examination shows a painful arc of abduction, usually occurring between 70 and 120°, and then when lowering the shoulder there is often a 'catch' of pain as impingement occurs. Passive movement is usually full and pain-free if there is adequate muscle relaxation. In the older patient there is often involvement of the acromioclavicular joint.

Pain felt in the upper arm may suggest involvement of the cervical root, although this is more likely to present with upper trapezius or suprascapular pain referred down into the arm. Pain on active arm movement and impingement testing helps in confirming a rotator-cuff tendinitis, and the preservation of passive range of movement helps differentiate a tendinitis from a capsulitis. In patients under the age of 25 years, tendinitis is usually due to an underlying instability, which can be confirmed on examination.

Some cases of supraspinatus tendinitis are associated with calcific deposits visible on radiographs (Fig. 3). The exact mechanism responsible for the deposition of the calcium hydroxyapatite crystals in the tendon is unclear (Uthoff and Sarkar 1989). The deposits may remain asymptomatic or produce chronic symptoms with nagging discomfort in the region of the affected tendon. The crystals may also be extravasated into the subacromial bursa, causing acute bursitis with intense shoulder pain, loss of movement, severe tenderness, swelling, and muscle spasm. Fever, sweating, and other systemic symptoms may be present, mimicking gout or septic arthritis.

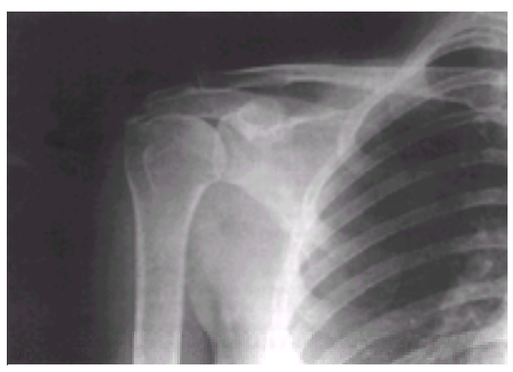


Fig. 3 Radiograph showing calcification of the supraspinatus tendon in the shoulder.

Infraspinatus and subscapularis lesions

These are less common and do not tend to calcify. While the symptoms are usually similar to those of supraspinatus tendinitis, pain induced by resisted external or internal rotation (Table 14) usually allows the correct diagnosis to be made.

Management

This is often difficult because of the patient's continuing participation in aggravating activities. Rest is necessary to prevent the condition becoming chronic. Initial treatment should aim to reduce inflammation with non-steroidal anti-inflammatory drugs and physical methods such as ultrasound. When there is no response, a subacromial injection of corticosteroid can be used.

In addition to improving pain, treatment should also aim to restore the range of movement and the normal biomechanics of shoulder movement, particularly a normal scapulohumeral rhythm. Then a strengthening programme of exercises should be given, with particular attention to restoring the rotator-cuff muscles to their function of stabilizing and depressing the humeral head.

The younger patient with instability requires a rehabilitation exercise programme and rarely requires injection. The older patient with a degenerative rotator cuff and associated pathology of the acromioclavicular joint may not be helped without surgery. The major indication for surgery is pain and, in the presence of an intact rotator cuff, failure to respond to conservative treatment within 1 year. Surgery consists of subacromial decompression by resecting the coracoacromial ligament, and an anterior acromioplasty. If there is pathology of the acromioclavicular joint, this also requires attention. Full-thickness tears can be demonstrated by arthrography, ultrasonography, or magnetic resonance imaging. Arthroscopy is also useful.

Tears of the rotator cuff

Tears of the rotator cuff may be acute or chronic, partial or full thickness, and are most common in patients over the age of 50 years. They usually occur at the enthesis; degenerative changes and alterations in collagen composition may explain the susceptibility to tearing at this site (Kumagai *et al*. 1994). The mechanism of injury is a fall on to the outstretched arm. Partial tears may occur after trauma at any age and may present as a painful arc syndrome; full active range of movement may be preserved. Complete tears are associated with marked weakness of abduction and external rotation, or of flexion, depending on the tendon involved; passive movement is full. Pain can be severe and there may be no history of injury. Partial tears may be difficult to differentiate from tendinitis, but there is often an inability to maintain the arm in abduction when lowering it from a raised position. Atrophy of supra- and infraspinatus muscles often follows, and weakness of external rotation reflects the size of the tear. Rupture of the tendon of the long head of the biceps is frequently associated with chronic disease of the rotator cuff.

Chronic, full-thickness tears are found in up to 25 per cent of cadavers at autopsy. Cuff arthropathy occurs when there is superior migration of the humeral head against the under surface of the acromion, which happens when the incompetent rotator cuff fails to stabilize and depress the humeral head and therefore counteract the pull of the deltoid.

Radiographs show degeneration and sclerosis of the rotator cuff, and cystic changes at the greater tuberosity. Osteophytes may be present at the inferior margin of

the acromioclavicular joint. The subacromial space is narrowed if there is a complete tear.

Treatment

Partial tears of the rotator cuff should initially be managed conservatively in a similar way to tendinitis of the cuff, although corticosteroid injections should be avoided within 6 weeks of injury. Acute ruptures in the young or active patient should have early surgery. In the older or less active patient it is usual to give a trial of conservative treatment. The treatment of large tears is controversial. It now seems that adequate decompression combined with anterior acromioplasty and debridement is preferred over extensive surgical procedures (Burkhart *et al*. 1994). Surgery should also be considered where there is associated rupture of the biceps tendon, as these patients are more likely to develop a cuff arthropathy.

Bicipital syndromes

The biceps tendon is seldom involved in isolation and involvement usually occurs with tendinitis or impingement of the rotator cuff, or with glenohumeral instability. As with rotator-cuff tendinitis, the young patient should be assessed for joint instability. Acute rupture is usually a result of overuse such as in weight-lifting.

Pain is felt over the anterior aspect of the shoulder, with localized tenderness over the tendon in the bicipital groove. Pain may be reproduced on resisted elbow flexion, although various provocation tests are inconsistently positive. Rupture of the tendon leads to a characteristic increase in the muscle belly of the biceps on resisted elbow flexion.

The tendon of the long head of the biceps tendon may be involved at its attachment to the superior glenoid labrum or as it runs in the bicipital groove.

Treatment

This consists of rest, physical procedures including laser therapy, and non-steroidal anti-inflammatory drugs. Care should be taken not to inject the tendon with corticosteroid. It is essential to assess whether the tendinitis is primary or secondary to pathology or instability of the rotator cuff, as failure to treat these causes will lead to a recurrence.

In chronic resistant cases, surgery may be required; either subacromial decompression or tenodesis, depending on the cause. Rupture of the tendon is usually treated conservatively, except in the occasional young patient where upper-arm strength is essential for their sport.

Subacromial bursitis

In most cases the bursa becomes inflamed as part of the impingement process and coexists with an underlying rotator-cuff tendinitis. In chronic cases the bursa becomes fibrotic and surgical excision or debridement may be necessary.

Acute traumatic bursitis may be differentiated from rotator-cuff tendinitis by the presence of increased tenderness and fluid at the subacromial space. Rest and physical treatment usually result in resolution.

Acromioclavicular syndrome

This is seen as an acute condition following trauma, usually after falls or contact sports. Septic arthritis can also rarely occur. Disruption of the joint may be seen in association with fractures of the clavicle. More common are injuries to the joint itself, which are graded I to III depending on the degree of disruption of the joint capsule and supporting ligaments.

Pain is well localized to the top of the shoulder, and the joint is often tender and swollen. Abduction is often limited. In complete disruption of the joint a visible step deformity is seen.

Treatment

For most injuries treatment is largely symptomatic with analgesics and a sling for a few days. Controversy exists over the management of complete dislocation of the joint. Surgical stabilization is usually unnecessary and has a high failure rate.

Patients with persistent pain at the acromioclavicular joint should be treated with an intra-articular corticosteroid and non-steroidal anti-inflammatory drugs. Long-term treatment is as for osteoarthritis of the joint.

Frozen shoulder (adhesive capsulitis)

This condition may occur spontaneously but can follow other lesions of the rotator cuff or trauma. In addition, conditions that produce pain (e.g. the referred pain of myocardial infarction) or immobility (e.g. from stroke or polymyalgia rheumatica) of the shoulder or arm can predispose to the development of a frozen shoulder. Early arthrography may reveal a small, shrunken, thickened capsule and some insist that these changes need to be present to make a diagnosis of adhesive capsulitis (Neviaser and Neviaser 1987).

Primary capsulitis can be defined as a condition of unknown aetiology in which there is a painful, global restriction of active and passive glenohumeral movement in all planes, in the absence of joint degeneration. Underlying conditions associated with this condition include diabetes mellitus, thyroid disease, and pulmonary disorders such as tuberculosis (Risk and Pinals 1982). Until recently, few pathological studies had been performed; Bunker and Anthony (1995) found features compatible with fibromatosis.

Onset under the age of 40 is uncommon; frozen shoulder is slightly more common in women than men and involvement of the contralateral shoulder occurs in up to 17 per cent of patients over the subsequent 5 years. There is severe night pain and pain on movement; improvement is gradual and spontaneous but may take 1 to 3 years. The extent of the recovery is variable and a clinically detectable limitation of shoulder movement can be seen in up to 15 per cent of patients (Bulgen *et al*. 1984).

Diagnosis is largely on clinical grounds as few abnormalities are found on investigation. Differentiation from rotator-cuff tendinitis is possible as there is global restriction of passive movement rather than simply loss of abduction and flexion.

Patients with even minor degrees of frozen shoulder may develop a secondary, reflex-sympathetic dystrophy syndrome—the shoulder–hand syndrome. This consists of a symptom complex characterized by an immobile painful shoulder associated with a swollen, painful, cold, and dystrophic-looking hand. The lesion may progress until the patient is left with a painful, tender, useless hand.

Treatment

The emphasis in the early stages should be on pain relief and the prevention of joint restriction. Analgesics are more effective than non-steroidal anti-inflammatory drugs; physiotherapy with interferential can reduce pain and reduce muscle spasm, and exercises within the limits of pain to maintain joint mobility are encouraged.

Intra-articular corticosteroid injections have improved pain and range of movement, although no long-term benefit has been shown. Oral corticosteroids have improved pain but not range of movement. No treatment has been consistently shown to affect rate of recovery or limit restriction of movement.

Distension of the subscapular bursa by arthrography and forced mobilization can result in immediate pain relief (Nobuhara *et al*. 1994). Manipulation under anaesthetic is sometimes advocated to restore joint movement by rupturing the inferior capsule. Care is required and very active early rehabilitation is necessary in

the postmanipulation period to maintain joint mobility. If this treatment is contemplated, it is recommended that it be reserved for the adhesive stage and not the early, painful phase.

This condition is very painful and disability prolonged. It is essential that the patient understands this and for their expectations to be appropriate.

Glenoid labrum tears

There has been considerable recent interest in tears of the glenoid labrum. The labrum is a fibrocartilaginous and fibrous structure attached to the glenoid that supports the stability of the joint by serving as an attachment site for the glenohumeral ligaments. Clinical features of tears include shoulder pain, especially with overhead activities, and snapping and catching on movement. Because the biceps tendon originates at the superioglenoid with its labral attachment, pain may be present in the anterior shoulder on resisted forward flexion with the elbow extended and forearm supinated (Speed's test) (Payne 1994).

The painful elbow

Pain round the elbow is commonly caused by soft-tissue lesions (Table 15), but care must be taken to exclude pain referred from the cervical spine, brachial plexus, and shoulder, and examination of the neck and shoulder is important in the assessment of elbow pain.

Humeral epicondylitis
Lateral:
Tennis elbow
Medial:
Golfer's elbow
Biceps and triceps tendinitis
Tear of brachialis muscle
Olecranon bursitis
Traumatic:
Student's elbow
Secondary to inflammatory joint disease
Friction neuritis of ulnar nerve

Table 15 The painful elbow

Humeral epicondylitis

Lateral involvement (tennis elbow) is much more common than medial involvement (golfer's elbow). In spite of their sporting connotations, both occur more frequently in those performing repetitive movements with their arms, for example operating machinery, using a screwdriver or doing housework, although some 40 per cent of tennis players do suffer from tennis elbow.

In tennis elbow there is pain over the lateral aspect of the elbow with localized tenderness near the lateral epicondyle. In general, tenoperiosteal lesions can be separated from intra-articular conditions because movement of the related joint is full, there is tenderness at the tenoperiosteal junction, and contraction of the muscle attached to the affected tendon reproduces the pain. About 1 to 3 per cent of adults are affected by it; usually they are aged between 40 and 60 years, the dominant arm being most frequently affected (Hamilton 1986). Resisted dorsiflexion of the wrist exacerbates the pain with the elbow in extension, and there is a reduction in grip strength. Tenderness is usually maximal over the lateral epicondyle. Thermography usually shows a discrete 'hot spot' on the side of the elbow. The range of movement of the elbow is usually normal. In golfer's elbow there is a tender spot at the medial epicondyle owing to a lesion of the common flexor tendon, and pain is induced by flexing the wrist against resistance with the elbow fully extended.

Pain on resisted flexion alone indicates the rarer brachialis muscle lesion, with pain and tenderness that is less well localized and found behind the biceps tendon. Although uncommon, such a tear is particularly prone to develop myositis ossificans, which in the early stages produces a warm, firm mass that can be mistaken for a tumour. It is unusual to have a lesion at the site of the triceps insertion into the olecranon. Lesions usually occur in this tendon at the musculotendinous junction. Ligamentous lesions do not usually occur in isolation; they are usually associated with traumatic synovitis of the joint.

Radial tunnel syndrome or compression of the posterior interosseous nerve can produce lateral elbow and forearm pain. These nerve entrapments are due to compressive lesions caused by fibrous bands in front of the radial head, or an abnormal origin of extensor carpi radialis brevis.

It is thought that the majority of patients have a musculotendinous lesion of the common extensor tendon at the attachment to the lateral epicondyle, especially that portion derived from extensor carpi radialis brevis. Macroscopic tears in the extensor tendon are occasionally found at operation, but these may have been caused by repeated steroid injections. Some cases show mesenchymal transformation suggestive of a chronic traction effect. Age is an important factor since lateral epicondylitis is uncommon before 30 years. Sometimes there may be bilateral involvement, either due to increased stress placed on the unaffected arm or as a general tendency to soft-tissue lesions in that individual.

Management (Table 16)

Early/mild	Rest, splinting
	Anti-inflammatory drugs/gels
Persistent	Local corticosteroid injections
	Ultrasound
Resistant	Manipulation?
	Surgery—lateral release

Table 16 Management of lateral epicondylitis

Reduced activity may result in improvement, particularly in early cases. Corticosteroid injections have been widely used; hydrocortisone is preferable to longer-acting preparations, which may lead to skin atrophy. Approximately 90 per cent of patients respond, but there may be increased pain for up to 48 h following injection and a significant number of cases recur. The injection may be repeated once after 4 weeks.

Non-steroidal anti-inflammatory drugs are often used but there is little evidence of their efficacy. A cock-up wrist splint reduces tension on forearm extensors and may help resolution. Numerous physical forms of treatment have been used but the efficacy of most is unproven (Chard and Hazleman 1989). Ultrasound, by its ability to cross myofascial planes and concentrate near bone, has theoretical advantages; in one study the rate of relapse was less with ultrasound than after corticosteroid injection (Binder *et al*. 1985).

Up to 40 per cent of patients have recurrent symptoms and some minor discomfort can persist for years. Early treatment may improve prognosis; firm strapping of the forearm muscles just distal to the elbow joint or the use of commercial elbow splints may be helpful. Patients should be advised to avoid straining the arm for some 2 months. Graded exercises to strengthen the forearm muscles may be advised, especially in sporting lesions.

Up to 10 per cent of patients fail to respond to physiotherapy. Surgical treatment, which attempts to correct the presumed pathological changes, can be helpful. Excision of tissues around the epicondyle or removal of a synovial fringe of the radiohumeral joint are the most common procedures.

Olecranon bursitis

The superficial bursa over the olecranon process is commonly involved in rheumatoid arthritis (with nodule formation) or gout, but can also be affected by trauma or infection. In the acute stage it distends with fluid, with prominent signs of acute inflammation. If it becomes chronic the wall can be greatly thickened. As the posterior wall of the bursa is so close to the periosteum of the olecranon, pain can be felt down the border of the ulna.

Treatment

Aspiration of the bursa reduces symptoms and allows examination of the fluid to exclude infection and the presence of crystals. If no infection is suspected, local steroid injections are effective; a compressive elastic bandage may help prevent the recurrence of swelling. If recurrent, non-infective bursitis occurs then surgical excision may be necessary. The presence of infection requires appropriate antibiotic treatment and surgical drainage may be necessary.

The painful wrist and hand (Table 17)

Dupuytren's contracture
Tenosynovitis—including De Quervain's
Stenosing tenovaginitis
Rupture of tendons
Ganglion
Median nerve compression—carpal tunnel syndrome
Ulnar nerve compression—in Guyon's canal

Table 17 The painful wrist and hand

Both seropositive and seronegative arthritides have a predilection for inflammatory involvement of the synovial structures of the tendons and joints in the wrist and hand. Tenosynovitis denotes an inflammation of the synovial lining of the tendon sheath, usually accompanied by inflammation of the contained tendon. The clinical manifestations are pain, tenderness, and swelling, with 'crepitus' that is palpable when the tendon moves within the inflamed sheath.

The flexor tendon sheaths enclose the tendons of flexor digitorum superficialis and profundus to their insertions on the middle and distal phalanges, respectively. The tendon sheath of the thumb flexor pollicis longus extends proximally to the carpal tunnel. The flexor sheath of the little finger is often continuous with the common flexor tendon sheath in the wrist. Segmental condensations in the digital flexor sheaths prevent bowstringing of the tendons.

Stenosing tenovaginitis

Stenosing tenovaginitis is primarily a fibrosis of the tendon sheaths with intrathecal narrowing of the lumen, especially involving sites near bony prominences where tendons pass through fibrous rings. This more commonly affects the flexor than extensor tendons in the hand. If a fibrous nodule develops in the flexor tendons, a 'trigger finger' can result, which further limits function. The finger often locks in flexion. Extension can be forced with difficulty and is often painful. Palpation during muscle action may reveal a mobile nodule within a tendon sheath of a finger or palm. Its incidence as an isolated lesion following overuse is low. The most common cause of a trigger finger or thumb is overuse from repetitive grasping activities.

Management

The management of stenosing tenovaginitis consists of modification of hand activity, gentle exercises, and non-steroidal anti-inflammatory drugs. Extension splinting of the affected digit at night prevents painful flexion during sleep. One or two corticosteroid injections into the affected tendon sheaths are effective in the majority. Surgical transection of the fibrous annular pulley of the finger or thumb flexor sheath is rarely required.

De Quervain's tenosynovitis

This common lesion, caused by repeated minor trauma, results from involvement of the tendon sheaths of abductor pollicis longus and extensor pollicis brevis. The patient complains of pain on using the thumb or wrist. Tenderness is maximal in the 'snuffbox' area between the two tendons, and there is often a visible tender swelling about the radial styloid. Pain can be elicited by forced ulnar deviation after placing the patient's thumb in the palm (Finkelstein's sign).

The tendon sheath, which is normally about 0.75 mm thick, increases in thickness three- or fourfold, and there is cellular infiltration with increased vascularity of the sheath, inflammatory proliferation of the epitendon, and expansion of part of the tendon to form a nodule.

It has been reported for years that De Quervain's is more common in assembly-line workers involved in repeated grasping movements between finger and thumb, and with rapid pronation–supination movements of the forearm, although most studies cannot demonstrate an association with occupation. It is 10 times more prevalent in women.

Treatment

The symptoms often resolve spontaneously with rest, but can be recurrent or persistent. Immobilization of the wrist and thumb by thermoplastic splinting is often helpful. Attention to hand activities with avoidance of tasks that require repetitive thumb movements or pinch grasping is important. In patients with more severe or persistent pain, one or more local corticosteroid injections are often helpful, giving relief in some 70 per cent. Surgical decompression of the first extensor compartment, with or without tenosynovectomy, is indicated in those with persistent or recurrent symptoms for more than 6 months.

Tenosynovitis and peritendinitis crepitans

By definition tenosynovitis is a disorder of the tenosynovium. It has been described after trauma in an industrial context for many years, affecting the long extensors of the fingers at the wrist and less commonly the long flexors. Tenosynovitis occurs after unusually active use of the wrist over a period of days or weeks. Tenosynovitis and peritendinitis crepitans are two distinct syndromes. Both present with pain, particularly on resisted movement, and there is usually localized swelling, tenderness, and crepitus. The distinguishing feature is the site of the lesion. In tenosynovitis the swelling and tenderness are confined to the synovial sheaths and the wrists, and respond to rest and local steroid injections. Peritendinitis crepitans presents with pain, tenderness, and swelling at the musculotendinous junction above the upper limit of the tendon sheaths in the forearm. It usually responds to rest, and to ultrasound.

Since 1947 tenosynovitis has been a United Kingdom Department of Health (and Social Security) prescribed industrial disease. It was intended to compensate manual workers suffering from tenosynovitis induced by excessively rough or arduous work. More recently, tenosynovitis has come to be used incorrectly as a generic term covering not only traumatic tenosynovitis and paratendinitis crepitans but a wide variety of non-specific aches and pains in the forearm, which trades union health-care advisers and members may often simply call 'teno'. The epidemiological evidence for the accepted overuse syndromes is substantial, although often clouded in recent years by differences in the terms used to describe these disorders and also by the emotional arguments pertaining to repetitive strain disorders (see above).

There is little doubt that the common occupational disorder known as peritendinitis crepitus is related to overuse. Biopsies were first carried out by Von Frisch in 1909; he found oedema with congestion of the peritendinous tissues, mainly at the musculotendinous junction and often around the muscle (Thompson *et al*. 1951). In more recent descriptions it is clear that local anatomical considerations are important (Williams 1977).

Dupuytren's contracture

This condition, of unknown aetiology, produces progressive thickening of the palmar fascia and causes flexion contracture predominantly affecting the ring and little fingers. It is commonly bilateral and can also involve the plantar fascia. The palm of the hand becomes indurated and lines of fibrosis with nodules and skin puckering run along the tendons, causing progressive fixed flexion of the metacarpophalangeal and proximal interphalangeal joints.

The rate of progression is variable and surgical fasciectomy should be performed only when disability is severe, but should be considered before amputation becomes the only alternative.

Management

Treatment depends entirely on the rate of progression and severity of the lesions. In patients with mild disease, local heat, stretching exercises, and the use of protective padded gloves during heavy manual tasks are often helpful. In more severe lesions with pain and inability to straighten the fingers, intralesional steroids may be beneficial. In those with progressive digital contracture of more than 30°, and functional impairment, a palmar fasciectomy with or without a skin graft is indicated. The risk of recurrence is greater in those with a family history or active bilateral disease.

Ganglia

Ganglia are tense, uni- or multilocular, cystic swellings that develop in relation to a joint capsule or tendon sheath and contain a clear, jelly-like substance. They vary in size and can be so tense that they may be mistaken for a bony swelling. They are sometimes provoked by injury or arthritis, but often occur spontaneously.

Entrapment neuropathy (carpal tunnel syndrome)

Nerve compression can occur at any site where a peripheral nerve passes through an opening in fibrous tissue or an osseofibrous canal. The clinical diagnosis, if in doubt, may be confirmed by electromyography, although the electromyogram can be normal in up to 8 per cent (Wilcox and Bilbao 1993), the results becoming abnormal only if significant demyelinization or axonal loss occurs.

Entrapment of the median nerve in the carpal tunnel at the wrist is the most common entrapment lesion. It is more frequent in women. Most cases are idiopathic or caused by unaccustomed repetitive use, but compression in the carpal tunnel may be associated with rheumatoid arthritis, myxoedema, acromegaly, pregnancy, and the contraceptive pill. Fracture, deformity, or dislocation of the carpal bones can cause similar problems. Early symptoms include painful tingling in the wrist and hands, mainly affecting the thumb, index, and middle fingers, but often symptoms are poorly localized. The pain may occasionally extend up the arm well above the wrist and often interferes with sleep. Patients try to obtain relief by hanging the affected limb out of the bed. Later, weakness (first in the abductor pollicis brevis) develops. Diminution of sensation to touch and pin-prick, usually over the palmar aspects of the distal phalanges of the index and middle fingers, may be found. Untreated, wasting of the muscles of the thenar eminence develops. Electromyography is helpful where the clinical diagnosis is in doubt. A delay in motor or sensory conduction velocity across the wrist confirms the diagnosis, although no electrical abnormality may be detected in early cases.

Treatment (Table 18)

1. Splinting–may be curative; especially if of recent onset
2. Local corticosteroid injection; second injection can be given if improvement not complete
3. Non-steroidal anti-inflammatory drugs; can be helpful if underlying inflammatory lesions such as tenosynovitis
4. Surgical release of transverse carpal ligament is definitive treatment; indicated when response to conservative measures inadequate, when there are progressive or persistent neurological changes, or when there is muscle atrophy

Table 18 Management of carpal tunnel syndrome

Conservative measures may suffice when symptoms are of short duration. Patients with progressive increases in distal motor latency times on repeated electromyograms should be considered for surgery. Most patients, however, do not need surgical decompression, and certainly not as initial treatment (Harter *et al*. 1993).

Splinting

This is simple, and if successful in relieving symptoms it often helps to confirm the diagnosis. A volar wrist splint that keeps the wrist in a vertical position is helpful at night, and splinting alone may be sufficient to relieve symptoms.

Local corticosteroid injection

This is also effective when the disease is of short duration and where there is slight muscle wasting. A second injection can be used but more frequent injections may lead to tendon rupture. Gelberman *et al*. (1980) found that the best response to injection was in patients who had had symptoms for less than a year and in whom there was no significant muscle atrophy or weakness.

Surgery

Surgical release of the transverse carpal ligament is the definitive treatment. It is indicated when response to conservative treatment is inadequate, if there are progressive or persistent neurological changes, or there is muscle atrophy. If there is also marked tenosynovitis, then a tenosynovectomy may also be necessary. Surgery relieves symptoms in about 90 per cent. Pain for more than 5 years preceding surgery indicated a poor prognosis, and diabetics do less well (Haupt *et al*. 1993).

Ulnar tunnel syndrome

Entrapment of the ulnar nerve can occur in Guyon's canal at the wrist. The most common causes are a ganglion and inflammatory joint disease. When compression occurs in the proximal portion of the ulnar tunnel the patient will have combined sensory and motor deficits, paraesthesias in the hypothenar region and the fourth and fifth fingers, as well as weakness of the intrinsic muscles of the hand. More distal lesions may present with either motor or sensory impairment. Electrodiagnostic studies are helpful in determining the site of entrapment and in defining which branches are involved.

Treatment

If conservative measures such as avoidance of trauma to the palm do not result in improvement, then surgical decompression may be necessary.

The painful hip

Soft-tissue lesions around the hip are common, particularly in sportsmen and -women. Some of the lesions are listed in Table 19.

Muscle strains:
Adductor
Iliopsoas
Rectus femoris
Rectus abdominis
Osteitis pubis
Inflamed bursae:
e.g. iliopsoas; trochanteric
Nerve entrapment:
e.g. ilioinguinal (T12–L1); lateral femoral cutaneous (meralgia paraesthetica)

Table 19 The painful hip

Meralgia paraesthetica

This is caused by entrapment of the lateral femoral cutaneous nerve, which is entirely sensory and arises from the second and third lumbar roots. It passes under the lateral aspect of the inguinal ligament just medial to the anterior superior iliac spine, the most common site of entrapment.

This condition may arise from direct pressure from belts or tight-fitting garments and may be associated with pregnancy, diabetes, and trauma. Patients complain of a burning pain and dysaesthesias in the sensory distribution of the nerve. The diagnosis can be confirmed by electrodiagnosis or a nerve block.

Treatment

It is often self-limiting, and removal of the aggravating factor is usually enough. Local injection of corticosteroid may be helpful. Surgical decompression can be carried out if the symptoms are chronic, although the results can be disappointing.

Groin pain

This occurs particularly in runners, and 5 per cent of football injuries are of the groin (Harris 1974). It may lead to prolonged disability. Pain arising from the symphysis pubis may account for about 5 per cent of hip pain. The onset of pain is usually insidious and in most cases there is no history of acute trauma. The pain is suprapubic, medial inguinal or proximal medial thigh. The pain settles with rest and returns with activity. The sensation of a 'click' on getting up from a sitting position indicates some instability at the symphysis.

On examination there is tenderness over the symphysis; adductor spasm and any discrepancy in leg length should be looked for. Any pain lasting for more than 4 weeks requires radiological assessment; radiographic changes include widening, irregularity, and sclerosis of the margins of the symphysis.

Treatment

Treatment is conservative, with strengthening of the abdominal musculature and increasing mobility at the hip joints the most important measures. Osteitis pubis is a disabling condition and its treatment may take 12 months or longer. Any predisposing factors should be identified and treated. In the acute phase, rest is essential, non-steroidal anti-inflammatory drugs are commonly prescribed, and steroid injections into the symphysis pubis may be of benefit. Results of fusing the symphysis have been disappointing—the procedure tends to place additional stresses on the hips and sacroiliac joints.

Bursitis and tendinitis

Trochanteric bursitis is the most common cause of pain about the hip joint. The soft tissues that cross the bony posterior portion of the greater trochanter are protected from it by the trochanteric bursa. Patients suffer a deep aching pain on the lateral aspect of the hip, which increases with activity. Active abduction of the hip when lying on the opposite side typically accentuates the discomfort.

The iliopectineal bursa is deep to the iliopsoas tendon over the front of the hip joint and communicates with it in 15 per cent of cases. The symptom is pain in the groin that irradiates along the anterior aspect of the thigh.

Adductor tendinitis is frequently seen in sportsmen and -women. Pain is localized to the medial aspect of the groin, and there is tenderness and pain exacerbated by movement. The differential diagnoses include tumour, stress fractures, and soft-tissue infections.

Treatment

This consists of non-steroidal anti-inflammatory drugs, local steroid injections, and physical forms of therapy including ultrasound. Only a few chronic cases require surgery.

Snapping hip

Some patients complain of hip pain associated with an audible and palpable snapping sensation of the lateral aspect of their hip. It is an uncommon syndrome noted most frequently in adolescents and young women, and it occurs in activities such as climbing or walking upstairs. The treatment is similar to that of trochanteric bursitis.

Piriformis syndrome

The piriformis is a deep external rotator of the hip. Symptoms include a deep posterior hip pain with shooting, burning symptoms, often confused with sciatica owing to

the anatomical proximity of the muscle to the sciatic nerve.

Ischiogluteal syndrome

A strain of the muscles originating from the ischial tuberosity presents as a posterior pain around the hip. Physiotherapy is the mainstay of treatment.

The painful knee and shin (Table 20)

Patellar tendinitis
Rupture of quadriceps apparatus
Apophysitis of tibial tubercle—Osgood–Schlatter disease
Bursitis:
Prepatellar
Infrapatellar
Popliteal cysts:
semimembranosus
Baker's cyst
Anserine
Infrapatellar fat-pad lesions
Medial and lateral ligament injuries
Tendon lesions
Anterior tibial syndrome

Table 20 The painful knee

Tendons, ligaments, and bursae are the soft-tissue structures around the knee from which pain may originate. When assessing the knee, the clinical history is more valuable than the findings of clinical examination. The patient's account of the symptoms and their onset will usually make it possible to decide which structures in the knee have been damaged and how seriously.

Tendinitis

The insertions of the tendons around the knee may become inflamed due to overuse or in association with an inflammatory arthropathy. It may be difficult to differentiate tendinitis from the medial ligament syndrome and bursal inflammation.

Patellar tendinitis

The ligamentum patellae may become painful and tender at its attachment to the upper or lower pole of the patella or at its distal attachment to the tibia. The condition is due to increased degeneration of the patellar tendon. Prolonged and heavy exercise, such as jogging or jumping, predisposes to this condition, which usually occurs in young adults. The majority of cases respond to rest and non-steroidal anti-inflammatory drugs. Steroid injections may be effective but should be used cautiously to avoid tendon damage; surgical exploration is occasionally necessary.

Pes anserinus tendinitis

The pes anserinus insertion may become inflamed, but this may be indistinguishable from anserine bursitis and the medial ligament syndrome. In cases of anserine tendinitis, rest is usually curative but in persistence a local injection of steroid may be required.

Popliteus tendinitis

There is pain over the lateral aspect of the joint produced by overuse and leading to inflammation of the popliteus tendon and paratendon. Symptoms usually settle with rest and non-steroidal anti-inflammatory drugs.

The biceps tendon is attached to the head of the fibula, and inflammation will cause pain around the tendon insertion, felt particularly when the knee is flexed against resistance and the area pressed firmly.

Pain in the back of the knee may arise from a hamstring injury. The pain usually comes on acutely after sudden activity without adequate warming-up.

Osgood–Schlatter disease

Apophysitis of the tibial tubercle at the insertion of the patellar tendon occurs predominantly in adolescent children and is another common cause of pain in front of the knee. The condition is more common in boys than girls; symptoms include localized tenderness and swelling at this site. Pain can be reproduced by resisted knee extension; radiographs may be abnormal and show an isolated spicule of bone at the site of the swelling. In most patients the condition improves spontaneously with rest but sometimes the pain may persist for several years, and injection, plaster immobilization, or even excision of the bony spicule may then be required. A similar, but rare, condition can occur at the lower pole of the patella (Sinding–Larsen syndrome).

Bursitis

Some bursae are in direct contact with the knee joint while others are separate. As with bursae elsewhere, acute or chronic inflammation, infection, and involvement in systemic diseases such as gout and arthritis can occur.

Acute bursitis is characterized by classical signs of acute inflammation, with intense, localized tenderness and marked restriction of movement. Chronic bursitis may follow acute episodes but is much more commonly associated with repeated trauma. The bursal lining becomes thickened and cells degenerate; adhesions, villi, and calcareous deposits eventually develop. The degree of inflammation, muscle weakness and wasting, and limitation of movement can vary widely, making diagnosis difficult.

Prepatellar bursitis (housemaid's knee)

The bursa is subcutaneous and inflammation usually results from repeated kneeling, but can follow a fall on to the patella. Infection of the prepatellar bursa gives a characteristic red, shiny appearance over the knee and is often mistaken for an infected knee joint. When chronic, 'melon seed' bodies and fibrous bands form in the thickened, enlarged bursa.

The majority of cases settle with rest and the avoidance of kneeling. Occasionally, aspiration is necessary and the fluid should be sent for microscopy and culture. Recurrent episodes of inflammation may require surgical excision of the bursa.

Infrapatellar bursitis

This small, deep bursa occupies the space between the upper part of the tibial tuberosity and ligamentum patellae, separated from the synovium by a fat pad. When inflamed, the fluid obliterates the depression on each side of the ligament, the fluctuant swelling being most marked when the knee is actively extended. The treatment is the same as for prepatellar bursitis.

Popliteal cysts (Baker's cysts)

Synovial cysts in the popliteal fossa are usually referred to as Baker's cysts. They may arise from the semimembranosus bursa when they communicate with the knee joint or from a posterior rupture of the knee joint capsule. Cysts are common in children and are of no serious significance, although in young adults there may be quite severe pain after exercise. The most common course is gradual resolution. If the cyst does not resolve, it may extend into the calf, or burst, with fluid tracking down the fascial planes of the calf, mimicking an acute deep venous thrombosis. The signs are similar but differentiation can usually be made from the history of swelling in the popliteal fossa and the sudden onset of pain. Arthrography of the knee joint establishes the diagnosis but must be performed early after the onset of symptoms as the leak may seal off after a few days. CT scanning and ultrasonography can also aid diagnosis.

Treatment in children should be conservative as they usually settle spontaneously and there is a high recurrence rate after surgery. Treatment in adults should be aimed at the underlying pathology. Aspiration and injection of steroid can be helpful but reaccumulation is frequent. A large cyst that is causing pain or interfering with knee movement should be excised.

Anserine bursitis

The tendons of sartorius, gracilis, and semitendinosus all cross the lower medial side of the femur and attach by a common tendon beneath which lies the bursa. Pain is felt diffusely on the medial aspect of the knee, but tenderness can usually be localized to the area of the bursa. Knee range is normal but contracting the hamstrings induces pain. The bursa is often inflamed in elderly, obese women with a valgus deformity and is also a common source of pain in inexperienced joggers, often being mistaken for joint injury.

The conditions of anserine bursitis, medial ligament syndrome, and pes anserinus tendinitis may be difficult to separate. Treatment is initially rest and non-steroidal anti-inflammatory drugs. In persistent cases, local steroid injections are usually curative.

Ligament injuries

If a patient has injured the knee in such a way that the structures on the medial side have been stressed, damage to the medial ligament must be suspected. Clinical examination may reveal tenderness localized to the upper and lower attachment of the ligament to the femur or tibia, respectively. The diagnosis can be confirmed by stressing the medial ligament; this will reproduce the pain or demonstrate ligament laxity. It is more common in females. Treatment is initially conservative with rest and heat. Persistent symptoms respond to a local steroid injection.

Pellegrini–Stieda disease

Pain over the femoral insertion follows injury to the medial collateral ligament. The area is tender and pain occurs on stressing the ligament. There is calcification at the ligament insertion, which probably forms in a haematoma following injury. The condition is usually self-limiting with rest but local steroid injections relieve persistent cases.

Anterior knee pain

This is common in adolescence, and the term should be applied to patients in which no specific cause can be found. Conditions such as patellar malalignment, Osgood–Schlatter disease, and trauma should be excluded. It is usually bilateral and occurs in adolescent girls. The pain is often present at night and worse after sport.

Many patients have poor quadriceps strength but no demonstrable patella instability or malalignment. Some patients have tight hamstrings and wearing high-heeled shoes may also cause the symptoms. The condition is usually self-limiting.

Treatment is conservative and involves isometric quadriceps-strengthening exercises and hamstring stretching exercises. Wearing high-heeled shoes should be avoided.

External-compartment syndromes

Exercise-induced pain of the lower extremity may prove difficult to diagnose and treat, and can be due to several conditions including stress fracture, periostitis, claudication from entrapment of the popliteal artery, peripheral nerve entrapment, and the compartment syndromes.

Exercise increases compartment width and volume by 20 per cent and patients with this syndrome have higher pressures before, during, and after exercise (Clinton and Solcher 1994). Patients complain of pain over the affected compartment that increases during exercise, is relieved by rest, and is occasionally associated with neurological symptoms. The diagnosis is confirmed by measuring intracompartmental pressure during exercise. Surgical release by fasciotomy usually relieves symptoms; it is most successful in the anterior compartment.

Anterior tibial syndrome

This condition consists of severe pain in the anterior aspect of the leg, associated with foot drop. occurring after exercise, and relieved by rest. It is thought to be the result of a tight fascia compressing the muscles in the anterior tibial compartment of the leg. There are many predisposing factors, of which exercise is the most common. Surgery gives relief.

The painful heel (Table 21)

Pain behind the heel
Achilles tendon lesions:
Rupture—partial/total
Central core degeneration
Ossification
Peritendinitis (Achilles tendinitis)
Bursitis:
Subachilles (retrocalcaneal)
Subcutaneous (postcalcaneal)
Retrocalcaneal apophysitis (Sever's disease)
Pain under the heel
Plantar fasciitis
Calcaneal apophysitis and spurs
Tender heel pad

Table 21 The painful heel

Painful disorders of the ankle and foot are common. Many are due to inappropriate footwear, foot deformities, and weak muscles. The diagnosis can usually be made after taking a detailed history and careful examination of the joints, soft tissues, nerve and blood supply, and of the lumbar spine.

Pain can arise from either the posterior or plantar aspects of the heel. Most lesions occur in people who walk or stand a great deal and are particularly common in athletes. Whatever the cause, the pain tends to be aggravated by walking and relieved by rest. Accurate localization of tenderness is important in diagnosis.

(a) Pain behind the heel

Tendon rupture

Complete or partial rupture of tendon may occur after vigorous activity. In the young, the musculotendinous junction tends to be the site of rupture, while in the old the tendon itself is at risk. It occurs in patients with pre-existing tendinitis, or in those receiving corticosteroids or following local corticosteroid injections. Complete rupture is easily recognized by loss of anatomical continuity and function. Partial rupture is more difficult to diagnose. Active plantar flexion of the ankle may be preserved but is painful. Initially, swelling and marked tenderness are noted just above the tendon insertion, with exquisite pain on movement. Later, a more irregular swelling due to fibrous tissue develops, with continuing pain on movement.

Orthopaedic treatment with immobilization and/or surgery is required (Plattner 1989).

Central core degeneration of the Achilles tendon

This must be differentiated from the more common peritendinous lesion, as tendon rupture is much more likely and local steroid injections are contraindicated. Onset is less dramatic than with partial rupture, but other clinical findings can be similar.

Pain gradually increases as the day progresses. Careful examination also reveals a localized tender nodule or thickening, 3 to 6 cm above the insertion to the calcaneum.

Peritendinitis (Achilles tendinitis)

This is the common cause of chronic, persistent, and often annoying pain at the back of the heel. It occurs most commonly but by no means exclusively in athletes, especially long-distance runners. In the acute stage there is diffuse swelling and tenderness on both sides of the tendon, sometimes accompanied by crepitus. Later, signs become less obvious but pain and tenderness recur with exercise. The symptoms usually develop gradually, distinguishing this from rupture of the tendon, and occasionally the tendon may calcify. Inflammation of the peroneal tendons behind the lateral malleolus causes pain when the muscle is contracted, as when walking over rough ground. Tendinitis here may be the presenting feature of rheumatoid arthritis.

Treatment

This consists of rest, avoiding the provocative occupational or athletic activity, shoe modification, and a heel raise to reduce tendon stretching during walking. Physiotherapy, including local heat and gentle stretching exercises, and occasionally a splint in slight plantar flexion, all help. Corticosteroid injections should not be given as they may predispose to rupture (Da Cruz *et al*. 1988). Surgical excision of the inflamed peritendinous tissue is rarely required.

Bursitis

Subachilles (retrocalcaneal) bursitis may also cause pain at the lower end of the Achilles tendon. Physical examination reveals the bulging, tender bursa on either side of the Achilles tendon with normal movement of the ankle joint. Dorsiflexion of the foot aggravates the pain by compressing the bursa. Frequently, the patient can recall trauma involving the foot. The diagnosis can be confirmed by radiography (showing obliteration of the retrocalcaneal recess).

Treatment

Rest, modification of activities, and a heel pad to elevate the heel are usually sufficient. A walking cast and/or a corticosteroid injection can be helpful. Surgery is rarely required.

Subcutaneous (postcalcaneal) bursitis

A more superficial bursitis can develop over the tendon attachment as a result of poorly fitting shoes; here the swelling is lower down, larger and fluctuant. Inflammatory changes in the overlying skin are common.

Treatment

This consists of rest, padding, and wearing a soft, non-restrictive shoe. Local corticosteroid injections should be avoided.

Retrocalcaneal apophysitis (Sever's disease)

The insertion of the Achilles tendon into the calcaneum can become inflamed, resulting in symptoms of pain and tenderness behind the heel; this usually occurs in boys aged between 9 and 15 years. A heel pad and restriction of activities usually result in improvement.

(b) Pain under the heel

Plantar fasciitis

In weight bearing, stress is placed on the long plantar ligament and fascia that supports the longitudinal arch of the foot. An enthesopathy can develop at the point of attachment of the ligament to the heel. Pain and tenderness may be confined to the point of attachment but can be much more widespread; it is always aggravated on walking. The presence of a calcaneal spur on the radiograph is not necessarily significant as it is also commonly seen on films of asymptomatic heels. Plantar fasciitis may be associated with the seronegative arthritides, such as Reiter's syndrome and ankylosing spondylitis, but is usually seen in the middle aged or obese as a result of repetitive trauma from athletic activities, occupations that involve excessive walking or standing, or from changes in footwear.

Tender heel pad

This causes pain in the hind part of the heel on standing or walking. The tough fibrofatty pad beneath the prominent weight-bearing part of the calcaneus is tender to finger palpation. The area of tenderness is well localized and may be a result of simple contusion; in most cases there is no history of trauma and in these patients the tenderness may result from obesity combined with excessive walking in unsatisfactory footwear, or from mild inflammation of uncertain origin.

Treatment

Treatment is similar to that for plantar fasciitis and also for a painful foot pad. Rest, reducing the weight-bearing pressure by using a soft heel pad, or weight reduction in obese patients are usually sufficient. Local injections of corticosteroids are often helpful. Occasionally a moulded orthosis or heel cup is required. Surgery is rarely indicted (Lester and Buchanan 1984).

The painful foot (Table 22)

Anterior flat foot—dropped transverse arch
Injuries to the spring ligament
Plantar digital neuritis—Morton's metatarsalgia
Tarsal tunnel syndrome
Plantar warts and callosities
Hallux valgus (bunion)
Dorsal exostoses with bursae

Table 22 The painful foot

'Flat feet' (pes planus)

'Flat feet' rarely cause symptoms in the young. In adolescents they may be part of a general postural defect. At this age, especially in boys, spasmodic flat foot is caused by peroneo–extensor spasm. The foot is rigid and painful, and symptoms are worse on walking.

Pain usually results from unaccustomed physical activity, weight gain or from ill-fitting shoes, causing ligamentous strain. Physical findings include a loss of the medial longitudinal arch on weight bearing, with medial and plantar displacement of the navicular and talar head. In severe cases the calcaneus is everted (valgus). A callosity can develop over the prominent talar head. In chronic pes planus deformity, secondary changes in the tarsal bones, spring and plantar ligaments, and in Achilles tendon, are common.

Treatment

No treatment is required for the asymptomatic flat foot. Pain and impaired function are indications for treatment. A soft arch support often gives relief. Exercises to strengthen the intrinsic muscles and good-fitting shoes are often helpful. In patients with hyperpronated foot a moulded orthotic device is often necessary.

Pes cavus

This is characterized by a high medial arch and is usually a hereditary condition. Shoe modifications to accommodate the high instep and claw toes with cushioned arch supports are often helpful. Stretching exercises of the toe extensors may also be taught.

'Spring' ligament

The 'spring' or plantar calcaneonavicular ligament supports the talonavicular joint from below and may be strained, leading to deep-seated pain on weight bearing.

Metatarsalgia

Patients with anatomical abnormalities of the feet, such as claw toes or equinus deformity, are liable to develop pain under the metatarsal heads on walking. Metatarsalgia, or pain around the metatarsal heads, is not itself a diagnosis. Between the metatarsal heads and the skin lie the flexor tendons and their sheaths, and, in the case of the great toe, the sesamoids in the flexor brevis hallucis. Inflammatory joint disease can produce inflammation of the tendon sheaths and causes generalized pain.

Soft or firm metatarsal pads placed proximal to the metatarsal heads, weight reduction in obese patients, or strengthening of the intrinsic muscles by toe flexion exercises are all helpful in treatment. If symptoms persist a metatarsal osteotomy and/or resection of the metatarsal head are indicated.

Hallux valgus

Hallux valgus (bunion) is an example of an adventitious bursa resulting from prolonged pressure over a bony prominence as a result of valgus deformity of the hallux. It is frequently inflamed. Initially, treatment consists of padding to reduce pressure and advice on suitable footwear. If these measures prove unhelpful, surgical treatment may be necessary.

Tibialis anterior pain

Pain in the mid-foot may arise around the insertion of the tibialis anterior after strenuous walking. Pain on the medial side of the foot as the point of insertion of the tibialis posterior tendon on the navicular is more common, and is particularly severe if the patient has an accessory navicular bone.

Dorsal exostoses

Some patients develop a bony prominence on the dorsum of the foot at the joint between the navicular and the cuneiform bone. Excessive friction may lead to the formation of a bursa.

Entrapment syndromes in the foot

Tarsal tunnel syndrome

This syndrome occurs more frequently than is suggested by isolated reports in the literature (Goodchild *et al*. 1965). Compression of the posterior tibial nerve under the flexor retinaculum below the medial malleolus can lead to burning pain and paraesthesiae under the medial side of the longitudinal arch of the foot. Symptoms are often mistaken for other foot disorders. Pain may awaken the patient at night and relief may be obtained by walking about. There is often tenderness over the nerve, and there may be vasomotor changes and weakness of toe flexion. Electrodiagnostic studies confirm the diagnosis.

Treatment includes corticosteroid injections and orthotics. Neither is consistently effective. Surgical decompression gives excellent results in the majority.

Lateral popliteal nerve lesions

The lateral popliteal branch of the sciatic nerve lies superficially against the head and neck of the fibula, where it is susceptible to compression especially when sitting with legs crossed, kneeling or wearing knee pads. Anaesthesia and periods of unconsciousness present particular risks. The symptoms are pain and tingling in the lateral aspect of the leg and dorsum of the foot, associated with weakness of dorsiflexion and eversion of the foot. Wasting of tibialis anterior, peroneal, and extensor digitorum brevis can occur. The high-stepping gait is typical. Electromyography can confirm the diagnosis.

Morton's metatarsalgia (plantar digital neuritis)

Patients complain of a sharp pain in the forefoot shooting through to the toes, usually the third and fourth. There may be tingling or even numbness in the adjacent sides of the third and fourth toes. The pain is localized between the metatarsal heads and the transverse metatarsal compression may produce a palpable click. The condition is caused by interstitial fibrosis compressing the digital nerve just before it divides into its two terminal branches and leading to the development of a

neuroma. In the early stages, well-fitting shoes with or without a metatarsal bar, give relief. Excision of the thickened tissue effects a cure, but may lead to numbness. One can decompress the nerve by release of the transverse metatarsal ligament.

Management of patients with soft-tissue rheumatism (general comments)

It is worth repeating that one feature common to all of the soft-tissue rheumatism syndromes is a tendency to spontaneous remission. Many of the syndromes improve in weeks and few persist significantly beyond 6 months. Therefore a clear diagnosis allows the physician to reassure the patient that arthritis is not present and that the prognosis is good. However, failure to resolve is often due to further injury. Few of the conditions require complete rather than selective rest. Detailed treatment has been discussed throughout the chapter.

Tenosynovitis around the wrist may respond to immobilization in a light forearm splint, maintaining the wrist in a position of optimum function (10–15° extension), and compression of the median nerve may respond to night splinting alone. The acutely painful shoulder also benefits from rest, mobility being maintained by pendular exercises. Physiotherapy in the acute stages only exacerbates the condition. Rest is essential for the painful heel and foot. Raising the heel of the shoe reduces tension on the Achilles tendon, and Elastoplast strapping will help immobilize the ankle. If there is flattening of the medial arch of the foot, then physiotherapy and an arch support may be helpful.

The short-term use of non-steroidal anti-inflammatory drugs may be necessary. Their immediate use is beneficial in limiting the pain and swelling of soft-tissue trauma. Many soft-tissue lesions respond to a local injection with steroid and anaesthetic. Some experience in the various techniques is necessary before reasonable success can be obtained. The correct choice of needle size makes a great deal of difference to the amount of pain experienced, and the amount that is injected varies according to the type of lesion.

The choice of steroid and local anaesthetic varies from user to user but, in practice, preparations differ significantly in their duration of action, and for most purposes the shorter-acting hydrocortisone acetate is suitable. The longer-acting steroids, methyl prednisolone acetate and triamcinolone hexacetonide, may cause skin atrophy when used for superficial lesions such as tennis elbow. After the injection there may be an increase in symptoms for up to 48 h and patients should be warned of this possibility. About 80 per cent of patients gain symptomatic benefit from these injections. It is important not to inject the tendoachilles itself, as this may predispose to rupture. In spite of decades of use of ultrasound in treating a wide spectrum of musculoskeletal disorders, its effectiveness remains unproven. Persistent lesions may require further injections but if conservative measures fail, surgery may be required.

Conclusion

About seven people in every hundred who visit their general medical practitioner seek help with symptoms arising from the soft tissues. The loss of work resulting from these symptoms is of the order of 11 million days annually. However, most soft-tissue lesions are eminently manageable, as discussed above.

Chapter References

Abrahamsson, S. O. (1991). Matrix metabolism and healing in the flexor tendon. Experimental studies on rabbit tendon. *Scandinavian Journal of Plastic Surgery of the Hand*, **25** (Suppl. 23), 1–51.

American College of Rheumatology (1990). Criteria for the classification of fibromyalgia: report of the multicentre criteria committee. *Arthritis and Rheumatism*, **33**, 160–72.

Awerbach, M. (1985). RSI or 'kangaroo paw'. *Medical Journal of Australia*, **142**, 237–328.

Beighton, P. B., Grahame, R., and Bird, H. A. (1989). *Hypermobility of joints*, (2nd edn). Springer, Berlin.

Binder, A., Hodge, G., Greenwood, A. M., Hazleman, B. L., and Page Thomas, D. P. (1985). Is therapeutic ultrasound effective in treating soft tissue lesions? *British Medical Journal*, **292**, 512–14.

Bjelle, A., Hagberg, M., and Michaelson, G. (1979). Clinical and ergonomic factors in prolonged shoulder pain among industrial workers. *Scandinavian Journal of Work and Environmental Health*, **5**, 205–10.

Bjelle, A., Allander, E., and Magi, M. (1990). Rheumatic disorders in the Swedish population and health care system. *Journal of Rheumatology*, **7**, 877–85.

Bulgen, D. Y., Binder, A., and Hazleman, B. L. (1984). Frozen shoulder: prospective clinical study with evaluation of three treatment regimes. *Annals of the Rheumatic Diseases*, **43**, 353–60.

Bunker, T. D. and Anthony, P. P. (1995). The pathology of frozen shoulder. *Journal of Bone and Joint Surgery* (B), **77**, 677–85.

Burkhart, S. S., Nottage, W. M., Ogilvie-Hans, D. J., Kohn, H. S., and Pacheli, A. (1994). Partial repair of irreparable rotator cuff lesions. *Arthroscopy*, 10, 363–70.

Carette S. (1995). Fibromyalgia 20 years later: What have we really accomplished. *Journal of Rheumatology*, **22**, 590–4.

Chard, M. D. and Hazleman, B. L. (1987). Shoulder disorders in the elderly (a hospital study). *Annals of the Rheumatic Diseases*, **46**, 684–7.

Chard, M. D. and Hazleman, B. L. (1989). Tennis elbow—a reappraisal. *British Journal of Rheumatology*, **28**, 186–90.

Chard, M. D., Gresham, A., and Hazleman, B. L. (1989). Age-related changed in the rotator cuff. *British Journal of Rheumatology*, **28** (Suppl. 1), 19.

Chard, M. D., Cawston, T. E., Riley, G. P., Gresham, G. A., and Hazleman, B. L. (1994). Rotator cuff degeneration and lateral epicondylitis—a comparative histological study. *Annals of the Rheumatic Diseases*, **53**, 30–4.

Child, A. H. (1986). Joint hypermobility syndrome: inherited disorders of collagen synthesis. *Journal of Rheumatology*, **13**, 239–42.

Cleland, L. G. (1987). RSI: a model of social iatrogenesis. *Medical Journal of Australia*, **147**, 236–7.

Clinton, T. and Solcher, B. (1994). Chronic leg pain in the athlete. *Clinical Sports Medicine*, **13**, 743–59.

Cohen, M. J. and Quintner, J. L. (1993). Fibromyalgia syndromes a problem of tautology. *Lancet*, **342**, 906–9.

Da Cruz, D. J., Geeson, M., Allen, M., and Phair, I. (1988). Achilles paratendinitis: an evaluation of steroid injection. *British Journal of Sports Medicine*, **22**, 64–5.

Dalton, S. E. (1989). Clinical examination of the painful shoulder. In *The shoulder joint (Baillière's clinical rheumatology*, Vol. 3, No. 3) (ed. B. L. Hazleman and P. A. Dieppe), pp. 453–74. Baillière Tindall, London.

Dalton, S. E., Cawston, T. E., Riley, G. P., Bayley, I. J. L., and Hazleman, B. L. (1995). Human shoulder tendon biopsy samples in organ culture produce procollagenase and TIMP. *Annals of the Rheumatic Diseases*, **54**, 571–7.

Department of Health and Social Security (1986). *Morbidity statistics from general practice. Third national study, 1981–1982*. HMSO, London.

Dixon, A. St. J. (1979). Soft tissue rheumatism. *Clinics in Rheumatic Diseases*, Vol. 5, No. 3, pp. 739–42. WB Saunders, Philadelphia.

Evans, E. J., Benjamin, M., and Pemberton, D. J. (1990). Fibrocartilage in the attachment zones of the quadriceps tendon and patellar ligament of man. *Journal of Anatomy*, **171**, 155–62.

Felson, D. T. and Goldenberg, D. L. (1986). The natural history of fibromyalgia. *Arthritis and Rheumatism*, **20**, 1522–6.

Garrick, J. G. and Requa, R. K. (1993). Ballet injuries: an analysis of epidemiology and financial outcome. *American Journal of Sports Medicine*, **21**, 586–90.

Gelberman, R. H., Aronson, D., and Weisman, M. (1980). Carpal tunnel syndrome. *Journal of Bone and Joint Surgery (A)*, **62**, 1181–4.

Goldenberg, D. (1995). Fibromyalgia: why such controversy? *Annals of the Rheumatic Diseases*, **54**, 3–5.

Goodchild, J., Kopell, H. P., and Sprelholz, N. I. (1965). The tarsal-tunnel syndrome. *New England Journal of Medicine*, **273**, 742–5.

Hagberg, M. (1981*a*). Muscular endurance and surface electromyogram in isometric and dynamic exercise. *Journal of Applied Physiology*, **51**, 1–7.

Hagberg, M. (1981*b*). Electromyographic signs of shoulder muscle fatigue in two elevated arm positions. *American Journal of Physical Medicine*, **60**, 111–21.

Hamilton, P. G. (1986). The prevalence of humeral epicondylitis: a survey in general practice. *Journal of the Royal College of General Practitioners*, **35**, 464–5.

Harris, N. H. (1974). Lesions of the symphysis in athletes. *British Medical Journal*, **4**, 211–14.

Harter, B. T., McKiernan, J. E., Kirzinger, S. S., Archer, F. W., Peters, C. K., and Harter, K. C. (1993). Carpal tunnel syndrome: surgical and non-surgical treatment. *Journal of Hand Surgery (A)*, **18**, 734–9.

Haupt, W. F., Wintzer, G., Schop, A., Lottger, J., and Pawlik, G. (1993). Tunnel decompression. *Journal of Hand Surgery (B)*, **18**, 471–4

Hocking, B. (1987). Epidemiological aspects of 'repetitive strain injury' in Telecom. *Medical Journal of Australia*, **147**, 218–22.

Hoppmann, R. A. and Patrone, N. A. (1989). A review of musculoskeletal problems in instrumental musicians. *Seminars in Arthritis and Rheumatism*, **19**, 117–26.

Ireland, D. C. R. (1988). Psychological and physical aspects of occupational arm pain. *Journal of Hand Surgery (B)*, **13**, 5–10.

Kannus, P. and Jozsa, L. (1991). Histopathological changes preceding spontaneous rupture of a tendon. *Journal of Bone and Joint Surgery (A)*, **73**, 1507–25.

Kellgren, J. H. (1939). On the distribution of pain arising from deep somatic structures with charts of segmental pain areas. *Clinical Science*, **4**, 35–46.

Kumagai, J., Sarkar, K., and Uhthoff, H. K. (1994). The collagen types in the attachment zone of rotator cuff tendons of the elderly: an immunohistological study. *Journal of Rheumatology*, **21**, 2096–100.

Labelle, H. *et al*. (1992). Lack of scientific evidence for the treatment of lateral epicondylitis of the elbow. *Journal of Bone and Joint Surgery*, **74**, 646–51.

Lester, D. K. and Buchanan, J. R. (1984). Surgical treatment of plantar fasciitis. *Clinical Orthopedics*, **186**, 202–4.

Littlejohn, G. O. (1986). Repetitive strain syndrome: an Australian experience. *Journal of Rheumatology*, **13**, 1004–6.

Littlejohn, G. O. (1989). Medicolegal aspects of fibrositis syndrome. *Journal of Rheumatology*, **16** (Suppl. 19), 169–74.

Maigne, R. (1972). Sémiologie clinique des dérangements intervertébraux mineurs. *Annales de Medecine*, **15**, 175–292.

Miller, H. W. (1973). *Plan and operation of the Health and Nutrition Examination Survey: United States 1973–5. Vital and health statistics*, Series 1, No. 10. Public Health Service, US Department of Health, Washington DC.

Moldofsky, H., Scarisbrick, P., England, R., and Smythe, H. (1975). Musculoskeletal symptoms and non-REM sleep disturbance in patients with 'fibrositis syndrome' and healthy subjects. *Psychosomatic Medicine*, **37**, 341–51.

Neer, C. S. (1983). Rotator cuff athropathy. *Journal of Bone and Joint Surgery (A)*, **65**, 1232–44.

Neviaser, R. J. and Neviaser, T. J. (1987). The frozen shoulder. Diagnosis and management. *Clinical Orthopedics*, **223**, 59–64.

Nobuhara, K., Supapo, A. R., and Himo, T. (1994). Effects of shoulder distension in shoulder disease. *Clinical Orthopedics*, **304**, 25–9.

Payne, L. Z. (1994). Tears of the glenoid labrum. *Orthopedics Review*, **23**, 577–83.

Plattner, P. F. (1989). Tendon problems of the foot and ankle. *Postgraduate Medicine*, **86**, 155–70.

Ramel, E. and Moritz, V. (1994). Self-reported musculoskeletal pain and discomfort in professional ballet dancers in Sweden. *Scandinavian Journal of Rehabilitation Medicine*, **26**, 11–16.

Riley, G. P., Harrall, R. L., Constant, C. R., Chard, M. D., Cawston, T. E., and Hazleman, B. L. (1994). Glycosaminoglycans of human rotator cuff tendons: changes with age and in chronic rotator cuff tendinitis. *Annals of the Rheumatic Diseases*, **53**, 367–76.

Risk, T. E. and Pinals, R. S. (1982). Frozen shoulder. *Seminars in Arthritis and Rheumatism*, **11**, 440–52.

Russell, I. J. *et al*. (1994). Elevated cerebrospinal fluid levels of substance P in patients with fibromyalgia syndrome. *Arthritis and Rheumatism*, **37**, 1593–601.

Scheuler, W. *et al*. (1988). The alpha-sleep pattern: quantitative analysis and functional aspects. In *Sleep* (ed. W. P. Koelta, F. Obal, H. Schultz, and P. Visser). Fischer, Berlin.

Smythe, H. A. (1989). Fibrositis syndrome. A historical perspective. *Journal of Rheumatology*, **16** (Suppl. 19), 1–6.

Thompson, A. R., Pleavis, L. W., and Shaw, E. G. (1951). Peritendinous crepitans and simple tenosynovitis: a clinical study of 544 cases in industry. *British Journal of Industrial Medicine*, **8**, 150.

Travell, J. and Simons, D. G. (1983). *Myofascial pain and dysfunction. The trigger point manual*. Williams and Wilkins, Baltimore.

Uthoff, H. K. and Sarkar, K. (1989). Calcifying tendinitis. In *The shoulder joint* (*Baillière's clinical rheumatology*, Vol. 3, No. 3) (ed. B. L. Hazleman and P. A. Dieppe), pp. 567–81. Baillière Tindall, London.

Vogel, K. G., Ordog, A., Pogany, G., and Olah, J. (1993). Proteoglycans in the compressed regions of human tibialis, posterior tendon and in ligaments. *Journal of Orthopedic Research*, **11**, 68–77.

Wilcox, M. S. and Bilbao, A. (1993). Sensitivity of electrophysiological studies and the carpal tunnel syndrome. *Muscle and Nerve*, **16**, 1265–6.

Williams, J. G. P. (1977). Surgical management of traumatic non- infective tenosynovitis of the wrist extensors. *Journal of Bone and Joint Surgery (B)*, **59**, 408–70.

Wood, P. H., Sturrock, A. W. and Badley, E. M. (1979). Soft tissue rheumatism in the community. In *Soft tissue rheumatism* (*Clinics in Rheumatic Diseases*, Vol. 5, No. 3), pp. 743–53. WB Saunders, Philadelphia.

World Health Organization (1988). *Identification and control of work-related diseases. Report of a WHO expert committee*. World Health Organization Technical Report Series No. 174. WHO, Geneva.

Yamanaka, K. and Fukuda, H. (1991). Ageing process of the human supraspinatus tendon with reference to rotator cuff tears. In *Surgical disorders of the shoulders* (ed. M. S. Watson), pp. 247–58. Churchill Livingstone, Edinburgh.

5.15 Osteoarthritis

Michael Doherty and Adrian Jones and T. E. Cawston

Introduction

Osteoarthritis is the commonest condition to affect the joints of man. As such it is a major cause of locomotor pain, the single most important rheumatological cause of disability and handicap, and an important health care challenge with major resource implications (Steven 1992; Yelin 1992; Badley *et al*. 1994). Indeed it has been estimated that 15 per cent of the United Kingdom population greater than 55 years of age have symptomatic osteoarthritis of the knee (Fig. 1). Osteoarthritis has, however, only recently become a significant focus of clinical interest and research. Previously considered a boring 'wear and tear', 'degenerative' disease that must be accepted as the inevitable consequence of trauma and ageing, osteoarthritis is increasingly viewed as a dynamic, essentially reparative process with potential for health intervention and prevention.

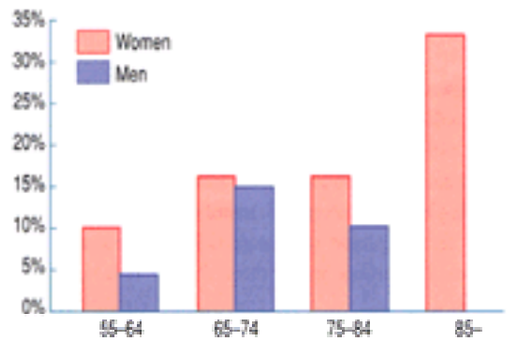


Fig. 1 Prevalence of symptomatic knee osteoarthritis by age group and gender in a United Kingdom community population. Derived from McAlindon *et al*. (McAlindon *et al*. 1992).

Historical perspective

At the turn of the century, pathologists and radiologists differentiated two main categories of chronic arthritis—atrophic and hypertrophic. The former was characterized by synovial inflammation with erosion or atrophy of cartilage and bone; it encompassed several disease entities including rheumatoid arthritis. Hypertrophic arthritis, by contrast, was characterized by more focal cartilage loss, minimal evidence of inflammation, and by hypertrophy of adjacent bone and soft tissues; this group became synonymous with osteoarthritis (Goldthwaite 1904) (Fig. 2). Since there were recognized associations with ageing and previous joint trauma, this led to ready acceptance of the alternative term 'degenerative joint disease'. With the major rheumatological interest focusing on inflammatory arthropathies, osteoarthritis was often used as a 'non-inflammatory' control disease, and even as a surrogate for normal joint tissue, in clinical and laboratory research. The term 'osteoarthrosis' was often used, therefore, to emphasize absence of overt inflammation.

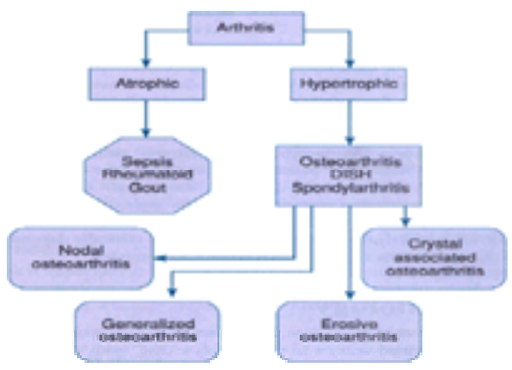


Fig. 2 The increasing differentiation of arthritis subtypes.

Advances in cartilage biochemistry and the recognition of calcium crystal-associated disease subsets proved important factors in renewing interest in this condition. Although clinical and histological inflammation is not as florid as in rheumatoid arthritis or seronegative spondylarthropathies, it is an undoubted component in many cases. Thus the term osteoarthritis is now generally preferred.

Definition of osteoarthritis

Concepts of osteoarthritis are still changing and there is no universal agreement on its definition. A current working definition of osteoarthritis is:

a condition of synovial joints characterized by cartilage loss (chondropathy) and evidence of accompanying periarticular bone response.

It is noteworthy that chondropathy may occur without hypertrophic bone response (e.g. polychondritis, rheumatoid arthritis), and periarticular new bone may develop without chondropathy (e.g. 'traction' spurs); it is only when the two occur together in synovial joints that the term osteoarthritis is appropriate.

The drawbacks of this working definition include:

1. exclusion of joints with early (initial) change—the very joints we need to study in order to elucidate aetiopathogenic processes;
2. emphasis on cartilage and bone—even though all other joint components (synovium, capsule, entheses, muscle) demonstrate change (Fig. 3);

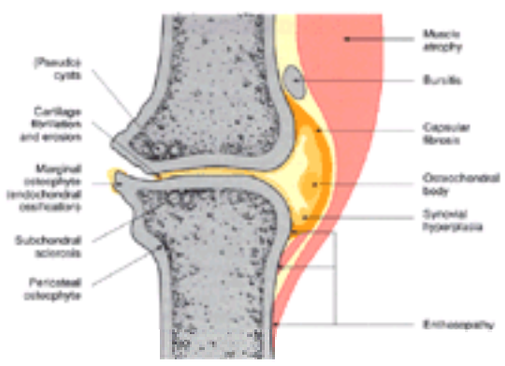


Fig. 3 Joint tissues affected in osteoarthritis. Redrawing from Doherty (ed.) *Color Atlas and Text of Osteoarthritis*. Wolfe, London. (Doherty).

3. structural rather than physiological emphasis—with no consideration of biological, symptomatic, and functional consequences.

Nevertheless, given these caveats, such a definition is a practicable starting point for examination of existing clinical, epidemiological, and experimental data. The American College of Rheumatology has devised criteria for classification of symptomatic osteoarthritis of knee (Altman *et al*. 1986) (Fig. 4), hand (Altman *et al*. 1990), and hip (Altman *et al*. 1991) that incorporate clinical, laboratory, and/or radiological features. Although these criteria will distinguish osteoarthritis from other painful joint conditions (the basis of their development) there remain questions about their use in other settings. In population studies they may be insensitive and do not detect asymptomatic disease. Many questions thus remain to be resolved. It is likely that future definitions will be critically dependent on the populations and use for which they are intended.

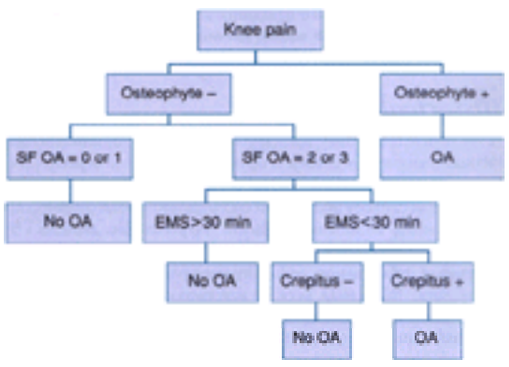


Fig. 4 Decision tree based classification of knee osteoarthritis according to the American College of Rheumatology classification criteria. Derived from Altman *et al*. (Altman *et al*. 1986).

Epidemiology

Current epidemiological data in osteoarthritis is beset by a lack of generally accepted criteria to distinguish a 'case' from a 'non-case'. Histopathological changes are obviously not appropriate for living subjects, and most surveys rely on radiographic features for definition and assessment of severity. Of the various radiographic criteria, the most widely employed are those of Kellgren and Lawrence (Kellgren and Lawrence 1957). Although differing in detail by joint site and by publication, these grade osteoarthritis into four categories depending on the presence and degree of various features (Fig. 5). Although these features purport to measure various pathological changes occurring in cartilage and subchondral bone (Table 1), there are problems with both validity and reproducibility. Particular problems include:

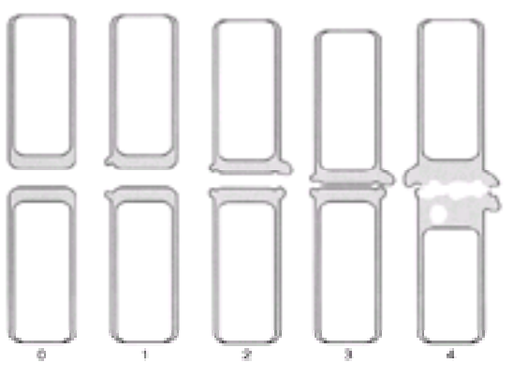


Fig. 5 Basis of Kellgren and Lawrence grading scheme. Grade 0, normal; Grade 1, minimal osteophyte, normal joint space; Grade 2, definite osteophyte, possible joint space narrowing; Grade 3, definite osteophyte and joint space narrowing; Grade 4, definite osteophyte and joint space narrowing with sclerosis and abnormal joint contour.

Pathological change	Radiographic abnormality
Cartilage fibrillation, erosion	Decrease in interosseous distance (localized)
Subchondral new bone formation	Sclerosis
New cartilage formation and endochondral ossification	Osteophyte
Fibrous-walled pseudocysts resulting from fluid intrusion or myxoid degeneration	Subchondral cysts
Trabecular compression	Bone collapse/attrition
Fragmentation of osteochondral surface; cartilage and bone metaphasis in synovium	Osseous ('loose') bodies

Table 1 Radiographical–pathological correlates in osteoarthritis

1. lack of distinction between isolated and chondropathy-associated osteophyte, the aetiological factors for which may vary;
2. the ignoring of clinical status, symptoms, and function;
3. inappropriate assumptions about disease progression in grading systems which combine several features, such as that of Kellgren and Lawrence.

Although attempts to define osteoarthritis clinically have been made (Claessens *et al*. 1990; Hart *et al*. 1991; Altman 1991) the lack of a suitable gold standard, poor correlation with currently accepted surrogate 'gold standards' (e.g. radiology), and poor reproducibility have hampered their widespread adoption.

Such problems in the definition of osteoarthritis, radiograph interpretation, and clinical measurement make the epidemiology of osteoarthritis difficult to analyse (Peyron 1979; Spector and Cooper 1993) but several broad conclusions may now be drawn.

Descriptive studies

Autopsy studies suggest that the majority of subjects over age 65 have evidence of osteoarthritis in at least one joint site at the time of death. Prevalence estimates from such studies tend to be higher than those from radiographic surveys, perhaps because the whole joint surface is available for study (Rogers *et al*. 1990).

Radiographic studies also report high prevalence in the middle-aged and elderly. This prevalence varies according to joint site and age, giving different patterns of distribution at different ages. For example hand interphalangeal joints and first metatarsophalangeal joints are affected commonly and at a relatively young age, whereas glenohumeral and shoulder joints are affected less commonly and principally in the elderly. From available large radiographic studies (Mikkelsen *et al*. 1970; Lawrence 1977; Felson 1988; Van Saase *et al*. 1989) the following generalizations can be made.

Age

This is a major determinant of prevalence at all important sites including hands, knees, and hips. Prevalence is low under age 45 years. Polyarticular osteoarthritis (³5 sites involved) is also rare in those aged less than 45 years. Prevalence increases up to age 65, when there is involvement of at least one joint group in at least 50 per cent of the population. Continuing increase in prevalence in those over age 65 is less clear cut, and indeed a plateauing out of prevalence in the very old has been suggested (Bagge *et al*. 1992) (Fig. 6).

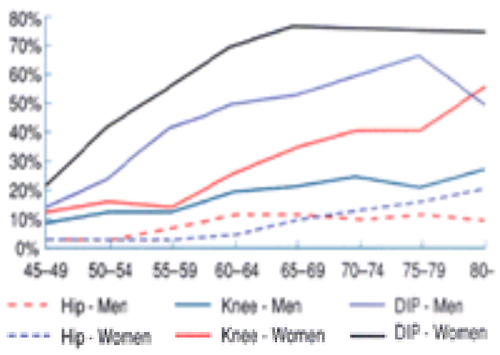


Fig. 6 Prevalence of radiological osteoarthritis by age group and gender in a Netherlands community population. Derived from van Saase *et al*. (van Saase *et al*. 1989).

Gender

This is important at some, but by no means all, joint sites. Although there is little or no gender difference in the prevalence of mild osteoarthritis, a female preponderance becomes more apparent:

1. for severe grades of osteoarthritis;
2. in older age groups;
3. for osteoarthritis of the hands and knees.

There is also a polyarticular form of hand osteoarthritis that has a predilection for perimenopausal women—so-called nodal generalized osteoarthritis. This is discussed further below.

Ethnic group

Given the difficulties of representative sampling and reproducible assessments, comparison between populations shows surprising similarity in age-specific prevalence by joint site (Van Saase *et al*. 1989). Possible exceptions are:

1. hip osteoarthritis, which shows substantially lower prevalence among black and Oriental populations than among whites;
2. polyarticular hand osteoarthritis, which appears to be less common in African and Malaysian populations.

Geographic variation—'endemic osteoarthritis'

Several forms of disabling polyarticular osteoarthritis occur with high frequency in certain geographical locations (Sokoloff 1985). The best described are Kashin-Beck disease (south-eastern Siberia, northern China, North Korea) and Mselini disease (limited to the Zulu and Tonga tribes of south-east Africa). Other forms of endemic osteoarthritis occur in Malnad and elsewhere in India. These conditions share in common:

1. onset during the first or second decade;
2. variable growth restriction;
3. characteristics of acquired rather than inherited disease;
4. involvement of impoverished rural communities.

The conditions radiographically most closely resemble those of spondyloepiphyseal dysplasia, and histologically show chondronecrosis as a likely initial event (Fig. 7). The cause(s) of endemic osteoarthritis remain unknown, though mycotoxins elaborated by moulds on local grains and abnormalities of trace elements (e.g. selenium) in soil or water have been investigated.

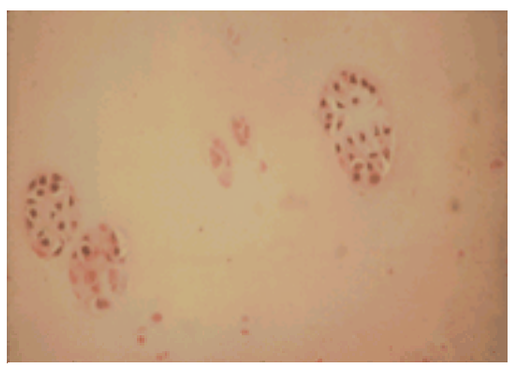


Fig. 7 Radiograph of coronal slab section of femoral head in Mselini hip disease. From Sokoloff in *Color Atlas and Text of Osteoarthritis* (ed. M Doherty). Wolfe, London (Sokoloff).

Analytical studies

Individual risk factors for osteoarthritis may conveniently be viewed as:

1. those influencing or marking a generalized predisposition to the condition;
2. those resulting in abnormal biomechanical loading at specific sites (Fig. 8).

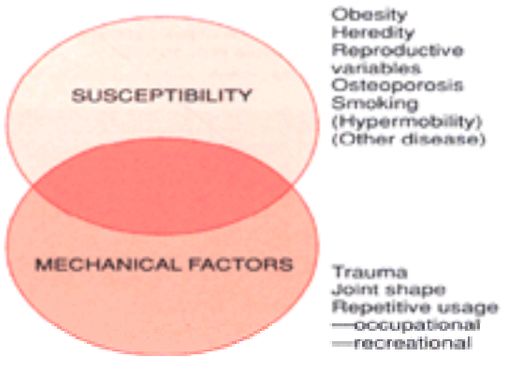


Fig. 8 Susceptibility to osteoarthritis: interaction between generalized and local factors.

Generalized susceptibility

Obesity

This is closely associated with knee osteoarthritis (odds ratio 4.5 for men, 9.0 for women, for those 50 per cent above ideal body weight compared to those at ideal body weight) but interestingly not at the hip (Lawrence *et al*. 1990; Felson *et al*. 1992). At the knee, obesity precedes rather than follows knee osteoarthritis and indeed weight loss may prevent the development of knee osteoarthritis (Felson *et al*. 1992) (Fig. 9). It is associated with radiographic change irrespective of whether symptoms are present or not, arguing against obesity resulting from more sedentary lifestyle of those with painful knees. It is also true, however, that obesity increases the risk of knee pain for a given degree of structural change. The mechanism relating obesity and osteoarthritis remains speculative although the apparent lack of association with hip osteoarthritis suggests metabolic or systemic factors rather than a purely mechanical explanation.

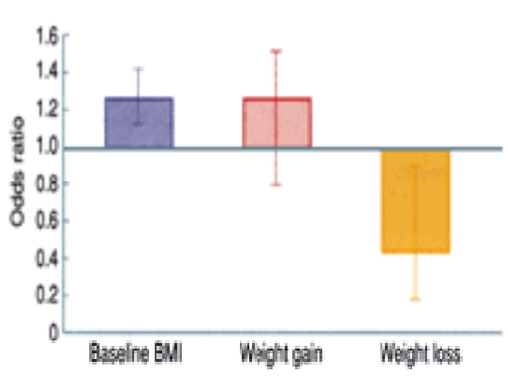


Fig. 9 Risk of symptomatic knee osteoarthritis per 2 units of body mass index (BMI) expressed as odds ratio and 95 per cent confidence interval. Derived from Felson *et al*. (Felson *et al*. 1992).

Genetic factors

A strong familial tendency is recognized for nodal generalized osteoarthritis, a polyarticular form of osteoarthritis occurring mainly in women of perimenopausal age. This is confirmed by family and twin studies (Lawrence 1977). Heberden's nodes appear to be inherited independently as an autosomal dominant trait with greater penetrance in women (Stecher 1953). The genetics and pathogenic mechanism remains unknown but associations have been reported with HLA A1, B8 haplotypes and with the a-1-antitrypsin MZ phenotype (Pattrick *et al*. 1989a). In addition, an increased incidence of IgG rheumatoid factors (Hopkinson *et al*. 1992) and a high frequency of immune complexes in cartilage and synovium of hips removed from nodal generalized osteoarthritis compared to pauciarticular osteoarthritis patients (Cooke 1985) has led to speculation of an autoimmune aetiology (Doherty *et al*. 1990). It is hypothesized that an unidentified 'single-shot' insult occurs in a genetically predisposed individual. This triggers an immunological response, leading to polyarticular damage and initiation of the osteoarthritis process (Fig. 10). However, because it is a single temporal insult the repair process wins through, 'compensates', and results in a good outcome, other than for leaving the structurally abnormal

joints that we recognize as 'nodal generalized osteoarthritis' (Doherty *et al*. 1990). These findings have not yet been confirmed.

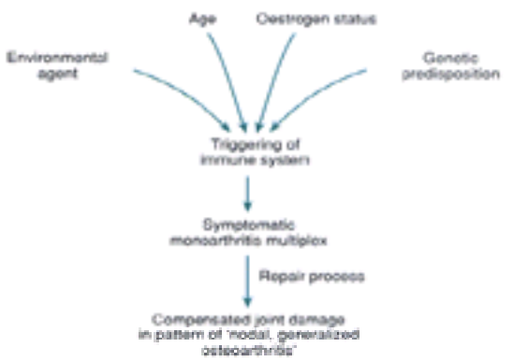


Fig. 10 Hypothesis explaining nodal generalized osteoarthritis as an 'autoimmune' condition.

An association with the COL2A1 gene has been demonstrated for hereditary forms of premature polyarticular osteoarthritis with mild dysplasia (Knowlton *et al*. 1990). It has also been found in some, but not all, forms of hereditary osteo-opthalmoarthropathy (Stickler's syndrome), an hereditary disease with variable ocular and midfacial abnormalities and a premature osteoarthritis-like syndrome (Williams and Jimenez 1993). Similar associations of type II collagen defects with nodal generalized and sporadic osteoarthritis have not been found.

Reproductive variables

Polyarticular osteoarthritis has a strong female predominance, a frequent onset around the menopause, and reported associations with previous hysterectomy, gynaecological surgery, and possible alterations of sex hormone binding globulin. This has led to the suggestion that hormonal factors are important in this subgroup (Spector and Campion 1989). Manipulation of sex hormones in animal models of 'osteoarthritis' and identification of oestrogen receptors on chondrocytes lend some support to a role for hormonal modulation. Similar treatment in humans has been unsuccessful in established osteoarthritis (Kellgren and Moore 1952).

Bone density

A negative association is reported between osteoporosis and osteoarthritis at certain sites (Lane and Nevitt 1994). The strongest evidence relates to the hip, where several studies support a negative correlation between hip osteoarthritis and risk of femoral neck fracture. In patients with polyarticular osteoarthritis, studies of bone density have produced conflicting results, with obesity a frequent confounding factor. One explanation for this negative relationship, if true, is that weak bone may absorb excessive impact loading and thus protect joint cartilage from damage and subsequent osteoarthritis. Certainly, in the converse rare situation of osteopetrosis, where the skeleton is diffusely sclerotic, a high incidence of premature polyarticular osteoarthritis is reported.

Cigarette smoking

A protective influence of smoking on knee osteoarthritis, after correction for possible confounding factors, is reported from the Framingham study (Felson 1988). Other studies have produced both confirmatory and contradictory results and the issue is not yet resolved. The observation is intriguing, however, but the mechanism is unclear. Smoking is a risk factor for osteoporosis, has antioestrogenic properties, and has many effects on cell function. It is unclear which, if any, of these is important.

Other suggested factors

Less definite associations are reported with diabetes, hypertension, and hyperuricaemia, which are independent of obesity. Clearer associations with acromegaly and haemochromatosis are evident (Jones *et al*. 1992). An increased frequency of osteoarthritis in subjects with generalized hypermobility has also been suggested, due either to associated connective tissue abnormality or joint trauma, but rigorous epidemiological support for this is still required.

Local mechanical factors

Trauma

Major direct injury is accepted as a predisposing cause of osteoarthritis (Wright 1990). Intra-articular fracture affecting the articular surface is probably associated with osteoarthritis although much of the evidence is retrospective nd uncontrolled. Major injury, particularly fracture, may also alter mechanical loading and predispose to osteoarthritis at distant sites, as with fractures of the femoral shaft (hip osteoarthritis) (Fig. 11), scaphoid (wrist osteoarthritis), tibia (ankle osteoarthritis), or humerus (shoulder osteoarthritis). Although trauma may be a predisposing factor that determines the site of osteoarthritis, mechanical insult alone is usually an insufficient cause for its development. For example following total meniscectomy not everyone develops knee osteoarthritis. Furthermore, the increased frequency of postmeniscectomy osteoarthritis in subjects with a generalized predisposition to osteoarthritis (i.e. those with distal interphalangeal joint osteoarthritis) compared to those with no such predisposition supports interaction between generalized (constitutional) and local (mechanical) factors (Doherty *et al*. 1983).

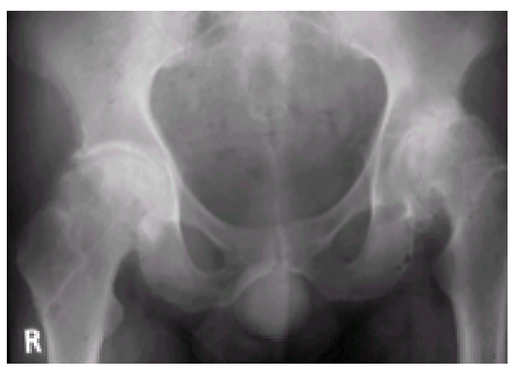


Fig. 11 Asymmetric hip osteoarthritis following previous trauma on the right side in a patient (Ledingham *et al*. 1992).

Joint shape

Abnormalities of articular contour, that may lead to abnormal load transmission across the joint, have particularly been linked with predisposition to osteoarthritis at the hip and knee. It is well established that childhood hip disorders such as Perthes' disease, slipped capital epiphysis, and congenital dislocation lead to premature hip osteoarthritis. It has been suggested that lesser degrees of acetabular dysplasia account for a proportion of hip osteoarthritis amongst younger subjects, though the impact of such mild developmental abnormality in causing later hip osteoarthritis is questionable, with a recent study suggesting it is of little importance (Croft *et al*. 1991). Cooke has suggested that mild, often unrecognized, dysplasia of the femoral condyles may similarly predispose to knee osteoarthritis via mechanical effects (Cooke 1985). Abnormal contour following intra-articular fractures may also contribute to premature osteoarthritis.

Occupational and recreational activities

Data in this area were conflicting but a consensus is now emerging (Cooper 1995). For example repetitive impact loading and trauma have been implicated in reported increased frequencies of osteoarthritis in miners (knees, spine), cotton workers (distal interphalangeal joints), and pneumatic drillers (elbows), though conversely parachutists, ballet dancers, and runners show no obvious increased risk of osteoarthritis. It may be that at some sites trauma is more important in selecting the site and severity of the condition than in determining whether osteoarthritis will develop or not. For example examination of patterns of osteoarthritis in the hand suggests that usage may influence distribution of osteoarthritis in the hand (Hadler *et al*. 1977), though evidence for an increase in frequency from repetitive usage is lacking.

At the hip, several studies have now shown a convincing increased incidence of osteoarthritis in farmers. The mechanisms underlying this association are unclear but repetitive trauma is suggested. At the knee there is increasing evidence that repetitive knee bending may be harmful (Felson *et al*. 1991). This association has only been clearly demonstrated in men.

Clinical features

Symptoms and impact of osteoarthritis

The principal clinical features of osteoarthritis are symptoms (pain, stiffness), functional impairment, and signs (primarily anatomical change). Though interrelated, there is often marked discordance between these three.

Symptoms

Although pain is the chief complaint its origin is not at all clear. Hyaline cartilage is aneural and this means that metabolic or structural alteration in this tissue is unlikely to be directly perceived as painful. Several other mechanisms of symptom production have been suggested (Fig. 12). These include:

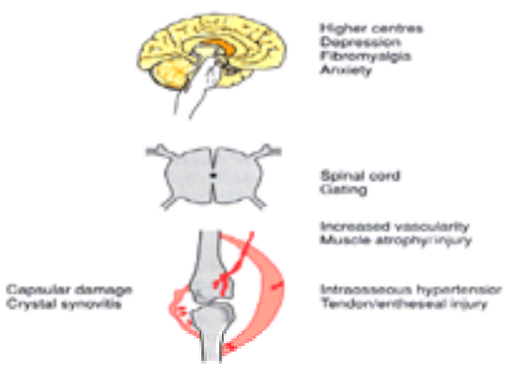


Fig. 12 Potential sites of pain perception and/or modification in osteoarthritis. Redrawn from Jones and Doherty (1992).

1. stimulation of capsular pain fibres and mechanoreceptors by intra-articular hypertension consequent upon synovial hypertrophy, increased fluid production, and decreased joint compliance;
2. inflammatory mediators stimulating pain fibres in the synovium and capsule;
3. stimulation of periosteal nerve fibres by intraosseous hypertension accompanying osteoarthritis;
4. perception of subchondral microfractures, painful enthesopathy, and bursitis that accompany structural alteration, muscle weakness, and altered usage;
5. maladaptive changes in the spinal cord and brain leading to persistent pain perception.

It has been suggested that these different mechanisms may produce different pain characteristics with, for example, pain predominantly:

1. on usage—being due to mechanical or enthesopathic problems;
2. at rest—being inflammatory in origin;
3. at night—being due to intraosseous hypertension.

The last of these may be a particularly poor prognostic factor and indicates severe damage. Pain in osteoarthritis may, however, be a transient feature and can be absent in spite of severe joint damage (Fig. 13). Correlation between pain and radiographic change varies according to site. It is best at the hip and then knee, with the poorest correlation occurring in the hands and spinal apophyseal joints. Joints with more severe radiographic change are more likely to be symptomatic than those with mild change (Fig. 14) but irrespective of structural change, pain is more common in women (Lawrence *et al*. 1990). As with any locomotor or other pain, the subjective magnitude and perception of pain may be greatly influenced by factors such as personality, anxiety, and/or depression. Indeed, several recent studies have demonstrated that these psychological factors may be much more important in determining symptomatic outcome. One study has even gone further and suggested that factors associated with radiographic outcome may be independent of and distinct from those associated with knee pain and disability (Davis *et al*. 1992). This has led at least one author to question whether we should in fact ignore the entity we try to describe as osteoarthritis and instead concentrate on knee pain (Hadler 1992).

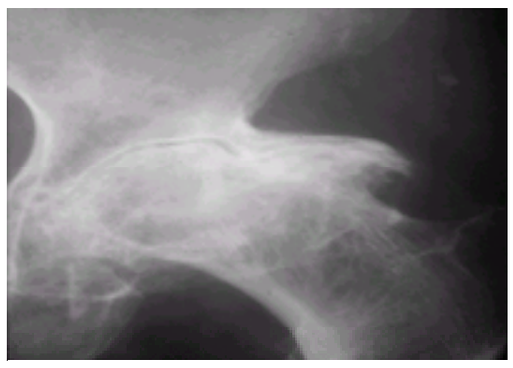


Fig. 13 Radiograph of an asymptomatic individual with severe radiographic osteoarthritis change.

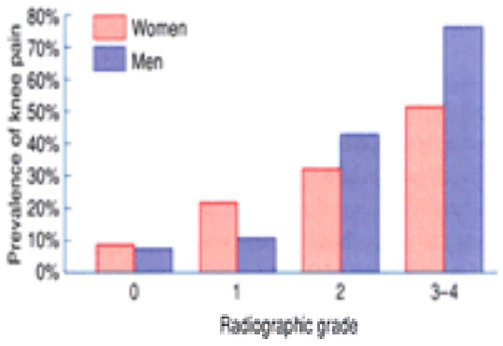


Fig. 14 Prevalence of knee pain by radiographic grade and gender in a United States community population. Derived from Davis *et al*. (Davis *et al*. 1992).

Stiffness is the other chief complaint. This is often described as 'gelling' of the joint after inactivity with difficulty in initiating movement. Prolonged morning or inactivity stiffness, often taken as a reflection of inflammation, is uncommon but may occur, particularly in patients with chronic pyrophosphate arthropathy.

Some patients may complain of joint swelling and deformity (particularly of hands), and coarse crepitus, even in the absence of other symptoms.

Functional impairment

Disability results from reduced range and control of movement and from pain. Handicap, of course, will vary according to individual patient requirements and aspirations. The pain and functional consequences of osteoarthritis are responsible for the huge burden of morbidity in the community. Severe knee and (less commonly) hip disease result in a massive health care problem to a generally older and otherwise fitter population. In addition to morbidity, cumulative mortality rates among subjects aged 55 to 74 in the National Health and Nutrition Examination Survey (NHANES-I) were significantly greater (relative risk 1.1) for women but not men with knee osteoarthritis (Lawrence *et al*. 1990). An increased mortality has also been associated with knee osteoarthritis in Sweden (Danielsson and Hernborg 1970).

Signs

Several features may occur in any combination and primarily reflect altered joint structure. These include:

1. crepitus, presumed due to an irregular articular surface;
2. bony enlargement, due to osteophyte and remodelling;
3. deformity;
4. instability;
5. restricted movement;
6. stress pain.

Varying degrees of synovitis (warmth, effusion, synovial thickening) may accompany joint line tenderness. Muscle weakness and wasting may also be apparent. Periarticular sources of pain, demonstrated by point tenderness away from the joint line and by stress testing, are commonly identified at the knee and hip.

Osteoarthritis 'subsets'

Since osteoarthritis is a process that may be triggered by diverse constitutional and environmental factors, a wide spectrum of clinical expression and outcome is to be expected. Attempts to define and classify osteoarthritis as a single disease entity have not been entirely successful and attempts have been made to separate osteoarthritis into more homogeneous groupings or 'subsets' so as better to define aetiological factors and prognosis.

Osteoarthritis was initially classified as primary (no cause identified) or secondary (an obvious cause identified, such as trauma or dysplasia). Indeed such a distinction is still retained in the American College of Rheumatology criteria (Altman *et al*. 1991). Such artificial separation has often proved unsatisfactory due to:

1. frequent lack of an identifiable cause, resulting in a large heterogeneous primary group;
2. overlap between subsets, as shown for example by the influence of predisposition to 'primary' generalized osteoarthritis in determining development of postmeniscectomy 'secondary' osteoarthritis (Doherty *et al*. 1983).

In addition to identifiable predisposing factors, the following more objective features have therefore often been used as a further basis of subset differentiation:

1. joint site involved (hip, knee, hand);
2. site within a joint (medial tibiofemoral, lateral tibiofemoral, patellofemoral);
3. number of joints involved (one, few, many);
4. presence of associated crystal deposition;
5. presence of marked clinical inflammation;
6. radiographic bone response (atrophic, hypertrophic).

A number of 'subsets' have emerged which differ in a number of such characteristics. It is important to note, however, that sharp distinction between subsets does not exist. Many of the above characteristics represent different aspects of the osteoarthritis process (the balance between damage and repair), and may dominate the clinical picture at just one phase in the evolution of the condition. One 'subset' may thus evolve into another, and different 'subsets' may exist at different sites within the same individual. Possibly the most important distinction is simply by site and number of joints involved; predisposing factors are increasingly being associated with specific joint sites, or to polyarticular as opposed to pauciarticular involvement. It follows that knowledge concerning pathogenesis, risk factors, or treatment success of osteoarthritis cannot necessarily be extrapolated from one site to another.

Nodal generalized osteoarthritis

This is perhaps the best recognized subset, characterized by:

1. polyarticular finger interphalangeal involvement;
2. Heberden's and Bouchard's nodes;
3. female preponderance;
4. peak onset around the menopause;
5. good functional outcome for hands;
6. predisposition to osteoarthritis of the knee, hip, and spine;
7. marked familial predisposition.

The typical patient is a woman in her forties or fifties who develops discomfort followed by swelling of a single finger interphalangeal joint. A few months later another interphalangeal joint becomes painful, then another producing a 'stuttering' onset polyarthritis of distal and proximal interphalangeal joints ('monoarthritis multiplex'). Affected interphalangeal joints may feel very stiff, be tender, and show tight posterolateral swelling with overlying erythema. Aspiration of such swellings may reveal viscous, clear, hyaluronate-rich 'jelly': these cysts represent mucoid transformation of periarticular fibroadipose tissue and may communicate with the joint. Each interphalangeal joint tends to go through a symptomatic phase while swelling and deformity become established, resulting in perhaps 1 to 3 years of episodic discomfort and stiffness. In almost all cases symptoms then subside, leaving the patient with typical posterolateral firm Heberden's (distal interphalangeal joint) and

Bouchard's (proximal interphalangeal joint) nodes, characteristic lateral deviations of interphalangeal joints (Fig. 15), and radiographic evidence of osteoarthritis (Fig. 16). In addition to finger interphalangeal joints, the first carpometacarpal, metacarpophalangeal, and interphalangeal joints of the thumb are commonly affected: other joint involvement in the hand and wrist, however, is usually restricted to the index and middle metacarpophalangeal, scaphotrapezoid, and pisiform-triquetral articulations. The prognosis seems to be good.

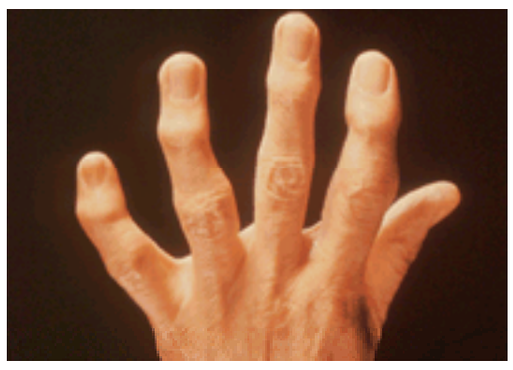


Fig. 15 Typical appearance of multiple Heberden's and Bouchard's nodes in nodal generalized osteoarthritis.

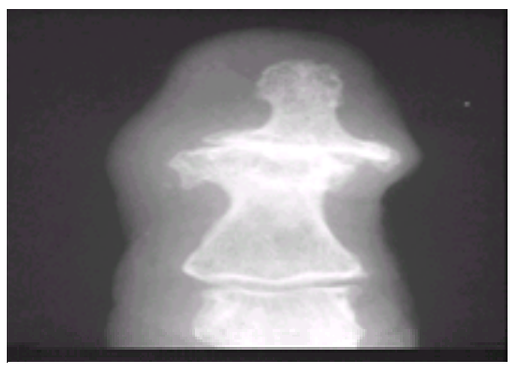


Fig. 16 Radiograph of distal interphalangeal joint showing many of the typical radiographic features of osteoarthritis (joint space narrowing, sclerosis, osteophyte).

Nodal generalized hand osteoarthritis is associated with an increased frequency of osteoarthritis at other sites, particularly knees, hips, first metatarsophalangeal joints, and cervical and lumbar apophyseal joints (Kellgren and Moore 1957). This concept of 'generalized osteoarthritis', with hand involvement as the marker for predisposition, was first described by Haygarth in 1805 and is supported by several studies (Roh *et al*. 1973; Acheson and Collart 1975; Solomon 1983). One such survey (Acheson and Collart 1975) suggests division into two groups (although such distinction is not supported by others):

1. nodal generalized osteoarthritis;
    a. nodes
    b. distal interphalangeal joint involved more than proximal interphalangeal joint
    c. marked female preponderance
    d. familial aggregation
2. non-nodal generalized osteoarthritis;
    a. proximal interphalangeal joint involved more than distal interphalangeal joint
    b. more equal sex distribution.

The inevitable problem that arises is that one or just a few Heberden's nodes and limited interphalangeal joint osteoarthritis are common, often asymptomatic findings in the elderly, and when should the title 'nodal generalized osteoarthritis' be applied? Although criteria for this have not been defined, some studies suggest that involvement of even a single interphalangeal joint may be important (Croft *et al*. 1992). Polyarticular hand osteoarthritis has been reported to be associated with certain intra-articular patterns of large joint osteoarthritis, for example concentric as opposed to superior pole osteoarthritis of the hip (Marks *et al*. 1979). This evidence supports the notion of nodal generalized osteoarthritis as a distinct condition.

Erosive ('inflammatory') osteoarthritis

This relatively uncommon condition is characterized by the following:

1. hand interphalangeal joint involvement;
2. often florid inflammatory component;
3. radiographic subchondral erosive change;
4. tendency to interphalangeal joint ankylosis.

The condition clinically resembles nodal generalized osteoarthritis in beginning as an additive polyarthritis of finger joints (Ehlich 1972). Inflammatory symptoms and signs, however, are often marked though episodic. Unlike nodal generalized osteoarthritis, proximal and distal interphalangeal joints are equally involved and, less frequently, index and middle metacarpophalangeal joints. Interphalangeal joint instability, which is rare in nodal generalized osteoarthritis, is common and, since there is also occasional spontaneous ankylosis of one or a few interphalangeal joints, the prognosis for hand function is less favourable than nodal generalized osteoarthritis (Pattrick *et al*. 1989b). The principal hallmark of this condition is the presence of subchondral erosive change (Fig. 17) which may lead to a 'gull's wing' appearance as remodelling occurs (Fig. 18). Early, florid subchondral erosive change is easily recognized, but lesser degrees of erosion, particularly in established cases, may prove difficult to distinguish from cysts and subchondral bony change of nodal generalized osteoarthritis. Indeed, although described as a separate clinical subset, recent data has questioned whether this is really the case or whether erosive osteoarthritis is merely an extreme end of the spectrum of nodal osteoarthritis (Cobby *et al*. 1990) although there is no evidence for predisposition to generalized osteoarthritis in this subset.

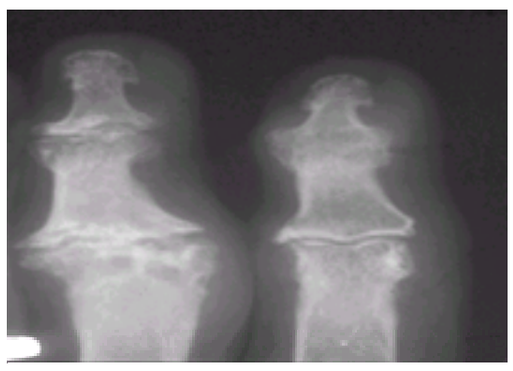


Fig. 17 Subchondral erosive change in erosive osteoarthritis.

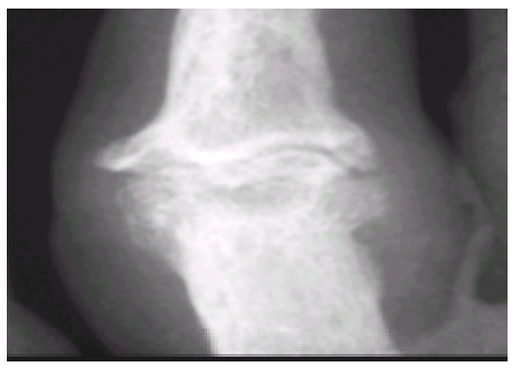


Fig. 18 'Gull's wing' deformity in erosive osteoarthritis.

Microscopically the synovium is infiltrated with lymphocytes and monocytes, and pannus may be seen. The nature of this inflammatory, destructive condition is unknown although it is of interest that similar hand changes may occur in some patients with Sjögren's syndrome. Of note, perhaps, is that 15 per cent of 170 patients presenting with Sjögren's syndrome and erosive osteoarthritis subsequently developed seropositive rheumatoid arthritis (Ehlich 1975).

Large joint osteoarthritis

Knee osteoarthritis

The knee is a commonly affected site. Involvement is usually bilateral, particularly in women and in the elderly. Although osteoarthritis may affect the knee as a mono- or pauciarticular problem, this site shows strong association with hand osteoarthritis (Kellgren and Moore 1952; Acheson and Collart 1975; Cushnaghan and Dieppe 1991). Population data demonstrates most frequent involvement of the medial tibiofemoral compartment (Fig. 19), severe bone and cartilage attrition at this site giving rise to the characteristic varus deformity (Fig. 20). The patellofemoral compartment, however, is often omitted from large studies, but when included it appears to be equally if not more commonly involved than the medial compartment, both symptomatically and radiographically (McAlindon *et al*. 1992) (Fig. 21).

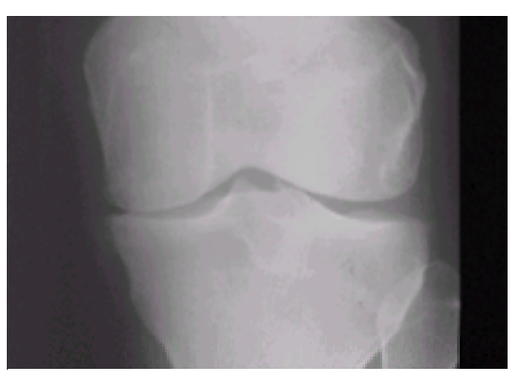


Fig. 19 Predominant medial compartment osteoarthritis at the knee (standing anterioposterior radiograph).

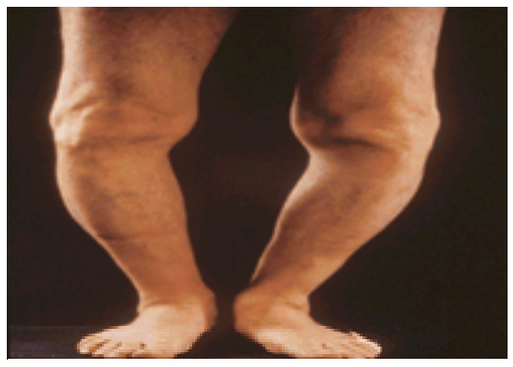


Fig. 20 Typical varus deformity of knee osteoarthritis.

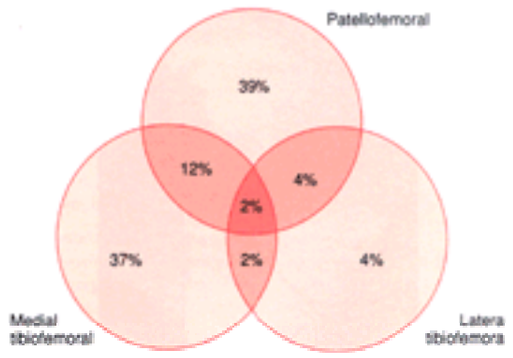


Fig. 21 Sites of radiographic involvement of knee osteoarthritis in patients with radiographic change in a community survey in the United Kingdom. Derived from McAlindon *et al*. (McAlindon *et al*. 1992).

Risk factors for development of knee osteoarthritis include previous trauma (e.g. meniscectomy), obesity (Felson 1988), generalized osteoarthritis, distal femoral dysplasia (Cooke 1985), and female gender and repetitive occupational knee bending (Felson *et al*. 1991). Smoking may be protective at this site (Felson 1988). Prognosis is discussed further below.

Recently, differing patterns of risk factors have been described in different compartments within the same joint. For example in some studies tibiofemoral disease has been more strongly linked to previous trauma and male gender whereas patellofemoral disease is more commonly symmetrical and occurs in women.

Hip osteoarthritis

Study of this joint has been particularly bedevilled by diverse classification systems which have reflected the differing interests of both orthopaedics and rheumatology at this site. Although hip involvement may occur in the context of nodal generalized osteoarthritis or involvement at other large joint sites, subdivision is usually made primarily on the basis of local radiographic patterns. Two groups have particularly been emphasized:

Superior pole osteoarthritis

The commonest pattern, this is characterized by focal cartilage loss in the superior part of the joint (Fig. 22). Osteophyte formation is most prominent at the lateral acetabular and medial femoral margins, often in combination with thickening (buttressing) of the cortex of the medial femoral neck (Fig. 23). Subchondral sclerosis and cyst formation on both sides of the narrowed joint may be marked. Originally it was suggested that this pattern is:

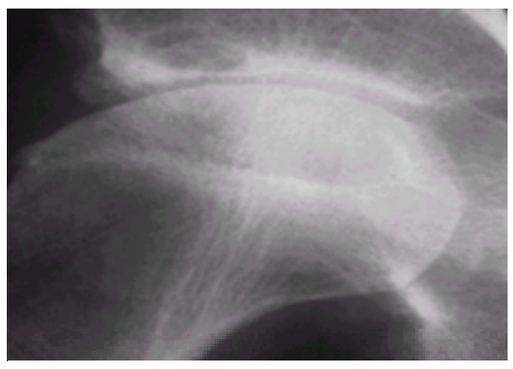


Fig. 22 Superior pole pattern of hip osteoarthritis (early) showing localized reduction in interosseous distance with focal sclerosis at the superior part of the joint, minor femoral osteophyte, and acetabular roof cyst.

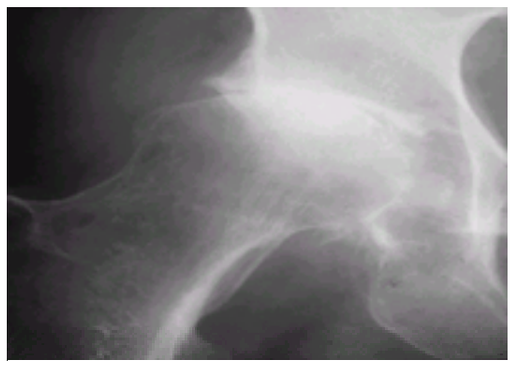


Fig. 23 Superior pole osteoarthritis (late) showing superolateral femoral head migration, bone attrition, and 'buttressing' medial periosteal osteophyte of the femoral neck.

1. more common in men;
2. mainly unilateral at presentation;
3. likely to progress, with superolateral (Fig. 13) or superomedial femoral migration;
4. commonly secondary to local structural abnormality.

Recent studies have questioned whether there is often an underlying structural abnormality (Croft *et al*. 1991). In a hospital based population, progression has been confirmed to be more likely than in other patterns (Ledingham *et al*. 1993).

Central (medial) osteoarthritis

This less common pattern shows more central joint space loss, with less prominent femoral neck buttressing. It is suggested that this pattern is:

1. more common in women;
2. commonly bilateral at presentation;
3. the pattern that particularly associates with nodal generalized osteoarthritis;
4. less likely to progress (with axial or medial migration) (Fig. 24).

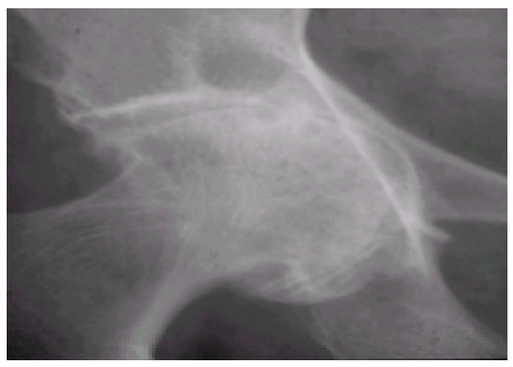


Fig. 24 Medial pole osteoarthritis, showing progression to a protrusio acetabuli deformity.

Other patterns are described (e.g. concentric) and many patients have 'indeterminate' radiographic patterns. Most disease is symmetrical unless there is a structural alteration in joint loading (Ledingham *et al*. 1992). Differentiation of these patterns does appear warranted since there may be differences in the association with hand osteoarthritis (Solomon 1976; Marks *et al*. 1979), and in prognosis (Danielsson 1964; Ledingham *et al*. 1992).

Suggested risk factors for development of hip osteoarthritis include previous hip disease (e.g. Perthes', slipped femoral epiphysis), acetabular dysplasia, avascular necrosis of the femoral head, severe trauma, generalized osteoarthritis, and occupation, particularly farming. The natural history of symptomatic hip osteoarthritis, as with osteoarthritis in general, is poorly documented, and it is impossible to predict outcome for the individual.

Crystal-associated subsets (see also Chapter 5.16)

A number of particles are commonly identified in synovial fluid and other tissues from osteoarthritis joints, most notably calcium pyrophosphate dihydrate and apatite (i.e. carbonated hydroxyapatite and other basic calcium phosphates) (Fig. 25). The origin and role of such particles in the osteoarthritis process remain unknown (Doherty and Dieppe 1988; Dieppe *et al*. 1988). By analogy with urate crystals in gout, it was initially assumed that such crystals were injurious and the cause of specific 'crystal deposition disease' (McCarty 1976). Certainly calcium pyrophosphate dihydrate and apatite are demonstrably inflammatory agents; being particulate they might also exert deleterious mechanical effects by deposition within cartilage and by acting as wear particles at the surface. However, the not uncommon occurrence of these crystals in asymptomatic, otherwise normal joints, the lack of association with specific locomotor disease, and the critical dependence of crystal identification on the technique used to detect them (Swan *et al*. 1994) has questioned such a direct pathogenic role. Multiple factors may influence deposition of these crystals, and it currently seems that calcium pyrophosphate dihydrate and apatite, in the context of osteoarthritis, most likely reflect underlying metabolic or physical facets of the process, thus potentially acting as markers for differing forms of joint response. In some situations (e.g. pseudogout) the crystals may provoke inflammation; in most others the crystals exert little phlogistic or mechanical effect, and are protected by protein coating from direct interaction with cell membranes and mediators of inflammation.

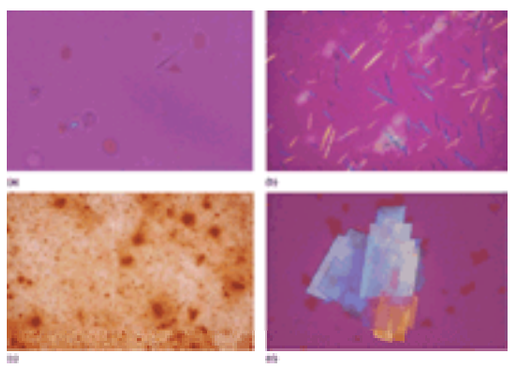


Fig. 25 (a) calcium pyrophosphate dihydrate, (b) monosodium urate monohydrate, (c) apatite, and (d) cholesterol crystals.

Pyrophosphate arthropathy

Calcium pyrophosphate dihydrate crystal deposition is the commonest cause of chondrocalcinosis (calcification of fibro- and hyaline cartilage). Although McCarty has proposed a complex clinical classification of 'pseudo-' syndromes (McCarty 1976) (Table 2), the simpler term 'pyrophosphate arthropathy' is often used to encompass those cases with accompanying arthritis. The three common clinical presentations are as:

Type	Pseudosyndrome	Presentation
A	Pseudogout	Acute or subacute synovitis
B	Pseudorheumatoid	Subacute attacks often with chronic systemic upset
C	Pseudo-osteoarthritis	Osteoarthritic change with superimposed acute synovitis
D	Pseudo-osteoarthritis	Osteoarthritic change without superimposed acute synovitis
E	Asymptomatic chondrocalcinosis	Asymptomatic radiographic finding
F	Pseudoneuropathic	Rapidly destructive often atrophic joint disease

Table 2 Classification of calcium pyrophosphate dihydrate related 'pseudosyndromes'

1. acute synovitis;
2. chronic arthropathy;
3. an incidental finding.

Other presentations are rare.

Acute synovitis ('pseudogout')

This is a common cause of acute monoarthritis in the elderly. Acute attacks may be the only manifestation of otherwise asymptomatic calcium pyrophosphate dihydrate deposition but in older patients, particularly women, they may often be superimposed upon a background of chronic symptomatic arthropathy. Any joint may be involved including the first metatarsophalangeal joint, 'pseudopodagra,' but the knee is by far the commonest site, followed by the wrist, shoulder, ankle, and elbow. Concurrent attacks in more than one joint are uncommon, and polyarticular attacks are unusual.

The typical attack develops rapidly with severe pain and swelling which is maximal within 6 to 24 h of onset. Overlying erythema is common and examination reveals a very tender joint with signs of marked synovitis (increased warmth, tense effusion, joint line tenderness, and restricted movement with stress pain). Fever is common, and elderly patients particularly may appear unwell and mildly confused, especially with knee or multiple joint involvement. Aspirated synovial fluid is inflammatory (low viscosity, turbid) and often blood stained. The cell count is usually very high, with predominance of neutrophils (>90 per cent). The diagnosis is confirmed by demonstration of synovial fluid crystals particularly if these are intracellular. Calcium pyrophosphate dihydrate crystals are poorly visualized by plain light microscopy, but can be seen under compensated polarized light microscopy (× 400). Calcium pyrophosphate dihydrate crystals (Fig. 25) are recognized by:

1. morphology—predominantly rhomboid or rod-shaped, occasionally needle-shaped;
2. size—approximately 2 to 10 μm long;
3. weak positive birefringence;
4. inclined extinction—15 to 20°.

'Twinning' of crystals, leaving a chip at one corner, is occasionally seen. Calcium pyrophosphate dihydrate crystals are less easily identified and often less numerous than urate crystals. They may often be missed unless carefully sought. Examination of a spun deposit may increase the detection rate. As with other synovial fluid particles, the use of polarized light microscopy is associated with a false positive and negative rate but identification by more definitive analytical means, such as infra-red spectrophotometry, electron microscopy, and X-ray diffraction, although ideal, is impractical for routine diagnostic purposes. The main differential diagnosis of pseudogout is sepsis and indeed these may coexist. Synovial fluid gram stain and culture should always be undertaken.

Acute attacks are self limiting and usually resolve within 1 to 3 weeks. Most episodes develop spontaneously but several provoking factors are recognized which may precede the attack by 1 to 3 days—the commonest being stress response to intercurrent illness (Table 3). Calcium pyrophosphate dihydrate crystals principally form in fibro- and hyaline cartilage, and it is 'shedding' of preformed, naked (i.e. not protein coated) crystals into the joint space, rather than acute crystallization, that is thought to be the mechanism of the attack. This is the one clear instance where calcium pyrophosphate dihydrate crystals are the likely cause of inflammation and

arthritis.

Intercurrent illness (e.g. chest infection)
Direct trauma to joint
Surgery (especially parathyroidectomy)
Blood transfusion, parenteral fluid administration
Institution of thyroxine replacement therapy
Joint lavage

NB Most cases develop spontaneously.

Table 3 Factors that may trigger acute pseudogout

Chronic pyrophosphate arthropathy

This common subset has several characteristics including:

1. predominance in elderly females;
2. an often florid inflammatory component possibly with superimposed acute attacks;
3. particular involvement of the knee;
4. frequent involvement of joints and joint compartments uncommonly affected by 'sporadic osteoarthritis';
5. frequent 'hypertrophic' radiographic appearance;
6. calcification of articular structures;
7. synovial fluid calcium pyrophosphate dihydrate crystals.

Large- and medium-sized joints are principally involved, with the knees being the most usual and severely affected site, followed by wrists, shoulders, elbows, hips, and midtarsal joints. In the hand, metacarpophalangeal joints (particularly index and middle) are the commonest, most severely affected. Symptoms are usually restricted to just a few joints, though single or multiple joint involvement also occurs; acute attacks may be superimposed upon chronic symptoms. Affected joints show signs of osteoarthritis (bony swelling, crepitus, restricted movement) and varying degrees of inflammation. Synovitis may be marked and is usually most evident at the knee, radiocarpal, or glenohumeral joints. Knees typically demonstrate bi- or tricompartmental involvement, with marked, usually predominant patellofemoral disease. In severe cases fixed flexion with either valgus or varus deformity may occur. Examination often reveals more widespread but asymptomatic joint abnormality; nodal generalized osteoarthritis, for example, is a common accompaniment.

The radiographic changes of this arthropathy are basically those of osteoarthritis with cartilage loss, sclerosis, cysts, osteophyte, and osteochondral bodies. Characteristics which may, however, permit distinction include:

1. atypical joint and intra-articular distribution compared to uncomplicated osteoarthritis;
2. often prominent, exuberant osteophyte and cyst formation (particularly at the knee).

These combined features may present a distinctive 'hypertrophic' appearance and distribution that suggest calcium pyrophosphate dihydrate even in the absence of radiographic chondrocalcinosis (Fig. 26). Many cases of pyrophosphate arthropathy, however, appear not dissimilar to 'uncomplicated' osteoarthritis. Furthermore, since nodal generalized osteoarthritis often coexists it is common to find otherwise typical osteoarthritis changes in some joints, with more distinctive changes of pyrophosphate arthropathy at others. It is probable, therefore, that pyrophosphate arthropathy is part of the spectrum of osteoarthritis rather than a truly distinct entity. As discussed further below the presence of crystals may tell us something about the underlying pathophysiological processes at work in the joint.

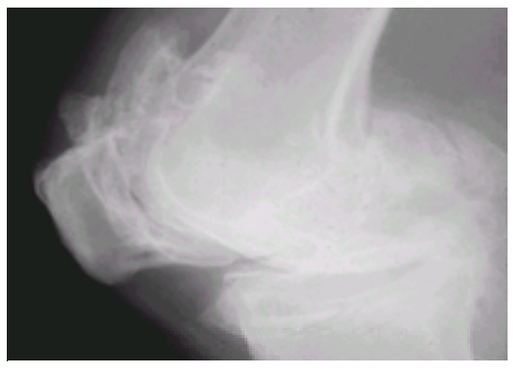


Fig. 26 Exuberant osteophytosis in patellofemoral compartment in pyrophosphate arthropathy.

P>Radiographic calcification may affect several joint tissues, but need not be present for the diagnosis of pyrophosphate arthropathy. Chondrocalcinosis most commonly affects fibrocartilage, particularly knee menisci, wrist triangular cartilage, and symphysis pubis, but it also occurs in hyaline cartilage, particularly in the knee, glenohumeral joint, and hip. It appears as thick linear deposits parallel to and separate from subchondral bone (Fig. 27). Chondrocalcinosis may be localized, for example to one knee, but usually affects several joints. If absent from knees, wrists, and symphysis pubis it is unlikely to be present elsewhere and thus views of these regions are often recommended for screening for its presence. Capsular and synovial calcification is less common, and usually most obvious at metacarpophalangeal joints (Fig. 28) and the knee. Calcium pyrophosphate dihydrate deposition can occur in tendons, particularly the Achilles, triceps, and obturators, and is typically linear and extensive. This is in comparison to apatite deposition which is typically discrete and nummular. Diffuse calcification of bursae such as the subacromial, olecranon, and retrocalcaneal is an occasional finding. Chondrocalcinosis and calcification are both dynamic features that may increase or decrease with time. Chondrocalcinosis may become less evident particularly if cartilage thickness is lost, or if crystals are 'shed' from cartilage during acute or recurrent inflammatory episodes.

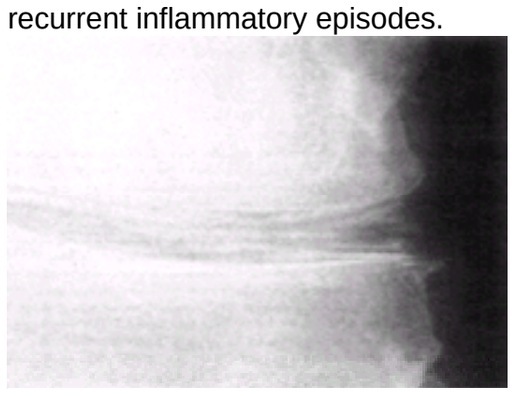


Fig. 27 Isolated chondrocalcinosis of both hyaline (linear, parallel to bone) and fibrocartilage (triangular calcification).

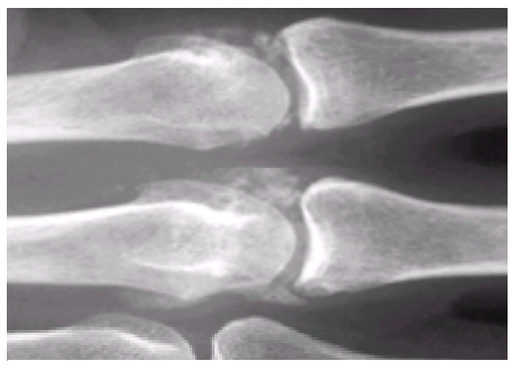


Fig. 28 Metacarpophalangeal joint radiograph in patient with pyrophosphate arthropathy showing calcification of synovium/capsule.

Incidental finding

Isolated chondrocalcinosis is a common age-associated phenomenon and as such is often observed as an incidental radiographic finding in the elderly. The only large population-based radiographic survey to study this is the Framingham study (Felson *et al*. 1989) which in the age range 63 to 93 showed an increasing overall prevalence with age ranging from 3 per cent in those aged <70 years, to 27 per cent in those >85 years (Fig. 29). A female preponderance was confirmed (relative risk 1.33), but was less pronounced than that derived from patient series. Although less well documented, it is likely that asymptomatic pyrophosphate arthropathy, like uncomplicated osteoarthritis, is common.

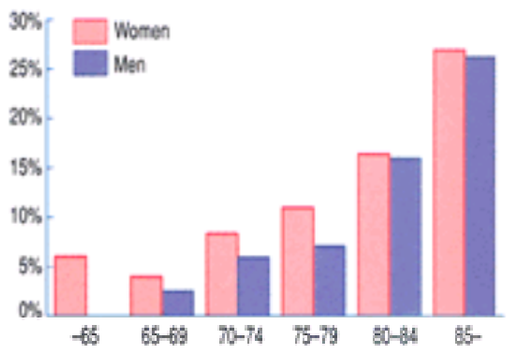


Fig. 29 Prevalence of radiographic chondrocalcinosis by age group and gender in a United States community population. Derived from Felson *et al*. (Felson *et al*. 1989).

Uncommon presentations

Polymyalgia rheumatica

This may be suggested if marked proximal stiffness accompanying glenohumeral and polyarticular involvement occurs.

Ankylosing spondylitis

Severe spinal stiffness, particularly in certain familial forms, may cause 'pseudo-ankylosing spondylitis'. True spinal ankylosis occurs in certain Chilean families.

Meningitis

Acute attacks in axial sites may cause self-limiting meningitic episodes. It is difficult to confirm these episodes although improvements in imaging has led to better recognition of such calcium pyrophosphate dihydrate deposits.

Tendinitis and tendon rupture

Acute inflammatory episodes relating to calcium pyrophosphate dihydrate deposition in tendons are described for triceps, flexor digitorum, and Achilles tendons, and tenosynovitis is reported for hand flexors and extensors. Tendon rupture is a rare complication.

Carpal tunnel syndrome

This may result from flexor tendon involvement or from wrist synovitis. Less frequently, combined median and ulnar nerve entrapment at the wrist may occur.

Bursitis

This may occur in the olecranon, infra-patellar, and retrocalcaneal bursae but is uncommon and is usually associated with widespread pyrophosphate arthropathy.

Tophi

Tophaceous (tumoral) calcium pyrophosphate dihydrate deposition is rare, but has been reported in both intra- and periarticular locations. Such deposits are usually solitary and develop in areas of chondroid metaplasia without predisposing metabolic abnormality or evidence of calcium pyrophosphate dihydrate deposition elsewhere. They produce spotty radiographic calcification and are commonly biopsied as malignant lesions, the diagnosis following histological examination (Fig. 30).

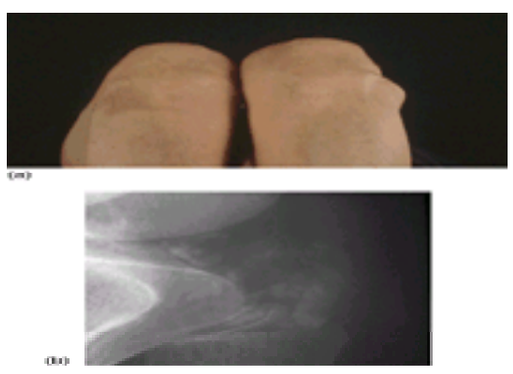


Fig. 30 Tophaceous pyrophosphate deposition superolateral to the tibiofibular joint; (a) clinical picture and (b) radiograph of the same site.

Associations of calcium pyrophosphate dihydrate deposition

A number of associations of calcium pyrophosphate dihydrate crystal deposition are recognized (Table 4). The mechanisms that underpin them are ill understood. Several possible mechanisms may be important for crystal formation including:

Positive	Negative
Ageing	Rheumatoid arthritis
Familial predisposition	
Joint insult, osteoarthritis	
Metabolic disease	

Table 4 Associations of calcium pyrophosphate dihydrate crystal deposition

1. changes, usually an increase, in relevant solute concentration;
2. presence of nucleating factors;
3. presence or absence of promotors or inhibitors of crystal growth;
4. the kinetics of crystal removal and dissolution.

In relation to calcium pyrophosphate dihydrate there is evidence to support mechanisms involving an increase in ionic product (calcium × pyrophosphate) and/or influence by promoters or inhibitors of crystal formation/growth in the cartilage matrix, the site of the majority of crystal formation. The precise nature of such matrix factors, however, is not known.

Age

Ageing is perhaps the single major predisposing factor. Indirect association could come from the positive correlation with osteoarthritis, but chondrocalcinosis commonly occurs in otherwise normal cartilage and this association remains unexplained. Inorganic pyrophosphate levels in synovial fluid from normal knees show no increase with age, suggesting that age-related alteration in matrix factors are most likely to be responsible.

Genetic

Chondrocalcinosis is reported from most countries and racial groups but only one study has defined a specific racial predisposition. Genetic predisposition, however, is well described from several countries and different ethnic groups including Czechoslovakia, Chile, Holland, France, Canada, Germany, Sweden, United States, Spain, Japan, Israel, and United Kingdom (Doherty *et al*. 1991a). Two main clinical phenotypes have been emphasized:

1. early onset in the third and fourth decades, florid polyarticular chondrocalcinosis, and varying severity of arthropathy, ranging from mild to severe, destructive;
2. late-onset oligoarticular chondrocalcinosis and arthritis, mainly affecting the knee, that is indistinguishable from sporadic pyrophosphate arthropathy.

The latter form may be more common than generally recognized since the late onset of clinical expression presents difficulties in confirming familial disease. An association with benign childhood fits is reported in one United Kingdom family with early onset polyarticular chondrocalcinosis. Autosomal dominant inheritance occurs in most families but the genetic basis has not been identified. A primary cartilage abnormality is supported by histological study of Swedish and Japanese cases, the former demonstrating proteoglycan depletion that precedes crystal formation, the latter showing hypertrophic lipid-laden chondrocytes in areas of calcium pyrophosphate dihydrate deposition. Conversely a generalized abnormality of pyrophosphate metabolism is suggested by two reports of increased intracellular pyrophosphate concentration (skin fibroblasts or transformed lymphocytes) in French and American kindreds. This finding is not present in five United Kingdom kindreds (Doherty *et al*. 1991a), thus suggesting differing genetic mechanisms in different kindreds.

Metabolic and endocrine

Although numerous metabolic associations are described, many reflect no more than chance concurrence of common age-related conditions (Table 5). The strongest evidence relates to hyperparathyroidism and haemochromatosis. Less certain evidence implicates hypothyroidism. With rare conditions convincing evidence is provided by occurrence of premature chondrocalcinosis in just a few cases, as with hypophosphatasia and hypomagnesaemia. The status of other previously implicated diseases such as Wilson's disease, ochronosis, and acromegaly remains unclear (Jones *et al*. 1992). These associations are rationalized through putative effects on pyrophosphate metabolism, extrapolated largely from *in vitro* data. Suggested mechanisms include:

	Chondrocalcinosis	Chronic arthritis
Hypophosphatasia	√	×
Hypomagnesaemia	√	×
Hyperparathyroidism	√	×
Haemachromatosis	√	√
Hypothyroidism	√	×
Gout	?	×
Acromegaly	?	×
Familial hypocalciuric hypercalcaemia	?	×
X-linked hypophosphataemic rickets	?	?
Wilson's disease	×	×
Ochronosis	×	×
Diabetes mellitus	×	×

Table 5 Metabolic diseases associated with calcium pyrophosphate dihydrate deposition

1. reduced breakdown of pyrophosphate by alkaline phosphatase due to;
   a. reduced enzyme levels, in particular hypophosphatasia
   b. presence of enzyme inhibitors such as calcium, iron, copper
   c. lack of enzyme cofactors such as magnesium
2. enhanced crystal nucleation e.g. by iron and copper;

3. increased calcium concentration e.g. hyperparathyroidism;
4. increased pyrophosphate production through parathyroid hormone stimulation of adenylate cyclase e.g. hyperparathyroidism;
5. decreased crystal solubility e.g. hypomagnesaemia;
6. altered cartilage matrix promoting crystal formation e.g. hypothyroidism.

Effects on pyrophosphate metabolism are certainly supported by elevated synovial fluid pyrophosphate levels in structurally normal knees of patients with hyperparathyroidism, haemochromatosis, or hypomagnesaemia, and elevated urinary pyrophosphate levels in hypophosphatasia (Doherty *et al*. 1991b). Mechanisms other than effects on pyrophosphate may also operate: bone and cartilage changes may be marked in these disorders, and in those with arthropathy, calcium pyrophosphate dihydrate deposition may be secondary to joint damage, mediated via alterations in matrix factors.

Joint insult

Several observations support a relationship between preceding joint insult and subsequent development of chondrocalcinosis. These include the high frequency of premature chondrocalcinosis localized to knees that have undergone menisectomy or surgery for osteochondritis dissecans as well as reports of localized pyrophosphate arthropathy as a late complication of juvenile chronic arthropathy, joint instability, and trauma. However, the apparent negative association between calcium pyrophosphate dihydrate deposition and rheumatoid arthritis, suggests that the primary association of calcium pyrophosphate dihydrate is with hypertrophic tissue response/osteoarthritis rather than with joint damage *per se*. Indeed those few patients with coexistent rheumatoid arthritis and calcium pyrophosphate dihydrate deposition often demonstrate an atypical, hypertrophic response with exuberant osteophytosis and remodelling suggestive of an essentially reparative response (Fig. 31).

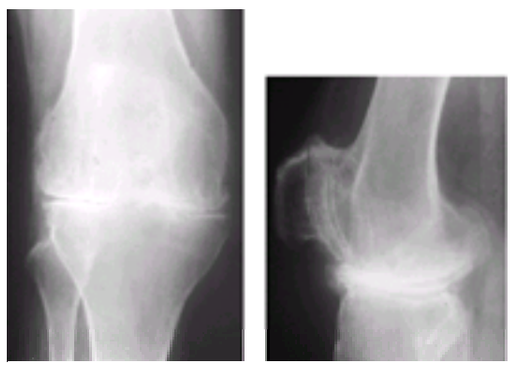


Fig. 31 Rheumatoid arthritis at the knee modified by calcium pyrophosphate dihydrate deposition. There is marked pan-compartmental joint space loss but, in addition, there is marked osteophytosis and little osteopenia.

Apatite associated arthropathy

This uncommon condition has the following characteristics:

1. confinement to elderly, predominantly female subjects;
2. localization to large joints;
3. rapid progression of arthropathy with instability;
4. marked attrition of cartilage and bone;
5. abundant apatite in synovial fluid and synovium.

This condition has a number of synonyms, including 'Milwaukee shoulder', 'basic calcium phosphate deposition disease', and 'analgesic hip'. Typical patients are elderly women with rapidly progressive arthropathy of the hip, shoulder, or knee. Usually only one or a few joints are affected. Onset is often quite sudden and within a few weeks or months the patient has severe rest and night pain, and shows large, cool effusions, with gross instability. Aspirated synovial fluid is often blood-stained and shows retained, high viscosity and only a modest increased cellularity. Alazarin red S staining at acidic pH shows multiple calcium containing aggregates (Fig. 25), confirmed as apatite, most commonly carbonate substituted hydroxyapatite, by more definitive means. The principal radiographic features are marked attrition of cartilage and bone, with a paucity of osteophyte and sclerosis, that is a markedly 'atrophic' appearance (Fig. 32). The differential diagnosis includes sepsis, an atrophic neuropathic joint, and late avascular necrosis.

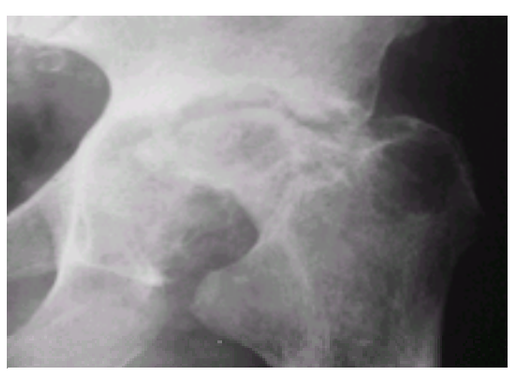


Fig. 32 Radiograph of apatite associated destructive arthritis of the hip, showing apparent increase in joint space (non-loaded film), marked loss of bone (femoral and acetabular components), and minimal bone response.

The pathogenesis of this condition, particularly as regards the role of the apatite aggregates, is controversial (Halverson and McCarty 1988). McCarty and colleagues have emphasized the presence of activated collagenase in synovial fluid, a proliferative response to crystals in the synovium, and the frequency of accompanying periarticular calcification. They therefore suggest that apatite that has been deposited in capsule and periarticular structures is enzymatically 'strip-mined', the free apatite then interacting with synoviocytes, resulting in further collagenase release, further strip-mining, and progression of arthropathy and instability via an 'amplification loop'. Others, however, have not confirmed increase in collagenase activities, and suggest that the apatite primarily originates from subchondral and marginal bone. The non-specific finding of apatite in varying quantities in other arthropathies, and even in small amounts in normal joints, is consistent with this interpretation. The speed of onset and progression, lack of overt inflammation, marked bone loss, specificity to certain anatomic sites, radiographic similarity to late avascular necrosis, and neuropathic joints support the contention that this arthropathy reflects a widespread nutritional catastrophe for the joint, possibly initiated by age-related compromise of subchondral bone blood flow. The large amount of observed apatite may thus simply reflect the rapidity of bone damage.

Other hypotheses seem less plausible; for example the suggestion that this condition results from non-steroidal anti-inflammatory drug usage and is a specific 'iatrogenic Charcot arthropathy', seems unlikely on current evidence (Doherty 1989).

Mixed crystal deposition

Comparison of clinical and radiographic features of typical calcium pyrophosphate dihydrate and apatite associated arthropathies shows marked contrasts, though each clearly falls within the spectrum of 'osteoarthritis'. The finding of both calcium pyrophosphate dihydrate and apatite in the joints of some patients is common

although there is little evidence that their combined presence is associated with particular clinical or radiographic characteristics, apart from a possible tendency to a more destructive, widespread osteoarthritis at the knee (Dieppe *et al*. 1988). Rather than classifying joints containing calcium pyrophosphate dihydrate, apatite, or calcium pyrophosphate dihydrate plus apatite as separate conditions it seems more reasonable to view these calcium crystals as markers of varying processes within the osteoarthritis joint, with calcium pyrophosphate dihydrate crystals marking a tendency to hypertrophic response, and plentiful apatite a tendency to atrophic response to insult (Fig. 33). If this is the case and osteoarthritis is regarded as a spectrum with varying balance between chronic injury and repair, frequent concurrence of both calcium crystals (in the same joint, in different joints of the same individual, at the same time or during differing phases) would be expected. In joints showing preponderance of either hypertrophic or atrophic processes (i.e. the two ends of the spectrum) the likelihood of a single crystal species would be higher. Although this may be true for typical cases, several reservations are necessary:

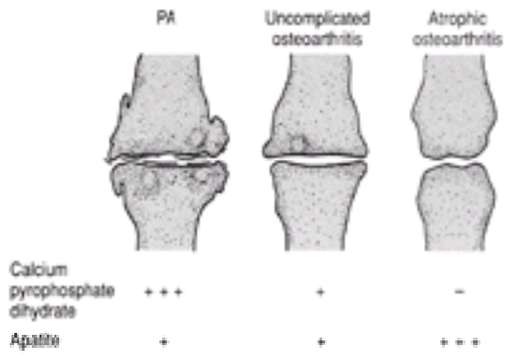


Fig. 33 Diagrammatic representation of calcium pyrophosphate dihydrate and apatite occurring as particulate 'markers' of a tendency to hypertrophic and atrophic osteoarthritis respectively.

1. Crystal identification is not an all-or-nothing phenomenon and is critically dependent on the experience and diligence of the observer as well as the methods employed (Swan *et al*. 1994).
2. There may not be distinct associations with different crystal types. In a study examining unselected osteoarthritis patients no real distinct characteristics, either clinical or radiographic, could be associated with the presence or absence of different crystal types (Pattrick *et al*. 1993).

Osteoarthritis at other joint sites

Selection of osteoarthritis for certain joint sites is striking. For example compared to osteoarthritis of interphalangeal joints, hips, or knees, involvement of the elbow, glenohumeral joint, or ankle is unusual and principally confined to the elderly.

Osteoarthritis of spinal apophyseal joints (particularly lower cervical and lower lumbar segments), first carpometacarpal, and/or first metatarsophalangeal joints is common and may occur as part of a pattern of generalized osteoarthritis (nodal or non-nodal) or as an isolated feature. In addition to nodal generalized osteoarthritis and trauma, suggested associations include:

1. metatarsus primus varus and the first metatarsophalangeal joint;
2. congenital structural anomalies, adjacent spondylosis deformans, and osteoarthritis at the apophyseal joints;
3. nodal generalized osteoarthritis, pyrophosphate arthropathy, and repetitive (principally occupational) trauma at the elbow, index, and middle metacarpophalangeal joints.

As with other arthropathies, the predilection of osteoarthritis for certain sites, with sparing of others, remains unexplained. However, one intriguing, unifying hypothesis is that human joints most commonly affected by osteoarthritis are in general those that have undergone the most rapid evolutionary change, particularly in regard to bipedal locomotion and oppositional grip (Hutton 1987). Such joints may not have had sufficient evolutionary time to fully adapt to the tasks demanded of them. They therefore have insufficient mechanical reserve, and thus fatigue and 'fail' more commonly than joints that have had longer to adapt to their new function.

Osteoarthritis as part of other disease

If, as we have discussed, osteoarthritis represents the inherent repair process of synovial joints, then osteoarthritic features would be expected to occur during certain phases of other defined arthropathies. For example in rheumatoid arthritis, osteophytosis, sclerosis, and remodelling may become prominent during later, less inflammatory periods of the disease. In such instances 'osteoarthritis' can be seen as an accompanying process of tissue response/repair rather than an acquired second condition.

The same considerations pertain to other inflammatory, metabolic, or structural arthropathies. Ochronosis and spondyloepiphyseal dysplasia, for example, are sometimes included within the umbrella of osteoarthritis since many of their radiographic features are typical of osteoarthritis. Similarly, endemic forms of osteoarthritis need consideration in their own right. For clinical purposes, however, it may still be useful to consider together conditions that may result in non-inflammatory arthropathy with radiographic changes predominantly of osteoarthritis (Table 6). An atypical presentation, however, should lead one to search for a defined, possibly treatable underlying cause. Many are rare, and have distinct clinical and radiographic features. In patients with 'osteoarthritis' one should consider specific predisposing factors when there is:

Generalized 'osteoarthritis'	(Spondylo-) epiphyseal dysplasia Ochronosis Haemochromatosis Wilson's disease Endemic osteoarthritis (e.g. Kashin–Beck disease)
Pauciarticular, large joint 'osteoarthritis'	Neuropathic joints syringomyelia—shoulders, wrists, elbows diabetes—hindfoot, midfoot tabes—knees, spine
	Acromegaly
	Avascular necrosis (mainly proximal and distal femur, proximal humerus)

Table 6 Principal conditions with presentations and radiographic changes that may simulate osteoarthritis

1. premature onset osteoarthritis, i.e. less than 45 years;
2. an atypical joint distribution, e.g. prominent metacarpophalangeal and radiocarpal involvement in haemochromatosis;
3. premature onset of chondrocalcinosis, i.e. less than 55 years;
4. florid polyarticular chondrocalcinosis at any age.

Investigations

Osteoarthritis is a diagnosis made on clinical and radiological grounds. To date there are no satisfactory criteria or specific laboratory tests. Furthermore, radiographic changes of osteoarthritis are commonly present but often asymptomatic. The problem is not usually to decide whether osteoarthritis is present but whether it is the

cause of the problem. Investigation plays little part in this decision, which should be made by a thorough clinical examination. Some investigations may be necessary to exclude alternative diagnoses or predisposing disease.

Osteoarthritis is not associated with extra-articular disease, synovitis is usually only mild or moderate, and overt immunological abnormality is not a feature. Changes reflecting an acute phase response (anaemia, thrombocytosis, elevated erythrocyte sedimentation rate and/or C-reactive protein) or overt immunological abnormality (autoantibodies, complement breakdown products) are therefore usually absent. However, evidence of a marked acute phase response occurs with pseudogout, and modest elevation of erythrocyte sedimentation rate and C reactive protein may occur in the 'inflammatory' subsets of osteoarthritis, particularly chronic pyrophosphate arthropathy. IgG rheumatoid factor (but not Rose Waaler or Latex) positivity may be more common in nodal generalized osteoarthritis than age-matched controls (Hopkinson *et al*. 1992). Presence of such modest or isolated abnormalities are therefore not against the diagnosis of osteoarthritis, particularly in elderly patients in whom coexistent disease is common. A search for predisposing metabolic or endocrine disease may be undertaken in selected patients with atypical distribution or premature-onset of osteoarthritis, atypical radiographic features, or early-onset or polyarticular chondrocalcinosis; such patients usually have other clinical or radiographic clues, and routine 'screening' for all known predisposing diseases is inappropriate.

The only investigations that are of importance in terms of determining the presence of osteoarthritis, the degree of structural and physiological change, and the presence of associated crystal deposition are plain radiographs and other forms of imaging. Synovial fluid analysis may also be useful. Here we shall only discuss selected aspects of imaging, and the potential for future biochemical 'markers' of the osteoarthritic process.

Imaging

Imaging aims to establish the diagnosis, to assess severity, delineate the likely pathology involved, and help exclude alternative diagnoses. Methods include plain radiographs, magnetic resonance imaging, ultrasound and radioisotope scanning. The objective of the user determines the most appropriate modality. Generally techniques that demonstrate structural change do so at the expense of demonstrating current biochemical/physiological activity and vice versa. As discussed below, a possible exception to this may be magnetic resonance imaging.

Plain films

Radiographs give an anatomical picture, that is they demonstrate past structural change rather than current 'disease activity'. Bony structures are readily defined but soft tissue imaging is more difficult. In particular, cartilage is not directly visualized and its thickness has to be inferred by assuming that it comprises the vast majority of the distance between two articulating bones, that is the interosseous distance. Focal cartilage loss is very difficult to detect and the bony changes reflect relatively late pathological abnormalities. Nevertheless, the availability of radiographs and their widespread current and historical use in studying osteoarthritis make them still the principal imaging modality for diagnosis, assessment, and follow-up.

Joint space narrowing, osteophyte, subchondral radiolucencies, and sclerosis are the classic radiological signs of osteoarthritis (Fig. 34). A spectrum of changes exist for each of these signs, there is no fixed correlation between one sign and another, and no one feature has really been shown to have particular discriminatory value, although osteophyte at the knee may be a relatively good predictor of knee pain in epidemiological studies (Spector *et al*. 1993). Additional features and complications may also be visualized including:

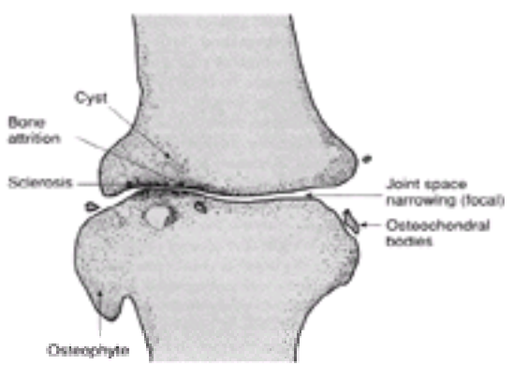


Fig. 34 Major radiographic features of osteoarthritis.

1. effusions;
2. the osseous component of osteochondral ('loose') bodies;
3. joint alignment;
4. subluxation.

Absence of features of other arthritides is expected although the radiographic diagnosis of osteoarthritis does not exclude other coexisting disease. Osteoarthritis may also be secondary to previous bone and joint disease.

Plain films facilitate subsetting of osteoarthritis by:

1. the distribution of osteoarthritis change within a joint, e.g. superior pole, medial, and concentric patterns of hip; patellofemoral and tibiofemoral knee osteoarthritis;
2. the emphasis on individual osteoarthritis features, e.g. hypertrophic versus atrophic appearance;
3. demonstration of additional features, e.g. calcification of cartilage and other joint structures in calcium pyrophosphate dihydrate deposition.

Some of these features may, however, be distorted by overlying structures and magnification effects resulting from the divergence of the X-ray beam. Attempts to improve the reproducibility and interpretability of plain radiographs have been made. Some involve standardization of positioning, the use of multiple observers, the use of different views and specialized radiographic techniques. Aspects of these are discussed further below.

Plain radiography of specific joints

Hand

The pattern of involvement of the interphalangeal joints helps to distinguish osteoarthritis and erosive osteoarthritis from rheumatoid and psoriatic arthropathy. For this purpose a single posteroanterior view of the hand may be satisfactory although it is possible such a view does not detect posterior osteophytes. The distinctive changes are (Fig. 35):

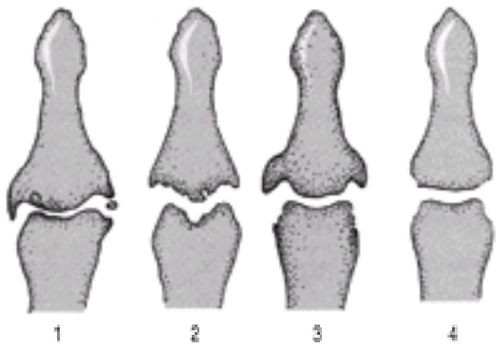


Fig. 35 The radiographic differences between (a) osteoarthritis, (b) erosive osteoarthritis, (c) psoriatic arthritis, and (d) rheumatoid arthritis.

1. osteoarthritis: bone sclerosis, focal narrowing, and lateral subluxation unaccompanied by erosion;
2. erosive osteoarthritis: changes of osteoarthritis plus erosions which are through the subchondral bone plate;
3. psoriatic arthritis: marginal erosions associated with a bone response, 'proliferative erosions'.
4. rheumatoid arthritis: marginal 'bare area' erosions unassociated with bone proliferation;

Knee

Weight-bearing films of the knee show the functional position of the limb and may allow more precise information on the extent of joint space narrowing (Messiah *et al*. 1990). To assess the less involved compartment, or to detect soft tissue laxity, stress views may be required. However, the changes of fibrillation, cartilage narrowing, and cratering are focal. Attempts to detect early change by using a semiflexed weight-bearing view have been suggested although difficulties with precise patient positioning are still evident. Recent evidence has confirmed that the patellofemoral joint is an important site of involvement of osteoarthritis (McAlindon *et al*. 1992). To date there is no consensus on the best plain radiographic method to be used to assess this joint and the choice lies between a lateral view or a 'skyline' view of the patella (Jones *et al*. 1993). Further work is necessary to establish the optimum view.

Hip

Traditionally this is taken as a non-weight-bearing anteroposterior view of the pelvis. This has the advantage of incorporating both hips on the same radiograph although gonadal dosage may be a problem. Evidence suggests that at the level of population studies such a view may be satisfactory with minimum joint space the best criteria to define a case. It is possible that this view does not provide full information of involvement of the hip since early change may be focal and posterior. The role of additional or stressed views is, however, not yet established.

Sacroiliac joints

In the sacroiliac joints, osteophyte and joint space loss may need to be distinguished from inflammatory sacroiliitis. The latter gives erosion and intra-articular ankylosis, whereas osteoarthritis gives more focal joint space narrowing and focal sclerosis with overlying osteophytes. These are usually anterosuperior or inferior, and may be identified by discontinuity of trabecular lines across the joint in contrast to continuous lines of ankylosis (Fig. 36). Usually a single anteroposterior view of the pelvis is obtained as for the assessment of the hip but such a view may be difficult to interpret since the joint margins may overlap. In this situation a coned posteroanterior view of the sacroiliac joints may help.

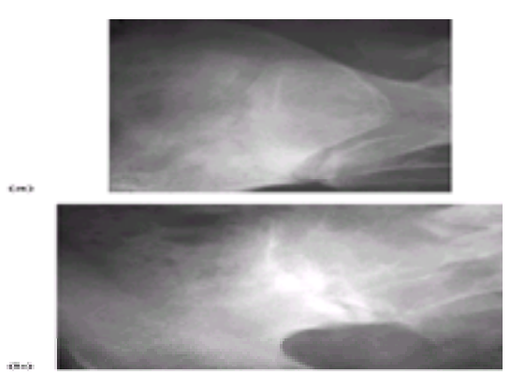


Fig. 36 (a) Sacroiliitis compared to (b) sacroiliac osteoarthritis.

Foot

A common site of osteoarthritis, this is not usually incorporated in radiographic surveys. A single posteroanterior radiograph of the foot is probably the simplest method required to demonstrate involvement of the first metatarsophalangeal joint. Specialized views would be required to demonstrate involvement of the tarsus and ankle joints but these are not usually employed in studies.

Spine

This can be a difficult area to assess for osteoarthritis. Changes suggestive of osteoarthritis, sclerosis, joint space narrowing, and osteophyte are virtually universal in middle-aged and elderly adults. These are particularly evident in the lower cervical and lower lumbar spine and may involve the facet joints and in the cervical region the uncovertebral joints (the joints of Luschka). In addition, similar changes are seen in the intervertebral discal joints. The precise views necessary to detect spinal osteoarthritis are not clear but lateral lumbosacral and cervical views are probably the most appropriate.

Disease of the spine, particularly if it involves the facet joints with subsequent osteophytosis, may result in either foraminal compression of the nerve roots or canal stenosis. Although plain views, particularly if oblique foraminal views are taken, may demonstrate narrowing (Fig. 37), more specialized techniques such as magnetic resonance imaging or computed axial tomography are usually required.

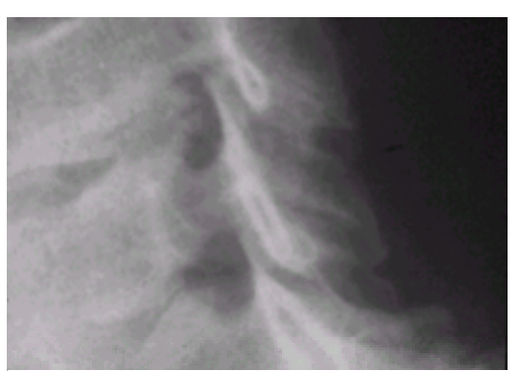


Fig. 37 Foraminal osteophytic encroachment in the neck of a patient who presented complaining of intermittent numbness and paraesthesia of the right arm.

Standardization of the assessment of plain radiographs

Attempts have been made to standardize and quantitate image analysis. Two approaches have been used. The first has used atlases of standard films against which to compare the study films (Kellgren and Lawrence 1957; Kallman *et al*. 1989; Burnett *et al*. 1994), the second has tried to assess radiographic features using automated approaches. Using an atlas, two further approaches have been employed.

The first has attempted to produce a single a grade of osteoarthritis from a combination of different radiographic features. The use of such a 'global osteoarthritis' grade is attractive since it simplifies statistical analysis and means a joint is assigned to only one of a few radiographic grades. There are dangers with such an approach. Firstly it assumes a hierarchy of change that may not exist. For example in the Kellgren and Lawrence grading system the presence of osteophyte is a *sine qua non* for the diagnosis of osteoarthritis. Yet, histologically and arthroscopically, osteoarthritis, in terms of typical focal cartilage loss, may occur in the absence of detectable osteophyte. Similarly, in moving between grades it is by no means clear that the 'developing' features are indicative of deteriorating osteoarthritis. For example the development of osteophyte has been suggested to be a potentially beneficial process acting to improve joint stability.

The alternative approach which has gained increasing favour is to grade for individual features of osteoarthritis such as osteophyte, joint space narrowing, cysts, and sclerosis. This has the advantage of not assuming a particular hierarchy of change or inter-relationship of features. However a large amount of descriptive data is generated for each joint which is difficult to combine. Simply summating scores for each feature is probably not valid. No consensus has been reached as to which approach is preferable.

In an attempt to reduce interobserver error, various manual and semiautomated methods have been evaluated particularly in the measurement of joint space. For an individual radiograph such an approach may be very reproducible particularly if an automated computerized method is used. Problems still remain regarding the obtaining of reproducible radiographs upon which to make the measurements (Spector 1995).

Modified radiographic techniques

Microfocal radiography is a magnification technique that allows higher resolution imaging, with the possibility of stereoscopic reconstruction. This technique appears particularly useful in showing early subchondral bone abnormalities and demonstrating small changes in joint space width over short periods of time (Buckland-Wright *et al*. 1990). Requirements of relatively long exposures and special equipment confine this method to a research tool but one which has particular potential in longitudinal studies (Spector 1995).

Some of the limitations of plain films can be overcome by tomography. This produces cross-sectional images that thus avoid problems of interpreting overlying structures. Images of spinal apophyseal joints are possible, and the pattern of sacroiliac disease is made clearer. In general this technique has been superseded by the newer technique of computed axial tomography.

Arthrography allows cartilage to be coated with contrast and thus be demonstrated, but this is invasive and unsuitable for routine or sequential use. It has a place, however, in demonstrating associated meniscal disease although this role is rapidly becoming superseded by magnetic resonance imaging.

Magnetic resonance imaging (see also Chapter 4.9.1)

This method uses the properties of hydrogen ions, principally those in water. Since cartilage contains a high proportion of water, it can allow cartilage to be imaged. The potential and limitations of this modality in respect to osteoarthritis are still being explored. There are several problems including:

1. limited resolution, since each image is reconstructed from information from a volume of tissue;
2. the magnetic effects of bone, particularly when juxtaposed to cartilage;
3. resolution of synovial fluid from cartilage, since both have a high water content.

However, by using different scan sequences it may be possible to resolve these problems as well as to obtain information on the process involved in osteoarthritis.

Radionuclide studies

Isotope bone scans give information on perfusion and bone activity. The mechanism of bone uptake is non-specific, but is more sensitive than radiographs at identifying involvement. Perhaps its most significant role in assessment of osteoarthritis is its ability to detect abnormalities before radiographic signs are identified, and to identify patients with 'active' disease who may go on to show progression (Hutton *et al*. 1986; Dieppe *et al*. 1993). This may prove useful in evaluation of drugs and other interventions before severe joint damage has occurred.

Ultrasound

This modality requires no radiation exposure and permits imaging of cartilage and tendons. Problems of access and sound distortion mean that its value is restricted and currently it remains a research tool. Nevertheless, it has been shown to give distinction between normal, fibrillated, and thin cartilage at the knee (Aisen *et al*. 1984). It may prove of value in the early detection of cartilage abnormalities.

Potential markers of tissue destruction, inflammation, and repair in osteoarthritis

There is considerable interest in measuring biochemical markers in synovial fluid or serum to allow the diagnosis of diseases, to follow progression of disease and response to treatment, and to determine disease prognosis and mechanism. Osteoarthritis involves active cellular processes of biosynthetic activity and matrix turnover and measurement of these parameters could allow these processes to be monitored. Such studies are based on two assumptions.

1. When cartilage is damaged there is a loss of matrix components, cytokines, or proteinases into synovial fluid, the lymphatic system, serum, and urine.
2. This loss is quantitative, with changes in concentration reflecting changes in rates of turnover.

The first of these assumptions is proven but data to support the second are confounded by unknown physiological variables. For example the concentration of fragments in synovial fluid depends not only on the rate of release from cartilage, but also on the volume and turnover of joint fluid and the speed of lymphatic drainage (all difficult to quantify). Similar problems apply to measurements in serum or urine: although they are more readily accessible compartments, the concentrations in them reflect release from all joints (normal and abnormal) and from other cartilage sites; selective, rapid elimination by the lymph nodes or liver, and the possible influence of renal and liver disease complicate interpretation. Finally, it is essential to know what a 'marker' signifies in terms of joint physiology: a reduction in the concentration of a cartilage marker after an intervention may be interpreted as reduced cartilage breakdown, it may signify inhibition of matrix synthesis or reflect the absence of any remaining cartilage within the joint (Fig. 38).

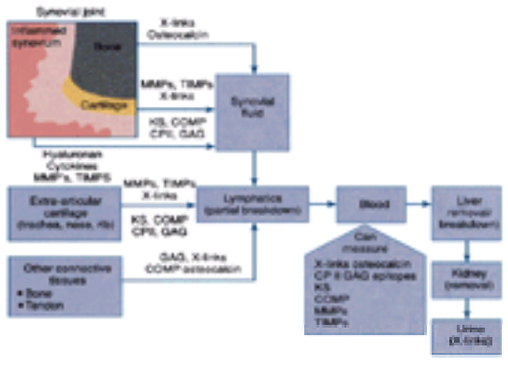


Fig. 38 Physiology of potential 'markers' of the osteoarthritic process.

Some of the biochemical products that have been investigated are listed, in Table 7, Table 8, and Table 9 along with the advantages and disadvantages of any particular parameter or method of measurement. The usefulness of these markers is not confined to osteoarthritis, as the principle of joint damage/repair is common to other conditions. At present there is no marker in biological fluids that can be used for diagnosis or for monitoring of the progression of osteoarthritis (Lohmander 1994).

Tissue	Markers	Indicates	Osteoarthritis		Rheumatoid arthritis	
			Serum	SF	Serum	SF
Synovium	Hyaluronan (hyaluronic acid, hyaluronate, HA)	Synovitis, HA synthesis stimulated by IL-1 and TNF-α	↑	ND	↑	↑
Cartilage	C-propeptide of type II collagen	Synthesis of cartilage type II collagen	↑	↕	↑	↑
	Type II collagen α chain fragments	Degradation of cartilage type II collagen	Studies in progress			
	Cartilage oligomeric protein	Synthesis and/or degradation?	ND	↕	↑	ND
	Cartilage matrix protein	Synthesis and/or degradation? only present in non-articular cartilages	ND	ND	↑	ND
Proteoglycan aggrecans	Keratan sulphate	Synthesis/degradation of aggrecan	↑	↕	↓	↑
	Core protein	Degradation	+	↕	+	↑
	Intact chondroitin sulphate epitopes e.g. 846, 3B3	Probably synthesis	↑	+	↑/0[a]	↑
Bone	Bone sialoprotein	Synthesis/degradation	Studies in progress			
	Osteocalcin	Synthesis	↑	↕	+	↑
	3-hydroxypyridinium (pyridinoline and deoxypyridinoline) cross-links	Degradation	Urine only			

ND=not determined, SF=synovial fluid, [a]acute disease only.

Table 7 Physiological correlates of matrix markers of tissue destruction, inflammation, and repair in osteoarthritis and rheumatoid arthritis (ND = not determined, SF = synovial fluid, [a] = acute disease only)

	Osteoarthritis	Rheumatoid arthritis
Hyaluronan	↑	↑
C-propeptide II	↑	↑
Keratan sulphate	↑	↓
Chondroitin sulphate epitopes	↑	?
Cross-links	↑	↑
Matrix metalloproteinases	↑	↑
Tissue inhibitors of metalloproteinases	↑	↑
Cartilage oligomeric protein	↑	↑

Table 8 Changes in matrix markers of tissue destruction, inflammation, and repair in osteoarthritis and rheumatoid arthritis

Assay	Advantages	Disadvantages
Where?		
Serum	Easy collection All patients	Dilution of marker Contribution from other tissues
Synovial fluid	Local to inflamed joint	Infrequent aspiration Small proportion of patients Low levels may result when large volumes of synovial fluid are produced
Urine		Difficult to collect 24-h urine Final level depends on rate of clearance in other compartments
How?		
Immunoassay	Rapid measurement of total level	Also measure non-functional protein
Activity assay	Measures biological activity	Often inhibitors also present, therefore sample may require fractionation

Table 9 Measurement of matrix markers of tissue destruction, inflammation, and repair (NB low levels may result when most of the cartilage tissue has been destroyed)

Cytokines

As little inflammation is present in osteoarthritic joints the measurement of cytokine levels is not helpful in osteoarthritis (Westacott *et al*. 1990).

Matrix components

Proteoglycan and glycosaminoglycan fragments

Because of their relatively rapid turnover many studies have estimated protein and carbohydrate moieties of the large, aggregating proteoglycans (aggrecan) of cartilage. Using immunoassays it has been shown that:

1. Radiographic loss of cartilage in rheumatoid arthritis correlates with a decline in the proteoglycan concentration of synovial fluid (Saxne *et al*. 1985; Poole 1993).
2. High concentrations of proteoglycan in synovial fluid occur in acute reactive arthritis and decline during remission after steroid injection (Saxne *et al*. 1986).
3. Large amounts of proteoglycan fragments occur in synovial fluid with transient synovitis and septic arthritis of the hip in children, and in adults after acute mechanical derangements of the knee (Lohmander *et al*. 1989).
4. Keratan sulphate is raised in synovial fluid in pseudogout and acute Reiter's disease, with significantly lower levels in osteoarthritis and chronic pyrophosphate arthropathy (Ratcliffe *et al*. 1988).

Thus, an increase in proteoglycan fragments in synovial fluid appears to occur only in acute inflammatory conditions, presumably reflecting accelerated breakdown.

Serum keratan sulphate is higher in osteoarthritis, particularly 'hypertrophic' osteoarthritis, than in normal controls (Thonar *et al*. 1987; Sweet *et al*. 1988), but overlap between normal and osteoarthritic individuals is marked. Studies with animal models of osteoarthritis have almost without exception failed to show significant changes in the serum keratan sulphate in early lesions. Nevertheless, the variability of the serum keratan sulphate is less within than between patients. Rather than reflecting the status of cartilage in osteoarthritic joints, elevated levels may thus signify a systemic increase in cartilage matrix turnover and ultimately prove useful in identifying individuals at risk of developing osteoarthritis.

Rather than attempting to detect the increased rate of a normal process (e.g. the release of keratan sulphate from cartilage), it would be preferable to identify a released antigen more distinctive of abnormal tissue. Epitopes that detect changed sequences in sulphation in chondroitin sulphate chains that are found infrequently in normal mature cartilage appear to be increased in experimental models. In experimental canine osteoarthritis, for example, the expression of a chain-terminal chondroitin sulphate epitope (3B3) is greatly increased in proteoglycan from operated joints as part of the pattern of increased synthesis and turnover that also results in longer chondroitin sulphate chains: the distribution of such epitopes varies greatly in chondroitin sulphates of different origin, and their expression is closely controlled during biosynthesis (Caterson *et al*. 1990). The biological significance of the chondroitin sulphate epitopes have yet to be determined. However, they may provide useful markers of processes occurring in cartilage before major joint damage.

Collagen fragments

The C-terminal propeptide of type II collagen is cleaved off at fibrillogenesis and not retained by interaction with other matrix constituents, and is therefore a molecule with the potential for identifying repair. In urine, the determination of collagen cross-links derived from cartilage (pyridinoline) and bone (pyridinoline and deoxypyridinoline) has provided a sensitive measure of tissue breakdown (Seibel *et al*. 1989). These cross-links are only present on mature collagen; thus relative amounts may indicate different forms of stages of arthropathy involving the degradation of bone and cartilage. Recent studies using immunoassays have shown that collagen crosslinks can also be shown to be raised in osteoarthritis in serum samples (Risteli *et al*. 1993; Robins *et al*. 1994).

Little progress has been made in following the release of type II, IX, or XI fragments into synovial fluid as a measure of the dismantling of the collagen fibrillar network. Caution has to be used in synovial fluid studies as often the centrifugation of the fluid to remove any neutrophils or other cells is sufficient to sediment some of the collagen fragments. Further work is needed to establish the usefulness of the measurement of collagen fragments in osteoarthritis.

Proteinases and inhibitors

Measurement of proteinases and inhibitors in serum is often difficult as levels can be low and sometimes below the level of detection (Lohmander *et al*. 1994). Collagenase, gelatinase, stromelysin, and TIMP are present in synovial fluid but are difficult to measure as the fluid has to be fractionated before measurement to allow enzyme activity to be measured whilst inhibitors are also present (Cawston *et al*. 1984). The development of immunoassays has shown that stromelysin (Lohmander *et al*. 1993), collagenase, and tissue inhibitor of metalloproteinase (TIMP) (Clark *et al*. 1994; Shinmei *et al*. 1992) are elevated in osteoarthritis synovial fluids although there is a large variation between patients and considerable overlap between disease groups. It is not yet known if high levels of proteinase or low levels of inhibitor correlate with excess joint damage. The results are also difficult to interpret as much of the metalloproteinases present are still in proenzyme form and consequently pose no threat to the joint structures until activated. Some studies have suggested that the measurement of proteinase-inhibitor complexes are more indicative of cartilage breakdown although these results are still to be confirmed in prospective studies. Raised levels of collagenase-TIMP levels in osteoarthritis are infrequent or at least below the level of detection for the immunoassays (Clark *et al*. 1994). Many studies have investigated single samples of synovial fluid from patients. Recently Lohmander *et al*. were able to show that levels of stromelysin and TIMP remained raised for long periods after traumatic injuries in knee joints (Lohmander *et al*. 1994) and this kind of study where individual patients are followed up sequentially with clinical, radiographic, and biochemical markers measured for individual joints represents the best way forward to evaluate these markers.

Diagnosis

The diagnosis of osteoarthritis is essentially clinical and many comments regarding this have already been made in the section on investigations. There are several different clinical problems.

The patient with pain

This is the commonest clinical situation. The priorities in this situation are to determine whether:

1. 'osteoarthritis' is the cause of the symptoms;
2. there is an articular or periarticular cause for the pain;
3. there are predisposing or adverse factors for the development or progression of osteoarthritis;
4. there is another underlying arthropathy;
5. how pain is being caused and modified.

Resolving these difficulties depends on a careful clinical history and examination; radiography and laboratory investigations play a relatively minor role. The examination aims to:

1. localize the site of the pain—joint line or periarticular;
2. detect the presence of clinical signs of osteoarthritis—crepitus, bony swelling, restricted range of motion;
3. define any adverse features—obesity, malalignment, abnormal usage (occupational or habitual);
4. rule out features of other arthropathy—e.g. rheumatoid arthritis, gout, seronegative spondylarthropathy;
5. determine factors aggravating the pain response—e.g. coexistent fibromyalgia, depression, sleep disturbance.

The radiograph merely serves to confirm the presence of the structural changes of osteoarthritis that may or may not be attributable to the symptoms. Since the radiograph is relatively insensitive, particularly in early disease, a normal radiograph does not rule out a diagnosis of osteoarthritis. Conversely, as has already been discussed, an abnormal radiograph is not necessarily the cause of symptoms.

The radiographic finding

It is not uncommon for patients to have radiographs taken following an acute episode of pain or trauma, particularly of the spine, and for osteoarthritic changes to be discovered. The difficulty then lies in knowing what, if anything to do regarding these findings. Undue emphasis on the structural changes can result in patients being unduly alarmed and adopting inappropriate illness behaviour. Again a careful clinical history and examination is necessary to determine the association between the current clinical problem and any structural change.

Neurological findings

Neurological finding are important, particularly in the lumbar and cervical spine where osteophytosis of the facet or apophyseal joint may lead to foraminal encroachment and subsequent nerve root compression. Peripheral nerve entrapment may also occur as a result of osteophytosis or synovitis. Possible sites for this are the ulnar nerve at the ulnar groove and the median nerve in the carpal tunnel. In such a situation, particularly in the spine, magnetic resonance imaging, computed axial tomography, and, now less commonly, contrast radiculography is necessary to fully elucidate the nature and presence of nerve compression.

As well as direct pressure on a nerve, vascular claudication may also occur, particularly in the lumbar spine, giving rise to the syndrome of spinal claudication. In this situation, exercise results in neurological signs and symptoms in the legs. The diagnosis is essentially based on the history but evidence of canal stenosis is sought, usually with computed axial tomography or magnetic resonance imaging.

In all situations, particularly in the spine, it is essential, in view of the high prevalence of abnormal structural findings in asymptomatic individuals, to ensure that there is a good match between clinical findings and demonstrated structural abnormalities.

Pathogenesis

The nature of osteoarthritis

It has been widely suggested that osteoarthritis is a process, rather than a disease, that shows variability in outcome. Support for this comes from several considerations:

1. osteoarthritis has accompanied man throughout his evolutionary history;
2. a similar process occurs in other animals that have fused epiphyses in the adult;
3. radiographic osteoarthritis is very common in adults, showing increased frequency with age;
4. in most instances osteoarthritis occurs without symptoms or disability, and its radiographic presence is not necessarily the explanation of locomotor pain;
5. symptoms relating to osteoarthritis are often phasic and may not necessarily be associated with a poor prognosis.

Such phylogenetic preservation, discordance between symptoms and structural change, and generally good outcome suggest that osteoarthritis reflects the inherent repair process of synovial joints (Fig. 39). In most cases this metabolically active process keeps pace with a variety of triggering insults and is non-progressive. In some, however, it fails to compensate, resulting in 'joint failure' (decompensated osteoarthritis) with perceived symptoms and disability. This interpretation partly explains the marked heterogeneity of osteoarthritis—a wide variety of 'insults' triggering a repair reaction but each resulting in a different pattern of involvement. As with any biological process, multiple constitutional and environmental factors may further modify the response, leading to variable outcome.

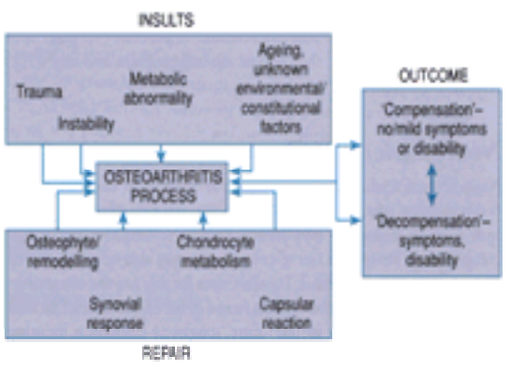


Fig. 39 Diagrammatic representation of osteoarthritis as the inherent repair process of synovial joints.

The mechanical, genetic and metabolic factors that may 'insult' the joint and thus trigger the osteoarthritis process have been discussed above. Indeed most evidence regarding the initiation of osteoarthritis in man derives from such studies since diagnosis of the early stages of sporadic osteoarthritis *in vivo* is currently not possible. The response of the joint to these insults has been studied by a variety of methods including *in vivo* animal models, *ex vivo* studies on tissue explants, and from pathological human tissue. The results from such studies and the insights which they give into the nature of osteoarthritis are discussed below.

Structure and metabolism of osteoarthritic tissues

Investigation of human osteoarthritis tissue is problematic: pathological studies often focus on late 'end-stage' osteoarthritis of surgically-derived large joints, whilst cadaveric studies examining 'early' osteoarthritis lack clinical correlates. All such studies are confounded by the heterogeneity of osteoarthritis and by restriction to single time point examination. Much work on pathophysiology therefore derives from animal models and *in vitro* experiments. Animal models often employ small quadrupeds, utilize invasive mechanical or chemical insult to stimulate joint response, or investigate hereditary forms of premature joint failure; the time scale is often short; and immature rather than mature animals may be used. Although some models (e.g. cruciate section in the adult dog) appear closer to human osteoarthritis than others, such contrasts with the clinical situation in man have raised questions as to their relevance. Rather than considering them as models of osteoarthritis, they are probably best viewed as a means of studying, in a well defined and controlled situation, dynamic biochemical and structural events at the earliest stages of joint insult. Thus, while animal models and human studies *in vitro* provide useful insights, caution is required in drawing together the available disparate data to develop a more complete understanding of pathophysiology relevant to human osteoarthritis.

Structural changes

The histological changes characteristic of osteoarthritis cartilage from humans and animal models are well described (Mankin 1974; Brandt 1988) and include:

1. reduction in stainable proteoglycan;
2. fibrillation;
3. collagen crimping;
4. chondrocyte multiplication or migration (cloning);
5. loss of cartilage (Fig. 40).

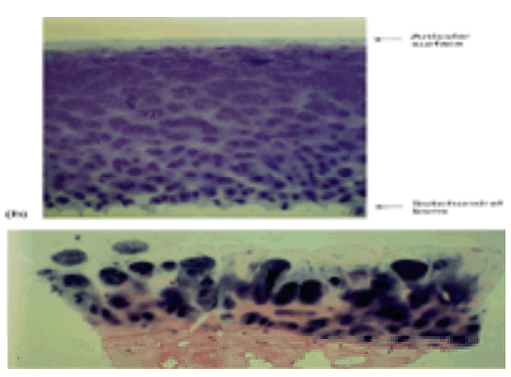


Fig. 40 (a) normal and (b) osteoarthritic human articular cartilage (toluidine blue). Note the extensive cell clustering, loss of metachromatic staining, fissuring of the tissue, and duplication of the 'tide-mark' in the osteoarthritis specimen.

Initially, localized areas of softening present a pebbled texture at the surface, followed by disruption along collagen fibre planes (tangential 'flaking', vertical 'fibrillation'). At the earliest stage of surface erosion and irregularity cartilage appears moderately hypercellular with alteration in staining quality of the matrix; the tidemark may also show irregularities and violation by blood vessels. As increasingly deep clefts form in cartilage, nearby matrix is depleted of metachromatic material indicating loss of proteoglycans. Microscopic changes are also apparent in cells, and necrosis ('ghosting') is often present. More common, however, is focal proliferation producing clumps ('clones') of chondrocytes, often surrounded by intense metachromatic material indicating increased proteoglycan. Such cell proliferation and metabolic activity represents attempted 'intrinsic' repair. With continuing movement fibrillated cartilage in habitually loaded areas may abrade to expose underlying bone, with progression to variable degrees of structural damage. It is noteworthy, however, that fibrillation itself is not necessarily progressive.

Indeed, it is a normal finding in adult human joints (e.g. hip, knee, humeroradial joint) in areas of cartilage that are habitually unloaded. In addition to intrinsic repair, 'extrinsic' repair commonly occurs by formation of new cartilage at the joint margin or in subchondral bone; such cartilage is generally more cellular than pre-existing hyaline cartilage and the chondrocytes are evenly distributed throughout the matrix. Extrinsic repair is with fibrocartilage, containing predominently type I collagen. Proliferating nodules of fibrocartilage arising from underlying marrow spaces may protrude through defects in the bone surface, occasionally coalescing to form a near continuous layer of replacement tissue.

In parallel with the cartilage changes, new bone formation occurs in subchondral bone and at the joint margins (central and marginal osteophyte), and occasionally beneath adjacent periosteum (e.g. femoral neck 'buttressing'—periosteal osteophyte). Osteophyte forms through the process of endochondral ossification, either by vascular penetration of existing cartilage or from marginal foci of cartilaginous metaplasia, particularly at capsular and ligamentous insertion sites. The location of osteophyte is characteristic for individual joints. Proliferation of subchondral bone is most apparent beneath areas of cartilage erosion and fibrillation. In such areas the tidemark is generally more irregular, thickened, and reduplicated (showing often three or more parallel tidemarks); new bone and thickening of existing trabeculas gives rise to sclerosis seen radiographically. With gross cartilage loss repeated motion may polish the bone ('eburnation') and, as a result of increased local stress, surface bone additionally may undergo focal pressure necrosis. Subarticular cysts (more correctly 'pseudocysts') predominate where overlying cartilage is thinned or absent. Pathologically, cysts show features of bone necrosis (loss of trabeculas, fibromyxomatous degeneration of marrow), frequently contain dead bone, cartilage, and amorphous material, and are surrounded by a rim of reactive new bone and fibrous tissue. They are thought to result from high intra-articular pressure transmitted through defects (microfractures) in the overlying cortex, or from intraosseous hypertension generated through abnormal loading and force transmission in the mechanically altered joint. As with fibrillation, both marginal osteophyte and cysts may occur in the absence of other features of osteoarthritis, reflecting bone response to mechanical stimulation.

Separated fragments of cartilage and bone may form 'loose bodies', undergo dissolution, or become incorporated into the synovium and proliferate locally. As they grow, their centres necrose and calcify, and periodic extensions give rise to a concentric ringed appearance: endochondral ossification of these bodies may follow vascular invasion. Osteochondral bodies may also arise in synovium by chondroid metaplasia of fibroblastic cells.

The synovium becomes both hypertrophic and hyperplastic, and the capsule thickens and contracts. In the synovium lymphoid follicles, as well as more diffuse infiltration by T and B lymphocytes and macrophages (DR positive), may be identified (Revell *et al*. 1988), often with accompanying diffuse and perivascular fibrosis. Synovial extension onto the articular surface (i.e. pannus) is common, particularly at the hip, but both synovitis and pannus are less extensive and aggressive than in rheumatoid arthritis, synovitis usually being confined to synovium rimming the cartilage. Haemosiderin staining of synovium, reflecting previous intra-articular bleeding, is common in large joints, and occasionally is marked. The role of particulate debris (osteochondral fragments, calcium crystals) in producing chronic inflammation is uncertain but, in general, synovitis is regarded more as a secondary, usually late phenomenon than a primary, early event in osteoarthritis.

Calcification is an integral part of new bone formation, and many osteoarthritis joints show evidence of calcium crystal deposition in cartilage, with presumed secondary uptake in synovium. Carbonate-substituted hydroxyapatite is the commonest particulate identified, particularly adjacent to hypertrophic and degenerating chondrocytes (Ohira and Ishikawa 1987). Calcium pyrophosphate dihydrate crystal deposition also associates with osteoarthritis at certain sites, particularly in the elderly (Felson *et al*. 1989). The relationship between calcium crystal formation and osteoarthritis is unclear.

Despite loss of bone and cartilage in some parts of the joint, the net effect of new cartilage and bone formation is an increase in joint size and remodelling of shape. The balance between degradative and reparative features is variable and leads to varying consequences with respect to joint congruity, stability, and load transmission. Associated periarticular abnormalities (muscle atrophy, bursitis, enthesitis) are a common accompaniment to established osteoarthritis. The radiographical–pathological correlations seen in osteoarthritis are outlined in Table 1.

Metabolic changes

Though changes occur in all joint tissues in osteoarthritis, most experimental work has investigated cartilage (Maroudas and Urban 1980; Hardingham and Bayliss 1990). The extracellular matrix of cartilage is complex, and its composition and turnover vary:

1. in different joints;
2. at different locations within the same joint;
3. through the depth of the tissue.

The initiating insult (e.g. mechanical, metabolic) also varies and age-related changes occur. It is sometimes, therefore, difficult to identify common events in osteoarthritis (Table 10).

Table 10 Metabolic changes in osteoarthritic cartilage

Analytical studies

The changes in composition of osteoarthritis cartilage, which are markedly different from those due to ageing, include:

1. an increase in water content;
2. a loosening of the 'collagen network' with a reduction in collagen fibre size, though the collagen concentration and phenotypes appear normal;
3. a reduction in proteoglycan concentration with a smaller proportion of total proteoglycan in aggregates, and alteration in proteoglycan structure.

Whereas ageing human cartilage undergoes some degree of dehydration, osteoarthritis cartilage has an increased water content. This increase is one of the earliest changes detected in animal models—initially within loaded regions but eventually involving all of the cartilage. This change probably reflects a defect in the arrangement of collagen fibres that allows proteoglycans to swell. The marked swelling of osteoarthritic cartilage when placed in hypotonic solutions *in vitro* is consistent with a loosening of the collagen network, and ultrastructural studies using animal models confirm early loss of orientation among superficial collagen fibrils, with individual fibrils being more widely spaced than normal. This basic structural change may arise in type II collagen fibres themselves (Maroudas *et al*. 1986), or more likely in cross-link molecules such as type IX collagen, a decrease of which (with little change in production) is reported in the rabbit postmeniscectomy model and in human osteoarthritis. A recent study (Bonassar *et al*. 1995) has shown that this increase in swelling can be mimicked by treatment of cartilage with stromelysin.

A decrease in proteoglycan content is the most consistent feature found in all studies of osteoarthritis. The main finding for glycosaminoglycans in human osteoarthritic cartilage is a decrease in keratan sulphate content. This is a real event, not merely reflecting a decrease in cartilage thickness (the deeper layers of normal cartilage are richer in keratan sulphate than the surface zones). The concomitant decrease in chondroitin-6-sulphate relative to chondroitin-4-sulphate gives an overall composition akin to that of immature cartilage, suggesting that, in osteoarthritis, chondrocytes revert to a chondroblastic state and synthesize fetal-like, immature proteoglycan. Because of the uncertainty over the production of the proteoglycan, however, it remains unclear in established human osteoarthritis if such

changes are biosynthetic or arise primarily from catabolic events.

Changes in keratan sulphate:chondroitin-6/-4-sulphate ratios are an inconsistent finding in animal models, and such changes may principally occur late in the pathogenesis. Specific modifications in the sulphation of chondroitin sulphate chains, however, appear early in animal models (at 3 months in Pond–Nuki model) and may relate to larger chain size; in addition, proteoglycans from cartilage with activated chondrocytes react with monoclonal antibodies to novel sequences of sulphation on chondroitin sulphate chains (similar epitopes being detected in proteoglycan from human osteoarthritis and normal immature cartilage). The significance of these structural changes in chondroitin sulphate is unknown, though one function may be to increase matrix binding of growth factors, thereby increasing the local pool of these agents.

There is little evidence from animal or human studies to suggest that proteoglycan is lost from the matrix through any abnormalities in its ability to aggregate. Aggregation of proteoglycan appears to be normal, and although the hyaluronan content of osteoarthritic cartilage is reportedly low, there appears sufficient to accommodate the reduced concentration of proteoglycan (Brocklehurst *et al*. 1984). The hyaluronan-binding regions of most proteoglycan monomers in osteoarthritic cartilage appear fully functional. In most animal models this also applies to newly synthesized proteoglycans, even though they are often larger than normal. In cartilage from late-stage human osteoarthritis, the assembly of newly synthesized proteoglycans into aggregates in the extracellular matrix appears faster than in normal adult joints, and more similar to that in immature cartilage. This aspect of proteoglycan structure could profoundly affect turnover rates and complicate further our understanding of mechanisms of cartilage repair.

Growth factors and cytokines

Matrix synthesis

The compositional changes occur as chondrocytes respond to a variety of growth factors and cytokines and so alter the rate of proliferation or either synthesis and/or degradation of these matrix components (Goldring 1993). A number of cytokines and growth factors are listed together with the possible effect shown on matrix metabolism (Table 11).

Growth factor	Major function in cartilage
TGFβ	Chondrocyte proliferation, promotes matrix synthesis, modulates IL-1 effects, increases proteinase inhibitors
PDGF	Proliferation of chondrocytes
bFGF	Proliferation and differentiation of chondrocytes, MMP production
IGF-1	Proliferation of chondrocytes, increases GAG synthesis
IL-1	Increases production of MMPs, PGE_2, and other cytokines; inhibits GAG synthesis
TNF-α	Similar catabolic effects as IL-1
IL-6	Increases proteinase inhibitor production, proliferation of chondrocytes

TGFβ, Transforming growth factor β; IL-1, interleukin-1; PDGF, platelet derived growth factor; IGF-1, insulin growth factor; TNF-α, tumour necrosis factor α; MMPs, Matrix metalloproteinases; PGE_2, prostaglandin E_2; GAG, glycosaminoglycans; IL-6, interleukin-6.

Table 11 Effect of cytokines and growth factors on chondrocyte metabolism

Chondrocytes can increase their biosynthesis to counteract increased loss, particularly early rather than late in the process. For example in the Pond–Nuki model (section of the anterior cruciate ligament in the mature canine stifle joint), chondrocytes from macroscopically normal cartilage show increased incorporation of ^{35}S-sulphate as an early feature that predates fibrillation, and other areas of cartilage that do not proceed to fibrillation may also show this change. In other models early hypermetabolic activity cannot be demonstrated, and whether increased synthesis occurs in early human osteoarthritis remains speculative.

In established human osteoarthritis, the variability in proteoglycan synthesis reported from different centres (i.e. increased or normal) may reflect sampling differences as much as variability of osteoarthritis itself. For example in reports of increased ^{35}S-sulphate incorporation by human osteoarthritic cartilage, with correlation between synthetic activity and histological grading (Ryu *et al*. 1984), the high-scoring samples could be mainly mid- and deep-zone cartilage, which in normal tissue contains cells with higher synthetic rates. After 'correcting' for topographical and zonal sampling, no difference in proteoglycan synthesis is apparent between osteoarthritic and normal, age-matched cartilage at the hip or knee (Brocklehurst *et al*. 1984). There is some evidence that proteoglycan synthesis is decreased in human osteoarthritis and it is also possible that these synthetic rates are affected by the drugs taken by patients. In late human osteoarthritis it is therefore unclear whether chondrocytes are metabolically hyperactive or whether their potential reparative response has been inhibited or lost.

Chondrocytes show an increased rate of synthesis of type II collagen, as with proteoglycans, early in animal models. The long turnover time of type II collagen in adult cartilage might suggest that mature chondrocytes have little chance of even minimal repair of defective collagen network. Nevertheless, an increase in collagen synthesis in human osteoarthritic cartilage has been demonstrated (Lipiello *et al*. 1977). When type II collagen fibrils form in the extracellular matrix, the N- and C-propeptides are removed by specific proteases and are lost from the matrix. Increased C-propeptide (CP-II) is found in human osteoarthritic compared to normal cartilage, mainly in the lower mid- and deep zones rather than the surface and upper mid zones where collagen degradation is more prevalent (Dodge and Poole 1989). There may thus be the potential for limited repair in deep zones, but an effective response in late osteoarthritis, when major disruption in collagen architecture has occurred, seems less likely.

Type X collagen, with its presumed role in mineralization, is regarded as a unique marker for hypertrophic chondrocytes, which are also rich in alkaline phosphatase activity. In osteoarthritic cartilage the alkaline phosphatase activity is very high, not just in the deep but also in the mid-zones (Fig. 41), and deposition of type X collagen once again becomes evident. Such characteristics are reminiscent of immature cartilage, and suggest that chondrocytes throughout the cartilage are resuming their potential to mineralize.

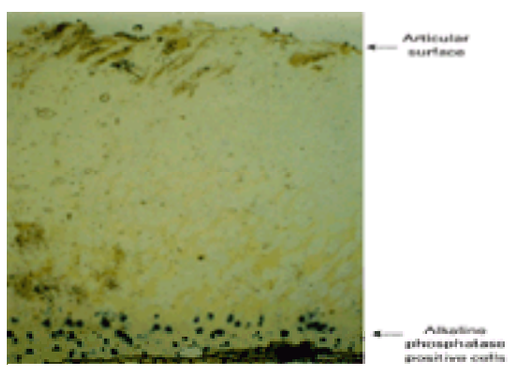


Fig. 41 Localization of alkaline phosphatase in human osteoarthritic cartilage (55-year-old patient).

Matrix degradation

Just as anabolic growth factors can influence matrix synthesis, the proinflammatory cytokines can increase matrix degradation. Both interleukin 1 and tumour necrosis factor are present in cartilage which is degrading extracellular matrix (Table 11) (Hammerman 1989; Goldring 1993). These cytokines when added to cartilage rapidly cause the release of proteoglycan with the later release of collagen fragments (Ellis *et al*. 1994). At the same time the synthesis of matrix components is also down regulated. It is not known exactly how proteoglycan and collagen turnover is increased in response to these cytokines but the levels of the matrix metalloproteinases

are increased (Murphy *et al*. 1991). This family of zinc and calcium dependant proteinases can degrade the proteins of the extracellular matrix at neutral pH and are divided into three main groups—the collagenases, the stromelysins, and the gelatinases (Fig. 42) (Woessner 1992; Brinkerhoff 1991). These enzymes contain common sequences of amino acids; the N-terminal domain contains a characteristic sequence of zinc-binding histidines and contains the catalytic zinc whilst the C-terminal domain determines the differences in substrate specificity. The matrix metalloproteinases can be controlled at different points (Fig. 43) which include stimulation of synthesis and secretion by cytokines, activation of proenzyme forms, and the inhibition by specific inhibitors called TIMPs. These enzymes are found in osteoarthritic cartilage and osteoarthritic cartilage appears deficient in endogenous protease inhibitors (Woessner 1992). This imbalance between matrix metalloproteinases and their inhibitors is likely to play a part in the accelerated breakdown of the matrix (Hammerman 1989; Dean *et al*. 1989). Specific inhibitors of matrix metalloproteinases are able to prevent both proteoglycan and collagen release from resorbing cartilage *in vitro* and in animal models (Andrews *et al*. 1992; Cawston *et al*. 1994). Other proteinases are also implicated in the destruction of cartilage in osteoarthritis (Buttle *et al*. 1993).

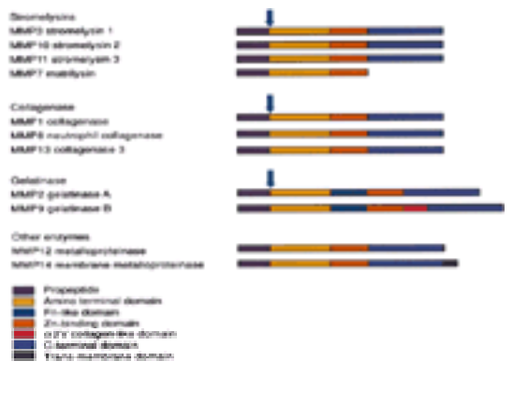


Fig. 42 Metalloproteinase domain structure.

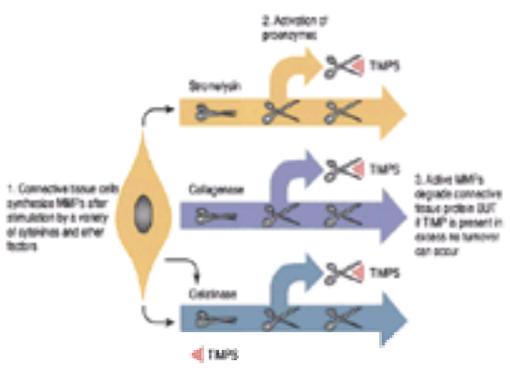


Fig. 43 Control of metalloproteinase activity.

Recently, the analysis of the fragments released from resorbing cartilage have been studied in order to determine degradative mechanisms. Proteoglycans released from osteoarthritic cartilage during culture *in vitro* cannot interact with hyaluronan and presumably have lost their hyaluronan-binding region. Analysis of the proteoglycan fragments suggest that an unidentified metalloproteinase called aggrecanase is responsible, as the cleavage site in the aggrecan molecule is not cleaved by any known enzyme (Sandy *et al*. 1992). Other studies have used a unique epitope hidden in native collagen but exposed by proteolytic cleavage to follow collagen turnover. In osteoarthritis, cartilage collagen cleavage can be shown by immunohistochemistry in both the superficial and intermediate zones (Hollander *et al*. 1994; Poole 1993).

Debate still continues over the relative importance of cartilage and bone changes in the initiation and progression of osteoarthritis. It is recognized that physiology and structure are initimately linked and that all joint tissues (including synovium, capsule, and periarticular tissues) interact together. Thus 'weakening' of cartilage and surrounding tissues and 'stiffening' of subchondral bone will each be deleterious in all tissues. All locomotor tissues require continuing physical stimulation to sustain normal development, nutrition, and adaptation. Interestingly, many of the metabolic and structural responses of cartilage and bone seen in animal models only develop if joint loading continues, suggesting that the potential repair process of joints is similarly driven by this physiological need. A more pressing question, of course, is what stops all joints progressing once cartilage loss and altered biomechanics have occurred. Articular cartilage has only a limited ability to 'repair' whereas bone is able to remodel and readily adapt to changing biochemical requirements. With respect to progression or non-progression in osteoarthritis, it may be that bone response and less investigated aspects (e.g. neuromuscular control, capsular fibrosis) will also prove to be important.

Management

There is no proven, effective treatment for the osteoarthritis process. Nevertheless, considerable benefit to the patient can be achieved by often very simple interventions, and surgery for severe 'end stage,' particularly large joint, osteoarthritis is excellent. The principal goals of management are:

1. education of the patient about osteoarthritis;
2. pain relief;
3. achieving and maintaining optimal joint and limb function.

A wide variety of modalities are available to realize these goals (Table 12), and an approach to management is outlined in the Management Summary (Box 1). As with any branch of medicine, an 'holistic' approach is appropriate and more likely to succeed. The frequent discordance between structural change, symptoms, and function is pertinent; since pain and physical handicap are the main clinical problems arising from osteoarthritis, their causes and treatment usually need consideration as independent issues.

Goals of management
Education
Relief of symptoms
Optimization of function
Available modalities
Counselling
Educational literature
Physical therapy
Occupational therapy
Drugs
Complementary medicine
Coping strategies
Surgery

Table 12 Current management of osteoarthritis

Box 1 Management summary

- Education concerning osteoarthritis
- Protect compromised joints from excessive loading, e.g.:
  - Reduce obesity
  - Modify inappropriate daily/occupational activities
  - Use a walking stick
  - Shock-absorbing insoles
  - Correct leg length discrepancy
- Maintain aerobic fitness, e.g., with swimming and/or walking
- Maintain joint motion and stability
  - Regular movement - 'little and often'
  - Muscle strengthening exercises
- Reduce pain and stiffness, e.g.:
  - Physiotherapy
- Intermittent analgesics
  - Consider topical NSAIDs
  - Consider occasional courses of NSAIDs
  - Consider peri- and intra-articular injection
  - Consider TENS and nerve blocks for severe pain
- Reduce impact of pain and disability
  - Treat depression, anxiety, fibromyalgia
  - Consider coping strategies
  - Modify patient environment to reduce handicap
- Consider surgery for:
  - Persistent severe pain
  - Disability

General approach

Successful management centres on careful questioning and examination of the patient. The history should yield clear information particularly about:

1. symptoms and their impact on the patients life;
2. functional disability;
3. functional requirements and thus the level of handicap;
4. patient expectations both from osteoarthritis and its treatment;
5. psychological factors including specific concerns and depression.

Careful examination should determine the:

1. extent of locomotor abnormality;
2. origin of current pain, either articular and/or periarticular;
3. degree of accompanying synovitis;
4. presence of instability;
5. local and general muscle condition;
6. evidence of accompanying fibromyalgia, neurological, or other relevant medical disease.

An accurate diagnosis can thus be made, the extent of the problems assessed, and any associated factors identified. An individual management programme can then be constructed. The results of such interventions need to be reviewed and the programme modified as the patient's characteristics and requirements change. As symptomatic osteoarthritis is a potentially complex problem, it often requires a co-ordinated approach.

Helping the patient understand osteoarthritis

The myth that osteoarthritis is an inevitable, progressive wear and tear disease of old age persists. This leads to negative attitudes in both patients and doctors and may encourage inappropriate action, for example reduced activity for fear of 'wearing the joints out'. It is important that all those involved in patient management have similar concepts so that conflicting information is avoided. It is also important to address the questions that the patient wants answered, and to respond in a manner understandable to the patient. For example most patients want to know about diet, exercise, and factors in their life that may have brought on their arthritis; explanation of cartilage and bone changes may have little relevance for them. Although the natural history of osteoarthritis is poorly documented, available data suggests that it is often considerably better than patients expect. This is not only true for hand involvement in nodal generalized osteoarthritis but also at the hip and knee. A reasonably optimistic, though not unrealistic, approach is therefore justified.

Addressing mechanical factors

Protecting compromised osteoarthritis joints from excessive or unusual loading often reduces pain. Obesity increases loading on many, not just 'weight-bearing', joints during daily activities, and obese patients are best advised to lose weight. Some prospective data suggest that such weight loss may prevent the development of osteoarthritis (Felson *et al*. 1991). Its role in established osteoarthritis, though logical, is unproven. An initially reducing, though well balanced and 'healthy' diet should therefore be encouraged. Appropriate use of a walking stick for hip or knee osteoarthritis will help reduce loading through the affected joint, and shock-absorbing footwear (e.g. 'trainers') may reduce impact; both manoeuvres may benefit symptoms. Altered mechanics due to leg length discrepancy may produce pain which is commonly periarticular and correction with a heel raise should be undertaken.

Occupational therapy advice concerning 'joint protection' and appropriate manoeuvres for activities of daily living may markedly reduce unnecessary overloading of upper and lower limb joints. Modification of both workplace and home activities should be considered if appropriate. Appropriate advice regarding sexual activity may also be helpful, particularly for women with hip osteoarthritis and their partners.

The integrity of articular tissues is maintained, and repair of damaged cartilage facilitated, by normal movement and loading of the joint. Patients should therefore be advised to keep active, within the bounds of common sense. 'Little and often' is prudent advice and particular emphasis should be on aerobic exercise since this has been demonstrated to be of symptomatic benefit in osteoarthritis (Minor *et al*. 1989; Kovar *et al*. 1992). Bland and Cooper recommend a comprehensive 'rest–exercise programme' based on the rationale that activity aids repair, whereas rest allows rehydration and recovery of cartilage after loading (Bland and Cooper 1984). It is difficult to obtain direct evidence for this in man, but there are considerable experimental data to support this concept. Data on the efficacy of muscle strengthening regimes in improving osteoarthritis are surprisingly sparse and often poorly controlled. The finding, however that quadriceps weakness is closely associated with disability in knee osteoarthritis (McAlindon *et al*. 1993) has led to renewed impetus to explore this area. In contrast, more evidence exists that aerobic exercise that improves fitness as well as improving self-efficacy may be beneficial in osteoarthritis (Minor *et al*. 1989; Kovar *et al*. 1992). All patients with osteoarthritis should be encouraged to maintain activity.

Direct attempts to relieve symptoms

A variety of physical measures, such as local heat, cold, massage, or hydrotherapy, may give temporary pain relief. Such modalities are usually administered initially by physiotherapists, and there is considerable interpatient variability as to which may help.

Simple analgesics and non-steroidal anti-inflammatory drugs are commonly used with good benefit. They should be regarded as an adjunct, rather than substitute, for other treatments. Comparative studies suggest that non-steroidal anti-inflammatory drugs are just marginally, or even no more effective than analgesics (Bradley *et al*. 1992; Williams *et al*. 1993; March *et al*. 1994). Non-steroidal anti-inflammatory drugs are, however, associated with a significant number of side-effects, particularly in the elderly. Simple analgesics should therefore be tried first, including a trial at maximal regular dosage. If unsuccessful, sequential trials of several non-steroidal anti-inflammatory drugs may be considered. There is currently no convincing evidence that one non-steroidal anti-inflammatory drug is superior to another in the symptomatic relief of osteoarthritis. There are, however, emerging differences in the side-effect profiles of different non-steroidal anti-inflammatory drugs (see Chapter 3.5.1). It would therefore seem prudent to begin with a drug with a lower incidence of gastrointestinal side-effects such as ibuprofen or nabumatone before using alternative agents. Due to marked interpatient variability several non-steroidal anti-inflammatory drugs may have to be tried to find which best suits the individual. The patient should be made aware that the aim of these drugs is to reduce, rather than abolish, pain and that they need only be taken when symptoms are bad. Symptoms in osteoarthritis are often phasic: repeat prescribing is to be avoided, and if a patient is taking an non-steroidal anti-inflammatory drug with apparent benefit they

should still regularly experiment by stopping the drug to see whether it is still needed. The possibility that non-steroidal anti-inflammatory drugs may affect the osteoarthritis process remains controversial.

Topically applied non-steroidal anti-inflammatory drugs offer considerable theoretical advantage over oral preparations in terms of reduced side-effects. Although there may be benefit for small superficial joints (e.g. hands) and ossibly the knee, deep-seated large joints (particularly hip and glenohumeral joints) are less amenable to this approach. It remains unclear whether such agents are any more effective than simple rubefacients.

If pain is thought to be predominantly periarticular in origin, due to enthesopathy or bursitis, then local injection of corticosteroid, possibly with local anaesthetic to the tender site, may prove helpful. Indeed one study has even suggested that peripatellar injection is as effective as intra-articular injection in knee osteoarthritis (Sambrook *et al*. 1987). Even temporary symptom improvement may permit involvement in other aspects of the management programme (e.g. exercise), and engender a more positive approach by the patient.

In patients with symptoms unresponsive to other measures, local injection with corticosteroid may produce temporary benefit, permitting involvement in physiotherapy and exercise, or allowing the patient to undertake a 'special event', for example a holiday or a family occasion. Anecdotally single injection at some sites, notably the thumb base, may provide surprisingly prolonged relief. Intra-articular injection of steroid, however, is controversial. Although studies in knee osteoarthritis show benefit over saline injection this effect is short-lived, probably lasting less than 6 weeks (Gaffney *et al*. 1995). Furthermore, the possibility of steroid-induced cartilage attrition (derived from animal work) is often cited when the issue of injection frequency is raised. The data from animal models is, however, conflicting. Individual patients, often those too infirm or otherwise unsuitable for alternative approaches, undoubtedly derive considerable benefit from occasional injection.

Patients with chronic pyrophosphate arthropathy and persistent synovitis who respond only temporarily to steroid injection may gain more prolonged control of synovitis and symptoms from intra-articular radiocolloid (Doherty and Dieppe 1981). Whether successful control of synovitis alters long term outcome is unclear.

Intra-articular hyaluronate preparations are now available for use in some countries. Their symptomatic benefit has been demonstrated in some but not all clinical trials. Their role is currently unclear. Not all patients derive benefit and most regimens require weekly injections over a period of 3 to 5 weeks. Although putative 'chondroprotective' effects have been demonstrated *in vitro* this has yet to be demonstrated in man. Whether such agents will become more widely available and used is to our minds unclear. Other intra-articular agents, such as the superoxide dismutase inhibitor orgotein, have also undergone promising clinical trials but as yet they remain experimental therapies.

Joint, principally hip, distension with saline is claimed to offer quite prolonged symptomatic benefit. It is a technique used principally in Europe. The proposed rationale is that capsular stretching, with or without rupture, results in reduced intra-articular hypertension.

For patients with troublesome knee synovitis joint lavage with saline either arthroscopically or via percutaneous irrigation may offer benefit, sometimes for several months. The mechanism of this non-specific treatment is unclear.

Local nerve blocks, particularly suprascapular and obturator blocks for glenohumeral and hip osteoarthritis, are also worthy of consideration. These have less side-effects, and are generally more effective than major, centrally acting analgesics. A word of caution has been raised by retrospective studies that have suggested more rapid deterioration in patients treated in this manner. Since these procedures are usually used in patients unfit for or awaiting replacement arthroplasty this fear may be unwarranted.

Coping with osteoarthritis

The ability to cope with chronic pain and disability varies greatly between individuals, and depends on multiple factors. Depression, often unrecognized, is certainly common in patients with osteoarthritis and is perhaps too often overlooked, preventing appropriate treatment. Fibromyalgia (chronic fatigue syndrome) may also be common, and again needs recognition and incorporation into the management programme. Even in the absence of depression, anxiety, or fibromyalgia, 'coping strategies' may improve the individual's response to pain and disability without altering the nature of the condition itself. Certain aspects of complementary medicine (e.g. herbal and dietary additives, magnets, charms) may fall into this category, as can meditation, yoga, psychotherapy (group or individual), and religion. Many patients turn to such supportive activities themselves, though it is apparent that many doctors, whilst perfectly happy with the concept of 'placebo effect', still underestimate and therefore underuse such strategies.

Surgery (see also Chapter 6.1)

The biggest revolution in osteoarthritis therapy undoubtedly has been the treatment of severe disease by surgery. The three principal surgical interventions for hips and knees are osteotomy, arthroplasty, and arthroscopy.

Osteotomy provides immediate pain relief, possibly by reduction in raised intraosseous pressure, and long-term benefit, presumably by alteration of mechanical forces and correction of joint deformity. These observations demonstrate that pain relief can occur without altering joint damage, and that the osteoarthritis joint surface can repair with fibrocartilage, which can function adequately for everyday use. Osteotomy needs careful planning and precision for success, and is still used at the knee, and by enthusiasts at the hip.

The overt success of hip arthroplasty has encouraged widespread use of this procedure. Surgeons, however, now face the growing challenge of revision and salvage operations for failed arthroplasties, and this is one aspect where osteotomy carries major advantage. Given the unpredictable natural history of large joint osteoarthritis, the frequency of the condition, and cost-benefit ratio in relation to non-operative therapy, the indications and timing of replacement surgery present difficulties. As new medical interventions become available the role of surgery will require continual reappraisal.

Arthroscopy is increasingly used for osteoarthritis. It allows direct inspection of the articular surface enabling the detection and assessment of minor degrees of articular cartilage damage. It also allows visualization and surgical correction of ligamentous and meniscal injury. These are increasingly being recognized as important accompaniments and possibly even initiating factors in osteoarthritis. Finally it is also a means whereby joint lavage and removal of 'loose bodies' can be performed. The precise role of arthroscopy in the management of osteoarthritis still remains to be defined (Casscells 1990).

Can we modify osteoarthritis ?—the concept of 'chondroprotection'

The possibility of therapeutic manipulation of the osteoarthritic process (in favour of repair) has gained momentum in the last decade. Much interest focuses on the already available non-steroidal anti-inflammatory drugs, though hyaluronate, sulphated glycosaminoglycans, and cartilage extracts have also been examined in this respect. At present, convincing human data on such compounds are lacking; nevertheless, this is a growing and potentially exciting field that has relevance beyond the sphere of osteoarthritis.

There is considerable *in vitro* and *in vivo* evidence that different non-steroidals may:

1. variably influence several aspects of cartilage metabolism;
2. show either detrimental or protective effects in spontaneous or induced animal models of 'osteoarthritis' (Ghosh 1988).

The mechanisms of such actions remain unexplained but appear largely independent of prostaglandin inhibition. They could well affect either the amount of cytokine produced or the response of chondrocytes to cytokines. Interestingly, susceptibility to influence by non-steroidals appears greater for osteoarthritic than for normal cartilage. This could relate to:

1. increased drug delivery from the hypervascular synovium and breaching of the calcified zone by subchondral vessels;
2. enhanced drug penetration due to the increased surface area of fissured cartilage and its altered charge characteristics;
3. increased susceptibility of stimulated chondrocytes.

Notwithstanding certain problems with such experimental data (Doherty 1989), it is apparent that many non-steroidals have suppressive effects on proteoglycan

synthesis and other aspects of cartilage metabolism that may be considered potentially detrimental; conversely, others have little or no suppressive effects at concentrations usually attained in man, and may be beneficial to compromised cartilage. Such observations cannot be directly extrapolated to the clinical situation of human osteoarthritis, and unfortunately there are few studies to date that have directly addressed whether non-steroidals influence (beneficially or detrimentally) the process of osteoarthritis in man.

The possibility that non-steroidal anti-inflammatory drugs are detrimental to osteoarthritis of the hip is supported by two studies reporting greater radiographic destructive change in patients taking indomethacin (Ronningen and Langeland 1979) or regular non-steroidals (Newman and Ling 1985) than in those receiving no indomethacin or infrequent non-steroidals. Both studies, however, can be criticized in terms of retrospective design, small numbers, radiographic assessment, and lack of control for factors that may influence progression. Different conclusions were reached in a prospective study of both osteoarthritic and rheumatoid hips, which implicated obesity, but not non-steroidals in the rate of loss of femoral head height (Watson 1976). A more recent prospective study of patients with osteoarthritis of the hip (Rashad *et al*. 1989) found more rapid radiographic progression and a shorter interval before surgery in patients taking indomethacin compared to those taking azapropazone. However, this study is again readily criticized, for example, in terms of radiographic assessment, lack of control for risk factors for progression, questionable criteria for surgical intervention, and lack of blinding in respect to treatments. It is known that several different non-steroidal anti-inflammatory drugs exert profound inhibition of heterotopic new bone formation after hip arthroplasty—an effect that would not be expected to be beneficial to the remodelling of osteoarthritic joints (Doherty 1990).

Some recent progress has been made with the design of low molecular weight inhibitors of the matrix metalloproteinases that have been shown to be effective *in vitro* and in animal models *in vivo* at preventing cartilage breakdown (Cawston 1995). Such compounds are under development and need to be tested in man to determine if cartilage and bone can be protected from degradative enzymes (Vincenti 1994). These trials are likely to be lengthy and costly as the destruction of tissue is relatively slow but such compounds could protect osteoarthritic cartilage.

The problem remains that, because of the heterogeneity of osteoarthritis, the chronicity of the condition, the discordance between anatomical, functional, and symptomatic manifestations, and unresolved difficulties of assessment, studies seeking to demonstrate even marked differences between drugs prove impracticable to organize. An alternative strategy may be to follow two treatment cohorts of patients with unilateral osteoarthritis of the knee (at high risk of developing this condition on the other side) and observe whether one cohort demonstrates less recruitment of osteoarthritis at new sites than the other.

Though cartilage may be the best understood component and the usual focus of interest, it should be remembered that other joint tissues (bone, capsule, synovium, muscle) may also influence the outcome of osteoarthritis. The term 'chondroprotection' is thus misleading; therapeutic modifications (good or bad) of all joint tissues requires consideration.

Prognosis

For such a common disease, the prognosis of osteoarthritis is largely unknown. From the data that is available one can be relatively optimistic about outcome for most patients. Attempts have been made to define the outcome for different subsets of osteoarthritis. By and large this has been difficult to do. The data will now be discussed by joint site and by osteoarthritis subset.

Knee

There is some data available regarding prognosis for knee osteoarthritis. The natural history of knee osteoarthritis may be less favourable than that for hips. Hernborg and Nilsson observed clinical and radiographic deterioration in the majority of cases followed for 10 to 18 years, with varus deformity, earlier age of onset, and being female relating to worse prognosis (Hernborg and Nilsson 1977). Isolated osteophytosis alone was shown not to associate with subsequent development of osteoarthritis change. In a long-term (12-year) study investigating cartilage loss in the general population (Schouten *et al*. 1992) obesity, presence of generalized osteoarthritis, age and varus/valgus knee deformity were all associated with progressive loss of cartilage. In a study of hospital-referred patients, only the presence of inflammation and calcium pyrophosphate deposition was associated with progression (Ledingham *et al*. 1995). In addition to increased morbidity, knee osteoarthritis is associated with increased mortality (Danielsson and Hernborg 1970; Lawrence *et al*. 1990).

Hip

The general view of inevitable progression in the majority of hip osteoarthritis patients may be unwarranted. In a 10-year, follow-up study, Danielsson found deterioration in symptoms in only 17 per cent of hip osteoarthritis subjects, symptoms improving over this period in 59 per cent, and completely resolving in 12 per cent; radiographic changes similarly showed progression in only a minority of cases, principally those with initial superolateral migration (Danielsson 1964). Furthermore patients with apparently progressive osteoarthritis who then improve with spontaneous 'healing' on radiographs (remodelling and partial restoration of joint space) are well described (Perry *et al*. 1972). Possible risk factors for progression (rather than development) include: superior pole pattern (Danielsson 1964); obesity (Watson 1976); presence of chondrocalcinosis at other sites (Menkes *et al*. 1985); and possibly non-steroidal anti-inflammatory drug usage (Ronningen and Langeland 1979; Newman and Ling 1985; Rashad *et al*. 1989). Evidence for the latter, however, is far from convincing (Doherty 1989).

Hand

The prognosis for hand osteoarthritis is generally good. Symptoms and hand function of nodal generalized osteoarthritis patients examined two or more decades after onset is no worse than that of similarly aged subjects with no hand osteoarthritis (Pattrick *et al*. 1989b). In contrast, in erosive osteoarthritis, in which bony ankylosis and instability is more common, long-term functional outcome may be worse (Pattrick *et al*. 1989b).

Functional outcome for thumb base involvement, carpometacarpophalangeal joint, and scaphotrapezoid disease is less clear-cut but again may be relatively poor.

Spine

Prognosis for osteoarthritis of the spine is unclear. This is largely because of the difficulty of correlating symptoms with structural change. In cases of either cord or nerve root compression prognosis is unclear, with deterioration often being slow.

Chronic pyrophosphate arthropathy

The natural history of chronic pyrophosphate arthropathy is poorly documented. Despite often severe symptoms and structural change at presentation, one 5-year, hospital-based, prospective study (Doherty and Dieppe 1988) suggests that most patients run a benign course, particularly with respect to small and medium-sized joint involvement. As expected symptomatic deterioration occurred mainly in large lower limb joints, but even in severely affected knees (the usual presenting site), two-thirds of patients showed stabilization or improvement of symptoms. The commonest radiographic change is an increase in osteophyte with bone remodelling, rather than progressive cartilage and bone attrition. Nevertheless, severe, progressive 'destructive pyrophosphate arthropathy' may occasionally occur, particularly at the knee, shoulder, and hip. This is virtually confined to elderly women or in association with haemochromatosis and may cause problematic recurrent haemarthrosis and a radiographic appearance of marked destruction resembling a Charcot or neuropathic joint.

Apatite associated destructive arthropathy

The prognosis of this form is seemingly poor with the majority having marked joint destruction requiring joint replacement.

Chapter References

Acheson, R.M. and Collart, A.B. (1975). New Haven Survey of joint diseases. XVII. Relationships between some systemic characteristics and osteoarthrosis in a general population. *Annals of the Rheumatic Diseases*, **34**, 379–87.

Aisen, A.M., McCune, W.J., MacGuire, A., Carson, P.L., Silver, T.M., Jafri, S.Z., and Martel, W. (1984). Sonographic evaluation of the cartilage of the knee. *Radiology*, **153**, 781–4.

Altman, R.D. (1991). Criteria for classification of clinical osteoarthritis. *Journal of Rheumatology*, **18** (suppl. 27), 10–12.

Altman, R., Asch, E., Block, D., *et al*. (1986). Development of criteria for the classification and reporting of osteoarthritis: classification of osteoarthritis of the knee. *Arthritis and Rheumatism*, **29**, 1039–49.

Altman, R., Alarcón, G., Appelrouth, D., *et al*. (1990). The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hand. *Arthritis and Rheumatism*, **33**, 1601–10.

Altman, R., Alarcón, G., Appelrouth, D., *et al*. (1991). The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hip. *Arthritis and Rheumatism*, **34**, 505–14.

Andrews, H.J., Plumpton, T.A., Harper, G.P., and Cawston, T.E. (1992). A synthetic peptide metalloproteinase inhibitor, but not TIMP, prevents the breakdown of proteoglycan within articular cartilage *in vitro*. *Agents and Actions*, **37**, 147–54.

Badley, E.M., Rasooly, I., and Webster, G.K. (1994). Relative importance of musculoskeletal disorders as a cause of chronic health problems, disability, and health care utilisation: findings from the 1990 Ontario Health Survey. *Journal of Rheumatology*, **21**, 505–14.

Bagge, E., Bjelle, A., and Svanborg, A. (1992). Radiographic osteoarthritis in the elderly. A cohort comparison and a longitudinal study of the '70-year old people in Göteborg'. *Clinical Rheumatology*, **11**, 486–91.

Bland, J.H. and Cooper, S.M. (1984). Osteoarthritis: a review of the cell biology involved and evidence for reversibility. Managemnt rationally related to known genesis and pathophysiology. *Seminars in Arthritis and Rheumatism*, **14**, 106–33.

Bonassar, L.J, Frank, E.H., Murray, J.C., Paguio, C.G., Moore, V.L., Lark, M.W., Sandy, J.D., Wu, J.-J., Eyre, D.R., and Grodzinsky, A.J. (1995). Changes in cartilage composition and physical properties due to stromelysin degradation. *Arthritis and Rheumatism*, **38**, 173–83.

Bradley, J.D., Brandt, K.D., Katz, B.P., Kalasinski, L.A., and Ryan, S.I. (1992). Comparison of an anti-inflammatory dose of ibuprofen, an analgesic dose of ibuprofen, and acetaminophen in the treatment of patients with osteoarthritis of the knee. *New England Journal of Medicine*, **325**, 87–91.

Brandt, K.D. (1988). Osteoarthritis. *Clinical Geriatric Medicine*, **4**, 279–93.

Brinkerhoff, C.E. (1991). Joint destruction in arthritis: metalloproteinases in the spotlight. *Arthritis and Rheumatism*, **34**, 1073–5.

Brocklehurst, R., Bayliss, M.T., Maroudas, A., Coysh, H.L., Freeman, M.A.R., Revell, P.A., and Ali, S.Y. (1984). The composition of normal and osteoarthritic articular cartilage from human knee joints. *Journal of Bone and Joint Surgery (America)*, **66A**, 95–106.

Buckland-Wright, J.C., MacFarlane, D.G., Lynch, J.A., and Clark, B. (1990). Quantitative microfocal radiographic assessment of progression in osteoarthritis of the hand. *Arthritis and Rheumatism*, **33**, 57–65.

Burnett, S., Hart, D.J., Cooper, C., and Spector, T.D. (1994). *A radiographic atlas of osteoarthritis*. Springer-Verlag, London

Buttle, D.J., Handley, C.J., Ilic, M.Z., Saklatvala, J., Murata, M., and Barrett, A.J. (1993). Inhibition of cartilage proteoglycan release by a specific inactivator of cathepsin B and an inhibitor of matrix metalloproteinases; evidence for two converging pathways of chondrocyte-mediated proteoglycan degradation. *Arthritis and Rheumatism*, **36**, 1709–17.

Casscells, S.W. (1990). What, if any, are the indications for arthroscopic debridement of the osteoarthritic knee. *Arthroscopy*, **6**, 169–70.

Caterson, B., Mahmoodian, F., Sorrell, J.M., *et al*. (1990). Modulation of native chondroitin sulphate structure in tissue development and disease. *Journal of Cell Science*, **97**, 411–17.

Cawston, TE. (1995). Proteinases and connective tissue breakdown. In *Mechanisms and models in rheumatoid arthritis* (ed. Henderson, Edwards, and Pettipher), pp. 333–59. Academic Press.

Cawston, T.E., Mercer, E., De Silva, M., and Hazleman, B.L. (1983). Metalloproteinases and collagenase inhibitors in rheumatoid synovial fluid. *Arthritis and Rheumatism*, **27**, 64l–6.

Cawston, T.E., Plumpton, T.A., Curry, V.A., Ellis, A., and Powell, L. (1994). The role of TIMP and MMP inhibition in preventing connective tissue breakdown. *Annals of the New York Academy of Science*, **732**, 75–83.

Claessens, A., Schouten, J., van den Ouweland, F., and Valkenburg, H. (1990). Do clinical findings associate with radiographic osteoarthritis of the knee? *Annals of the Rheumatic Diseases*, **49**, 771–4.

Clark, I.M., Powell, L.K., Ramsey, S., Hazleman, B.L., and Cawston, T.E. (1993). The measurement of collagenase, TIMP and collagenase-TIMP complex in synovial fluids from patients with osteoarthritis and rheumatoid arthritis. *Arthritis and Rheumatism*, **36**, 372–9.

Cobby, M., Cushnaghan, J., Creamer, P., Dieppe, P., and Watt, I. (1990). Erosive osteoarthritis: is it a separate disease entity? *Clinical Radiology*, **42**, 258–63.

Cooke, T.D.V. (1985). Pathogenic mechanisms in polyarticular osteoarthritis. In *Clinics in rheumatic diseases* (ed. L Sokoloff), Vol. 11 (2), pp. 203–38. WB Saunders, London.

Cooper, C. (1995). Occupational activity and the risk of the osteoarthritis. *Journal of Rheumatology*, **22** (suppl. 43), 10–12.

Croft, P., Cooper, C., Wickham, C., and Coggon, D. (1991). Osteoarthritis of the hip and acetabular dysplasia. *Annals of the Rheumatic Diseases*, **50**, 308–10.

Cushnaghan, J. and Dieppe, P. (1991). Study of 500 patients with limb joint osteoarthritis. I. Analysis by age, sex, and distribution of symptomatic joint sites. *Annals of the Rheumatic Diseases*, **50**, 8–13.

Danielsson, L.G. (1964). Incidence and prognosis of coxarthrosis. *Acta Orthopaedica Scandinavica*, **66** (suppl.), 1–114.

Danielsson, L.G. and Hernborg, J. (1970). Morbidity and mortality of osteoarthritis of the knee (gonarthrosis) in Malmo, Sweden. *Clinical Orthopaedics and Related Research*, **69**, 224–6.

Davis, M.A., Ettinger, W.H., Neuhaus, J.M., Barclay, J.D., and Segal, M.R. (1992). Correlates of knee pain amongst US adults with and without radiographic knee osteoarthritis. *Journal of Rheumatology*, **19**, 1943–9.

Dean, D.D., Martel-Pelletier, J., Pelletier, J.P., Howell, D.S., and Woessner, J.F. (1989). Evidence for metalloproteinase and metalloproteinase inhibitor imbalance in human osteoarthritic cartilage. *Journal of Clinical Investigation*, **84**, 678–85.

Dieppe, P.A., Campion, G., and Doherty, M. (1988). Mixed crystal deposition. *Rheumatology Disease Clinics of North America*, **14**, 415–26.

Dieppe, P., Cushnaghan, J., Young, P., and Kirwan, J. (1993). Prediction of the progression of joint space narrowing in osteoarthritis of the knee by bone scintigraphy. *Annals of the Rheumatic Diseases*, **52**, 557–63.

Dodge, G.R. and Poole, A.R. (1989). Immunohistochemical detection and immunochemical analysis of type II collagen degradation in human normal, rheumatoid and osteoarthritic articular cartilages and in explants of bovine articular cartilage cultured with interleukin 1. *Journal of Clinical Investigation*, **83**, 647–61.

Doherty, M. (1989). Chondroprotection by NSAIDs. *Annals of the Rheumatic Diseases*, **48**, 619–21.

Doherty, M. (ed.). (1994). *Color atlas and text of osteoarthritis*. Wolfe, London.

Doherty, M. and Dieppe, P.A. (1981). Effect of intra-articular yttrium-90 on chronic pyrophosphate arthropathy of the knee. *Lancet*, **2**, 1243–6.

Doherty, M. and Dieppe P.A. (1988). Clinical aspects of calcium pyrophosphate dihydrate crystal deposition. *Rheumatology Disease Clinics of North America*, **14**, 395–414.

Doherty, M., Watt, I., and Dieppe, P. (1983). Influence of primary generalised osteoarthritis on development of secondary osteoarthritis. *Lancet*, **ii**, 8–11.

Doherty, M., Pattrick, M., and Powell, RJ. (1990). Hypothesis—nodal generalised osteoarthritis is an auto-immune disease. *Annals of the Rheumatic Diseases*, **49**, 1017–20.

Doherty, M., Hamilton, E., Henderson, J., Misra, H., and Dixey, J. (1991 *a*). Familial chondrocalcinosis due to calcium pyrophosphate dihydrate crystal deposition in English families. *British Journal of Rheumatology*, **30**, 10–15.

Doherty, M., Chuck, A., Hosking, D., and Hamilton, E. (1991 *b*). Inorganic pyrophosphate in metabolic diseases predisposing to calcium oyrophosphate dihydrate crystal deposition. *Arthritis and*

Rheumatism, **34**, 1297–303.

Ehlich, G.E. (1972). Inflammatory osteoarthritis: I. The clinical syndrome. *Journal of Chronic Diseases*, **25**, 317–28.

Ehlich, G.E. (1975). Osteoarthritis beginning with inflammation definitions and correlations. *Journal of the American Medical Association*, **232**, 157–9.

Ellis, A.J., Curry, V.A., Powell, E.K., and Cawston, T.E. (1994). The prevention of collagen breakdown in bovine nasal cartilage by TIMP, TIMP-2 and a low molecular weight synthetic inhibitor. *Biochemical Biophysical Research Communications*, **201**, 94–101.

Felson, D.T. (1988). Epidemiology of hip and knee osteoarthritis. *Epidemiologic Reviews*, **10**, 1–28.

Felson, D.T., Anderson, J.J., Naimark, A., Kannel, W., Meenan, R.F. (1989). The prevalence of chondrocalcinosis in the elderly and its association with knee osteoarthritis: the Framingham study. *Journal of Rheumatology*, **16**, 1241–5.

Felson, D.T., Hannan, M.T., Naimark, A., Berkely, J., Gordon, G., Wilson, P.W.F., and Anderson, J. (1991). Occupational physical demands, knee bending, and knee osteoarthritis: results from the Framingham study. *Journal of Rheumatology*, **18**, 1587–92.

Felson, D.T., Ahange, Y., Anthony, J.M., Naimark, A., and Anderson, J.J. (1992). Weight loss reduces the risk for symptomatic knee osteoarthritis in women. *Annals of Internal Medicine*, **116**, 535–9.

Gaffney, K., Ledingham, J., and Perry, J.D. (1995). Intra-articular triamcinolone hexacetonide in knee osteoarthritis: factors influencing the clinical response. *Annals of the Rheumatic Diseases*, **54**, 379–81.

Ghosh, P. (1988). Anti-rheumatic drugs and cartilage. *Clinical Rheumatology*, **2**, 309–38.

Goldring, M.B. (1993). Degradation of articular cartilage in culture: regulatory factors In *Joint cartilage degradation* (ed. J.F. Woessner and D.S. Howell). Marcel Dekker, New York.

Goldthwaite, J.E. (1904). The differential diagnosis and treatment of so-called rheumatoid diseases. *Boston Medical Surgery Journal*, **151**, 529–34.

Hadler, N.M. (1992). Knee pain is the malady—not osteoarthritis. *Annals of Internal Medicine*, **116**, 598–9.

Hadler, N.M., Gillings, D.B., Imbus, H.R., Levitin, P.M., Makuc, D., Utsinger, P.D., Yount, W.J., Slusser, D., and Moskovitz, N. (1977). Hand structure and function in an industrial setting. Influence of three patterns of stereotyped, repetative usage. *Arthritis and Rheumatism*, **20**, 1019–25.

Halverson, P.B. and McCarty, D.J. (1988). Clinical aspects of basic calcium phosphate crystal deposition. *Rheumatology Disease Clinics of North America*, **14**, 427–39.

Hammerman, D. (1989). The biology of osteoarthritis. *New England Journal of Medicine*, **320**, 1322–30.

Hardingham, T. and Bayliss, M. (1990). Proteoglycans of articular cartilage: changes in ageing and in joint disease. *Seminars in Arthritis and Rheumatism*, **20** (suppl. 1), 12–33.

Hart, D.J., Spector, T.D., Brown, P., Wilson, P., Doyle, D.V., and Silman, A.J. (1991). Clinical signs of early osteoarthritis: reproducibility and relation to x-ray changes in 541 women in the general population. *Annals of the Rheumatic Diseases*, **50**, 467–70.

Hernborg, J.S. and Nilsson, B.E. (1977). The natural course of untreated osteoarthritis of the knee. *Clinical Orthopaedics and Related Research*, **123**, 130–7.

Hollander, A.P., Heathfield, T.F., Webber, C., Iwata, Y., Bourne, R., Rorabeck, C., and Poole, A.R. (1994). Increased damage to type II collagen in osteoarthritic articular cartilage detected by a new immunoassay. *Journal of Clinical Investigation*, **93**, 1722–32.

Hopkinson, N., Powell, R.J., and Doherty, M. (1992). Auto-antibodies, immunoglobulins and Gm allotypes in nodal generalised osteoarthritis. *British Journal of Rheumatology*, **31**, 605–8.

Hutton, C. (1987). Generalised osteoarthritis: an evolutionary problem. *Lancet*, **i**,1463–5.

Hutton, C.W., Higgs, E.R., Jackson, P.C., Watt, I., Dieppe, P.A. (1986). TcHMDP bone scanning in generalised nodal osteoarthritis. II. The four hour bone scan image predicts radiographic change. *Annals of the Rheumatic Diseases*, **45**, 622–6.

Jones, A.C. and Doherty, M. (1992). The treatment of osteoarthritis. *British Journal of Clinical Pharmacology*.

Jones, A.C., Chuck, A.J., Arie, E.A., Green, D.J., and Doherty, M. (1992). Diseases associated with calcium pyrophosphate deposition disease. *Seminars in Arthritis and Rheumatism*, **22**, 188–202.

Jones, A.C., Ledingham, J., McAlindon, T., Regam, M., Hart, D., MacMillan, P.J., and Doherty, M. (1993). Radiographic assessment of patellofemoral osteoarthritis. *Annals of the Rheumatic Diseases*, **52**, 655–8.

Kallman, D.A., Wigley, F.M., Scott, W.W., Hochberg, M.C., and Tobin, J.D. (1989). New radiographic grading scales for osteoarthritis of the hand. *Arthritis and Rheumatism*, **32**, 1584–91.

Kellgren, J.H. and Lawrence, J.S. (1957). Radiological assessment of osteoarthritis. *Annals of the Rheumatic Diseases*, **16**, 494–502.

Kellgren, J.H. and Moore, R. (1952). Generalised osteoarthritis and Heberden's nodes. *British Medical Journal*, **1**, 181–7.

Knowlton, R.G., Katzenstein, P.L., Moskowitz, R.W., Weaver, E.J., Malemud, C., Pathria, M., Jimenez, S., and Prockop, D.J. (1990). Demonstration of genetic linkage of a polymorphism in the type II procollagen gene (Col2A1) to primary osteoarthritis associated with mild chondrodysplasia. *New England Journal of Medicine*, **322**, 526–30.

Kovar, P.A., Allegrante, J.P., MacKenzie, C.R., Peterson, M.G.E., Gutin, B., and Charlson, M.E. (1992). Supervised fitness walking in patients with osteoarthritis of the knee. *Annals of Internal Medicine*, **116**, 529–34.

Lane, N.E. and Nevitt, M.C. (1994). Osteoarthritis and bone mass. *Journal of Rheumatology*, **21**, 1393–6.

Lawrence, J.S., Bremner, J.M., and Beir, F. (1966). Osteoarthrosis. Prevalence in the population and relationship between symptoms and x-ray changes. *Annals of the Rheumatic Diseases*, **25**, 1–24.

Lawrence, R.C., Everett, D., Hochberg, M.C. (1990). Arthritis. In *Health status and well-being of the elderly: national health and nutrition examination-I Epidemiologic followup survey* (ed. R. Huntley and J. Cornoni-Huntley), pp. 136–51. Oxford University Press, New York.

Ledingham, J., Dawson, S., Preston, B., Milligan, G., and Doherty, M. (1992). Radiographic patterns and associations of osteoarthritis of the hip. *Annals of the Rheumatic Diseases*, **51**, 1111–6.

Ledingham, J., Dawson, S., Preston, B., Milligan, G., and Doherty, M. (1993). Radiographic progression of hospital referred osteoarthritis of the hip. *Annals of the Rheumatic Diseases*, **52**, 263–7.

Ledingham, J., Regam, M., Jones, A., and Doherty, M. (1995). Factors affecting radiographic progression of knee osteoarthritis. *Annals of the Rheumatic Diseases*, **54**, 53–8.

Lipiello, L., Hall, D., and Mankin, H.J. (1977). Collagen synthesis in normal and osteoarthritic human cartilage. *Journal of Clinical Investigation*, **59**, 593–600.

Lohmander, L.S., (1994). Articular cartilage and osteoarthritis. The role of molecular markers to monitor breakdown, repair and disease. *Journal of Anatomy*, **184**, 477–92.

Lohmander, L.S., Dahlberg, L., Ryd, L., and Heinegard, D. (1989). Increased levels of proteoglycan fragments in knee joint fluid after injury. *Arthritis and Rheumatism*, **32**, 1434–42.

Lohmander, L.S., Hoerrner, L.A., and Lark, M.W. (1993). Metalloproteinases, tissue inhibitor and proteoglycan fragments in knee synovial fluid in human osteoarthritis. *Arthritis and Rheumatism*, **36**, 181–9.

Lohmander, L.S., Roos, H., Dahlberg, L., Hoerrner, L.A., and Lark, M.W. (1994). Temporal patterns of stromelysin-1, tissue inhibitor, and proteoglycan fragments in human knee joint fluid after injury to the cruciate ligament or meniscus. *Journal of Orthopaedic Research*, **12**, 21–8.

Mankin, H.J. (1974). The reaction of cartilage to injury and osteoarthritis. Parts I and II. *New England Journal of Medicine*, **291**, 1285–92; 1335–40.

March, L., Irwig, L., Schwarz, J., Simpson, J., Choch, C., and Brooks, P. (1994). N of 1 trials comparing a non-steroidal anti-inflammatory drug with paracetamol in osteoarthritis. *British Medical Journal*, **309**, 1041–6.

Marks, J.S., Stewart, I.M., and Hardinge, K. (1979). Primary osteoarthrosis of the hip and Heberden's nodes. *Annals of the Rheumatic Diseases*, **38**, 107–11.

Maroudas, A. and Urban, J.P.G. (1980). Metabolism of cartilage. In: *Studies in joint disease* (ed. A. Maroudas and E.J. Holborow), pp. 87–116, Vol 1. Pitman Medical, London.

Maroudas, A., Katz, E.P., Wachtel, E.J., *et al*. (1986). Physiochemical properties and functional behaviour of normal and osteoarthritic human cartilage. In *Articular cartilage biochemistry* (ed. K.E. Kuettner, R. Schleyerbach, and V. Hascall), pp. 311–30. Raven Press, New York.

McAlindon, T.E., Snow, S., Cooper, C., and Dieppe, P.A. (1992). Radiographic patterns of osteoarthritis of the knee joint in the community: the importance of the patellofemoral joint. *Annals of the Rheumatic Diseases*, **51**, 844–9.

McAlindon, T.E., Cooper, C., Kirwan, J.R., and Dieppe, P.A. (1993). Determinants of disability in osteoarthritis of the knee. *Annals of the Rheumatic Diseases*, **52**, 258–62.

McCarty, D.J. (1976). Calcium pyrophosphate dihydrate crystal deposition disease 1975. *Arthritis and Rheumatism*, **19** (suppl.), 275–86.

Menkes, C.-J., Decraemere, W., Postel, M., and Forest, M. (1985). Chondrocalcinosis and rapid destruction of the hip. *Journal of Rheumatology*, **12**, 130–3.

Messiah, S.S., Fowler, P.J., and Munro, T. (1990). Anteroposterior radiographs of the osteoarthritic knee. *Journal of Bone and Joint Surgery (British)*, **72-B**, 639–40.

Mikkelsen, W.M., Duff, I.F., and Dodge, H.J. (1970). Age, sex specific prevalence of radiographic abnormalities of the joints of the hands, wrists, and cervical spine of adult residents of the Tecumseh, Michigan community health study area, 1962–1965. *Journal of Chronic Diseases*, **23**, 151–9.

Minor, M.A., Hewett, J.E., Webel, R.R., Anderson, S.K., and Kay, D.R. (1989). Efficacy of physical conditioning exercise in patients with rheumatoid arthritis and osteoarthritis. *Arthritis and Rheumatism*, **32**, 1396–405.

Murphy, G., Docherty, A.J.P., Hembry, R.M., and Reynolds, J.J. (1991). Metalloproteinases and tissue damage. *British Journal of Rheumatology*, **30**, 25–31.

Newman, N.M. and Ling, R.S.M. (1985). Acetabular bone destruction related to non-steroidal anti-inflammatory drugs. *Lancet*, **ii**, 11–14.

Ohira, T. and Ishikawa, K. (1987). Hydroxyapatite deposition in osteoarthritic articular cartilage of the proximal femoral head. *Arthritis and Rheumatism*, **30**, 651–60.

Pattrick, M., Manhire, A., Milford-Ward, A., and Doherty, M. (1989 *a*). HLA AB antigens and alpha-1-antitrypsin phenotypes in nodal generalised osteoarthritis and erosive osteoarthritis. *Annals of the Rheumatic Diseases*, **48**, 470–5.

Pattrick, M., Aldridge, S., Hamilton, E., Manhire, A., and Doherty, M. (1989 *b*). A controlled study of hand function in nodal and erosive osteoarthritis. *Annals of the Rheumatic Diseases*, **48**, 978–82.

Pattrick, M., Hamilton, E., Wilson, R., Austin, S., and Doherty, M (1993). Association of radiographic changes of osteoarthritis, symptoms and synovial fluid particles in 300 knees. *Annals of the Rheumatic Diseases*, **52**, 97–103.

Perry, G.H., Smith, M.J.G., and Whiteside, C.G. (1972). Spontaneous recovery of the joint space in degenerative hip disease. *Annals of the Rheumatic Diseases*, **31**, 440–8.

Peyron, J.G. (1979). Epidemiologic and aetiologic approach to osteoarthritis. *Seminars in Arthritis and Rheumatism*, **8**, 288–306.

Poole, R.A. (1993). Immunochemical markers of joint inflammation, skeletal and repair where are we now? *Annals of the Rheumatic Diseases*, **53**, 305.

Rashad, S., Revell, P., Hemingway, A., Low, F., Rainsford, K., and Walker, F. (1989). Effect of non-steroidal anti-inflammatory drugs on the course of osteoarthritis. *Lancet*, **ii**, 519–22.

Ratcliffe, A., Doherty, M., Maini, R.N., and Hardingham, T.E. (1988). Increased levels of proteoglycan components in the synovial fluids of patients with acute but not chronic joint disease. *Annals of the Rheumatic Diseases*, **47**, 826–32.

Revell, P.A., Mayston, V., Lalor, P., and Mapp, P. (1988). The synovial membrane in osteoarthritis: a histological study including the characterisation of the cellular infiltrate present in inflammatory osteoarthritis using monoclonal antibodies. *Annals of the Rheumatic Diseases*, **47**, 300–7.

Risteli, J., Glorian, I., Nierin, S., Naramo, A., and Risteli, L. (1993). Radioimmunoassay for the pyridinoline cross-linked C terminal telopeptide of type 1 collagen. A new serum marker of bone collagen degradation. *Clinical Chemistry*, **39**, 635–40.

Robins, S.P., Woitge, H., Hesley, R., Ju, J., Seyedin, S., and Seibel, M.J. (1994). Direct, enzyme-linked immunoassay for urinary deoxypyridinoline as a specific marker for measuring bone resorption. *Journal of Bone and Mineral Research*, **9**, 1643–9.

Rogers, J., Watt, I., and Dieppe, P. (1990). Comparison of visual and radiographic detection of bony changes at the knee joint. *British Medical Journal*, **300**, 367–8.

Roh, Y.S., Dequecker, J., and Mulier, J.C. (1973). Osteoarthrosis at the hand skeleton in primary osteoarthrosis of the hip and in normal controls. *Clinical Orthopedics*, **90**, 90–4.

Ronningen, H. and Langeland, N. (1979). Indomethacin treatment in osteoarthritis of the hip joint. Does the treatment interfere with the natural course of the disease? *Acta Orthopaedica Scandinavica*, **50**, 169–74.

Ryu, J., Treadwell, B.V., and Mankin, H.J. (1984). Biochemical and metabolic abnormalities in normal and osteoarthritic human articular cartilage. *Arthritis and Rheumatism*, **27**, 49–56.

Sambrook, P.N., Champion, G.D., Browne, C.D., Cairns, D., Cohen, M.L., Day, R.O., *et al*. (1987). Corticosteroid injection for osteoarthritis of the knee: peripatellar compared to intra-articular route. *Clinical and Experimental Rheumatology*, **7**, 609–13.

Sandy, J.D., Flannery, C.R., Neame, P.J., and Lohmander, L.S. (1992). The structure of aggrecan fragments in human synovial fluid: Evidence for the involvment in osteoarthritis of a novel proteinase. *Journal of Clinical Investigation*, **89**, 1512–6.

Saxne, T., Heinegard, D., Wollheim, F.A., *et al*. (1985). Difference in cartilage proteoglycan level in synovial fluid in early rheumatoid arthritis and reactive arthritis. *Lancet*, **ii**, 127–8.

Saxne, T., Heinegard, D., and Wollheim, F.A. (1986). Therapeutic effects on cartilage metabolism in arthritis as measured by release of proteoglycan structures into the synovial fluid. *Annals of the Rheumatic Diseases*, **45**, 491–7.

Schouten, J.S.A.G., van den Ouweland, F.A., and Valkenburg, H.A. (1992). A 12 year follow up study in the general population on prognostic factors of cartilage loss in osteoarthritis of the knee. *Annals of the Rheumatic Diseases*, **51**, 932–7.

Seibel, M.J., Duncan, A., and Robins, S.P. (1989). Urinary hydroxy-pyridinium crosslinks of collagen reflect cartilage and bone involvement in rheumatic joint diseases. *Journal of Rheumatology*, **16**, 964–70.

Shinmei, M., Inamori, Y., Yoshihara, Y., Kikuchi, T., and Hayakawa, T. (1992). The potential of cartilage markers in joint fluid for drug evaluation. In *Articular cartilage and osteoarthritis* (ed. K. Kuettner *et al*.). Raven Press, New York.

Sokoloff, L. (1985). Endemic forms of osteoarthritis. *Clinics in the Rheumatic Diseases*, **11**, 187–202.

Sokoloff, L. (1994). In *Color atlas and text of osteoarthritis* (ed. M. Doherty). Wolfe, London.

Solomon, L. (1976). Patterns of osteoarthritis of the hip. *Journal of Bone and Joint Surgery (British)*, **58-B**, 176–83.

Solomon, L. (1983). Osteoarthritis, local and generalised: a uniform disease? *Journal of Rheumatology*, **10** (suppl. 9), 13–15.

Spector, T.D. (1995). Measuring joint space in knee osteoarthritis: position or precision? *Journal of Rheumatology*, **22**, 807–8.

Spector, T.D. and Campion, G.D. (1989). Generalised osteoarthritis: a hormonally mediated disease. *Annals of the Rheumatic Diseases*, **48**, 523–7.

Spector, T.D. and Cooper, C. (1993). Radiographic assessment of osteoarthritis in the general population: whither Kellgren and Lawrence? *Osteoarthritis and Cartilage*, **1**, 203–6.

Spector, T.D., Hart, D.J., Byrne, J., Harris, P.A., Dacre, J.E., and Doyle, D.V. (1993). Definition of osteoarthritis of the knee for epidemiological studies. *Annals of the Rheumatic Diseases*, **52**, 790–4.

Stecher, R.M. (1953). Heberdens nodes. A clinical description of osteoarthritis of the finger joints. *Annals of the Rheumatic Diseases*, **14**, 1–10.

Steven, M.M. (1992). Prevalence of chronic arthritis in four geographical areas of the Scottish Highlands. *Annals of the Rheumatic Diseases*, **51**, 186–96.

Swan, A., Chapman, B., Heap, P., Seward, H., and Dieppe, P. (1994). Submicroscopic crystals in osteoarthritic synovial fluids. *Annals of the Rheumatic Diseases*, **53**, 467–70.

Sweet, M.B.E., Coelho, A., Schnitzler, C.M., Schnitzer, T.J., Lenz, M.E., Jakin, I., Kuettner, K.E., and Thonar, E.J.M.A. (1988). Serum keratan sulphate levels in osteoarthritis patients. *Arthritis and Rheumatism*, **31**, 648–52.

Thonar, E.J.-M.A., Schitzner, T., and Kuettner, K.E. (1987). Quantification of keratan sulphate in blood as a marker of cartilage catabolism. *Journal of Rheumatology*, **14** (suppl. 14), 23–4.

van Saase, J.L.C.M., Van Romunde, L.K.J., Cats, A., Vandenbrouke, J.P., and Valkenburg, H.A. (1989). Epidemiology of osteoarthritis: Zoetermeer survey. Comparison of radiological osteoarthritis in a Dutch population with that in 10 other populations. *Annals of the Rheumatic Diseases*, **48**, 271–80.

Vincenti, M.P., Clark, I.M., and Brinkerhoff, C.E. (1994). Using inhibitors of metalloproteinases to treat arthritis. Easier said than done? *Arthritis and Rheumatism*, **37**, 1115–26.

Watson, M. (1976). Femoral head height loss: a study of the relative significance of some of its determinants in hip degeneration. *Rheumatology Rehabilitation*, **15**, 264–9.

Westacott, C.I., Whicher, J.T., Barnes, I.C., Thompson, D., Swan, A.J., and Dieppe, P.A. (1990). Synovial fluid concentrations of five cytokines in rheumatic diseases. *Annals of the Rheumatic Diseases*, **49**, 676–81.

Williams, C.J. and Jimenez, S.A. (1993). Heriditary, genes and osteoarthritis. *Rheumatology Disease Clinics of North America*, **19**, 523–43.

Williams, W.J., Ward, J.R., Eggar, M.J., Neuner, R., Brooks, R.H., Clegg, D.O., *et al*. (1993). Comparison of naproxen and acetaminophen in a two-year study of treatment of osteoarthritis of the knee. *Arthritis and Rheumatism*, **36**, 1196–206.

Woessner, J. (1991). Matrix metalloproteinases and their inhibitors in connective tissue remodeling. *FASEB Journal*, **5**, 2145–54.

Wright, V. (1990). Post-traumatic osteoarthritis—a medico-legal minefield. *British Journal of Rheumatology*, **29**, 474–8.

Yelin, E. (1992). Arthritis: the cumulative impact of a common chronic condition. *Arthritis and Rheumatism*, **35**, 489–97.

5.16 Crystal arthropathies

Ann K. Rosenthal

The crystal arthropathies include gout, calcium pyrophosphate dihydrate (**CPPD**) deposition disease, the basic calcium phosphate (**BCP**)-associated syndromes, and calcium oxalate arthritis. The most common forms of crystal-associated arthritis are readily diagnosed at the bedside, and yet continue to present many challenges to the clinician.

Gout

History

Gout is one of the oldest known forms of arthritis, and holds a unique place in the history of medicine. Famous physicians such as Hippocrates (1886) and Sydenham (Copeman 1964) made observations about gout which remain true centuries later. Gout has served as a paradigm for the study of other types of crystal-associated arthritis and as a prime example of the way in which understanding the pathophysiology of a disease leads to the development of effective therapy.

Uric acid was first identified in a renal calculus in the late 1700s by Scheele (1776), but it was not until 1899 that Freudweiler (1899) showed that intra-articular injection of uric acid crystals mimicked attacks of gouty arthritis. Decades later, uric acid crystals were identified in synovial fluids (McCarty and Hollander 1961). Based on the prescient work of Garrod (1876) in the 19th century, the link between gout and hyperuricaemia was established. Progress was aided by the discovery that an enzyme involved in purine metabolism was responsible for the Lesch–Nyhan syndrome (Seegmiller *et al*. 1967). Subsequently, from knowledge of purine and uric acid metabolism, effective drugs for hyperuricaemia were developed.

Definition

The term gout is derived from the latin word 'guta' meaning a drop, and originally may have referred to a drop of poison or evil humor. However, gout remains poorly defined as a clinical descriptor (Simkin 1993). In its most general sense, it is a group of diseases characterized by hyperuricaemia and uric acid crystal formation. These clinical syndromes include gouty arthritis, tophaceous gout, uric acid nephrolithiasis, and gouty nephropathy. In its more narrow and perhaps more commonly used definition, gout refers to arthritis caused by uric acid crystals.

Epidemiology

Gout is not an uncommon disease. Recent statistics from the United States show a self-reported prevalence of 13.6/1000 persons in adult males and 6.4/1000 persons for females (Collins 1988). In older studies utilizing physician diagnoses, gender specific rates were 5 to 6.6/1000 in men and 1 to 3/1000 in women (Lawrence *et al*. 1989). Similarly, in a study of patients in the Scottish Highlands, the overall prevalence of physician-documented gout was 3.4/1000 (Steven 1992). With an estimated prevalence of 1 per cent among adults in the United States, gout is the most common cause of inflammatory arthritis in men (Lawrence *et al*. 1989). Despite effective therapy, it is estimated that 37 million working days per year in the United States have been lost to gout (Hochberg 1991).

Incidence rates for gout are more difficult to ascertain. A longitudinal study of American medical students showed an incidence of 1.7 cases per 1000 person-years of follow-up (Roubenoff *et al*. 1991). Similar results were obtained from the Framingham data (Abbott *et al*. 1988). Gout may be more common in certain areas of the world such as Java (Darmawan *et al*. 1992), and among certain races (Lennane *et al*. 1960).

There is good evidence to suggest that average uric acid levels have been slowly rising over the last two decades (Gresser *et al*. 1990). Self-reported prevalence studies also show a parallel increase in the prevalence of gout (Lawrence *et al*. 1989)

The known risk factors for gout are well characterized and mirror risk factors for hyperuricaemia (Table 1). Gout is clearly more common in men and is rare in women prior to menopause. Overall, less than 5 per cent of patients with gout are female (Lally *et al*. 1986). Mean urate levels rise at puberty in boys and remain 1 to 2 mg/dl higher in men than women until menopause, when gender differences in uric acid levels diminish. Serum urate levels increase with age in both men and women (Becker 1988). The peak incidence of gout in men is between the fourth and sixth decade (Hall *et al*. 1967). In women, it is between the sixth and seventh decade (Puig *et al*. 1991).

Male gender
Age >40 years
Obesity
Family history
Alcohol use
Renal insufficiency
Hypertension

Table 1 Risk factors for primary gout

Primary gout is associated with a variety of medical conditions. Some such as obesity, renal insufficiency, and diuretic use clearly cause hyperuricaemia. Serum urate levels are directly correlated with body weight (Seidell *et al*. 1986). Obesity is a particularly important risk factor for gout in men (Roubenoff *et al*. 1991). Similarly, alcohol use increases uric acid levels by providing a dietary source of purines (Gibson *et al*. 1984). Ethanol increases lactic acidaemia thus interfering with the excretion of uric acid (Lieber *et al*. 1962), and increases adenine nucleotide catabolism via acetate intermediaries (Puig and Fox 1984) At one time, alcohol (particularly port wine and moonshine or unbonded whisky) was also an excellent source of lead (Ball and Sorenson 1969).

Other risk factors for gout such as hypertension, family history, and coronary artery disease may not be causally associated with high uric acid levels. Hypertension occurs in 25 to 50 per cent of gout patients (Grahame and Scott 1970), and a similar percentage of untreated hypertensives will have hyperuricaemia (Wyngaarden and Kelley 1976). Gout is also associated with hypertriglyceridaemia (Scott 1977). The association between gout and coronary artery disease is somewhat controversial (Myers *et al*. 1968; Emmerson 1974; Abbott *et al*. 1988). Gout may be associated with atherosclerosis only because of the high prevalence of obesity and hypertension in gout patients (Abbott *et al*. 1988). Alternatively, gout may be an independent risk factor for coronary artery disease (Gertler *et al*. 1931). Lastly, between 6 and 18 per cent of patients with gout will have a family history of gout (Neel 1947).

Risk factors for gout in women are not as clearly defined as those for men. As in men, age, renal insufficiency, and diuretic use are certainly linked with gout in women. However, other male risk factors such as body weight and alcohol intake may not play as important a role in gout in women (Puig *et al*. 1991).

Pathophysiology

Delineating the pathophysiology of gout requires an understanding of purine metabolism, as uric acid is an end product of purine biosynthesis. Hyperuricaemia is a necessary but not a sufficient condition for the development of gout; and although the mechanisms of excess uric acid accumulation are well defined, the subsequent phases of crystal formation and release into tissues remain less well characterized.

Purine metabolism

Purines are derived from two sources. They are ingested in food or are generated via a complex *de novo* synthetic pathway (Wyngaarden and Kelley 1976). The synthetic pathway is outlined in Fig. 1. Components of the purine ring are complexed to the donor substrate phosphoribosylpyrophosphate (**PRPP**). These are then taken through a 10-step process culminating in purine nucleotide formation. PRPP is also used as a substrate for pyrimidine and pyridine synthesis. Thus, the first committed step in purine synthesis is catalysed by the enzyme amidophosphoribosyl transferase (**amidoPRT**). The *de novo* synthetic pathway requires heavy energy consumption in the form of ATP (Holmes 1978). Consequently, numerous enzymes for salvaging and interconverting premade purines exist to recycle these energy-rich compounds. Two salvage enzymes are particularly important in gout. These are hypoxanthine guanine phosphoribosyl transferase (**HPRT**) and adenine phosphoribosyl transferase (**APRT**).

Fig. 1 Simplified scheme of normal purine metabolism in man. PRPP synthestase, phosphoribosylpyrophosphate synthetase; amidoPRT, amidophosphoribosyl transferase; HGPT, hypoxanthine guanine phosphoribosyl transferase; 5'-NT, 5'-nucleotidase.

Control of normal purine metabolism in man is well understood. The first commited step in purine biosynthesis is rate limiting. AmidoPRT is allosterically activated by its substrate PRPP and inhibited by purine nucleotides, its end products (Holmes *et al*. 1973). The enzyme PRPP synthetase is similarly regulated, but is less sensitive to small changes in end-product concentrations (Yen *et al*. 1981; Becker and Kim 1987). Thus, control of purine metabolism is negatively affected by purines themselves and positively influenced by PRPP (Becker and Kim 1987; Itakura *et al*. 1981).

Uric acid metabolism

Uric acid is ultimately formed from purine nucleotides through the intermediate compounds xanthine, hypoxanthine, and guanine by the enzyme xanthine oxidase. It is a terminal product as no mammalian uricase exists. Uric acid is made primarily in the liver. The average pool size is 1200 mg in men and 600 mg in women. In both men and women, about two-thirds of the total uric acid pool is turned over each day (Scott *et al*. 1969). Uric acid pools in patients with gout are always larger than normal, usually in the range of 2000 to 4000 mg. In patients with tophi, uric acid burdens can be as high as 30 000 mg. (Benedict *et al*. 1949; Bishop *et al*. 1951; Sorenson 1959).

Two-thirds of uric acid is renally excreted (Buzzard *et al*. 1955; Sorenson 1960). The remainder is degraded by gut bacteria via the process of 'intestinal uricolysis' (Wyngaarden and Stetten 1953). The renal handling of uric acid is complex and not fully understood. It entails a four-step process beginning with glomerular filtration, followed by active reabsorption in the proximal tubule, tubular secretion at some distal site, and ending with postsecretory tubular reabsorption (Levinson and Sorenson 1980) (Fig. 2).

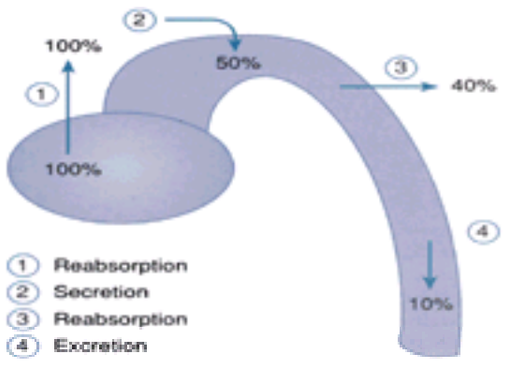


Fig. 2 Simplified scheme of renal excretion of urate. (1) reabsorption after filtration, (2) secretion in the proximal tubule, (3) reabsorption at some distal site, and (4) excretion in the urine.

For the sake of discussion, we will divide the known metabolic causes of hyperuricaemia associated with gout into three catagories: causes of primary gout, defined inborn errors of metabolism, and aetiologies of secondary gout.

Mechanisms of hyperuricaemia in primary gout

Primary gout is simply defined by the absence of any identifiable underlying disease causing hyperuricaemia. This criterion defines the largest group of patients with gout. Most of these patients are older men and 80 to 85 per cent are hyperuricaemic on the basis of underexcretion of uric acid (Wyngaarden and Kelley 1985). There is no difference in rates of intestinal uricolysis in patients with primary gout compared with controls (Sorenson 1962). Thus, the site of the abnormality in patients with primary gout who underexcrete is most likely to be at the kidney. Simkin (1977b) has shown that most patients with primary gout have low fractional uric acid excretion rates. The mechanism of underexcretion remains to be elucidated but is most likely a defect in secretion or reabsorption rather than in filtration (Reiselbach *et al*. 1970).

A minority of patients with primary gout have high urinary uric acid levels and excessive *de novo* purine synthesis. The best evidence to date supports a role for increased PRPP availability or decreased purine nucleotide concentrations (thus diminishing feedback inhibition of the synthetic enzymes) in patients with primary gout who overproduce (Levinson and Becker 1993).

A recent study of purine metabolism in women with primary gout showed similar abnormalities to men with primary gout (Puig *et al*. 1994).

Inborn errors of metabolism causing hyperuricaemia

The enzyme defects associated with gout often present as precocious gout in childhood or early adulthood in the setting of a strong family history of gout. There are three well-characterized enzyme defects causing hyperuricaemia. Together these account for less than 5 per cent of cases of gout. These are summarized in Table 2 and include HPRT deficiency, PRPP synthetase superactivity, and glucose 6-phosphatase (**G6P**) deficiency.

Syndrome	Pattern of inheritance	Mechanism of hyperuricaemia
HPRT deficiency (Lesch–Nyhan)	X-linked	HPRT deficiency increases PRPP
Increased PRPP synthetase	X-linked	Overactivity of PRPP synthetase
von Gierke's disease (G6P deficiency)	Autosomal recessive	Increased activity of amidoPRT, decreased renal excretion due to acidaemia

Table 2 Inherited metabolic disorders causing gout

HPRT deficiency produces hyperuricaemia by increasing *de novo* synthesis of purine nucleotides through increased availability of the substrate PRPP, which stimulates synthesis. In its most complete form, HPRT deficiency results in the Lesch–Nyhan syndrome (Lesch and Nyhan 1964). This syndrome presents in early childhood with severe mental retardation, self-mutilation, choreoathetosis, spasticity, hyperuricaemia, and premature gout. Although it is linked to the X chromosome, there are two reported cases in girls (Yukawa *et al*. 1992). Partial defects of HPRT result in hyperuricaemia alone without the severe neurological consequences of complete HRPT deficiency (Kelley *et al*. 1967). Some patients with partial HPRT deficiencies may have subtle neurological impairments (Kelley *et al*. 1969). The diagnosis of HPRT deficiency can be made by measuring HPRT activity in erythrocytes (Kelley *et al*. 1969). In partial deficiencies, activity levels vary from less than 1 to 70 per cent of normal values. Over 26 different genetic mutations have been described in HPRT deficiency (Wilson *et al*. 1983).

Increased PRPP synthetase activity was described by Sperling *et al*. (1972) in two brothers. Overactivity of this enzyme results in increased PRPP levels and causes profound premature hyperuricaemia. The trait is inherited as an X-linked dominant condition (Yen *et al*. 1978). Three kindreds each with a different genetic mutation have been well characterized. Gout and renal stones develop in the second and third decades in affected males.

The third well-characterized enzyme defect associated with gout is glucose 6-phosphatase deficiency (Alepa *et al*. 1967). This is also known as glycogen storage disease type I, or von Gierke's disease. It presents in childhood with characteristic short stature, hepatomegaly, and recurrent hypoglycaemia. Less frequently, a bleeding diathesis may accompany the syndrome. Affected patients cannot release glucose from premade glycogen stores. Subsequent hypoglycaemia results in ATP catabolism, lactic acidaemia, and elevated levels of free fatty acids, pyruvate, and triglycerides. Hyperuricaemia in this syndrome is due to two effects. Renal excretion of uric acid is diminished because other organic anions compete for transport in the kidney. Overproduction via the *de novo* synthetic pathway due to decreased feedback inhibition of amidoPRT also occurs (Itakura *et al*. 1981). Hyperuricaemia is often present from infancy, with gout occurring as early as 10 years of age. In its classic form, the inheritance is as an autosomal recessive disorder. Some forms of partial enzyme deficiency have been described (Nuki and Parker 1979).

Causes of secondary gout

The causes of secondary gout are well defined. Some may result in hyperuricaemia on the basis of overproduction of uric acid, such as tumour lysis syndrome, myeloproliferative disease, haemolytic amaemia, and psoriasis. All of these conditions are characterized by increased cell turnover with a subsequent increase in purine synthesis and catabolism.

Alternatively, conditions such as renal failure and many drugs produce hyperuricaemia and gout by promoting undersecretion of uric acid. Although mild renal insufficiency from any cause is a risk factor for gout (Berger and Yu 1960), certain forms of kidney disease such as autosomal dominant polycystic kidney disease may be preferentially associated with gout (Mejias *et al*. 1989). Interestingly, endstage renal disease often produces hyperuricaemia, but these patients rarely develop gout (Richet *et al*. 1965; Sarre and Mertz 1965).

Lead exposure is a cause of secondary gout. Saternine gout, which is associated with heavy lead exposures (often from the ingestion of lead-laden whisky), is rare today. Lead interferes with renal excretion of uric acid by altering the tubular transport of urate. It affects purine metabolism by altering purine nucleotide turnover

(Ludwig 1957; Forkas *et al*. 1978). Lead levels are increased in gout patients. Whether this phenomenon is primary or secondary to the renal insufficiency that often accompanies gout remains to be determined. Industrial lead exposure occurs in plumbers, pipe fitters, painters, steel workers, battery plant employees, and vehicle mechanics (Hochberg 1991). Recent changes in environmental safety standards may reduce lead exposure in these settings.

Drugs commonly associated with hyperuricaemia are listed in Table 3.

Drugs causing overproduction	Drugs causing underexcretion
Ethanol	Ethanol
Fructose	Salicylates (< 2 g/day)
Cytotoxic drugs	Cyclosporin
Vitamin B_{12}	Diuretics
Warfarin	Ethambutol
	Pyrazinamide
	Levodopa
	Angiotensin
	Vasopressin
	Laxatives
	Nicotinic acid
	Nitroglycerin
	Methoxyflurane

Table 3 Drugs associated with hyperuricaemia and gout

Urate crystal formation (Fig. 3)

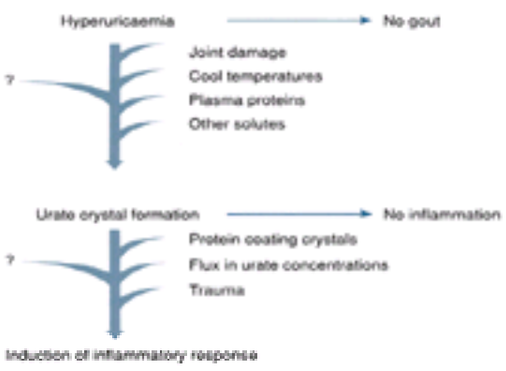


Fig. 3 Overview of factors influencing the development of gouty arthritis.

We know little about why uric acid crystals form *in vivo*. Yet, because the vast majority of patients with hyperuricaemia never develop gout (Hall *et al*. 1967), it is these processes of crystal formation and release that define clinical gout.

Our current understanding of urate crystal formation is based on knowledge of the biochemistry of urate, *in vitro* studies of crystallogenesis, and the histopathology of gout. Urate is present in two forms in the body. At neutral and alkaline pH, the monosodium salt (**MSU**) predominates. At acidic pH, such as in urine, the primary form is uric acid. Both forms are relatively insoluble, although uric acid is less soluble than its salt (Peters and Van Slyke 1946). A solution is supersaturated with monosodium urate at 37°C at concentrations greater than 6.4 mg/dl (Peters and Van Slyke 1946). Cooler temperatures decrease urate solubility (Loeb 1972), thus predicting the association of gout with distal joints.

In vitro studies of crystallogenesis have yielded conflicting results in regard to the influences of serum, synovial fluid, and many individual plasma proteins, including albumin and proteoglycans, on crystal formation (Katz and Erlich 1986; McGill and Dieppe 1991b). Crystal formation is probably enhanced by IgG and type I collagen (McGill and Dieppe 1991b). In addition, urate crystals themselves initiate further crystal formation. McGill and Dieppe (1991a) and others (Tak I>et al. 1980; Burt and Dutt 1986) demonstrated increased crystallogenesis in the presence of 'particulate-free' synovial fluid from patients with gout compared with those with rheumatoid arthritis or CPPD deposition disease, suggesting the presence of as yet uncharacterized promoters of crystal formation in gouty fluid.

Urate crystals have a unique distribution in the body. They prefer sites rich in connective tissue, such as synovium, cartilage, tendon, skin, and the renal interstitium (Sokoloff 1957). Gouty arthritis often develops in previously damaged joints. Disturbances of proteoglycan metabolism or collagen structure may promote crystal formation and release (Katz and Erlich 1974; Wilcox and Khalaf 1975; Perricone and Brandt 1978). Joint effusions may also affect urate crystal formation. Because the diffusion of urate through the synovial membrane is less than that of water (Simkin 1977a), resorption of water in a joint with an effusion during recumbency would increase the effective urate concentration and promote crystal formation.

Pathogenesis of the acute attack (Fig. 4)

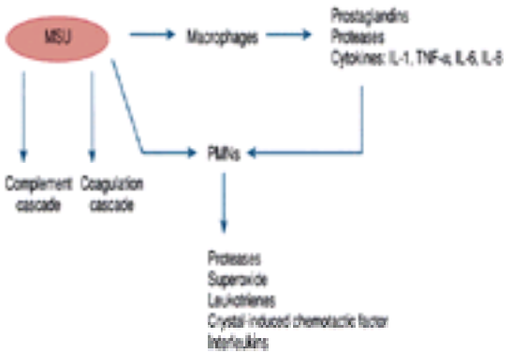


Fig. 4 The mechanisms through which monosodium urate (MSU) crystals cause inflammation in the joint.

Once monosodium urate crystals form in the joint, they may induce an acute attack of gouty arthritis. Urate crystals clearly cause inflammation (Seegmiller *et al*. 1962); yet we are unable to explain why they are present in synovial fluids from asymptomatic uninflamed joints (Pascual 1991). The type and quantity of protein coating the crystals may affect their ability to induce inflammation. IgG is found in association with MSU crystals and increases cell activation (Russell *et al*. 1983). Terkeltaub *et al*. (1991) demonstrated a role for the lipoprotein apoE in inhibiting crystal-induced inflammation.

The effects of uncoated MSU crystals on cells are well characterized. The cells that urate crystals first encounter in the joint are most probably macrophage-like synovial cells. Here they induce the release of vasoactive prostaglandins, proteases, and proinflammatory cytokines including interleukin 1 (**IL-1**), IL-6, and IL-8 which initiate a vigorous inflammatory response (DiGiovine *et al*. 1987). Recruited polymorphonuclear leucocytes release proteases, superoxides, leukotrienes, and interleukins when exposed to MSU crystals. (Abramson *et al*. 1982; Cheung *et al*. 1983b; Bhatt and Spillberg 1988). MSU crystals activate complement via the

classical pathway (Giclas *et al*. 1979). They also promote the release of Hageman factor and the subsequent activation of bradykinin, kallikrein, and other coagulation factors (Ginsberg *et al*. 1980).

The mechanisms through which they stimulate these cells remain to be elucidated. Certainly, some crystals are phagocytosed and cause lysis of the phagolysosome, release of its toxic contents, and death of the cell (Schumacher and Phelps 1971). Other effects may be mediated through cell membrane perturbations (Mandel 1976). The highly charged surfaces of MSU crystals may bind to and cross-link membrane receptors, thus mediating some of the crystals' immediate effects (Burt and Jackson 1990).

The factors that terminate an acute attack remain unclear. One hypothesis suggests changes in crystal size and protein coating render the crystals less inflammatory. These changes may be mediated through generation of oxygen free radicals by polymorphonuclear leucocytes (Marcalongo *et al*. 1988).

Tophus formation

Tophi are soft tissue deposits of urate. We know little about how and why they form. Palmer *et al*. (1989) proposed that tophi are urate-lowering organs. Based on histological studies, they proposed that acini of macrophages develop in areas of high local urate concentrations. These organized cells actively transport urate from the interstitial fluid to the centre of the acinus. They grow to a certain size and then fuse with other acini, eventually forming tophi. Further work is necessary to confirm this theory.

Clinical features

We will discuss the clinical features of acute and chronic gout, and renal syndromes associated with uric acid.

Articular gout is often divided into four clinical stages. The first stage is defined by asymptomatic hyperuricaemia. This is followed by acute gouty arthritis and then by another asymptomatic phase termed intercritical gout. When allowed to proceed untreated, some patients will go on to develop chronic tophaceous gout.

Acute gout

The first clinical symptom of gout is usually an acute, self-limited, monoarticular inflammatory arthritis affecting the joints of the lower extremities. Gout has a predilection for the first metatarsophalangeal joint. As many as 50 to 70 per cent of first gout attacks occur in the big toe (Delbarre *et al*. 1967; Grahame and Scott 1970). Other frequently involved joints include those of the foot, ankle, knee, wrist, elbow, and the small joints of the hands. The large axial joints and those of the spine are uncommon sites for early acute gout attacks.

The onset of an attack occurs suddenly and often late at night or early in the morning. Patients will describe very severe pain, associated with swelling, extreme tenderness, and redness overlying the joint. Without intervention, the attack will usually subside within 5 to 7 days (Bellamy *et al*. 1987). Low-grade fever, malaise, and anorexia may occur. The attack may be preceded by brief twinges of pain (petit attacks) in the affected joint.

Common precipitants of acute attacks include excess alcohol intake, intercurrent illness, surgery (Bartles 1957), starvation, trauma, and the initiation of drugs that alter urate metabolism. All of these precipitants alter serum urate levels.

Physical examination shows signs of inflammation with erythema, warmth, and swelling over the joint, often extending to the overlying skin. There is exquisite tenderness over the affected joint. Not infrequently, an overlying cellulitis or accompanying tenosynovitis occurs. The skin may desquamate in the later days of an attack. Acute gout can also occur in bursas, and gout is a common cause of acute inflammatory olecranon bursitis (Canosa and Yood 1979).

After the attack resolves, the patient will be completely asymptomatic. This phase is referred to as intercritical gout. Most patients will go on to have an additional attack within 2 years of the first attack (Gutman 1973). In one study, 78 per cent had recurrent attacks within 2 years, and after 10 years, 93 per cent had had more than one attack. Untreated, the intercritical phases become shorter. Interestingly, they still present an opportunity for diagnosis, as many joints will still have urate crystals in the synovial fluid during the intercritical phase if they were involved in a previous attack, and urate-lowering therapy has not been initiated (Pascual 1991).

Chronic tophaceous gout

In the later stages of untreated disease, clinical manifestations characteristically change. Acute attacks are more often polyarticular. The intercritical stage shortens, and repeated joint damage results in permanent deformities, loss of motion, chronic pain, and tophi (Nakayama *et al*. 1984).

Polyarticular gout occurs in late stage disease, although some patients present earlier with polyarticular attacks (Raddatz *et al*. 1983; Lawry *et al*. 1988). Intercritical stages are short or non-existent and involvement of atypical sites including the upper extremities, the spine, and axial joints may ensue (Lagier and MacGee 1983; Varga *et al*. 1985). After repeated attacks in a single joint, deformity and loss of motion may occur.

Tophi are deposits of urate embedded in a matrix composed of amorphous urates, lipids, proteins, and calcific debris (Fig. 5). Tophi are usually subcutaneous, but they can occur in bone and other organs including the heart valves and the eye (Ferry *et al*. 1985; Gawoski *et al*. 1985). Classic sites include the pinna of the ear, bursas around elbows and knees, the dorsal surfaces of the metacarpophalangeal joints, and the Achilles tendon. Tophi are not distinguishable on physical examination from rheumatoid nodules or other subcutaneous nodules. There is no accompanying inflammation and they are usually painless. The overlying skin may be taut and shiny. A thick white or whitish-yellow exudate is seen if the skin integrity is compromised.

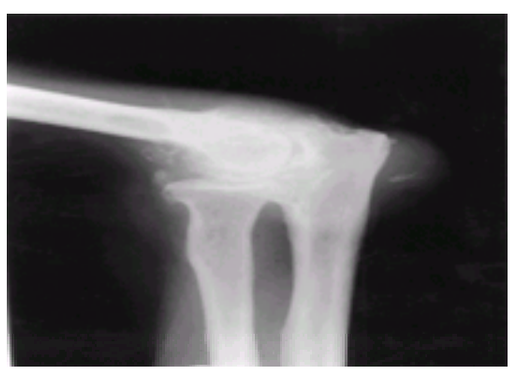


Fig. 5 A tophus in the olecranon bursa appears as a typical soft tissue density with a small amount of calcification.

Tophi or chronic polyarthritis may occur as early as 3 years or as late as 42 years after the first acute attack. In the pretreatment era, 50 per cent of patients with gout had tophi after 10 years of disease (Gutman 1973). Currently, about 5 per cent of patients with gout will have tophi (O'Duffy *et al*. 1975). Their occurrence is directly correlated with serum urate levels, and they identify a group of patients with severe and prolonged hyperuricaemia (Nakayama *et al*. 1984). Another group at risk of developing tophi and olyarticular gout are elderly women with primary nodal osteoarthritis on diuretic therapy (Macfarlane and Dieppe 1985).

Other articular manifestations of gout

Gout has been variably associated with avascular necrosis of the femoral head (Hunder *et al*. 1968; Stockman *et al*. 1980). Patients with gout may also have a higher prevalence of chondrocalcinosis (Dodds and Steinbach 1966).

Renal disease and gout

The relationship between kidney dysfunction and gout remains complex and confusing. Three renal syndromes are associated with gout. Urate crystals can form in the renal interstitium causing urate nephropathy. Uric acid can acutely precipitate in the collecting tubules resulting in uric acid nephropathy. Lastly, uric acid nephrolithiasis may occur.

Urate nephropathy

The pathological changes that define urate nephropathy are common. MSU crystals in the renal medulla are associated with a giant-cell inflammatory reaction (Sokoloff 1957). The clinical significance of these pathological findings, however, remains unclear. Renal insufficiency is unequivocally common in patients with gout, but controversy exists as to the aetiology of this renal dysfunction. Current dogma states that urate crystals themselves produce only a minor amount of renal damage. Most of the renal disease associated with gout is secondary to inadequately controlled hypertension and other comorbidities (Fessel 1967; Berger and Yu 1960).

Uric acid nephropathy

Excluding nephrolithiasis, the renal syndrome most often associated with uric acid today is acute uric acid nephropathy (Reiselbach *et al*. 1962; Frei *et al*. 1963). This often occurs in an acutely ill, dehydrated patient treated with cytotoxic drugs for a lymphoproliferative disorder. An acute obstructive uropathy ensues with oliguric renal failure. Uric acid crystals form in the collecting tubules and are found in the urine. Uric acid/creatinine ratios are often greater than 1.0. This complication can be avoided with adequate hydration and the prophylactic administration of allopurinol prior to initiating chemotherapy.

Uric acid nephrolithiasis

The association between gout and nephrolithiasis is well established (Yu and Gutman 1967). Prevalence figures for renal stones vary between 10 and 42 per cent of patients with gout. Of these stones, 84 per cent are uric acid stones. Risk factors for developing uric acid stones include elevated urinary uric acid levels, and low urine pH (Plante *et al*. 1968; Yu 1981). Half the patients excreting greater than 1100 mg uric acid per day will have stones (Yu and Gutman 1967). Interestingly, however, only 20 per cent of patients with uric acid stones are hyperuricaemic (Talbott 1957). Patients with gout also are at a higher risk for non-urate stones. Of the 16 per cent of stones in patients with gout that do not contain uric acid, 8 per cent are calcium oxalate, 4 per cent are calcium phosphates, and 4 per cent are mixed stones. It is postulated that uric acid serves as a nidus for calcium oxalate crystal growth. Alternatively, or in addition, levels of inhibitors of calcium oxalate stone formation may be decreased in the urine of patients with gout.

Laboratory investigation

Investigation of the patient with an acute attack

Synovial fluid analysis remains the single most important diagnostic study in the patient with suspected gout. Synovial fluid is usually easily obtained from a large joint, while often only a drop of fluid or blood from the joint or adjacent tissues is necessary to provide a sample for definitive diagnosis of gout in a small joint. Synovial fluid is typically inflammatory with a mean white cell count of 20 000 cells/mm ^{3}. Most cells are polymorphonuclear leucocytes. Viscosity is often poor. A definite diagnosis can be made if typical, negatively birefringent, needle-shaped crystals are seen in the fluid with a polarizing light microscope (Fig. 6). They may be extra- or intracellular. Rarely one may see spherules of uric acid in acute gout (Fiechtner and Simkin 1981).

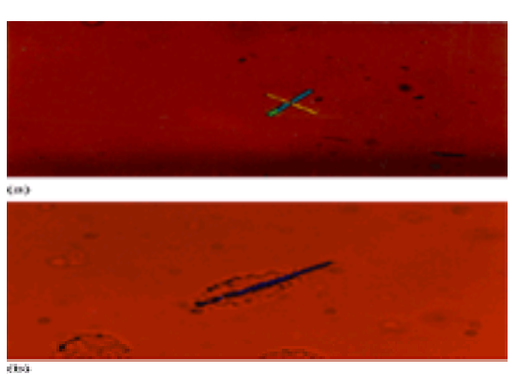


Fig. 6 Typical, needle-shaped, negatively birefringent crystals are seen under polazing light microscopy in the synovial fluid of patients with gouty arthritis. These crystals may be extracellular (a) or intracellular (b).

Few other laboratory studies are of significant clinical utility in diagnosing acute gout. Serum uric acid levels during the acute attack may not reflect pre-attack levels and can not be used to make a diagnosis of gout in the absence of urate crystals in the synovial fluid. One may see a peripheral leucocytosis, an elevated erythrocyte sedimentation rate, and increased levels of other acute-phase reactants during an acute attack. Synovial fluid cultures and Gram stains may help rule out concurrent infection. Radiographs are often normal during early episodes of gout. They may be useful to differentiate other problems such as fracture or infection from acute gout. Often soft tissue swelling is the sole radiographic finding in early gout.

Evaluation of the patient with recurrent attacks

Serum uric acid levels and 24-h urine collections for creatinine and uric acid may be helpful in evaluating the patient once the acute attack has subsided. Serum urate levels over 7 mg/dl define hyperuricaemia in most laboratories. Values of urinary uric acid over 1000 mg/day on an unrestricted diet define patients that overproduce, and may influence the choice of therapy. Patients suspected of a primary metabolic disorder should have a 24-h urine sample tested for creatinine, protein, and uric acid and a careful family history taken. If urinary uric acid levels are high, levels of enzyme activity for HPRT, PRPP synthetase, and glucose 6-phosphatase can be measured in specialist laboratories.

In later disease, gout has a typical radiographic appearance (Block *et al*. 1980; Barthelemy *et al*. 1984). The hallmarks of radiographic gout are due to the presence of tophi in or near the joint. In the soft tissues, tophi appear eccentric and nodular (Fig. 5). A small percentage of them calcify. Tophi may occur in or near a joint or distant from periarticular tissues. As a soft tissue tophus enlarges, it may encroach on the adjacent bone and produce a cortical erosion with focal periosteal new bone formation. This occurrence results in an erosion with a typical overhanging margin, present in 40 per cent of gouty erosions (Martel 1968) (Fig. 7). When in the joint, tophi produce marginal erosions with a characteristic 'punched out' or sclerotic border (Fig. 8). Erosions are particularly common on the medial and dorsal portions of the first metatarsal head. Similar changes can affect the digits of the hand. Bony abnormalities of the periosteum may occur in association with tophus formation. Specifically, a 'lace pattern' of finely striated periosteal reaction may develop adjacent to a tophus (Fig. 9). Rarely, bony proliferation may occur at the ends or shafts of long bones (mushrooming). Diaphyseal thickening may also occur. The joint space is characteristically well preserved until late in the disease. When joint space narrowing does occur it affects all joint compartments symmetrically, similar to other inflammatory joint disorders (Fig. 10). Bones may be osteopenic from disuse, but are usually well mineralized.

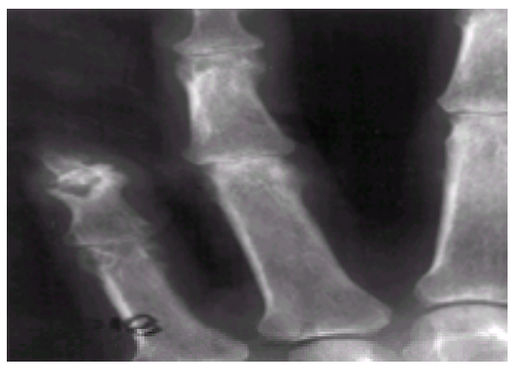


Fig. 7 A typical gouty erosion with overhanging margins is seen on the medial aspect of the proximal interphalangeal joint in this radiograph.

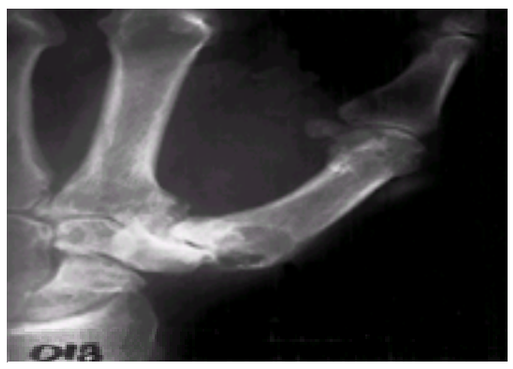


Fig. 8 An interosseus tophus appears as a 'punched out' or sclerotic erosion at the base of the proximal phalanx of the thumb.

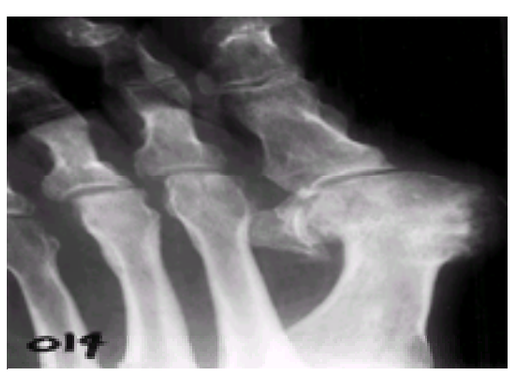


Fig. 9 On the lateral aspect of the distal metatarsal joint, a lace-like periosteal reaction adjacent to a tophus can be seen.

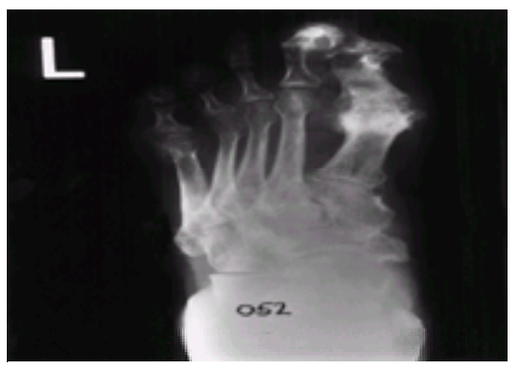


Fig. 10 Extensive bony destruction and deformities are seen with far-advanced gouty arthritis of the foot.

Diagnosis

A definitive diagnosis of gout can only be made by the identification of urate crystals in the synovial fluid of an affected joint. Identification of urate crystals in tophi also allows a definitive diagnosis to be made. In the absence of these findings, other clinical criteria may be used to make a putative diagnosis of gout (Wallace *et al*. 1977). As crystals may be present in the intercritical phase, one may aspirate an asymptomatic but previously affected joint to establish a definite diagnosis.

Many clinical conditions can mimic acute gout (Table 4). These include infectious arthritis, other crystal-associated arthropathies such as pseudogout or BCP-associated periarthritis, or trauma. Patients with palindromic rheumatism may give a similar history of self-limited monoarticular attacks associated with exquisite pain, tenderness, and erythema near the affected joint. Rarely, other causes of polyarticular inflammatory arthritis, particularly psoriatic arthritis or Reiter's syndrome, may present with monoarticular self-limited attacks of the lower extremities which may be confused with gout. Once tophi and deformities occur, gout can be misdiagnosed as rheumatoid arthritis. As many as 30 per cent of patients with gout will have positive serum rheumatoid factors (Kozin and McCarty 1977).

CPPD disease (pseudogout)
BCP arthritis
Cellulitis
Erythema nodosum arthritis
Trauma
Palindromic rheumatism
Reiter's syndrome
Psoriatic arthritis
Rheumatoid arthritis

Table 4 Clinical conditions that mimic gouty arthritis

Management of gouty arthritis

Management of gouty arthritis can be divided into three phases: treatment of acute gout, treatment of chronic or tophaceous gout, and preventive measures.

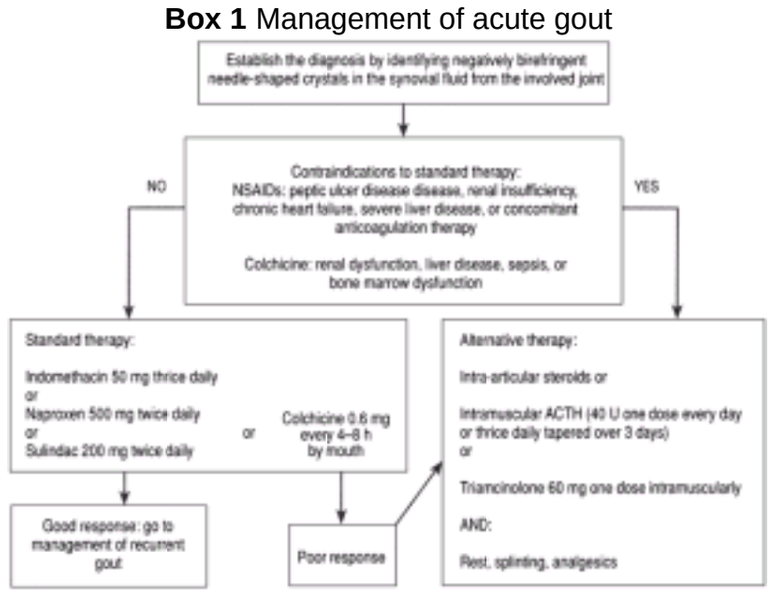
Preventive measures

Gout is a significant public health problem despite our excellent therapy. Hochberg (1991) suggests that we may be able to decrease the incidence of gout with preventive measures such as avoiding excess weight gain, reducing risks for hypertension, avoiding diuretic therapy, controlling alcohol intake, and minimizing occupational lead exposures. The efficacy of such interventions remains to be proven.

Management of acute gout

Traditional therapies for acute gout include non-steroidal anti-inflammatory drugs (**NSAIDs**), colchicine, and steroids. Rest and splinting of the affected joint may be helpful adjuncts to any pharmacological therapy. The management of acute gout is summarized in Box 1.

Box 1 Management of acute gout



NSAIDs

NSAIDs have replaced colchicine as the most commonly used drugs in the treatment of acute gout (Wallace 1977; Stuart *et al*. 1991). They interfere with the inflammation induced by MSU crystals (Abramson and Weissmann 1989). Traditionally, indomethacin has been the NSAID of choice in acute gout, but probably has no advantage over other NSAIDs (Sterling 1991). Sulindac may be better tolerated in those patients at high risk of renal side-effects from NSAIDs (Ciabottoni *et al*. 1984). In general, drugs with a shorter half-life achieve quicker therapeutic plasma levels and faster relief of pain.. With treatment, symptoms should subside within 3 to 5 days. NSAIDs are contraindicated in patients with significant renal insufficiency, peptic ulcer disease, concurrent warfarin therapy, or liver disease.

Colchicine

Colchicine is the oldest drug for gout, but its safety remains controversial. When used correctly, it can be very effective with a rapid onset of action. Colchicine tends to be much more effective in the early hours of an attack and loses efficacy with time. The mechanism of action of colchicine is unknown. It inhibits polymorphonuclear leucocyte function through its action on microtubules, but may also have very specific effects on crystal-induced inflammation (Roberge *et al*. 1993). Colchicine can be given in intravenous and oral forms.

Current recommendations for intravenous use are cautious (Moreland and Ball 1991) (Table 5). A dose of 1 to 3 mg diluted in 20 ml of normal saline can be slowly instilled into a large vein. Another 1-mg dose can be given 6 h later if the clinical response is incomplete. The maximum dose is 4 mg in 24 h. No additional doses should be given for 7 days after the initial dose. Intravenous colchicine is particularly useful for patients unable to take oral medications. Now that parenteral forms of NSAIDs are available, its use may decline further. Absolute contraindications to the use of intravenous colchicine include significant renal or hepatic compromise, bone marrow suppression, or sepsis. It should be used with great caution in patients with mild renal or hepatic disease, and, for example, those on daily oral colchicine prophylaxis. Side-effects range from venous sclerosis or tissue damage from extravasation of colchicine, to fatal bone marrow failure. Other side-effects include renal or hepatic failure, disseminated intravascular coagulation, and neuromuscular toxicity. Deaths from misuse of intravenous colchicine have been reported (Roberts *et al*. 1987).

(1) Intravenous colchicine should not be given in patients with:
- Creatinine clearance <10 ml/min
- Significant active liver disease or extrahepatic biliary obstruction
- Sepsis
- Bone marrow depression

(2) Intravenous colchicine should be used cautiously in patients on oral colchicine prophylaxis, with poor venous access, or with any degree of hepatic or renal compromise

(3) Recommended doses should not exceed 3 mg intravenous per 24 h or 5 mg orally per 24 h

Table 5 Recommendations for the use of colchicine

Oral colchicine is currently used more frequently than the intravenous form. It is given as an initial 0.5 to 0.6 mg dose which is repeated every 1 to 2 h until gastrointestinal symptoms ensue or pain resolves. Doses should not exceed 5 mg in 24 h. Similar side-effects are reported for oral and intravenous forms. Oral colchicine, however, has a higher incidence of gastrointestinal side-effects and a lower incidence of major toxicities, probably because the tolerated dose is lower.

Corticosteroids

Corticosteroid use in gout has endured much ebb and flow in popularity during recent years. Although ACTH has been used for many years to treat gout, textbooks of the last two decades cautioned against systemic steroids because of concerns about rebound symptoms and inconsistent results (Gutman and Yu 1952). More recently, there has been a resurgence of interest in their use in gout. Regimens for acute gout include intramuscular ACTH (Axelrod and Preston 1989; Ritter *et al*. 1993), intramuscular triamcinolone (Siegel *et al*. 1994), and oral steroids (Groff *et al*. 1990) (Table 6). These regimens are safe and well tolerated. Their efficacy remains to be proven. They may be particularly useful for patients in whom NSAIDs and colchicine are contraindicated (Ritter *et al*. 1993). The use of intra-articular steroids is less controversial. Although no studies of the efficacy of intra-articular steroids have been published, they remain a mainstay of therapy in patients unable to tolerate more traditional therapies (Hollander 1953).

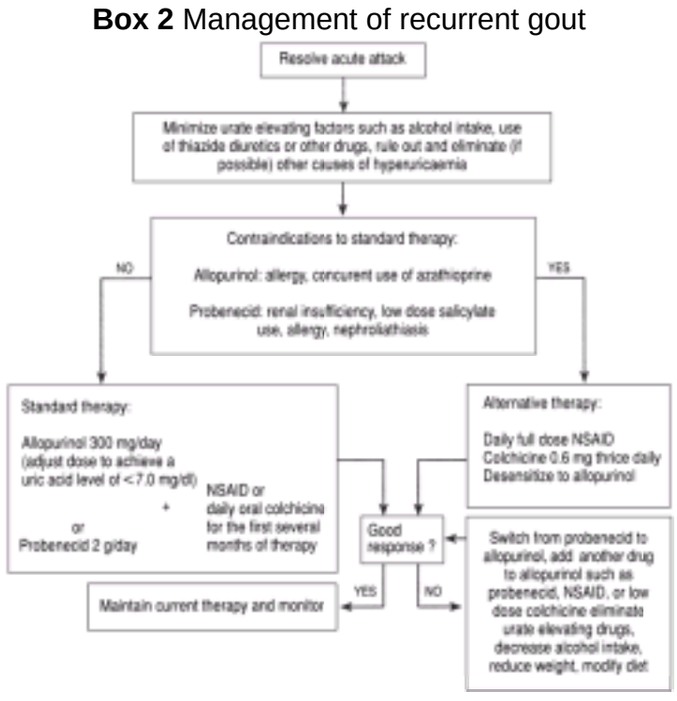
Drug	Dose	Route of administration	Length of use
Triamcinolone acetonide	60 mg	Intramuscular	Give once, repeat in 48 h if needed
Prednisone	20–50 mg with daily taper	Oral	3–20 days (mean 10 days)
ACTH	40 IU	Intramuscular	Give once
ACTH	40–80 IU	Intramuscular, intravenous or subcutaneous	Every 6 h then every 12 h then each day for 3 successive days

Table 6 Regimens for use of corticosteroids in acute gouty arthritis

Prophylactic therapy

Drugs which reduce serum uric acid levels such as allopurinol and probenecid are the standard therapies available for prophylaxis of gout attacks. Colchicine and NSAIDs are less commonly used as prophylactic drugs, but may also decrease attack frequency and severity when used alone or in combination with other therapies. Indications for the use of prophylactic drugs include recurrent attacks, tophi, severe or polyarticular disease, renal disease, or an inborn error of metabolism causing gout. Urate-lowering therapies are traditionally not started during an acute attack and are initiated concurrently with a 2-week course of low-dose colchicine or a NSAID. This regimen may avoid the risk of precipitating an acute attack by rapidly lowering uric acid levels. Goals of therapy are to decrease attack frequency, dissolve tophi, and maintain serum uric acid levels in the normal range. The management of recurrent gout is summarized in Box 2.

Box 2 Management of recurrent gout



Allopurinol

Allopurinol is the drug most commonly used for the prevention of acute gout and the treatment of chronic tophaceous gout (Stuart *et al*. 1991). It is a xanthine oxidase inhibitor and thus decreases uric acid production. It may also have other actions (Adriani and Naraghi 1985). It is very effective in lowering serum uric acid levels and is the drug of choice for patients with renal insufficiency, a history of nephrolithiasis, or tophi. It is also indicated in patients who clearly overproduce uric acid such as those with tumour lysis syndrome or primary metabolic defects. Allopurinol is usually given as a once daily dose of 300 mg. The dose should be adjusted downward in the presence of significant renal compromise. It may be increased to a maximum of 900 mg/day to achieve normal uric acid levels. The onset of action is rapid and effects can be seen as early as 4 days to 2 weeks (Wyngaarden *et al*. 1965; Rundle *et al*. 1966). Allopurinol interferes with the metabolism of azathioprine, potentiating its marrow-suppressive effects (Boston Collaborative Drug Surveillance Program 1974), and may augment anticoagulant effects of warfarin (Self *et al*. 1957). In general, the incidence of side-effects with allopurinol is low (less than 2 per cent). The most common side-effects include a hypersensitivity syndrome of rash and fever (McInnes *et al*. 1981; Singer and Wallace 1986). Life-threatening reactions, including fulminant hepatitis, interstitial nephritis, and toxic epidermal necrolysis, are even more unusual (Lupton and Odom 1979). Patients with renal disease may have a greater incidence of drug allergy (Singer and Wallace 1986). Allopurinol can be given cautiously to an allergic patient using available desensitization regimens (Northridge and Almack 1986; Fam *et al*. 1992).

Uricosurics

Probenecid and sulphinpyrazone are the most commonly used uricosuric drugs. They interfere with the renal handling of urate by altering organic anion transport, thus increasing urate excretion (Gutman and Yu 1957). Usual doses of probenecid are 0.5 to 2 g given in twice daily doses (Boger and Strickland 1955). Sulphinpyrazone is given in doses of 300 to 400 mg/day (Emmerson 1963). Both drugs are well tolerated and effective. Maximum doses of 3 g/day of probenecid and 800 mg/day of sulphinpyrazone are occasionally necessary to return the levels of serum uric acid to normal. Uricosurics are contraindicated in patients with nephrolithiasis and ineffective in patients with significant renal compromise, or those using acetylsalicylates. Side-effects include rare hypersensitivity reactions, rashes, gastrointestinal complaints, and nephrotic syndrome (Reynolds *et al*. 1957; Hertz *et al*. 1972).

Benzbromarone is a uricosuric drug available in Europe, that is effective in the presence of renal insufficiency. At doses of 25 to 125 mg/day, it may lower values of serum uric acid in patients with serum creatinine levels as high as 2.0 mg/dl (Zollner *et al*. 1970).

Colchicine

In patients who are unable to take allopurinol or a uricosuric drug, daily low-dose oral colchicine may be useful in preventing further attacks of gout (Yu and Gutman 1961). It is given in doses of 0.5 to 1.2 mg/day. Unfortunately, serious side-effects have occurred even on this low-dose regimen in patients with renal insufficiency or liver disease (Neuss *et al*. 1986).

Drug combinations and other therapies

Patients who continue to have gout attacks on a single prophylactic drug can be treated with a combination of medications. The combination of probenecid and colchicine is more effective than probenecid alone (Paulus *et al*. 1974). Small numbers of patients have been reported to respond to a combination of allopurinol and a uricosuric (Kuzell *et al*. 1966) when either one alone was ineffective. Other measures such as eliminating the concurrent use of urate-elevating drugs, decreasing alcohol intake, weight reduction, and improving medication compliance may also help achieve control of uric acid levels in refractory patients. Rarely, patients with long-standing uncontrolled disease have permanent joint dysfunction which is best treated surgically. Tophi may occasionally require surgical removal with a low incidence of recurrence (Smyth 1953).

Uricase has been used as a novel approach to gout therapy. This enzyme dissolves uric acid crystals and has been used successfully in a small number of patients (Chua *et al*. 1988; Rozenberg *et al*. 1993). In some cases, polyethylene glycol treatment of the uricase was used to reduce its immunogenicity (Chua *et al*. 1988).

Management of nephrolithiasis

Patients with uric acid stones are best managed with adequate hydration, urinary alkalinization (with bicarbonate or acetozolamide), and allopurinol. Potassium citrate may be used in place of allopurinol if necessary (Pak *et al*. 1986). This regimen is also effective in preventing calcium oxalate stones in patients with hyperuricaemia (Pak *et al*. 1981).

Gout in the transplant recipient

In the late 1980s, it was noted that an unusually large number of cases of gout were occurring in renal transplant patients. This coincided with the popular use of cyclosporin as an immunosuppressant. Hyperuricaemia occurs in 30 to 84 per cent and gout in 7 to 9 per cent of patients with renal transplants who are treated with cyclosporin (Lin *et al*. 1989). The incidence of gout in patients with heart and lung transplants on cyclosporin is also significantly increased (Burak *et al*. 1992). Gout in these patients is often more severe than primary gout. Polyarticular and tophaceous gout occurs earlier in cyclosporin-induced gout than in primary gout (Baethge *et al*. 1993). Post-transplant patients can be particularly difficult to treat because of contraindications to the use of NSAIDs and colchicine, resistance to systemic steroids which they take chronically, and concerns about interactions between allopurinol and azathioprine. These patients are frequently on diuretics, and their risk of infection is high because of their compromised immune status. In the acute management of gout in the transplant patient, an arthrocentesis is usually warranted to eliminate the possibility of infection. Once a diagnosis is established, intra-articular steroids and systemic pain medications may be helpful. Chronically, some of these patients may be treated with uricosurics. If allopurinol is indicated, azathioprine can be reduced to one-third of its usual dose and 50 to 100 mg of allopurinol can be added. Blood counts should be carefully monitored during the initiation of allopurinol therapy.

Prognosis

In the era of antihyperuricaemic therapy, the prognosis for patients with gouty arthritis is excellent. Moreover, despite the association of hyperuricaemia with heart disease, hypertension, and renal insufficiency, there is no evidence that patients with gout have decreased longevity compared with controls who do not have gout.

Calcium pyrophospate dihydrate (CPPD) disease

CPPD disease is the second most common form of the crystal-associated arthritides. Unlike gout, the pathophysiology of CPPD disease remains poorly understood, and consequently no specific therapies for this arthropathy exist.

History

CPPD disease was initially described by McCarty *et al*. (1962) in the early 1960s. While studying crystals from synovial fluid of patients with presumed gout, it was noted that some crystals were resistant to dissolution by uricase (McCarty *et al*. 1962). On further characterization, these crystals were composed of calcium pyrophosphate dihydrate (Kohn *et al*. 1962). Similar crystals were subsequently identified in the synovial fluid from patients with both acute and chronic arthritis (McCarty 1966). They were noted to be associated with advanced age, radiographic chondrocalcinosis, and a characteristic pattern of severe joint degeneration. Our understanding of the pathophysiology of CPPD crystal formation remains rudimentary.

Definition

No concensus as to the proper nomenclature of CPPD disease exists. CPPD disease or CPPD deposition disease are the terms most commonly used to refer to the arthritis caused by CPPD crystals. Other terms used similarly include CPDD disease and pyrophosphate arthropathy (McCarty 1982). The acute form of CPPD disease is commonly referred to as pseudogout. Chondrocalcinosis, a frequent associated finding, is defined by the radiographic presence of finely stippled calcification in articular hyaline and fibrocartilage. Although these deposits are usually composed of CPPD crystals, other mineral forms have been identified in pathological specimens with chondrocalcinosis (McCarty 1966).

Epidemiology

Because the clinical syndromes associated with CPPD crystals are heterogeneous and may mimic other rheumatic diseases, the prevalence and incidence of CPPD deposition disease is difficult to define. Small studies have suggested a prevalence rate as high as 0.9/1000 (O'Duffy 1976), a figure about half that of gouty arthritis. Gender predominance varies from study to study, but symptomatic CPPD disease may be more common in women (Bergstrom *et al*. 1986a; Bergstrom *et al*. 1986b; Felson *et al*. 1989). The average age in the largest collection of symptomatic patients with definite or probable CPPD disease is 72 years (Ryan and McCarty 1993).

Most studies on prevalence have relied on data from autopsies or are based on radiographic chondrocalcinosis. These studies clearly demonstrate that CPPD crystal deposition increases with age. Autopsy studies show prevalence rates of 20 per cent in knee joints from patients over the age of 60 years (Mitrovic *et al*. 1988). CPPD crystals are unusual in joints of patients under the age of 60, and may be present in about 50 per cent of nonagenarians (Mitrovic *et al*. 1988). Radiographic studies confirm this pattern. An overall prevalence of chondrocalcinosis of 8.1 per cent in the population between the ages of 63 and 93 was noted in Framingham, Massachusetts. Age-specific rates rose from 3.2 per cent in the 65 to 69 year age group to 27.1 per cent in patients over the age of 85 years (Felson *et al*. 1989). Similar results were described in Nottingham. Prevalence rates were 6 per cent in the 55 to 64 age group and 32 per cent in people over 75 years of age (Jones *et al*. 1992a).

CPPD disease usually occurs sporadically. However, familial cases in the Chiloean Islands of South America (Reginato *et al*. 1975), France (Gaucher *et al*. 1977), Spain (Balsa *et al*. 1990), Mexico, (Richardson *et al*. 1983), and Canada (Gaudreau *et al*. 1981) have been well characterized. Familial CPPD occurs prematurely and is often associated with very severe arthritis presenting in the second and third decades. Although the genetics of these kindreds are variable, most show an autosomal dominant pattern of inheritance (Ryan and McCarty 1993). The incidence of familial CPPD disease may be underestimated (Rodriguez-Valverde *et al*. 1988).

Like gout, CPPD disease is associated with a variety of other medical conditions. Unlike gout, the pathophysiological connections between CPPD disease and these disorders is not always apparent. Associated conditions can be divided into definite and possible associations. Definite associations include hypomagnesaemia (Milazzo *et al*. 1981; Salvarini *et al*. 1989), hypophosphatasia (O'Duffy 1970; Whyte *et al*. 1982; Chuck *et al*. 1989), haemochromatosis (Hamilton *et al*. 1981), Wilson's disease (Feller and Schumacher 1972; Golding and Walshe 1977), and hyperparathyroidism (Bywaters *et al*. 1963; McGill *et al*. 1984). Possible associations include gout (Stockman *et al*. 1980), ochronosis (Reginato *et al*. 1973; Rynes *et al*. 1975), familial hypocalciuric hypercalcaemia (Marx *et al*. 1981) and perhaps other causes of sustained hypercalcaemia, diabetes mellitus (Solnica *et al*. 1966; Boussina *et al*. 1971), and X-linked hypophosphataemic rickets (Taylor and Hothersall 1991). An association with hypothyroidism is controversial (Jones *et al*. 1992a). An excellent review of these associated diseases was recently published (Jones *et al*. 1992a)

Clinical features

CPPD disease is a clinically heterogeneous disorder causing both acute and chronic arthritis. Its presentations have been separated into seven syndromes based on clinical features (Ryan and McCarty 1993) (Table 7).

Type	Clinical description	Frequency (%)
A	Pseudogout/acute inflammatory monoarthritis	25
B	Pseudorheumatoid/polyarthritis with synovitis	5
C and D	Pseudo-osteoarthritis/joint degeneration with (type C) or without (type D) acute attacks	50
E	Lanthanic/asymptomatic	?
F	Pseudoneurotrophic/severe joint destruction with or without neuropathy	Rare
Others	Tophaceous CPPD deposits Spinal CPPD deposition Crowned dens syndrome Spinal stenosis	Rare

Table 7 Clinical presentations of CPPD disease

The most commonly recognized form of CPPD disease is the acute arthritis known as pseudogout (or CPPD disease type A). Pseudogout is often clinically indistinguishable from gout, hence its name. It is characterized by the acute onset of pain, warmth, erythema, and swelling usually affecting a single joint. The knee is most commonly involved. Other large joints such as shoulders, wrists, elbows, and ankles may be affected. Pseudogout occurs only rarely in small joints. Attacks are variable in duration, lasting from 2 to 108 days (O'Duffy 1976). Few other signs or symptoms are present. Precipitants of acute attacks include intercurrent illnesses, stroke, trauma, surgery (particularly parathyroidectomy; Glass and Grahame 1976), and rapid diuresis (O'Duffy 1973; O'Duffy 1976).

In the extremely elderly or ill patient, acute pseudogout may have a dramatic presentation. High fever, hypotension, and delerium may mimic sepsis or other systemic diseases (Bong and Bennett 1981).

CPPD disease may have a rheumatoid-like (type B) presentation. This form of CPPD disease accounts for less than 5 per cent of patients with CPPD disease. Symptoms include generalized joint pain and stiffness. The large joints such as knees, wrists, and elbows are commonly involved. Morning stiffness may be prolonged. Synovitis is noted on physical examination. Features which distinguish CPPD disease from rheumatoid arthritis include the absence of small joint involvement, negative serum rheumatoid factor, and the presence of CPPD crystals in the synovial fluid of involved joints. In addition, patients with rheumatoid arthritis have polyarthritis which flares in phase, while the patients with pseudorheumatoid disease will have joint flares out of phase with one another (Ryan and McCarty 1993).

More than 50 per cent of patients with CPPD disease present with an osteoarthritis-like syndrome. These patients are classified as having the type C and type D forms of the disease. They have pain, stiffness, and limited range of motion of the affected joints. Joints not commonly involved in osteoarthritis may be affected, such as shoulders, elbows, and wrists. Examination shows little or no synovitis. Radiographs may show severe degenerative joint disease. Patients with the type C disease differ from those with type D in that they have acute attacks superimposed on chronic symptoms. Patients with the type D disease do not have acute attacks.

CPPD disease type E is clinically silent. It is also referred to as lanthanic CPPD disease. It is picked up incidently as radiographic chondrocalcinosis. This may be the most common type of CPPD disease.

Type F (or neuropathic) CPPD disease remains controversial and is rare; although described in patients with tertiary syphilis and severe neurotrophic arthritis (McCarty 1966). A similar clinical picture can be seen even in the absence of significant neuropathy (Menkes *et al*. 1976).

Although this classification scheme is helpful in describing patients with CPPD disease, not all cases are easily catagorized. The scheme does serve to emphasize the clinical heterogeneity of this disorder and the fact that CPPD disease must be considered as a cause of acute and chronic, inflammatory and non-inflammatory arthritis.

Tophaceous deposits of CPPD crystals have been described, but are not commonly seen (Jones *et al*. 1992b). In the wrist they can cause median nerve compression (Rate *et al*. 1992). In the spine, they may involve the ligamentum flavum and have been reported to produce symptoms of central canal stenosis (Delamarter *et al*. 1993) as well as cervical myelopathy (Berghausen *et al*. 1987). The crowned dens syndrome involves the deposition of CPPD (or BCP) crystals around the atlantoaxial joint and is a cause of acute neck pain (Bouvet *et al*. 1985). CPPD deposits rarely occur in tendons, bursas, and bone (Jones *et al*. 1992b; Kanterewicz *et al*. 1993).

Laboratory investigation

CPPD disease is defined by the presence of CPPD crystals in the synovial fluid of affected joints. Synovial fluids should be collected in heparin-containing tubes and carefully examined under polarized-light microscopy for the positively birefringent rhomboidal crystals of calcium pyrophosphate dihydrate (Fig. 11). These crystals may be sparse or abundant and can be either intra- or extracellular. Their number is not correlated with the degree of inflammation in the joint (Ryan and McCarty 1993). Synovial fluid from patients with acute pseudogout may have cell counts as low as 500 cells/mm^3 or as high as 50 000 cells/mm^3. The mean cell counts are similar to those of acute gout (20 000 cells/mm^3) with greater than 90 per cent polymorphonuclear leucocytes (Ryan and McCarty 1993). Fluid may be haemorrhagic, particularly early in an attack (Ryan and McCarty 1993). Many synovial fluids will also contain BCP crystals (Rachow *et al*. 1988). During the acute attack, serum levels of acute-phase reactants may be elevated.

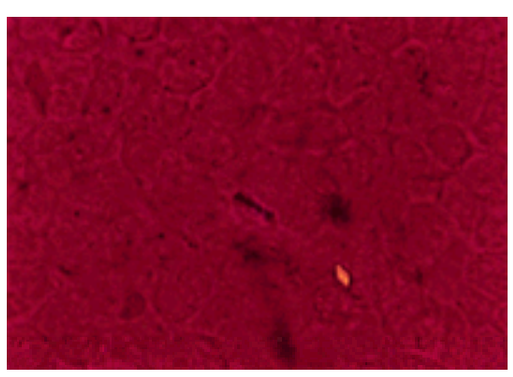


Fig. 11 Typical, positively birefringent crystals of calcium pyrophosphate dihydrate are seen here. Note the rhomboid shape and the weak birefringence.

Radiographs of the affected joint may not be helpful in establishing a diagnosis acutely. However, the presence of chondrocalcinosis increases the likelihood of CPPD disease. Chondrocalcinosis is found in both fibrocartilage and hyaline articular cartilage. Common sites for chondrocalcinosis are the menisci of the knee (Fig. 12), the triangular cartilage of the wrist (Fig. 13), and the symphysis pubis (Fig. 14) (Ryan and McCarty 1993). Other radiographic changes may provide helpful supportive data. Typical findings include eburnation, joint space narrowing, and subchondral cyst formation (Resnick *et al*. 1977). Osteophyte formation is variable, but may be less common than in osteoarthritis. Progressive destruction of the joint with bony collapse and loose body formation occurs frequently in CPPD disease.

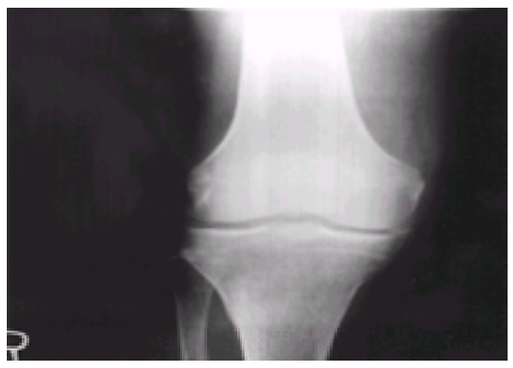


Fig. 12 Chrondrocalcinosis is seen as finely stippled calcification of the cartilage in the knee.

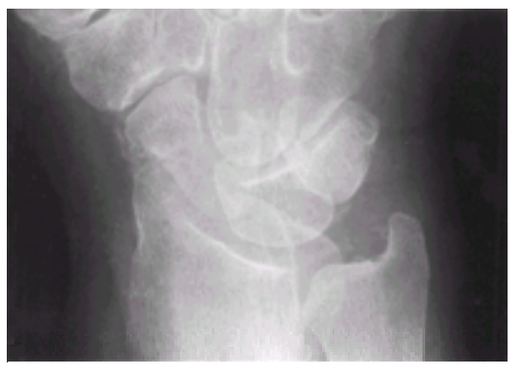


Fig. 13 Chrondrocalcinosis is seen in the triangular cartilage of the wrist.

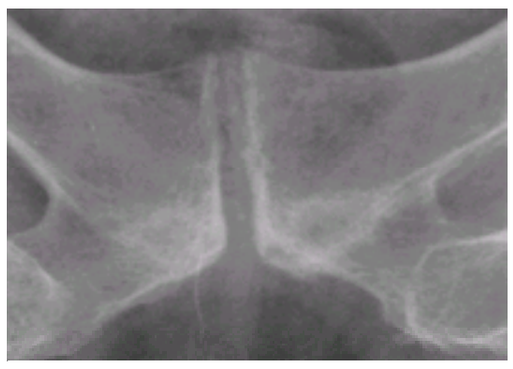


Fig. 14 Chrondrocalcinosis is seen in the pubic symphysis.

The distribution and pattern of radiographic involvement may be helpful in differentiating CPPD disease from osteoarthritis. CPPD disease affects joints not commonly involved in osteoarthritis. Axial involvement with intervertebral disc calcification, sacroiliac erosions, and subchondral cysts of the facet joints occur with CPPD disease. In the knee, tricompartmental involvement or isolated patellofemoral abnormalities are seen with CPPD disease more commonly than with osteoarthritis. Cortical erosions on the femur superior to the patella and osteonecrosis of the medial femoral condyle may also be diagnostic clues to CPPD disease (Lagier 1974; Watt and Dieppe 1983).

Patients with CPPD disease should be screened for one of the unusual metabolic disorders that promotes premature CPPD deposition. Laboratory evaluation should include serum levels of calcium, magnesium, phosphate, iron, and measurement of iron-binding capacity, and alkaline phosphatase.

Diagnosis

A definite diagnosis of CPPD disease can only be made when typical crystals are seen in synovial fluid with polarized-light miscroscopy. Phase-contrast microscopy may be a useful adjunct in detecting crystals. Radiographic changes can be used to establish a probable diagnosis. Diagnostic criteria based on crystal identification and radiological findings have been proposed (Ryan and McCarty 1993).

The differential diagnosis of pseudogout is much the same as that of gout. Infectious arthritis, trauma, and other crystal-associated arthropathies are the conditions that are most often mistaken for pseudogout. The chronic arthritis of CPPD disease can be mistaken for osteoarthritis, any one of the seronegative spondylarthropathies, rheumatoid arthritis, or neurotrophic arthritis.

Pathogenesis (Fig. 15)

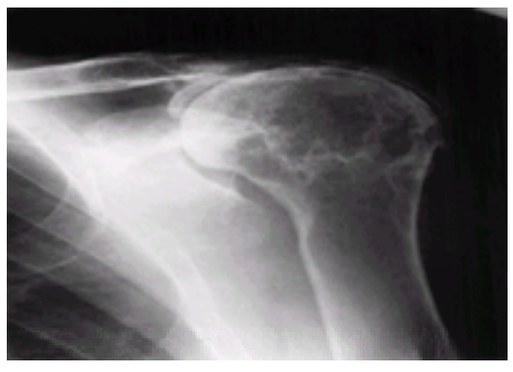


Fig. 15 Scheme of the pathogenesis of CPPD crystal deposition.

The pathogenesis of CPPD disease remains obscure. Conceptually, CPPD crystal deposition, like gout, can be divided into phases of crystal formation, crystal release, and the induction and cessation of an inflammatory response. Using gout as a paradigm, much of the work on the pathophysiology of CPPD deposition rests on the theory that various metabolic abnormalities lead via a final common pathway to CPPD crystal formation. Unlike gout, the nature of this final common pathway remains unclear. Both radiological and histological studies of affected joints implicate cartilage as the primary site of CPPD crystal deposition. Crystals form in both fibrocartilage and the mid-zone of hyaline articular cartilage. The smallest and presumably the earliest crystals are found adjacent to chondrocytes or may replace the

chondron (Pritzker *et al*. 1988). Crystals occur less commonly in synovium and tendon and may form in areas of chondroid metaplasia in these tissues (Beutler *et al*. 1993).

In cartilage, crystals are associated with large or 'hypertrophic' chondrocytes which contain unusual inclusions (Masuda *et al*. 1989). These are not seen in osteoarthritic cartilage. Fragmented collagen fibres and histochemical markers associated with degenerated cartilage such as type I collagen, S-100 protein, and the proteoglycan, dermatan sulphate, have been identified near crystal deposits (Masuda *et al*. 1989).

Theoretically, three components influence CPPD crystal formation. Changes in local concentrations of calcium and pyrophosphate as well as alterations of cartilage matrix might favour crystallogenesis. Despite the association between sustained hypercalcaemia and CPPD disease, no consistent changes in local calcium concentrations in affected joints have been documented (Ryan and McCarty 1993). Moreover, treatment of hyperparathyroidism, for example, does not usually ameliorate the arthritis (Pritchard and Jessop 1977). Studies of crystal formation in gels implicate a role for matrix in influencing crystal formation (Mandel and Mandel 1985), but in general the influence of matrix changes has been difficult to approach in the laboratory. For these and other reasons, much of the work on CPPD crystal formation has concentrated on pyrophosphate metabolism.

Pyrophosphate levels are elevated in the synovial fluids of patients with CPPD disease when compared with patients with other types of arthritis (Silcox and McCarty 1974). Moreover, skin fibroblasts from patients with familial CPPD disease have higher levels of pyrophosphate than those from normal controls (Lust *et al*. 1981). *In vitro*, normal hyaline and fibrocartilage elaborate extracellular pyrophosphate while other joint tissues do not (Ryan *et al*. 1981).

Thus, disordered pyrophosphate metabolism appears to have a crucial role in CPPD crystal formation. Whether this pyrophosphate is generated extracellularly through the activity of the family of ectoenzymes known as nucleoside triphosphate pyrophosphohydrolases (Huang *et al*. 1994), or is made inside the chondrocyte (Rosenthal *et al*. 1991), remains unclear. A role for transforming growth factor-b, a cartilage growth factor involved in repair, is supported by its unique ability to stimulate pyrophosphate elaboration by cartilage (Rosenthal *et al*. 1991), and to increase articular cartilage vesiculation (Derfus *et al*. 1994).

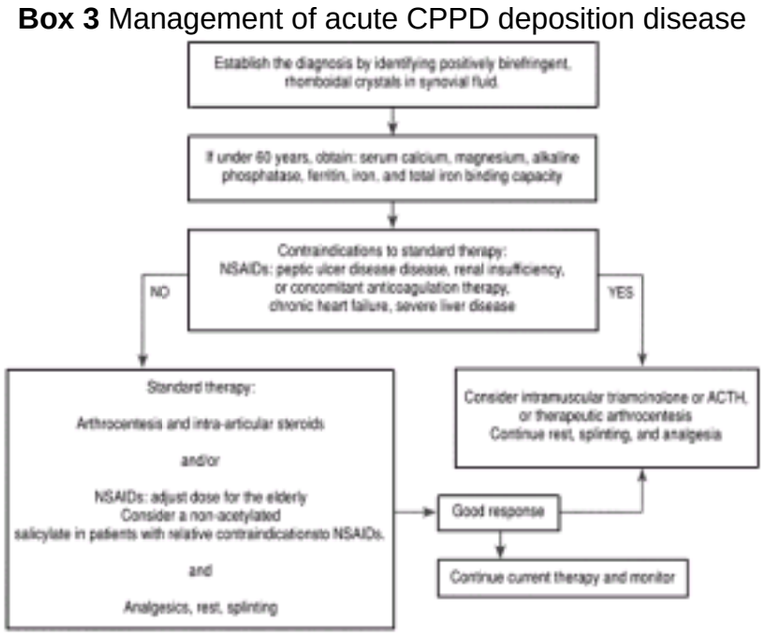
Understanding pyrophosphate metabolism may aid in determining the link between CPPD disease and its associated metabolic disorders. For example, magnesium is a cofactor of pyrophosphatase, which hydrolyses pyrophosphate to phospate. Thus, low magnesium would prevent action of this enzyme and favour elevated pyrophosphate concentrations in the joint. Similarly, hypophosphatasia which is characterized by low alkaline phosphatase activity could cause elevation of pyrophosphate levels by decreasing pyrophosphate hydrolysis.

When concentrations of pyrophosphate and calcium as well as matrix conditions are favourable, CPPD crystals form. Clinical symptoms arise when crystals are released into the synovial space. The mechanisms through which crystals are released into the joint space remain unknown. It has been postulated that trauma or sudden changes in crystal solubility may induce crystal release. Like MSU crystals, CPPD crystals are phlogistic and initiate an inflammatory response similar to that initiated by gout crystals (Malawista *et al*. 1985). The smallest crystals may be the most inflammatory (Ishikawa *et al*. 1987). Like gout, the factors terminating an acute attack remain speculative.

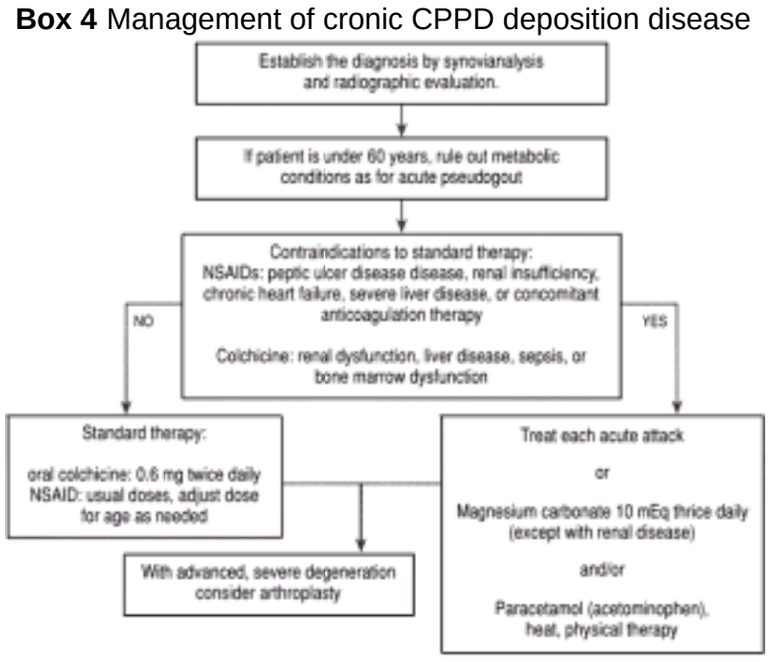
Management

Because our understanding of the pathophysiology of CPPD disease is inadequate, we have no specific treatments for this disorder. Standard therapies include NSAIDs, intra-articular corticosteroids, and colchicine. The management of acute and chronic CPPD deposition disease is summarized in Box 3 and Box 4.

Box 3 Management of acute CPPD deposition disease



Box 4 Management of cronic CPPD deposition disease



NSAIDs are the most commonly used therapy for CPPD disease. There are no controlled trials of their efficacy and no single NSAID is preferable over others. Sulindac may be renal sparing (Ciabottoni *et al*. 1984), and may be safer than other NSAIDs in this elderly population.

Joint aspiration with intra-articular instillation of corticosteroids remains a mainstay of therapy for patients with pseudogout. Although no prospective trials of intra-articular corticosteroids in CPPD exist, two retrospective studies support the their effectiveness in shortening the duration of an acute attack (O'Duffy 1976; Masuda and Ishikawa 1987). As in gout, there has been a resurgence of interest in using parenteral corticosteroids, particularly ACTH, for acute pseudogout. Doses vary between 40 IU and 80 IU given every 8 h with a 3-day taper (Ritter *et al*. 1993).

Intravenous colchicine is of proven efficacy in acute pseudogout (Spilberg *et al*. 1980). However, the age and comorbidities of patients with CPPD disease often preclude its use. Low-dose oral colchicine may be useful prophylactically. At doses of 1 mg/day, it reduced the frequency of acute attacks of pseudogout in one small study (Avarellos and Spilberg 1986).

Other less traditional treatments for CPPD disease have been proposed. Magnesium carbonate increases the solubility of CPPD crystals and acts as a cofactor for pyrophosphatases. At a dosage of 10 mEq thrice daily, magnesium carbonate decreased the severity of symptoms in patients with chronic CPPD deposition disease (Doherty and Dieppe 1983). Intra-articular Arteparon®, a glycosaminoglycan polysulphate, was demonstrated to decrease pain, stiffness, and immobility in one

uncontrolled trial (Sarkozi *et al*. 1988). Intra-articular yttrium-90 combined with steroids was also effective in patients with chronic knee pain from CPPD diseases (Doherty and Dieppe 1981). Intramuscular gold has been used in patients with a pseudorheumatoid presentation with reportedly good results (Ryan and McCarty 1993). Rest, splinting, and eventual joint replacement may be helpful adjunctive measures.

Prognosis

The prognosis for CPPD disease patients is not known. Most cases of acute pseudogout are alleviated with standard therapy. However, as no good prophylactic treatment exists, and the disease often coexists with severe joint degeneration, many patients have unrelenting symptoms from CPPD disease.

The basic calcium phosphate (BCP)-associated syndromes

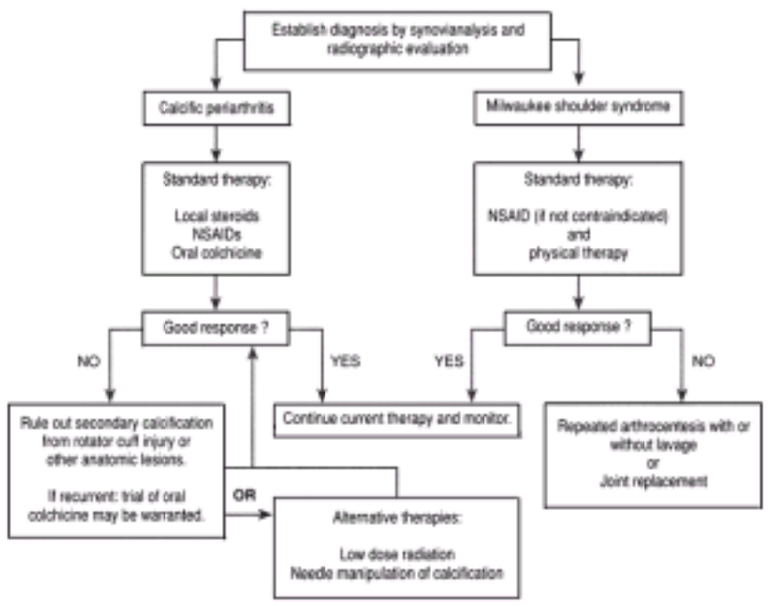
BCP crystals include hydroxyapatite, octacalcium phosphate, and tricalcium phosphate. These crystals are often (less accurately) referred to as apatite or hydroxyapatite crystals. They are associated with a wide variety of clinical syndromes as illustrated in Table 8. The management of BCP crystal diseases is summarized in Box 5.

In joints
- Milwaukee shoulder syndrome
- Osteoarthritis
- Erosive arthritis
- Acute inflammatory arthritis
- Mixed crystal deposition

Near joints
- Periarthritis:
  - Shoulder
  - Upper extremity in women
  - Pseudopodagra
- Calcific tendinitis and bursitis

Table 8 Rheumatic syndromes associated with BCP crystals

Box 5 Management of BCP crystal disease



History

Calcific periarthritis and tendinitis have been recognized for many years, and a clinical syndrome similar to Milwaukee shoulder syndrome was described by Robert Adams in 1857 (McCarty 1989). With the identification of BCP crystals in the synovial fluid of patients with arthritis (Dieppe *et al*. 1976), interest in these crystals and their biological effects grew. Although several BCP-associated syndromes have recently been characterized, there is still much to learn about the significance and distribution of BCP crystals.

Arthritis associated with BCP crystals

Milwaukee shoulder syndrome

This unusual form of arthritis is one of the better defined of the BCP crystal-associated syndromes (McCarty *et al*. 1981; Halverson *et al*. 1984; Halverson *et al*. 1990). Milwaukee shoulder syndrome was initially described in 1981 as a severe shoulder arthropathy of elderly women.

Thirty patients with Milwaukee shoulder syndrome have been carefully described. (Halverson *et al*. 1990). The syndrome is more common in women then men. Patients describe a gradual onset of mild to moderate shoulder pain that is often bilateral and worse at night. Symptoms are more severe on the dominant side. Knee pain, stiffness, and swelling may also occur. Instability, large effusions, and loose bodies are noted on examination of the shoulder. A history of trauma or overuse may antedate the development of this syndrome. Renal disease may also be a predisposing factor (Halverson *et al*. 1987).

Radiographs show exaggerated joint degeneration (Fig. 16). Glenohumeral joint space narrowing, deformities of the humeral head with focal osteoporosis, small osteophytes, loose bodies, and calcifications are particularly characteristic. Large and extensive rotator cuff tears are commonly seen on arthrograms. Knee involvement is common and differs from primary osteoarthritis in that the lateral compartment is often predominantly involved. Synovial fluids have low white blood-cell counts (less than 500 cells/mm^3). Particulate collagens and variable levels of active proteases have been identified in effusions from patients with Milwaukee shoulder syndrome. CPPD crystals are present in one-third of patients. There is currently no widely available method for detecting or quantifying BCP crystals. They can not be reliably identified with conventional or polarized-light microscopy. Stains such as Alizarin red are sensitive but not specific for BCP crystals (Halverson and McCarty 1979). Characteristic crystals can be seen under electron microscopy. At certain centres, a semiquantitative radiometric assay based on diphosphonate binding is in use (Halverson and McCarty 1979).

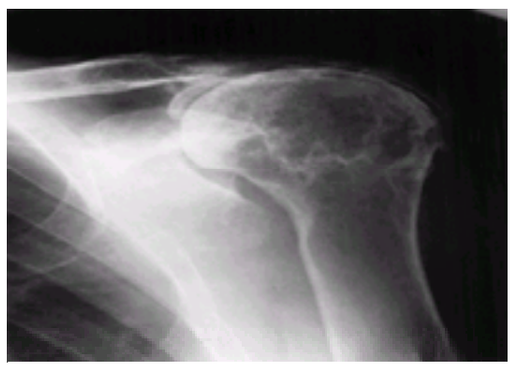


Fig. 16 This radiograph illustrates typical findings in Milwaukee shoulder syndrome. Note the abnormally high position of the humeral head.

The pathogenesis of Milwaukee shoulder syndrome remains an area of active research. BCP crystals are mitogenic and induce the elaboration of collagenase and other proteases from synovial fibroblasts (Cheung *et al*. 1981) and chondrocytes (Cheung *et al*. 1983a). These proteases may be responsible for the extensive destruction of joint structures seen in this syndrome.

No specific therapy is available for Milwaukee shoulder syndrome. Some patients may respond to conservative treatment with analgesics such as NSAIDs and repeated joint aspirations. The utility of intra-articular corticosteroid injections remains unclear. With far-advanced disease or collapse of the humeral head, shoulder arthroplasties may be indicated.

The prognosis for recovery of motion and decreased pain in patients with far-advanced joint degeneration is poor.

BCP crystals in osteoarthritis

BCP crystals are found in 30 to 60 per cent of synovial fluids from patients with osteoarthritis using widely available detection techniques (Dieppe *et al*. 1979; Gibilisco *et al*. 1985). Small or sparse submicroscopic crystals may be found in even higher percentages of osteoarthritic synovial fluids when carefully examined (Swan *et al*. 1994). The quantity of BCP crystal present correlates with the degree of radiographic degeneration. There is no association between synovial fluid cell count (Dieppe *et al*. 1979) or the pattern of radiographic appearance (Halverson and McCarty 1986) and BCP crystals. Thus, the significance of these crystals is unclear.

Erosive arthritis associated with BCP crystals

Several patients with erosive arthritis associated with BCP crystals have been described (Schumacher *et al*. 1981; Zwillich *et al*. 1988). These patients have peripheral or axial arthritis with acute attacks. Pre-existing renal disease, chronic arthritis, and tophus-like subcutaneous nodules have also been described. Radiographs show bony erosions and calcifications. ACTH and colchicine improve symptoms in some patients.

Acute arthritis associated with BCP crystals

A small number of patients with an acute gout-like syndrome associated with BCP crystals in the synovial fluid have been reported (Schumacher *et al*. 1977). These patients have elevated synovial fluid cell counts and normal radiographs. They may represent an early stage of the erosive arthritis described above.

Mixed crystal deposition

BCP and CPPD crystals frequently coexist in a single joint (Dieppe *et al*. 1977; Rachow *et al*. 1988). It is more common to see these crystals together than to identify either one alone. No clear radiographic or clinical syndromes have been characterized on the basis of the coexistence of these two crystals.

Syndromes associated with non-articular BCP crystals

Periarthritis associated with BCP crystals

BCP crystals often cause periarthritis. BCP periarthritis occurs most often around the shoulder joint of people between the ages of 30 and 60 years. Familial forms, involving multiple sites, have been described (Marcos *et al*. 1981). Renal failure may predispose to BCP deposition. Patients usually present with acute pain, warmth, erythema, and swelling in the affected area lasting for several weeks. The diagnosis is suggested by the presence of extrarticular calcium deposits on radiographs.

Two variants of BCP periarthritis have been characterized recently. The first is periarthritis involving the hand and elbow in young women (McCarthy *et al*. 1993; Yosipovitch and Yosipovitch 1993). This occurs in otherwise healthy young women without an antecedent history of trauma. Several patients had recently given birth and were breast feeding. The symptoms are acute and severe and are often misdiagnosed as gout or cellulitis. Sites of involvement include the lateral epicondyle of the elbow, the wrist, and the finger joints. Symptoms respond dramatically to intralesional corticosteroids. Follow-up radiographs demonstrate resolution of the calcific deposits within 7 to 36 days.

The term hydroxyapatite pseudopodagra has been used to refer to BCP periarthritis of the first metatarsophalangeal joint. These patients present with acute pain, swelling, warmth, and erythema of this joint (Fam and Rubenstein 1989). It often occurs in young women and is clinically indistinguishable from gout. Radiographs show amorphous periarticular calcifications. No patients had metabolic abnormalities, although pseudopodagra has been reported in pregnancy (McCarty 1991). Attacks are self-limited, lasting 1 to 3 weeks, and respond to conservative treatment with NSAIDs. Intralesional corticosteroids have also been used with good success (McCarty 1991).

Calcific tendinitis and bursitis

This syndrome frequently involves the shoulder. It is often clinically indistinguishable from traumatic tendinitis or bursitis. Patients present with acute onset of severe pain in the affected area lasting several weeks. The radiographic presence of amorphous calcific deposits and the absence of antecedent trauma suggest this diagnosis. Calcific deposits in the shoulder may also occur with chronic rotator cuff injuries. These deposits occur at the insertion site of the rotator cuff on the humerus and are not reabsorbed with time. In contrast, the calcifications of calcific tendinitis occur in metaplastic fibrocartilage within the rotator cuff and disappear with time (Sakar and Uthoff 1984). The pathophysiology of calcific tendinitis is unknown. Treatment with NSAIDs and intralesional corticosteroids usually results in rapid improvement. Colchicine may also be of some benefit (Swannell *et al*. 1970).

Calcium oxalate arthritis

Calcium oxalate crystals produce an unusual form of arthritis. Although seen most commonly in patients with renal failure on dialysis, calcium oxalate arthritis has also been described in patients with bowel disease and primary oxalosis, a rare inborn error of metabolism.

Epidemiology

There are no good studies of the incidence or prevalence of calcium oxalate arthritis in susceptible populations. Primary oxalosis is quite rare. In contrast, 90 per cent of patients on long-term haemodialysis have pathological evidence of calcium oxalate deposition in kidney and bone tissue (Fayemi *et al*. 1977).

Ascorbic acid supplementation is an added risk factor in patients on dialysis (Balcke *et al*. 1984), although oxalate arthritis has been reported in patients on dialysis who did not receive ascorbic acid supplements (Rosenthal *et al*. 1988). Other risk factors include short bowel syndrome from bowel bypass surgery or inflammatory bowel disease, dietary excesses of unusual foods such as rhubarb, and thiamine and pyridoxine deficiencies (Williams and Smith 1968) (Table 9).

Endstage renal disease on dialysis
Short bowel syndrome
Unusual diets rich in rhubarb, spinach, or ascorbic acid
Thiamine deficiency
Pyridoxine deficiency
Primary oxalosis

Table 9 Conditions associated with calcium oxalate arthritis

Clinical features

Oxalate arthritis most often occurs in the setting of renal failure. No large-scale studies of affected patients exist. Patients usually present with acute mono- or oligoarticular arthritis involving the small joints of the hands particularly the proximal interphalangeals and the metacarpophalangeals. Oxalate arthritis may be symmetric, and is often accompanied by tenosynovitis. Bursal involvement has also been described. Unlike gout, initial episodes may be prolonged and chronic arthritis may rapidly develop.

Primary oxalosis results from one of two defined enzyme deficiencies (Williams and Smith 1968). It is inherited as a recessive trait. Patients with primary oxalosis succumb to endstage renal disease in their early twenties. They have diffuse oxalate deposits at autopsy which may also involve articular tissues. Acute and chronic arthritis (Hockaday *et al*. 1964) as well as tenosynovitis (Cohen and Reid 1935) have been described in these patients, but identification of crystals in affected joints has been difficult (Hockaday *et al*. 1964).

Diagnosis

The diagnosis of calcium oxalate arthritis is made by identifying characteristic crystals in synovial fluid from affected joints. Calcium oxalate crystals may be of two morphologies (Hoffman *et al*. 1982; Reginato *et al*. 1986). The more commonly identified type is weddelite or calcium oxalate dihydrate. This is a positively birefringent bipyramidal crystal (Fig. 17). Less commonly, calcium oxalate monohydrate (whewellite) occurs. Whewellite is polymorphic and may be seen as chunks, rods, ovals, or microspherules. Scanning electron microscopy and X-ray diffraction are often necessary to confirm the identities of these crystals.

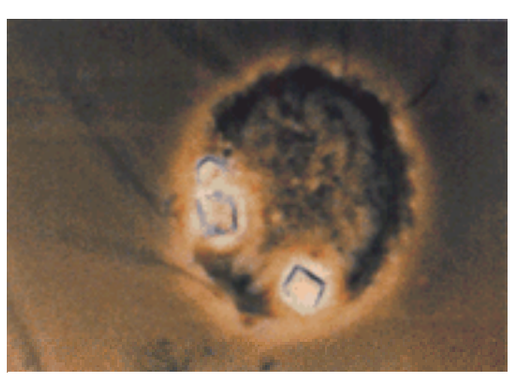


Fig. 17 The typical bipyramidal crystals of calcium oxalate dihydrate are seen here inside a cell.

Synovianalysis usually shows clear or bloody fluid with normal viscosity. Cell counts are typically low and often neutrophils predominate, although large mononuclear cells have also been described (Hoffman *et al*. 1982).

Radiographs are not diagnostic but may be helpful. Miliary calcific deposits in the soft tissue can be seen. Similarly, vascular calcification is a common finding. Less commonly, bony abnormalities such as localized or metaphyseal sclerosis, pseudoarthroses, and pathological fractures can be seen (Milgram 1974; Gherardi *et al*. 1980; Nartijn and Thijn 1982).

Pathophysiology

Oxalic acid is a metabolic end product of amino acid synthesis and ascorbic acid metabolism. It is well absorbed from the gut and is renally excreted. Hence oxalate may accumulate in tissues from excess absorption, overproduction, or underexcretion. Renal failure causes underexcretion of oxalate and levels are high in patients on dialysis. Neither haemo- nor peritoneal dialysis adequately clears oxalate from tissues (Op de Hock *et al*. 1980). Causes of overproduction include the enzyme defects of primary oxalosis, and thiamine and pyridoxine deficiencies. Dietary excesses of spinach, rhubarb, and ascorbic acid supplements may also raise oxalate levels. Excess intestinal absorption may be due to short bowel syndrome from any causes. The most common associated bowel disorders are bowel bypass syndrome from obesity surgery and inflammatory bowel disease.

Once oxalate crystals form in articular tissues, they are released by unknown mechanisms and initiate an inflammatory response (Faires and McCarty 1962).

Management

There is no specific treatment for calcium oxalate arthritis. In general, response to conventional therapies including NSAIDs, intra-articular corticosteroids, colchicine, and increased dialysis are poor. The grim outlook for patients with primary oxalosis may be brightened by early liver transplantation before renal failure has developed (Watts *et al*. 1991).

Other crystals

Other types of crystals can be identified in synovial fluid. These crystals are listed in Table 10. Commonly seen crystals are illustrated in Fig. 18, Fig. 19, and Fig. 20.

Crystal	Morphology	Setting	Significance
Lipid	Round	Trauma	Uncertain
Cholesterol	Plate-like	Unclear	Uncertain
Steroids	Any shape, very bright	After injection	Recent steroid injection
Cryoglobulin	Polygonal, positively birefringent	Cryoglobulinaemia	
Charcot-Leyden	Bipyramidal or hexagonal	Hypereosinophilia	
Cystine	Hexahedral, weak or bright	Cystinosis	
Xanthine	Rhomboidal or plate-like	Xanthinuria	

Table 10 Other crystals found in synovial fluid

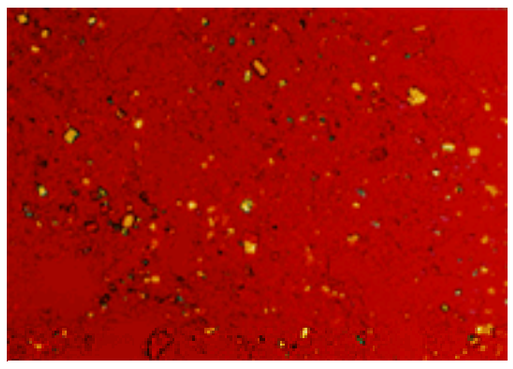


Fig. 18 Steroids crystals can be seen in synovial fluids after intra-articular corticosteroid injections.

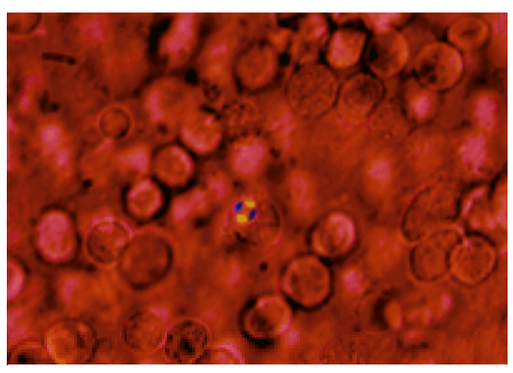


Fig. 19 The typical appearance of lipid crystals in synovial fluid.

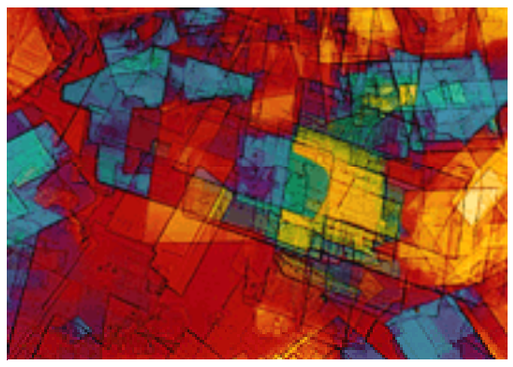


Fig. 20 The typical appearance of cholesterol crystals in synovial fluid.

Chapter References

Abbott, R.D., Brand, F.N., Kannel, W.B., and Castelli, W.P. (1988). Gout and coronary heart disease: the Framingham study. *Journal of Clinical Epidemiology*, **41**, 237–42.

Abramson, S.B. and Weissmann, G. (1989). The mechanism of action of nonsteroidal antiinflammatory drugs. *Arthritis and Rheumatism*, **32**, 1–9.

Abramson, S., Hoffstein, S.T., and Weismann, G. (1982). Superoxide anion generated by human neutrophils exposed to monosodium urate. *Arthritis and Rheumatism*, **15**, 174–80.

Adriani, J. and Naraghi, M. (1985). Allopurinol: actions, adverse reaction and drug interactions. *IM*, **6**, 114–23.

Alepa, F.P. *et al*. (1967). Relationship between glycogen storage disease and tophaceous gout. *American Journal of Medicine*, **42**, 58–63.

Avarellos, A. and Spilberg, I. (1986). Colchicine prophylaxis in pseudogout. *Journal of Rheumatology*, **13**, 804–6.

Axelrod, D. and Preston, S. (1989). Comparison of parenteral adrenocorticotropic hormone with oral indomethacin in the treatment of acute gout. *Arthritis and Rheumatism*, **31**, 803–5.

Baethge, B.A., Work, J., Landreneau, M.D., and McDonald, J.C. (1993). Tophaceous gout in patients with renal transplants treated with cyclosporine A. *Journal of Rheumatology*, **20**, 718–20.

Balcke, P. *et al*. (1984). Ascorbic acid aggravates secondary hyperoxalemia in patients on chronic hemodialysis. *Annals of Internal Medicine*, **10**, 344–5.

Ball, G.V. and Sorenson, L.B. (1969). Pathogenesis of hyperuricemia in sauternine gout. *New England Journal of Medicine*, **280**, 1199–202.

Balsa, A. *et al*. (1990). Familial articular chondrocalcinosis in the Spanish population. *Annals of the Rheumatic Diseases*, **49**, 531–5.

Barthelemy, C.R. *et al*. (1984). Gouty arthritis: a prospective radiographic evaluation of sixty patients. *Skeletal Radiology*, **11**, 1–8.

Bartles, E.C. (1957). Gout as a complication of surgery. *Surgical Clinics of North America*, **38**, 845–8.

Becker, M.A. (1988). Clincial aspects of monosodium urate monohydrate crystal deposition disease (gout). *Rheumatic Diseases Clinics of North America*, **14**, 377–94.

Becker, M.A. and Kim, M. (1987). Regulation of purine synthesis *de novo* in human fibroblasts by purine nucleotides and phosphoribosylpyrophosphate. *Journal of Biological Chemistry*, **262**, 14531–5.

Bellamy, K.N., Downie, W.W., and Buchana, W.W. (1987). Observations on spontaneous improvement in patients with podagra; implications for therapeutic trials fo non-steroidal anti-inflammatory drugs. *British Journal of Clinical Pharmacology*, **24**, 33–6.

Benedict, J.D., Forsham, P.H., and Stetten, D.W., Jr (1949). The metabolism of uric acid in normal and gouty human studied with the aid of istopic uric acid. *Journal of Biological Chemistry*, **181**, 183–93.

Berger, L. and Yu, T-F. (1960). Renal function in gout: an analysis of 524 gouty subjects including long-term follow-up studies. *Proceedings of the Royal Society of Medicine*, **53**, 522–6.

Berghausen, E.J. *et al*. (1987). Cervical myelopathy attributable to pseudogout. *Clinical Orthopaedics and Related Research*, **214**, 217–21.

Bergstrom, G. *et al*. (1986*a*). Joint disorders at ages 70, 75, and 79 years—a cross-sectional comparison. *British Journal of Rheumatology*, **25**, 333–41.

Bergstrom, G. *et al*. (1986*b*). Prevalence of rheumatoid arthritis, osteoarthritis, choncrocalcinosis and gouty arthritis at age 79. *Journal of Rheumatology*, **13**, 527–34.

Beutler, A., Rothfuss, S., Clayburne, G., Sieck, M., and Schumacher, H.R. (1993). Calcium pyrophosphate dihydrate crystal deposition in synovium: relationship to collagen fibers and chondrometaplasia. *Arthritis and Rheumatism*, **36**, 704–15.

Bhatt, A. and Spliberg, I. (1988). Purification of crystal-induced chemotactic factor from neutrophils. *Clinical Biochemistry*, **21**, 342–5.

Bishop, C., Garne, W., and Talbot, J.H. (1951). Pool size, turnover rate, and rapidity of equilibrium of injected isotopic uric acid in normal and pathological subjects. *Journal of Clinical Investigation*, **30**, 879–88.

Block, C., Hermann, G., and Yu, T.F. (1980). A radiologic re-evaluation of gout: a study of 2000 patients. *American Journal of Roentgenology*, **134**, 781–8.

Boger, W.P. and Strickland, S.C. (1955). Probenecid (Benemid) its use and side effects in 2502 patients. *Archives of Internal Medicine*, **95**, 83–92.

Bong, D. and Bennett, R. (1981). Pseudogout mimicking systemic disease. *Journal of the American Medical Association*, **246**, 1438–40.

Boston Collaborative Drug Surveillance Program (1974). Allopurinol and cytotoxic drugs. Interaction in relation to bone marrow depression. *Journal of the American Medical Association*, **227**, 1036–40.

Boussina, I. *et al*. (1971). Etude de la frequence du diabete sucre dans la chondrocalcinose articulaire. *Schweizerische Medizinische Wochenschrift*, **101**, 1413–17.

Bouvet, J. *et al*. (1985). Acute neck pain due to calcifications surrounding the odontoid process: the crowned dens syndrome. *Arthritis and Rheumatism*, **22**, 928–32.

Burak, D.A., Griffith, B.P., Thompson, M.E., and Kahl, L.E. (1992). Hyperuricemia and gout among heart transplant recipients receiving cyclosporine. *American Journal of Medicine*, **92**, 141–6

Burt, H.M. and Dutt, Y.C. (1986). Growth of monosodium urate monohydrate crystals: effect of cartilage and synovial fluid components on *in vitro* growth rates. *Annals of the Rheumatic Diseases*, **45**, 858–64.

Burt, H.M. and Jackson, J.K. (1990). Role of membrane proteins in monososium urate crystal-membrane interactions: 2. Effect of pretreatment with membrane permeable and impermeable protein crosslinking agents. *Journal of Rheumatology*, **17**, 1359–63.

Buzzard, J., Bishop, C., and Talbott, J.H. (1955). The fate of uric acid in the normal and gouty human being. *Journal of Chronic Diseases*, **2**, 42–9.

Bywaters, E.G., Dixon, A., St J., and Scott, J.T. (1963). Joint lesion of hyperparathyroidism. *Annals of the Rheumatic Diseases*, **22**, 171–87.

Canosa, J.J. and Yood, R.A. (1979). Acute gouty bursitis: report of 15 cases. *Annals of the Rheumatic Diseases*, **38**, 326–8.

Cheung, H.S., Halverson, P.B., and McCarty, D.J. (1981). Release of collagenase, neutral protease and prostaglandins from cultured mammalian synovial cells by hydroxyapatite and calcium pyrophosphate dihydrate crystals. *Arthritis and Rheumatism*, **24**, 1338–44.

Cheung, H.S., Halverson, P.B., and McCarty, D.J. (1983 *a*). Phagocytosis of hydroxyapatite or calcium pyrophosphate crystals by rabbit articular chondrocytes release of collagenase, neutral protease and prostaglandin E2 and F2. *Proceedings of the Society for Experimental Biology and Medicine*, **173**, 181–9.

Cheung, H.S., Bohon, S., and Kozin, F. (1983 *b*). Kinetics of collagenase and neutral protease release by neutrophils exposed to microcrystalline sodium urate. *Connective Tissue Research*, **2**, 79–85.

Chua, C.C., Greenberg, M.L., Viau, A.T., Nucci, M., Brenchman, W.D., and Hershfield, M.S. (1988). Use of polyethylene glycol-modified uricase (PEG-uricase) to treat hyperuricemia in a patient with non-Hodgkin lymphoma. *Annals of Internal Medicine*, **109**, 114–17.

Chuck, A.J. *et al*. (1989). Crystal deposition in hypophosphatasia: a reappraisal. *Annals of the Rheumatic Diseases*, **48**, 571–6.

Ciabottoni, G. *et al*. (1984). Effects of sulindac and ibuporfen in patients with chronic glomerular disease. *New England Journal of Medicine*, **310**, 279–83.

Cohen, H. and Reid, J.B. (1935). Tenosynovitis crepitans associated with oxaluria. *Liverpool Med Chirurg, J* **77**, 195–205.

Collins, J.G. (1988). *Prevalence of selected chronic conditions, United States 1983–85*. Advance data from Vital and Health Statistics of the National Center for Health Statistics, No. 155, USDHHS, Washington, DC.

Copeman, W.C.S. (1964). *A short history of the gout*. University of California Press. Los Angeles.

Darmawan, J., Valkenburg, H.A., Muriden, K.D., and Wigley R.D. (1992). The epidemiology of gout and hyperuricemia in a rural population of Java. *Journal of Rheumatology*, **19**, 1595–9.

Delamarter, R.B., Sherman, J.E., and Carr, J. (1993). Lumbar spinal stenosis secondary to calcium pyrophsophate crystal deposition. *Clinical Orthopedics*, **289**, 127–30.

Delbarre, F., Braun, S., and St. George-Chaumet, F. (1967). La Goutte: problems clinique, biologiques et therapeutices. *Semaines des Hôpitaux Paris*, **43**, 623–33.

Derfus, B.A., Rosenthal, A.K., and Ryan, L.M. (1994). TGFb1 and ascorbate each increase mineralizing vesicle elaboration from articular cartilage monolayers. *Arthritis and Rheumatism*, **35**, S75.

Dieppe, P.A. *et al*. (1976). Apatite deposition disease. A new arthropathy. *Lancet*, **i**, 266–76.

Dieppe, P.A. *et al*. (1977). Mixed crystal deposition disease and osteoarthritis. *British Medical Journal*, **1**, 150–60.

Dieppe, P.A. *et al*. (1979). Synovial fluid crystals. I>Quarterly Journal of Medicine, **192**, 533–53.

DiGiovine, F.S. *et al*. (1987). Interleukin 1(IL1) as a mediator of crystal arthritis. Stimulation of T cell and synovial fibroblast mitogenesis by urate crystal induced IL1. *Journal of Immunology*, **138**, 3213–18.

Dodds, W.J. and Steinbach, H.L. (1966). Gout associated with calcification of cartilage. *New England Journal of Medicine*, **275**, 745–9.

Doherty, M. and Dieppe, P.A. (1981). Effect of intra-articular yttrium-90 on chronic pyrophosphate arthropathy of the knee. *Lancet*, **i**, 1243–6.

Doherty, M. and Dieppe, P.A. (1983). Double blind placebo controlled trial of magnesium carbonate in chronic pyrophosphate arthropathy. *Annals of the Rheumatic Diseases*, **42** (Suppl), 106–7.

Doherty, M. and Dieppe, P. (1986). Crystal deposition in the elderly. *Clinics in the Rheumatic Diseases*, **12**, 97.

Emmerson, B.T. (1963). A comparison of uricosuric agents in gout, with special reference to sulfinpyrazone. *Medical Journal of Australia*, **1**, 839–44.

Emmerson, B.T. (1974). Atherosclerosis and urate metabolism. *Australian and New Zealand Journal of Medicine*, **3**, 410–12.

Faires, J.S. and McCarty, D.J. (1962). Acute synovitis in normal joints of man and dog produced by injections of microcrystalline sodium urate, calcium oxalate and corticosteroid esters. *Arthritis and Rheumatism*, **5**, 95.

Fam, A.G. and Rubenstein, J. (1989). Hydroxyapatite pseudopodagra: a syndrome of young women. *Arthritis and Rheumatism*, **32**, 741–7.

Fam, A.G., Lewtas, J., Stein, J., and Paton, T.W. (1992). Desensitization to allopurinol in patients with gout and cutaneous reactions. *American Journal of Medicine*, **93**, 299–302.

Fayemi, A.O., Ali, M., and Braun E.V. (1977). Oxalosis in hemodialysis patients. A pathologic study of 80 cases. *Archives of Pathology and Laborarory Medicine*, **103**, 58–62.

Feller, E.R. and Schumacher, H.R. (1972). Osteoarticular changes in Wilson's disease. *Arthritis and Rheumatism*, **15**, 259–66.

Felson, D.T., Anderson, J.J., Naimark, A., Kannel, W., and Meenan, R.F. (1989). The prevalence of chondrocalcinosis in the elderly and its association with knee osteoarthritis: the Framingham study. *Journal of Rheumatology*, **16**, 1241–5.

Ferry, A.P., Safir, A., and Melikian, H.E. (1985). Ocular abnormalities in patients with gout. *Annals of Opthalmology*, **17**, 632–5.

Fessel, J.W. (1967). Renal outcomes of gout and hyperuricemia. *American Journal of Medicine*, **42**, 27–37.

Fiechtner, J.J. and Simkin, P. (1981). Urate spherules in gouty synovia. *Journal of the American Medical Association*, **245**, 1533–6.

Forkas, W.K., Stanawitz, T., and Schneider, M. (1978). Saternine gout: lead-induced formation of guanine crystals. *Science*, **199**, 786–92.

Frei, E. *et al*. (1963). Renal complications of neoplastic diseases. *Journal of Chronic Diseases*, **16**, 757–76.

Freudweiler, M. (1899). Experimentelle Untersuchungen uber das wesen der Gichtnoter. *Deutsche Archiv der Klinische Medizin*, **63**, 266–355.

Garrod, A.B. (1876) *Treatise on gout and rheumatic gout* (*rheumatoid arthritis*) (3rd edn). Longmans Green and Co., London.

Gaucher, A. *et al*. (1977). Hereditary diffuse articular chondrocalcinosis. *Scandinavian Journal of Rheumatology*, **6**, 217–21.

Gaudreau, A. *et al*. (1981). Familial articular chondrocalcinosis in Quebec. *Arthritis and Rheumatism*, **24**, 611–15.

Gawoski, J.M., Balogh, K., and Landis, W.J. (1985). Aortic valve tophus: identification by X-ray diffraction of urate and calcium phosphates. *Journal of Clinical Pathology*, **38**, 873–6.

Gertler, M., Garn, S.M., and Levine, S.A. (1931). Serum uric acid in relation to age and physique in health and coronary artery disease. *Annals of Internal Medicine*, **34**, 1421–31.

Gherardi, G. *et al*. (1980). Bone oxalosis and renal osteodystrophy. *Archives of Pathology and Laboratory Medicine*, **104**, 105–11.

Gibilisco, P.A. *et al*. (1985). Synovial fluid crystals in osteoarthritis. *Arthritis and Rheumatism*, **28**, 511–15.

Gibson, T., Rodgers, A.V., Simmonds, H.A., and Toseland, P. (1984). Beer drinking and its effect on uric acid. *British Journal of Rheumatology*, **23**, 203–9.

Giclas, P.C., Ginsberg, M.H., and Cooper, N.R. (1979). Immunoglobulin G independent activation of the classical complement pathways by monosodium urate crystals. *Journal of Clinical Investigation*, **63**, 759–765.

Ginsberg, M.H., Jaques, B., Cochrane, C.G., and Griffin, J.H. (1980). Urate crystal-dependent cleavage of Hageman factor in human plasma and synovial fluid. *Journal of Laboratory and Clinical Medicine*, **95**, 497–506.

Glass, J.S. and Grahame, R. (1976). Chondrocalcinosis after parathyroidectomy. *Annals of the Rheumatic Diseases*, **35**, 521–5.

Golding, D.N. and Walshe, J.M. (1977). Arthropathy of Wilson's disease. *Annals of the Rheumatic Diseases*, **36**, 99–111.

Grahame, R. and Scott, J.T. (1970). Clincial survey of 354 patients with gout. *Annals of the Rheumatic Diseases*, **29**, 461–8,

Gresser, U., Gathof, B., and Zollner, N. (1990). Uric acid levels in southern Germany in 1989, a comparison with studies from 1962, 1971, and 1984. *Klinische Wochenschrift*, **68**, 1222–8.

Groff, G.D., Franck, W.A., and Raddatz, D.A. (1990). Systemic steroid therapy for acute gout: a clinical trial and review of the literature. *Seminars in Arthritis and Rheumatism*, **19**, 329–36.

Gutman, A.B. (1973). The past four decades of progress in the knowledge of gout with an assessment of the present status. *Arthritis and Rheumatism*, **16**, 431–45.

Gutman, A.B. and Yu, T-F. (1952). Current principles of management in gout. *American Journal of Medicine*, **13**, 744–59.

Gutman, A.B. and Yu, T-F. (1957). Protracted uricosuric therapy in tophaceous gout. *Lancet*, **ii**, 1258–60.

Hall, A.P., Barry, P.E., Dawber, T.R., and McNamara, P.M. (1967). Epidemiology of gout and hyeruricemia. *American Journal of Medicine*, **134**, 332–9.

Halverson, P.B. and McCarty, D.J. (1979). Identification of hydroxyapatite crystals in synovial fluid. *Arthritis and Rheumatism*, **22**, 389–95.

Halverson, P.B. and McCarty, D.J. (1986). Patterns of radiographic abnormalities associated with basic calcium phosphate and calcium pyrophosphate dihydrate crystal deposition in the knee. *Annals of the Rheumatic Diseases*, **45**, 603–5.

Halverson, P.B., McCarty, D.J., Cheung, H.S., and Ryan, L.M. (1984). Milwaukee shoulder syndrome: 11 additional cases with involvement of the knee in 7. *Seminars in Arthritis and Rheumatism*, **14**, 36–44.

Halverson, P.B., Cheung, H.S., and McCarty, D.J. (1987). Milwaukee shoulder syndrome (MSS): description of predisposing factors. *Arthritis and Rheumatism*, **30**, S131,

Halverson, P.B., Carrera, G.F., and McCarty, D.J. (1990). Milwaukee shoulder syndrome: fifteen additional cases and a description of contributing factors. *Archives of Internal Medicine*, **150**, 677–82.

Hamilton, E.B.D. *et al*. (1981). The natural history of arthritis in idiopathic haemochromotosis: progression of the clinical and radiologic features over ten years. *Quarterly Journal of Medicine*, **50**, 3321–9.

Hertz, P., Yager, H., and Richardson, J.A. (1972). Probenecid induced nephrotic syndrome. *Archives of Pathology*, **94**, 241–3.

Hippocrates (1886). *The genuine works of Hippocrates*. Vols 1 and 2. (trans. and annotated by F. Adams). Wood, New York.

Hochberg, M.C. (1991). Opportunities for the primary and secondary prevention of gout. *Mediguide to Inflammatory Diseases*, **10**, 1–5.

Hockaday, T.D.R., Clayton, J.E., Frederick, E.W., and Smith, L.H. (1964). Primary hyperoxaluria. *Medicine*, **43**, 315–45.

Hoffman, G.S. *et al*. (1982). Calcium oxalate micro-crystalline-associated arthritis in end stage renal disease. *Annals of Internal Medicine*, **97**, 36–42.

Hollander, J.L. (1953). Intra-articular hydrocortisone in arthritis and allied conditions. *Journal of Bone and Joint Surgery*, **35A**, 983.

Holmes, E.W. (1978). Regulation of purine biosynthesis *de novo*. In *Uric acid: handbook of experimental pharmacology*, Vol 51 (ed. W.N. Kelley and I.M. Weiner). Springer-Verlag, New York.

Holmes, E.W. *et al*. (1973). Human glutamine phosphoribosylpyrophosphate amidotransferase: kinetic and regulatory properties. *Journal of Biological Chemistry*, **248**, 144–50.

Huang, R. *et al*. (1994). Expression of the murine plasma cell nucleotide pyrophosphoshydrolase PC-1 is shared by human liver, bone and cartilage cells. Regulation of PC-1 expression in osteosarcoma cells by transforming growth factor-b. *Journal of Clinical Investigation*, **94**, 560–7.

Hunder, G.G., Worthington, J.W., and Bickel, W.H. (1968). Avascular necrosis of the femoral head in a patient with gout. *Journal of the American Medical Association*, **203**, 101–3.

Ishikawa, H., Ueba, Y., Isobe, T., and Hirohata, K. (1987). Interaction of polymorphonuclear leukocytes with calcium pyrophosphate dihydrate crystals deposited in chondrocalcinosis cartilage. *Rheumatology International*, **7**, 217–21.

Itakura, M., Sabina, R.L., Heald, P.W., and Holmes, E.W. (1981). Basis for the control of purine biosynthesis by purine ribonucleotides. *Journal of Clinical Investigation*, **67**, 994–1002.

Jones, A.C., Chuck, A.J., Arie, E.A., Green, D.J., and Doherty, M. (1992 *a*). Diseases associated with calcium pyrophosphate deposition disease. *Seminars in Arthritis and Rheumatism*, **22**, 188–202.

Jones, A., Baron, N., Pattrick, M., and Doherty, M. (1992 *b*). Tophaceous pyrophosphate deposition with extensor tendon rupture. *British Journal of Rheumatology*, **31**, 421–3.

Kanterewicz, E., Sanmarti, R., Panella, D., and Brugues, J. (1993). Tendon calcifications of the hip adductors in chondrocalcinosis: a radiologic study of 75 patients. *British Journal of Rheumatology*, **32**, 790–3.

Katz, W.A. and Ehrlich, G.E. (1968). The solubility of monosodium urate in serum and connective tissue fractions. *Arthritis and Rheumatism*, **11**, 492–7.

Kelley, W.N., Rosenbloom, F.M., Henderson, J.F., and Seegmiller, J.E. (1967). A specific enzyme defect in gout associated with overproduction of uric acid. *Proceedings of the National Academy of Sciences* (*USA*), **57**, 1735–9.

Kelley, W.N. *et al*. (1969). Hypoxanthine-guanine phosphoribosyl transferase deficiency in gout. *Annals of Internal Medicine*, **70**, 155–206.

Kohn N.N. *et al*. (1962). The signficance of calcium phosphate crystals in the synovial fluid of arthritis patients: the 'pseudogout syndrome.' II. Identification of crystals. *Annals of Internal Medicine*, **56**, 738–45.

Kozin, F. and McCarty, D.J. (1977). Rheumatoid factors in the serum of gouty patients. *Arthritis and Rheumatism*, **18**, 49–58.

Kuzell, W.C., Seebach, L.M., Glover, R.P., and Jackman, A.E. (1966). Treatment of gout with allopurinol and sulfinpyrazone in combination and with allopurinol alone. *Annals of the Rheumatic Diseases*, **25**, 634–42.

Lagier, R. (1974). Case report: rare femoral erosions and osteoarthritis of the knee associated with chondrocalcinosis. A histological study of this cortical remodeling. *Virchows Archiv* (A), **364**, 215–23.

Lagier, R. and MacGee, W. (1983). Spondylodiscal erosions due to gout: anatomicroradiologic study of a case. *Annals of the Rheumatic Diseases*, **42**, 350–3.

Lally, E.W., Ho, G., Jr, and Kaplan, S.R. (1986). The clinical spectrum of gouty arthritis in women. *Archives of Internal Medicine*, **146**, 2221–5.

Lawrence, R.C. *et al*. (1989). Estimates of the prevelance of selected arthritic and musculoskeletal diseases in the United States. *Journal of Rheumatology*, **16**, 427–41.

Lawry, G.V., Fan, P.T., and Bluestone, R. (1988). Polyarticular versus monoarticular gout: a prospective comparative analysis of clinical features. *Medicine*, **67**, 335–43.

Lennane, G.A.Q., Rose B.S., and Isdale, I.C. (1960). Gout in the Maori. *Annals of the Rheumatic Diseases*, **19**, 120–5.

Lesch, M. and Nyhan, W.L. (1964). A familial disorder of uric acid metabolism and central nervous system funcion. *American Journal of Medicine*, **36**, 561–70.

Levinson, D.J. and Becker, M.A. (1993). Clinical gout and the pathogenesis of hyperuricemia. In *Arthritis and allied conditions* (12th edn) (ed. D.J. MCarty and W.J. Koopman.), pp. 1773–805. Lea and Febiger, Philadelphia.

Levinson, D.J. and Sorenson, L.B. (1980). Renal handling of uric acid in normal and gouty subjects: evidence for a 4-component system. *Annals of the Rheumatic Diseases*, **39**, 173–9.

Lieber, C.S. *et al*. (1962). Interrelation of uric acid and ethanol metabolism in man. *Journal of Clinical Investigation*, **41**, 1863–70.

Lin, H-Y., Rocher, L.L., McQuillan, M.A., Schmaltz, S., Palella, T.D., and Fox, I.H. (1989). Cyclosporine induced hyperuricemia and gout. *New England Journal of Medicine*, **321**, 287–92.

Loeb, J.N. (1972). The influence of temperature on the solubility of monosodium urate. *Arthritis and Rheumatism*, **15**, 189–92.

Ludwig, G.D. (1957). Saternine gout. *Archives of Internal Medicine*, **100**, 802–10.

Lupton, G.P. and Odom, R.B. (1979). The allopurinol hypersensitivity syndrome: unnecessary morbidity and mortality. *Journal of the American Academy of Dermatology*, **1**, 365–74.

Lust, G., Faurem, G., Netter, P., and Seegmiller, J.E. (1981). Increased pyrophosphate in fibroblasts and lymphoblasts from patients with hereditary diffuse articular chondrocalcinosis. *Science*, **214**, 809–10.

Macfarlane, D.G. and Dieppe, P.A. (1985). Diuretic induced gout in elderly women. *British Journal of Rheumatology*, **24**, 155–7.

Malawista, S.E. *et al*. (1985). Crystal-induced endogenous pyrogen production. A further look at gouty inflammation. *Arthritis and Rheumatism*, **28**, 1039–46.

Mandel, N.S. (1976). The structural basis of crystal-induced membranolysis. *Arthritis and Rheumatism*, **19**, 439–55.

Mandel, N.S. and Mandel, G.S. (1985). Nucleation and growth of CPPD crystals and related species *in vitro*. In *Calcium in biologic systems* (ed. R.P. Rubin, G. Weiss, and J.W. Putner). Plenum, New York.

Marcalongo, R., Calabria, A.A., Lalumera, M., Gerli, R., Alessandrini, C., and Cavallo, G. (1988). The 'switch-off' mechanism of spontaneous resolution of acute gout attack. *Journal of Rheumatology*, **15**, 101–9.

Marcos, J.C. *et al*. (1981). Idiopathic familial chondrocalcinosis due to apatite crystal deposition. *American Journal of Medicine*, **71**, 557–563.

Martel, W. (1968). The overhanging margin of bone: a roentgenologic manifestation of gout. *Radiology*, **91**, 1755–6.

Marx, S.J. *et al*. (1981). The hypocalciuric or benign variant of familial hypercalcemia: clinical and biochemical features in fifteen kindreds. *Medicine*, **60**, 397–412.

Masuda, I. and Ishikawa, K. (1987). Clincial features of pseudogout attack: a survey of 50 cases. *Clinical Orthopaedics and Related Research*, **229**, 173–9.

Masuda, I., Ishikawa, K., and Usuku, G. (1989). A histologic and immunohistologic study of calcium pyrophosphate dihydrate crystal deposition disease. *Clinical Orthopedics and Related Research*, **263**, 272–87.

McCarthy, G.M., Carrera, G.F., and Ryan, L.M. (1993). Acute calcific periarthritis of the finger joints: a syndrome of young women. *Journal of Rheumatology*, **20**, 1077–9.

McCarty, D.J. (1966). Calcium pyrophosphate deposition disease. In *Arthritis and allied conditions* (ed. H.L. Hollander), pp. 947–64. Lea and Febiger, Philadelphia.

McCarty, D.J. (1982). The Hebreden Oration. Crystals, joints and consternation. *Annals of the Rheumatic Diseases*, **42**, 243–53.

McCarty, D.J. (1989). Robert Adams' rheumatic arthritis of the shoulder: Milwaukee shoulder syndrome revisited. *Journal of Rheumatology*, **16**, 668–70.

McCarty, D.J. (1991). Podagra due to calcific periarthritis in a young woman. *Internal Medicine*, **12**, 68, 75–76.

McCarty, D.J. and Hollander, J.L. (1961). Identification of urate crystals in gouty synovial fluid. *Annals of Internal Medicine*, **54**, 454–60.

McCarty, D.J., Kohn, N.N., and Faires, J.S. (1962). The significance of calcium phosphate crystals in the synovial fluid of arthritis patients: the 'pseudogout syndrome' I. Clinical aspects. *Annals of Internal Medicine*, **56**, 711–37.

McCarty, D.J., Halverson, P.B., Carrera, G.F., Brewer, B.J., and Kozin, F.K. (1981). Milwaukee shoulder syndrome: association of microspheroids containing hydroxyapatite crystals, active collagenase, and neutral protease with rotator cuff defects. I: Clinical aspects. *Arthritis and Rheumatism*, **24**, 464–73.

McGill, N.W. and Dieppe, P.A. (1991 *a*). Evidence for a promoter of urate crystal formation in gouty synovial fluid. *Annals of the Rheumatic Diseases*, **50**, 558–61.

McGill, N.W. and Dieppe, P.A. (1991 *b*). The role of serum and synovial fluid components in the promotion of urate crystal formation. *Journal of Rheumatology*, **18**, 1042–5.

McGill, P.E., Grange, A.T., and Royston, C.S.M. (1984). Chondrocalcinosis in primary hyperparathyroidism. *Scandinavian Journal of Rheumatology*, **13**, 56–8.

McInnes, G.T., Lawson, D.H., and Jick, H. (1981). Acute adverse reactions attributed to allopurinol in hospitalized patients. *Annals of the Rheumatic Diseases*, **40**, 245–9.

Mejias, E., Lluberes, R., and Martinez-Maldanado, M. (1989). Hyperuricemia, gout, and autosomal dominant polycystic kidney disease. *American Journal of the Medical Sciences*, **297**, 145–8.

Menkes, C.J., Simon, F., Delrieu, F., Forest, M., and Delbarre, F. (1976). Destructive arthropathy in chondrocalcinosis articularis. *Arthritis and Rheumatism*, **19**, 329–48.

Milazzo, S.C., Cleland, L.G., and Ahern, M.J. (1981). Calcium pyrophosphate dihydrate deposition disease and familial hypomagnesemia. *Journal of Rheumatology*, **8**, 767–77.

Milgram, J.W. (1974). Chronic renal failure, recurrent secondary hyperparathyroidism, multiple metaphyseal infractions, and secondary oxalosis. *Bulletin / Hospital for Joint Diseases*, **35**, 118–44.

Mitrovic, D.R. *et al*. (1988). The prevalence of chondrocalcinosis in the human knee joint: an autopsy survey. *Journal of Rheumatology*, **15**, 633–41.

Moreland, L.W. and Ball, G.V. (1991). Colchicine and gout. *Arthritis and Rheumatism*, **34**, 782–6.

Myers, A.R. *et al*. (1968). The relationship of serum uric acid to risk factors in coronary heart disease. *American Journal of Medicine*, **45**, 520–8.

Nakayama, D.A. *et al*. (1984). Tophaceous gout: a clincial and radiographic assessment. *Arthritis and Rheumatism*, **27**, 468–71.

Nartijn, A. and Thijn, C.J.P. (1982). Radiologic findings in primary oxaluria. *Skeletal Radiology*, **8**, 21–4.

Neel, J.V. (1947). The clinical detection of the genetic carriers of inherited disease. *Medicine*, **26**, 115–53.

Neuss, M.N., McCallum, R.M., Brenckman, W.D., and Silberman, H.R. (1986). Long-term colchicine administration leading to colchicine toxicity and death. *Arthritis and Rheumatism*, **29**, 448–9.

Northridge, D.B. and Almack, P.M. (1986). Allopurinol desensitisation. *British Journal of Pharmaceutical Practice*, **8**, 200–5.

Nuki, G. and Parker, J. (1979). Clinical and enzymatic studies in a child with type I glycogen storage disease associated with partial deficiency of hepatic glucose-6-phosphatase. *Advances in Experimental Medicine and Biology*, **122A**, 189–202.

O'Duffy, J.D. (1970). Hypophosphatasia associated with calcium pyrophosphate dihydrate deposits in cartilage. *Arthritis and Rheumatism*, **13**, 381–8.

O'Duffy, J.D. (1973). Pseudogout syndrome in hospitalized patients. *Journal of the American Medical Association*, **226**, 42.

O'Duffy, J.D. (1976). Clinical studies of acute pseudogout attacks. Comments on prevalence, predispositions, and treatment. *Arthritis and Rheumatism*, **19** (Suppl.), 349.

O'Duffy, J.D., Hunter, G.G., and Kelly, P.J. (1975). Decreasing prevalence of tophaceous gout. *Mayo Clinic Proceedings*, **50**, 227–8.

Op de Hoch, C.T. *et al*. (1980). Oxalosis in chronic renal failure. *Proceedings of the European Dialysis and Transplantation Association*, **17**, 730–5.

Pak, C.Y.C. *et al*. (1981). Is selective therapy of recurrent nephrolithiasis possible? *American Journal of Medicine*, **71**, 615–22.

Pak, C.Y.C., Sakhaee, K., and Fuller, C. (1986). Successful management of uric acid nephrolithiasis with potassium citrate. *Kidney International*, **30**, 422–8.

Palmer, D.G., Highton, J., and Hessian, P.A. (1989). Development of the gout tophus: a hypothesis. *American Journal of Clinical Pathology*, **91**, 190–5.

Pascual, E. (1991). Persistence of monosodium urate crystals and low-grade inflammation in the synovial fluid of patients with untreated gout. *Arthritis and Rheumatism*, **34**, 141–5.

Paulus, H.E., Schlosstein, L.H., Godfrey, R.G., Klinenberg, J.R., and Bluestone, R. (1974). Prophylactic colchicine therapy of intercritical gout: a placebo-controlled study of probenecid-treated patients. *Arthritis and Rheumatism*, **17**, 609–14.

Perricone, E. and Brandt, K.D. (1978). Enhancement of urate solubility by connective tissue: 1. effect of proteoglycan aggregates and buffer cation. *Arthritis and Rheumatism*, **21**, 453–60.

Peters, J.P. and Van Slyke, K.K. (1946). *Quantitiative clinical chemistry*, Vol 1. (2nd edn), p. 937. Williams and Wilkins Company, Baltimore.

Plante, G.E., Durivage, J., and Lemieux, G. (1968). Renal excretion of hydrogen in primary gout. *Metabolism*, **17**, 377–97.

Pritchard, M.H. and Jessop, J.D. (1977). Chondrocalcinosis in primary hyperparathyroidism. *Annals of the Rheumatic Diseases*, **36**, 146–51.

Pritzker, K.P.H., Cheng, P.T., and Renlund, R.C. (1988). Calcium pyrophosphate crystal deposition in hyaline cartilage. Ultrastructural analysis and implications for pathogenesis. *Journal of Rheumatology*, **15**, 828–35.

Puig, J.G. and Fox, I.H. (1984). Ethanol-induced activation of adenine nucleotide turnover. Evidence for a role for acetate. *Journal of Clinical Investigation*, **74**, 936–41.

Puig, J.G. *et al*. (1991). Female gout: clinical spectrum and uric acid metabolism. *Archives of Internal Medicine*, **151**, 726–32.

Puig, J.G. *et al*. (1994). Purine metabolism in women with primary gout. *American Journal of Medicine*, **97**, 332–8.

Rachow, J.W., Ryan, L.M., McCarty, D.J., and Halverson, P.B. (1988). Synovial fluid inorganic pyrophosphate concentration and nucleotide pyrophosphohydrolase activity in basic calcium phosphate deposition arthropathy and Milwaukee shoulder syndrome. *Arthritis and Rheumatism*, **31**, 408–13.

Raddatz, D.A., Mahowald, M.L., and Bilka, P.J. (1983). Acute polyarticular gout. *Annals of the Rheumatic Diseases*, **42**, 117–22.

Rate, A.J., Parkinson, R.W., Meadows, T.H., and Freemont., A.J. (1992). Acute carpel tunnel syndrome due to pseudogout. *Journal of Hand Surgery* (*British*), **17**, 217–18.

Reginato, A.J., Schumacher, H.R., and Martinez, V.A. (1973). Ochronotic arthropathy with calcium pyrophosphate dihydrate crystal deposition. *Arthritis and Rheumatism*, **16**, 705–14.

Reginato, A.J. *et al*. (1975). Familial chondrocalcinosis in the Chiloe Islands, Chile. *Annals of the Rheumatic Diseases*, **34**, 260–8.

Reginato, A.J. *et al*. (1986). Arthropathy and cutaneous calcinosis in hemodialysis oxalosis. *Arthritis and Rheumatism*, **29**, 1387–96.

Reiselbach, R.E. *et al*. (1962). Uric acid excretion and renal function in the acute hyperuricemia of leukemia. *American Journal of Medicine*, **37**, 872–84.

Reiselbach, R.E., Sorenson, L.B., Selp, W.D., and Steele, T.H. (1970). Diminished renal urate secretion per nephron as a basis for primary gout. *Annals of Internal Medicine*, **73**, 359–66.

Resnick, D. *et al*. (1977). Clinical, radiologic, and pathologic abnormalities in calcium pyrophosphate dihydrate deposition disease (CPPD) pseudogout. *Diagnostic Radiology*, **122**, 1–15.

Reynolds, E.S. *et al*. (1957). Fatal massive necrosis of the liver as a manifestation of hypersensitivity to probenecid. *New England Journal of Medicine*, **256**, 5.

Richardson, B.B. *et al*. (1983). Hereditary chondrocalcinosis in a Mexican–American family. *Arthritis and Rheumatism*, **26**, 1387–96.

Richet, G., Mignon, F., and Ardaillou, R. (1965). Goutte secondaire de nephropathies chroniques. *Presse Médicale*, **73**, 633–8.

Ritter, J., Kerr, L.D., Valeriano-Marcet, J., and Spiera, H. (1993). ACTH revisited: effective treatment for acute crystal induced synovitis in patients with multiple medical problems. *Journal of Rheumatology*, **21**, 696–9.

Roberge, C.J., Gaudry, M., deMedicis, R., Lussier, A., Poubelle, P.E., and Naccache, P.H. (1993). Crystal-induced neutrophil activation. IV. Specific inhibition of tyrosine phosphorylation by colchicine. *Journal of Clinical Investigation*, **92**, 1722–9.

Roberts, W.M., Liang, M.H., and Stern, S.H. (1987). Colchicine in acute gout. *Journal of the American Medical Association*, **257**, 1920–1.

Rodriquez-Valverde, V. *et al*., (1988). Hereditary articular chondrocalcinosis. *American Journal of Medicine*, **84**, 101–8.

Rosenthal, A.K., Ryan, L.M., and McCarty, D.J. (1988). Arthritis associated with calcium oxalate crystals in an anephric patient treated with peritoneal dialysis. *Journal of the American Medical Association*, **260**, 1290–2.

Rosenthal, A.K., Cheung, H.S., and Ryan, L.M. (1991). Transforming growth factor-b stimulates inorganic pyrophosphate elaboration by porcine cartilage. *Arthritis and Rheumatism*, **34**, 904–10.

Roubenoff, R., Klag, M.J., Mead, L.A., Liang, K-Y., Seidler, A.J., and Hochberg, M.C. (1991). Incidence and risk factors for gout in white men. *Journal of the American Medical Association*, **266**,

3004–7.

Rozenberg, S., Koeger, A-C., and Bourgeois, P. (1993). Urate-oxidase for gouty arthritis in cardiac transplant recipients. *Journal of Rheumatology*, **20**, 2171.

Rundle, R.W., Metz, E.N., and Silberman, H.R. (1966). Allopurinol in the treatment of gout. *Annals of Internal Medicine*, **64**, 842–7.

Russell, I.J. *et al*. (1983). Effects of IgG and C-reactive protein on complement depletion by monosodium urate crystals. *Journal of Rheumatology*, **10**, 425–33.

Ryan, L.M. and McCarty, D.J. (1993). Calcium pyrophosphate crystal deposition disease; pseudogout; articular chondrocalcinosis. In *Arthritis and allied conditions* (12th edn) (ed. D.J. McCarty and W.J. Koopman), pp. 1835–55. Lea and Febiger, Philadelphia.

Ryan, L.M., Cheung, H.S., and McCarty, D.J. (1981). Release of pyrophosphate by normal mammalian hyaline and fibrocartilage. *Arthritis and Rheumatism*, **24**, 1522–7.

Rynes, R.I., Sosman, J.L., and Holdsworth, D.E. (1975). Pseudogout in ochronosis. *Arthritis and Rheumatism*, **18**, 21–5.

Sakar, K. and Uhthoff, H.K. (1984). Rotator cuff tendinopathies with calcification. In *Calcium in biological systems* (ed. R.P. Rubin, G. Weiss, and J.W. Putney). Plenum, New York.

Salvarini, C. *et al*. (1989). Bartter's syndrome and chondrocalcinosis: a possible role for hypomagnesemia in the deposition of calcium pyrophosphate dihyrate (CPPD) crystals. *Clinical and Experimental Rheumatology*, **7**, 415–20.

Sarkozi, A.M., Nemeth-Csoka, M., and Bartosiewicz, G. (1988). Effect of glycosoaminoglycan polysulfate in the treatment of chondrocalcinosis. *Clinical and Experimental Rheumatology*, **6**, 3–8.

Sarre, H. and Mertz, D.P. (1965). Sekundare Gight bei Neireninsuffizenz. *Klinische Wochenschrift*, **43**, 1134–40.

Scheele, K.W. (1776). Examen chemicum calculi urinarii. *Opuscula*, **2**, 73.

Schumacher, H.R. and Phelps, P. (1971). Sequential changes in human polymorphonuclear leukocytes after urate crystal phagocytosis. *Arthritis and Rheumatism*, **14**, 513–26.

Schumacher, H.R. *et al*. (1977). Arthritis associated with apatite crystals. *Annals of Internal Medicine*, **87**, 411–16.

Schumacher, H.R., Miller, J.L., Ludivico, C., and Jessar, R.A. (1981). Erosive arthritis associated with apatite crystal deposition. *Arthritis and Rheumatism*, **24**, 31–7.

Scott, J.T. (1977). Obesity and hyperuricaemia. *Clinics in the Rheumatic Diseases*, **3**, 25–35.

Scott, J.T. *et al*. (1969). Studies of uric acid pool size and turnover rate. *Annals of the Rheumatic Diseases*, **28**, 366–73.

Seegmiller, J.D., Howell, R.R., and Malawista, S. E. (1962). Inflammatory reaction to sodium urate: its possible relationship to genesis of acute gouty arthritis. *Journal of the American Medical Association*, **180**, 469–76.

Seegmiller, J.E., Rosenbloom, F.M., and Kelley, W.M. (1967). An enzyme defect associated with a sex-linked neurologic disorder and excessive purine synthesis. *Science*, **155**, 1682–4.

Seidell, J.C. *et al*. (1986). Overweight and chronic illness—a retrospective cohort study with a follow-up of 6–17 years in men and women of intially 20–50 years of age. *Journal of Chronic Diseases*, **39**, 585–93.

Self, T.H., Evans, W.E., and Ferguson, T. (1957). Drug enhancement of warfarin activity. *Lancet*, **ii**, 557–8.

Siegel, L.B., Alloway, J.A., and Nashel, D.J. (1994). Comparison of adrenocorticotropic hormone and triamcinolone acetonide in the treatment of acute gouty arthritis. *Journal of Rheumatology*, **21**, 1325–7.

Silcox, D.C. and McCarty, D.J. (1974). Elevated inorganic pyrophosphate concentrations in synovial fluid in osteoarthritis and pseudogout. *Journal of Laboratory and Clinical Medicine*, **83**, 518–31.

Simkin, P.A. (1977*a*). The pathogenesis of podagra. *Annals of Internal Medicine*, **86**, 230–3.

Simkin, P.A. (1977*b*). Urate excretion in normal and gouty men. *Developmental and Experimental Medicine and Biology*, **76B**, 41–5.

Simkin, P.A. (1993). Towards a coherent terminology of gout. *Annals of the Rheumatic Diseases*, **52**, 693–4.

Singer, J.Z. and Wallace, S.L. (1986). The allopurinol hypersenstivity syndrome. Unnecessary morbidity and mortality. *Arthritis and Rheumatism*, **29**, 82–7.

Smyth, C.J. (1953). Gout. In *Comroe's arthritis and allied conditions* (5th edn) (ed. J.L. Hollanders). Lea and Febiger, Philadelphia.

Sokoloff, L. (1957). The pathology of gout. *Metabolism*, **6**, 230–43.

Solnica, J. *et al*. (1966). Les chondrocalinoses articulaires etude clinique, biologique crystallographique et etiologique. *Revue du Rhumatisme et des Maladies Osteo-Articulaires*, **33**, 93–9.

Sorenson, L.B. (1959). Degradation of uric acid in man. *Metabolism*, **8**, 687–703.

Sorenson, L.B. (1960). The elimination of uric acid in man studied by means of ^{14}C-labelled uric acid. *Scandinavian Journal of Clinical and Laboratory Investigation*, **12** (Suppl.), 1–14.

Sorenson, L.B. (1962). The pathogenesis of gout. *Archives of Internal Medicine*, **109**, 379–90.

Sperling, O. *et al*. (1972). Accelerated erythrocyte 5'-phosphoribosyl-1-pyrophosphate synthesis. A familial abnormality associated with excessive uric acid production and gout. *Biochemical Medicine*, **6**, 310–16.

Spilberg, I. *et al*. (1980). Colchicine and pseudogout. *Arthritis and Rheumatism*, **23**, 1062–3.

Sterling, L.P. (1991). The clinical management of gout. *American Pharmacy*, (NS) **31**, 368–74.

Steven, M.M. (1992). Prevalence of chronic arthritis in four geographical areas of the Scottish Highlands. *Annals of the Rheumatic Diseases*, **51**, 186–94.

Stockman, A., Darlington, L.G., and Scott, J.T. (1980). Frequency of chondrocalcinosis of the knee and avascular necrosis of the femoral head in gout. *Annals of the Rheumatic Diseases*, **39**, 7–11.

Stuart, R.A., Gow, P.J., Bellamy, N., Campbell, J., and Grigor, R. (1991). A survey of current prescribing practices of antiinflammatory and urate-lowering drugs in gouty arthritis. *New Zealand Medical Journal*, **104**, 115–17.

Swan, A., Chapman, B., Heap, P., Seward, H., and Dieppe, P. (1994). Submicroscopic crystals in osteoarthritic synovial fluids. *Annals of the Rheumatic Diseases*, **53**, 467–70.

Swannell, A.J., Underweeo, F.A., and Dixon, A.S. (1970). Periarticular calcific deposits mimicking acute arthritis. *Annals of the Rheumatic Diseases*, **29**, 380–5.

Tak, H-J., Cooper, S.M., and Wilcox, W.R. (1980). Studies on the nucleation of monosodium urate at 37°C. *Arthritis and Rheumatism*, **23**, 574–80.

Talbott, J.S. (1957). *Gout*. Grune and Stratton, New York.

Taylor, H.G. and Hothersall, T.E. (1991). Hypophosphataemic rickets and pyrophosphate arthropathy. *Clinical Rheumatology*, **10**, 155–7.

Terkeltaub, R.A. *et al*. (1991). Apolipoprotein E inhibits the capacity of monosodium urate crystals to stimulate neutrophils. *Journal of Clinical Investigation*, **87**, 20–6.

Varga, J., Giampaola, C., and Goldenberg, D.L. (1985). Tophaceous gout of the spine in a patient with no peripheral tophi. Case report and review of the literature. *Arthritis and Rheumatism*, **28**, 1312–15.

Wallace, S.L. (1977). The treatment of the acute attack of gout. *Clinics in the Rheumatic Diseases*, **3**, 133–43.

Wallace, S.L. *et al*. (1977). Preliminary criteria for the classification of the acute arthritis of primary gout. *Arthritis and Rheumatism*, **20**, 895–900.

Watt, I. and Dieppe, P. (1983). Medial femoral condyle necrosis and chondrocalcinosis; a causal relationship. *British Journal of Radiology*, **56**, 7–22.

Watts, W.E. *et al*. (1991). Combined hepatic and renal transplantation in primary hyperoxaluria type I: clinical report of 9 cases. *American Journal of Medicine*, **90**, 179–87.

Whyte, M.P., Murphy, W.A., and Fallon, M.D. (1982). Adult hypophosphatasia with chondrocalcinosis and arthropathy. *American Journal of Medicine*, **72**, 631–41.

Wilcox, W.R. and Khalaf, A.A. (1975). Nucleation of monosodium urate crystals. *Annals of the Rheumatic Diseases*, **34**, 332–9.

Williams, H.E. and Smith, H.E. (1968). Disorders of oxalate metabolism. *American Journal of Medicine*, **45**, 715–35.

Wilson, J.M., Young, A.S., and Kelley, W.N. (1983). Hypoxanthine-guanine phosphoribosyltransferase deficiency: the molecular basis of the clinical syndromes. *New England Journal of Medicine*, **309**, 900–10.

Wyngaarden, J.B. and Kelley, W.N. (1976). *Gout and hyperuricemia*. Grune and Sratton, New York.

Wyngaarden, J.B. and Kelley, W.N. (1985). Gout. In *The metabolic basis of inherited disease*. (5th edn) (ed. J.B. Stanbury *et al*.), pp. 1043–114. McGraw-Hill, New York.

Wyngaarden, J.B. and Stetten, D.W., Jr (1953). Uricolysis in normal man. *Journal of Biological Chemistry*, **203**, 9–21.

Wyngaarden, J.B., Rundle, R.W., and Metz, E.N. (1965). Allopurinol in the treatment of gout. *Annals of Internal Medicine*, **62**, 842–7.

Yen, R.C.K., Adams, W.B., Lazar, C., and Becker, M.A. (1978). Evidence for X-linkage of human phosphoribosylpyrophosphate synthetase. *Proceedings of the National Academy of Sciences* (*USA*), **75**, 482–5.

Yen, R.C.K., Raivio, K.O., and Becker, M.A. (1981). Inhibition of phosphoribosylpyrophosphate synthesis by 6-methylthionisoinate. *Journal of Biological Chemistry*, **256**, 1839–45.

Yosipovitch, G. and Yosipovitch, Z. (1993). Acute calcific periarthritis of the hand and elbows in women. A study and review of the literature. *Journal of Rheumatology*, **20**, 1533–8.

Yu, T-F. (1981). Urolithiasis in hyperuricemia and gout. *Journal of Urology*, **126**, 424–30.

Yu, T-F. and Gutman, A.B. (1961). Efficacy of colchicine prophylaxis. Prevention of recurrent gout arthritis over a mean period of five years in 208 gouty subjects. *Annals of Internal Medicine*, **55**, 179–85.

Yu, T-F., and Gutman, A.B. (1967). Uric acid nephrolithiasis in gout: predisposing factors. *Annals of Internal Medicine*, **67**, 1133–48.

Yukawa, T., Akazawa, H., Miyake, Y., Takahashi, Y., Nagao, H., and Takeda, E. (1992). A female patient with Lesch–Nyhan syndrome. *Developmental Medicine and Child Neurology*, **34**, 543–6.

Zollner, N., Griebsch, A., and Fink, J.K. (1970). Uber die Wirking von Benzbromarone auf den Serumharnsauerspiegel und die Harnsaureausscheidung des Gichtkranken. *Deutsche Medizinische Wochenschrift*, **95**, 2405–11.

Zwillich, S.H., Schumacher, H.R., Jr., Hoyt, T.S., Luger, A.M., Morrison, G., and Walder, S.E. (1988). Universal spondylodiscitis in a patient with erosive peripheral arthritis and apatite crystal deposition. *Journal of Rheumatology*, **15**, 123–8.

5.17.1 Osteoporosis and osteomalacia

A. K. Bhalla

Introduction

Disorders of bone that lead to a reduction in bone mass and strength cause much suffering in the population and often present to the rheumatologist with pain, fractures, and deformity. The most common bone disorder in Western nations is osteoporosis, although globally osteomalacia or rickets secondary to vitamin D deficiencies are more common. Unlike osteomalacia, in osteoporosis the mineralized bone, while reduced, is normal; the disorder can be defined as a reduction in bone mass and strength resulting in an increased risk of fracture with minimal or no trauma. The term 'osteopenia' is sometimes used to delineate a state of reduced bone mass, for example identified during radiological investigation or bone mass measurement, without fracture.

Osteoporosis

Definition

Until recently there was no internationally agreed definition of osteoporosis. A 1991 Consensus Development Conference defined osteoporosis as 'a disease characterised by low bone mass and micro architectorial deterioration of bone tissue, leading to enhanced bone fragility and a consequent increased risk in fracture' (Consensus Development Conference 1991).

This definition of osteoporosis has been accepted by the World Health Organization (**WHO**), which has now also defined osteoporosis on the basis of bone density. The categories of disease as defined by the WHO 1994 criteria based on bone mineral density are shown in Table 1.

Disease category	BMD or BMC value
Normal	Not less than 1 SD below young adult mean value
Low bone mass (osteopenia)	More than 1 SD below young adult mean value but not less than 2.5 SD below young adult mean value
Osteoporosis	More than 2.5 SD below young adult mean value
Severe osteoporosis (established osteoporosis)	More than 2.5 SD below young adult mean value in the presence of one or more low-trauma or fragility fractures

BMD, bone mineral density; BMC, bone mineral content.

Table 1 Categories of osteoporosis as defined by WHO based on bone mineral density (after Kanis *et al*. 1994*a*)

These diagnostic criteria for osteoporosis do have some limitations and are likely to be changed in time. It is important to remember that they apply to postmenopausal women and it is not yet agreed what criteria should be applied to men or to young individuals who have yet to attain skeletal maturity, or to special situations such as corticosteroid osteoporosis. Kanis *et al*. (1994*b*) suggest that, since in some populations the risk of fracture in men is substantially lower for bone mineral measurements within their own reference range, diagnostic criteria of 3–4 SD below a young reference mean may be more appropriate. In addition, different cut-off values may be appropriate for different communities since differences in the risk of hip fractures between communities cannot solely be explained on the basis of bone mineral density.

Epidemiology

In the Western world osteoporosis causes considerable suffering in individuals over the age of 50, and, because it predisposes to skeletal failure and fractures, is responsible for consuming an enormous amount of the budget devoted to health care. The main sites of fracture traditionally thought to be related to osteoporosis are the vertebral body, proximal femur, distal forearm, the proximal humerus, and pelvis, of which the first three are more common (Fig. 1) (Riggs and Melton 1986). The risk of an osteoporotic fracture is greater in women than men, and in both sexes the risk varies appreciably between countries. In women from the United States and Europe, the lifetime risk of a hip fracture at the age of 50 is 17.5 per cent whilst the lifetime risk of a hip fracture in a male is 6 per cent. The lifetime risk of a vertebral fracture is estimated at 15.6. per cent for women and 5 per cent for men, and the risk of having either a fracture of the femur or the distal forearm is approx. 40 per cent in women aged over 50 years and 13.1 per cent in men (Table 2) (Melton *et al*. 1992).

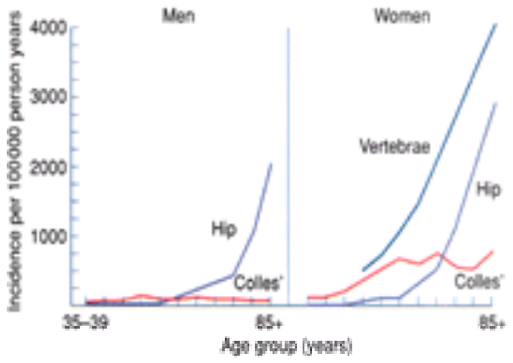


Fig. 1 Incidence rates for the three common osteoporotic fractures (Colles', hip, and vertebral) in men and women, plotted as a function of age at the time of the fracture (reproduced from Riggs and Melton 1986 with permission).

Fracture site	Lifetime risk (%)	
	Men	Women
Hip	6.0	17.5
Vertebral	5.0	15.6
Distal forearm	2.5	16.0
Any of the above	13.1	39.7

Table 2 Lifetime fracture risk in 50-year-old white men and women (Melton *et al*. 1992)

As discussed below, the three most common fractures have different sex ratios and incidence rates at different decades. The incidence of fractures of the distal forearm in women increases soon after the menopause, with a peak incidence between the ages of 50 and 65 (there is a peak in the risk of falls in women aged 45 to 60 years), while in adult males the incidence rates remain constant. Presumably the attempt to break the fall by stretching out the arm accounts for this fracture. In contrast, fractures of the proximal femur show an exponential increase in incidence much later in life, with a 20-fold increase between the ages of 65 to 85 years (where the incidence of falls increases by only twofold) and a male:female ratio of 1:2. This indicates that the probability of a fall resulting in a fracture of the femoral neck is modulated by other age-related factors. Such factors would include the progressive loss of bone with ageing (Winner *et al*. 1989), and a tendency, with increasing age, for a greater likelihood of a fall on the hip than on the hand, presumably due to neuromuscular impairment making it less likely that he or she has enough time to break the fall by throwing out an arm (Grimley Evans 1990).

Proximal femoral fracture

Hip fracture, the most serious complication of osteoporosis, usually follows a fall from a standing position, although it may occur spontaneously. The occurrence of hip fracture has a seasonal variation in both hemispheres. Hip fractures occur more frequently during the winter months; the majority follow falls indoors and therefore are not related to slipping on ice. In most Western countries the incidence of hip fractures has been rising steadily and not all the increase can be explained by an ageing population (Boyce and Vessey 1985; Obront *et al*. 1989). In Sweden the total number of hip fractures between 1950 and 1985 has increased sevenfold while the number of inhabitants over the age of 50 has doubled during this period (Fig. 2). A large part of the increase in hip fracture is due to an increase in the age-adjusted incidence of hip fractures (Obront *et al*. 1989).

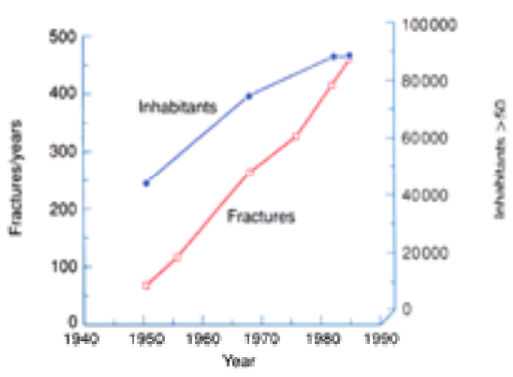


Fig. 2 The relation between the total number of fractures of the hip and the population over 50 years of age in Malmo, Sweden (reproduced from Obront *et al*. 1989 with permission).

In the United Kingdom the incidence of hip fractures rises exponentially with age in both sexes (Fig. 3) and with an approximate 2:1 female to male incidence. This fact, and because there is a large number of elderly women in the population, results in over 80 per cent of all hip fractures occurring in women over the age of 65.

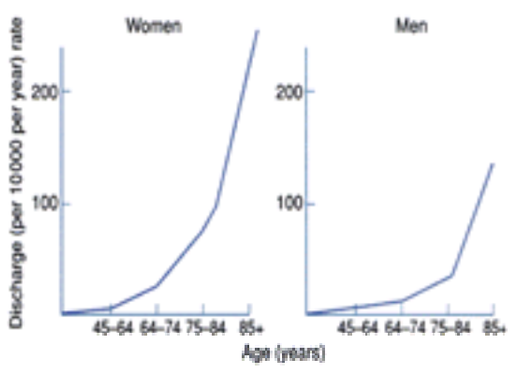


Fig. 3 Hospital discharge rates for fractured neck of femur by age and sex in England, 1985 (derived from Royal College of Physicians 1989).

There were 46 000 hip fractures in England and Wales in 1985 and current estimates suggest that 60 000 occur annually (Advisory Group on Osteoporosis 1994). In 1989 the Royal College of Physicians estimated that if the 1985 age- and sex-specific incidence rates continue to double every 30 years, then by the year 2016 there

will be 117 000 hip fractures annually. Using 1985 incidence rates the probability of a woman sustaining a hip fracture before the age of 85 in the United Kingdom is 12 per cent; for a man it is 5 per cent. In the United Kingdom the direct hospital cost for hip fractures in 1985 were about £160 million. The most recent estimate, using 1992/93 prices, suggests that in England alone the acute inpatient costs will be £237 million, and additional costs of long-term and community care could be a further £444 million (Advisory Group on Osteoporosis 1994).

In the United States there were 238 000 hip fractures in 1986 in adults over the age of 50 years, and, assuming that current age-specific incidence rates remain unaltered, then by the year 2020 there will be 347 000 hip fractures. The annual acute-care costs will rise from the current sum of $8 billion by four- to eightfold.

Hip fractures are recognized as major health problems in Western countries, but there is a belief that since most fractures occur in Caucasian populations the problem may not affect the nations of Asia, South America, and Africa. However, using current incidence rates of hip fractures in various parts of the world to projected populations in the year 2050, it has been estimated that the number of hip fractures occurring world-wide each year will rise from 1.6 million in 1990 to 6.26 million by 2050. Whilst in 1990 nearly half the hip fractures occurred in North America and Europe, with 31 per cent in Asia, by the year 2050 50 per cent of hip fractures will occur in Asia, with 13 per cent in Europe and 12 per cent in North America. Thus osteoporosis, particularly hip fractures, will become a world-wide problem requiring preventative strategies to be developed in regions where currently it is thought not to be a major problem (Cooper *et al*. 1992).

Hip fractures result in considerable mortality. Surveys suggests that between 10 and 40 per cent of all patients who fracture a hip die within 12 months of the fracture (Fig. 4) (Aitken 1987). The excess mortality may, in part, be due to the presence of more coexisting morbidity in patients with fractures than in the age-matched, non-fractured population. In a prospective study, patients admitted with a hip fracture alone has 0 per cent mortality at 1 year compared to 14 per cent and 24 per cent if two or three or more additional medical conditions were present (Svensson *et al*. 1996).

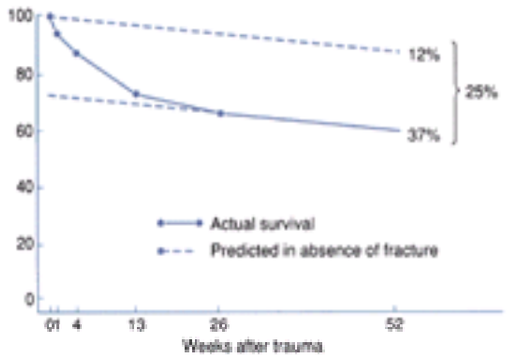


Fig. 4 Relation between survival and passage of time since femoral-neck fracture in women (reproduced from Aitken 1987 with permission).

Risk factors for hip fracture

Recently, two large, well-conducted studies have examined the role of risk factors in hip fracture (Table 3) (Cummings *et al*. 1995; Johnell *et al*. 1995). Cummings *et al*. (1995) followed 9516 women aged 65 or over for 4 years and identified 192 incidences of hip fractures. These investigators reported that a maternal history of hip fracture doubled the risk of hip fracture, with a relative risk of 2 falling to 1.8 after adjusting for bone density. Women with a history of any previous fracture since the age of 50 showed an increased risk of hip fracture (relative risk 1.5), confirming previous observations (Cooper *et al*. 1988). Other identifiable risk factors included height at the age of 25 (odds ratio 1.3 per 6 cm), a finding also noted by Meyer *et al*. (1993). The reason for this association may be that tall women have further to fall and that they also have a longer hip axis (the distance from the greater trocanter to the inner pelvic ring), which has been associated with a greater risk of hip fracture (Faulkner *et al*. 1993).

Maternal history of hip fracture
Low body-mass index
Height age 25
Short fertile period
Previous hyperthyroidism
Previous treatment with benzodiazepines or anticonvulsants
Poor mental score
Poor depth perception
Low physical activity

Table 3 Important risk factors for hip fractures in women

Johnell *et al*. (1995) found no significant adverse effect of smoking, unlike Cummings *et al*. (1995), who could show a significant effect of smoking in an age-adjusted model only, but not after multivariate adjustment for factors such as weight, poor health, and inadequate exercise, all of which were more pronounced in smokers. The relation between calcium intake and hip fractures remains controversial in view of the inconsistent findings in these two epidemiological studies. Cummings *et al*. 1995 found no relation between dietary calcium intake and hip fracture, while in the European study (Johnell *et al*. 1995) a low intake of milk was associated with a significantly increased risk. The increased risk was, however, confined to that 10 per cent of the population with a calcium intake of less than 240 mg daily.

The incidence of hip fracture rises depending on the number of risk factors present. For example, those with two or less risk factors had an incidence of 1.1 hip fractures per 1000 women years compared to an incidence of 19 in those with five or more risk factors. The risk of fracture is greatly increased if bone density is reduced (see Fig. 5) in association with other risk factors (Cummings *et al*. 1995). In the European study (Johnell *et al*. 1995), risk factors such as late menarche, poor mental score, low body mass index, reduced physical activity, low sunshine exposure, and low intake of calcium explained about 70 per cent of the total risk of fracture. Further studies are required to confirm these observations.

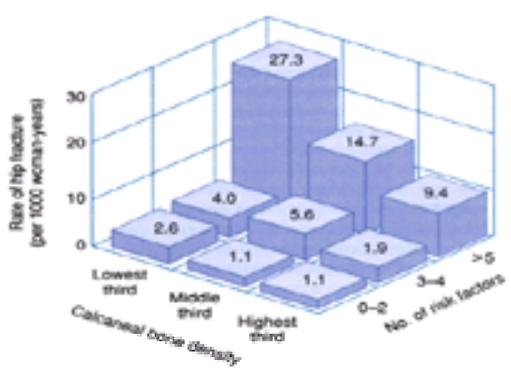


Fig. 5 Annual risk of hip fracture in relation to number of risk factors and calcaneal bone density (after Cummings *et al*. 1995).

Vertebral fractures

The prevalence of vertebral fractures is not accurately known since there is no universally agreed definition of vertebral fracture and they are often asymptomatic so that clinical diagnosis may be delayed until multiple fractures have occurred, leading to spinal deformity and loss of height. The socioeconomic costs of vertebral fracture are therefore unknown.

Epidemiological studies on the prevalence of vertebral fractures have often used differing criteria for defining a fracture and currently there is no accepted definition of vertebral fracture based on the degree of deformation. The criterion for defining a wedge fracture varies from a 20 to 30 per cent reduction in anterior or posterior height of the vertebra compared with adjacent vertebrae, and the lack of agreement may account for the discordant results from studies of drugs, such as fluoride, in established osteoporosis. A clinically significant vertebral fracture requires a deformation of at least 20 to 25 per cent and more minor deformities may reverse spontaneously (Kleerekoper 1992).

The incidence of vertebral fractures increases with age and about 20 to 30 per cent of women aged over 60 may have an asymptomatic vertebral fracture on routine radiography. In cross-sectional studies in the United States, it has been estimated that each year 538 000 postmenopausal women will sustain their first vertebral fracture (Riggs and Melton 1986). The prevalence of vertebral fractures in the United Kingdom has been estimated at 7.8 per cent and the prevalence rises with advancing age from 4.3 per cent at 55 to 59 years to 27.8 per cent at 80 years or over (Cooper *et al*. 1991). It has recently been estimated that 40 000 clinically diagnosed vertebral fractures occur annually in postmenopausal women in the United Kingdom. Studies of the Scandinavian population show not only an age-related increase in incidence of vertebral fractures, but also a fourfold increase in incidence and prevalence between 1950 and 1980 (Benger *et al*. 1988). The prevalence of vertebral fracture increases with decreasing bone density and approaches 50 per cent in those with vertebral bone density below 0.6 g/cm ^{2} as measured by dual-photon absorptiometry (Riggs and Melton 1986). A recent European epidemiological study suggests that the prevalence of vertebral deformities is 12 per cent in women (range 6–21 per cent) and 12 per cent in men (range 8–20 per cent). The prevalence increased with age in both sexes with a steeper rise in women. In women aged 76 to 79 the prevalence was 24.7 per cent compared to 18.1 per cent in men (O'Neill *et al*. 1996). Using the prevalence figures for the United Kingdom from the study, it is estimated that 900 000 men and 1 million women between the age of 50 and 79 years have vertebral deformity, and, even if only 1 out of 3 is symptomatic, 350 000 more will seek medical advice and help (O'Neill *et al*. 1996).

Distal radius fractures

The incidence of fracture of the distal radius (Colles' fracture) after a fall increases in women after the age of 40 and continues to rise until age 65, when it reaches a plateau. No such increase is observed in men (Fig. 1). The reason for the plateau in incidence rates in females after the age of 65 is unknown, but may relate to the incidence patterns of fall with advancing age. In older women, impaired neuromuscular co-ordination is more likely to lead to a fall on the hip than on the wrist whereas a younger woman may be able to break the fall by outstretching her arms. It is assumed that the fracture after trauma is the result of rapid loss of trabecular bone following the menopause. There is a peak in the incidence of falls between the ages of 45 and 60 years in women, which may contribute to the incidence of this fracture (Winner *et al*. 1989). The occurrence of a Colles' fracture may provide an early warning of osteoporosis in the affected individual.

Classification

There are several causes of osteoporosis and it may be either primary or secondary (Table 4).

Physiological or primary
Postmenopausal (type I)
Age-related (type II)
Idiopathic
Juvenile
Secondary
Endocrine
Cushing syndrome
Hyperparathyroidism
Hypogonadism
Thyroid toxicosis
Hypopituitarism
Insulin-dependent diabetes mellitus
Drugs
Corticosteroids
Excess thyroxine replacement
Heparin
Anticonvulsants
Methotrexate
Haematopoietic
Multiple myeloma
Lymphoma, leukaemia
Mastocytosis
Immobilization
Idiopathic hypercalciuria
Gastrectomy and other bowel disease
Osteoporosis of pregnancy
Inflammatory disorders
Rheumatoid arthritis
Ankylosing spondylitis
Congenital
Osteogenesis imperfecta

Table 4 Classification of osteoporosis

Physiological or primary osteoporosis has been divided into two main types by some as type I (postmenopausal) and type II (age-related) osteoporosis. Type I osteoporosis affects predominantly trabecular bone and is associated with an increased risk of wrist and vertebral fractures, while type II affects both trabecular and cortical bone leading to hip fractures. In practice there is considerable overlap between these groups, since patients with vertebral fractures have a greater risk of having a fractured hip and vice versa. In most patients with osteoporosis, distinguishing primary from secondary osteoporosis may be difficult since many factors may be present in one individual, for example a postmenopausal, elderly woman on corticosteroids.

Physiological osteopenia or age-related osteopenia affects all individuals and occurs in cortical and trabecular bone. It leads mainly to hip and vertebral fractures, but other fractures, such as in pelvis or humerus, are also common. Individuals with age-related osteoporosis have reduced bone mass at all sites, with values in the lower half of the distribution range (for further discussion see below under peak bone mass).

Postmenopausal osteoporosis (type I)

Bone loss in women aged between 40 and 65 years is primarily due to loss of ovarian function, as predicted by Albright (Albright *et al*. 1976). Most studies suggest that, in postmenopausal women, bone loss began perimenopausally and will decrease exponentially after 5 to 8 years to match the slower, age-related loss (Krolner and Nielsen 1982). Even though 80 per cent of the skeleton consists of cortical bone, loss occurs predominantly in bones of primarily trabecular structure since the surface area of trabecular bone is four times that of cortical bone. The rapid bone loss is due to increased resorption superimposed on the age-related loss.

The mechanism by which oestrogens protect the skeleton is unknown. An early hypothesis suggested that malabsorption of calcium was the primary event that leads to removal of calcium from the skeleton. This is unlikely since calcium alone cannot reduce postmenopausal bone loss. In addition, reduced absorption of calcium would be expected to activate the release of parathyroid hormone (**PTH**) and increase the production of 1,25(OH)$_2$D (calcitriol), whereas the opposite is usually found in postmenopausal women. An alternative theory suggested that the menopause is accompanied by a relative deficiency in calcitonin reserve or secretion (Reginster *et al*. 1987), leading to a partial removal of the inhibitory effects of this hormone on osteoclastic bone resorption. However, other studies have not observed impaired calcitonin secretion after the menopause, and patients who have been rendered calcitonin-deficient by thyroidectomy do not have an increased risk of osteoporosis, which would strongly argue against a major role for calcitonin in postmenopausal bone loss (Stevenson 1988). Recently, oestrogen receptors have been identified in human osteoblast-like cells (Eriksen *et al*. 1988) and osteocytes (Braidman *et al*. 1995), which suggests that oestrogen can act directly on bone. One effect of oestrogen on bone cells is to increase the release of anabolic cytokines such as transforming growth factor-b (TGF-b), a potent regulator of osteoblast proliferation and a potential coupling agent between bone resorption and formation in the remodelling cycle (Gowen 1991). Peripheral blood monocytes from women with 'high turnover' osteoporosis, histologically defined as having increased bone resorption and bone formation, synthesize higher amounts of the catabolic cytokine interleukin (**IL**) 1 than those from women with 'low turnover' osteoporosis, in whom bone formation is impaired (Pacifici *et al*. 1987). Furthermore, in the early years after ovarian failure, when bone resorption is rapid, the increased synthesis of IL-1 by monocytes can be blocked by oestrogen replacement therapy (Pacifici *et al*. 1989). Similarly, oestrogen decreases the release of tumour necrosis factor (**TNF**)-a from circulating monocytes of postmenopausal but not premenopausal women (Ralston *et al*. 1990). The increased production of IL-1 and TNF-a by *ex vivo* cultures of monocytes after a surgical oophorectomy was reversed by oestrogen replacement (Pacifici *et al*. 1991). It is likely that IL-1 and TNF play an important part in bone loss associated with oestrogen deficiency.

Two other cytokines, IL-6 and IL-11, may also be targets for the antiresorptive effect of oestrogens (Manolagas and Jilka 1995). To explore the hypothesis that osteoclasts and granulocyte/macrophages arise from common haemopoietic progenitor, the genesis of osteoclasts in animal models of osteoporosis was assessed by using bone marrow culture and counting granulocyte/macrophage colony-forming units. Ovariectomized mice had a high number of granulocyte/macrophage colony-forming units from bone marrow as well as from spleen than normal controls, which correlated with a greater number of osteoclasts present in their bone. This effect was prevented when animals were treated with either oestrogens or antibodies to IL-6 (Jilka *et al*. 1992). The hypothesis generated by this study was that IL-6 directs the genesis of osteoclasts in oestrogen-depleted animals. More recently the same group have also shown that IL-11 may also play a critical part in osteoclast differentiation and development in an oestrogen-depleted state (Girasole *et al*. 1994). Thus, oestrogens may maintain bone mass by inhibiting the release of cytokines that stimulate bone resorption, probably by an effect on osteoclast differentiation, while upgrading the synthesis of other cytokines involved in bone formation. This interrelation between oestrogens, cytokines, and bone resorption may turn out to be much more complicated, but also offers a potential for targeting therapy of osteoporosis.

Age-related ('senile') osteoporosis (type II)

Type II osteoporosis, as defined by Riggs and Melton (1986), affects men and women of 70 years or older, resulting in hip and vertebral fractures, although fractures at other sites are also common. The vertebral fractures are of the wedge variety leading to dorsal kyphosis, and often the deformity is painless, the result of gradual, age-related loss of trabecular bone. Unlike type I osteoporosis, the incidence of hip fractures in type II osteoporosis is only twice as great in women as in men, suggesting that the age-related bone loss is important in the pathogenesis of this disorder (Fig. 1). It has been suggested that the two important factors responsible for the slow phase of bone loss are defective osteoblast function and renal endocrine failure. The impaired production of 1,25(OH) $_2$D by the kidney leads to decreased calcium absorption and secondary hyperparathyroidism. Low concentrations of 1,25(OH) $_2$D have been found in some elderly patients with hip fractures but concentrations of PTH have not been found to be elevated in all studies. However, in a recent trial on the prevention of hip fractures by vitamin D and calcium supplementation, the effect of these two drugs was to reduce the high normal concentrations of PTH in the elderly population and at the same time reduce substantially the risk of hip fractures (Chapuy *et al*. 1992). In some studies, serum 1,25-dihydroxyvitamin D_3 is found to be lower in women with postmenopausal osteoporosis than in age-matched controls, and the density and function of vitamin D receptors in the intestine were also lower in such individuals. Since synthesis of 1,25 dihydroxy vitamin D_3 in response to PTH may be impaired in women with postmenopausal osteoporosis, one can generate a hypothesis that reduced calcium intake and vitamin D synthesis lead to increased activity of PTH, then producing an attempt to correct the deficiency by stimulating the renal la-hydroxylase to synthesize more 1,25 dihydroxy vitamin D_3. As this pathway is less efficient, calcium deficiency persists and the increased PTH leads to increased bone resorption.

Idiopathic osteoporosis

This defines the occurrence of osteoporosis in premenopausal women or men under the age of 60 who do not have an obvious secondary cause for the disease. The male:female ratio is usually 10:1 and presentation is usually with vertebral compression fractures. Some individuals may be found to have mild forms of osteogenesis imperfecta. A subgroup of males with idiopathic osteoporosis has 'high bone turnover' osteoporosis with hypercalcuria and may be treated with antiresorption drugs such as calcitonin and bisphosphonates (Perry *et al*. 1982); most of the others have defective osteoblast function with low rates of bone formation (Jackson *et al*. 1987).

Juvenile osteoporosis

This is an uncommon disorder that occurs before, or at the onset of, puberty and affects both sexes equally. The pathophysiology is unknown. The child usually presents with pain in the lower back, hip, and feet, and fractures of weight-bearing joints and collapsed vertebrae. There is no consistent biochemical abnormality, and radiographs may show osteopenia, fracture and vertebral abnormalities, and scoliosis. The differential diagnosis includes osteogenesis imperfecta and osteomalacia, and other secondary causes should be excluded. There is no specific treatment for this disorder and recovery occurs as puberty progresses, with a return of bone mass towards normal. In some patients, however, fractures lead to deformity of the spine or extremity, and it is important that once the condition is diagnosed, non-weight-bearing and supportive physical therapy be made available.

Endocrine causes

Corticosteroids

Endogenous Cushing syndrome or chronic treatment with glucocorticosteroids is a well-recognized cause for severe osteopenia and fractures (especially of the vertebrae and ribs, sites that contain predominantly trabecular bone) and sometimes occurs without other evidence of glucocorticoid excess. In endogenous Cushing syndrome the prevalence of osteoporosis may be as high as 50 per cent and the incidence of spontaneous fractures 19 per cent (Ross and Linch 1982). In iatrogenic disease the prevalence of fractures may vary from 2 to 17 per cent, but the incidence may depend on the dose of drug used and its duration (Reid 1990). Asthmatics on a mean dose of 12 mg prednisolone over 7 years had a 34 per cent prevalence of vertebral fractures (Luengo *et al*. 1991). In a retrospective study of patients with severe asthma on corticosteroids, vertebral fractures were found in 11 per cent of patients admitted to hospital (Adinoff and Hollister 1983). An even higher prevalence was found in a small prospective study of patients with asthma, with 42 per cent having vertebral and hip fractures (Adinoff and Hollister 1983). An increased incidence of fractures has also been noted in patients with rheumatoid arthritis treated with glucocorticoids. Michel *et al*. (1991) reported a 34 per cent prevalence of vertebral fractures in patients with rheumatoid arthritis on corticosteroids. A more recent study using morphometric techniques found a 12 per cent prevalence in postmenopausal women under 65 years of age with rheumatoid arthritis compared to 6 per cent in matched controls, with no increase in prevalence in those treated with doses of prednisolone below 7.5 mg (Spector *et al*. 1993). Postmenopausal females with rheumatoid arthritis on a mean daily dose of prednisolone 7 mg experienced twice as many vertebral fractures as patients not receiving prednisolone (Laan *et al*. 1992). A case-controlled study suggested a doubling of the risk of hip fractures in patients with rheumatoid arthritis receiving glucocorticoids (Cooper *et al*. 1995).

Individuals at greatest risk of glucocorticoid-induced osteoporosis are postmenopausal women and men over the age of 50 years, as well as prepubertal children, in whom corticosteroids also delay skeletal growth.

Reduced bone mass in patients treated with corticosteroids has been demonstrated at multiple skeletal sites, including primary cortical bone in, for example, the femoral head, using a variety of techniques including quantitative computed tomography and dual-energy X-ray absorptiometry. Trabecular bone with its larger surface area seems particularly prone to loss. Vertebral bone density may be reduced by as much as 40 per cent in patients treated with a moderate dose of corticosteroids (Laan *et al*. 1993*a*). Bone loss from trabecular regions may occur very rapidly, especially in the first year of treatment (Laan *et al*. 1993*b*).

The diagnosis of corticosteroid-induced osteoporosis is relatively easy in a young patient who sustains fractures whilst on corticosteroids. However, it is a more difficult diagnosis to make in an older patient where there may be other risk factors such as postmenopausal osteoporosis. Although osteoporosis has been defined on the basis of bone density using the WHO criteria, the definition of postmenopausal osteoporosis may not be relevant to corticosteroid-induced osteoporosis since there is evidence that fractures in patients treated with corticosteroids occur at higher bone density (Luengo *et al*. 1991). This difference makes it difficult to identify the individual at risk, particularly on the basis of pretreatment measurements of bone density.

The pathogenesis of glucocorticoid-induced osteoporosis is controversial since these drugs may affect calcium and phosphate metabolism, either directly or indirectly through actions on bone, kidney, and intestine (Fig. 6). Corticosteroids may indirectly enhance bone resorption and directly suppress bone formation by decreasing osteoblast function and maturation. The increased osteoclastic bone resorption by corticosteroids is secondary to reduced intestinal absorption of calcium leading to a rise in PTH, and in animals parathyroidectomy abolishes the osteoclastic response to steroids. The hypothesis that secondary hypoparathyroidism plays a part in the osteoporosis which results from glucocorticoid excess leading to a decrease in intestinal calcium absorption and increased urinary excretion of calcium has been put in some doubt by recent studies showing that, when intact PTH is measured, its concentrations do not change with the use of glucocorticoids.

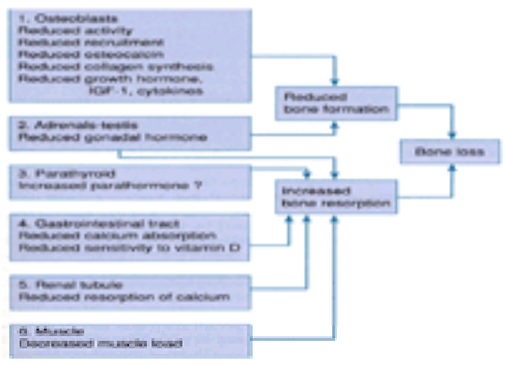


Fig. 6 Pathogenesis of glucocorticoid-induced osteoporosis.

Another mechanism for steroid osteoporosis may be through inhibition of secretion of calcitonin, a hormone whose major effect is to inhibit bone resorption. Decreased intestinal absorption of calcium by glucocorticoids was thought to be due to decreased production of 1,25-dihydroxyvitamin D_3 from its precursor, 25-hydroxyvitamin D_3. This is unlikely, since recent studies indicate that the circulating concentrations of these vitamin D metabolites are unaltered and that the defect in calcium absorption cannot be overcome entirely by dietary supplementation with them. The mechanism of action is probably indirect, and, since glucocorticoid receptors are present in the small intestine, glucocorticoids may interfere with the binding of 1,25-dihydroxyvitamin D_3 to its receptor in the intestine or affect the function of important enzymes in the mucosal cell needed for calcium absorption and transport.

Excessive exogenous use of corticosteroids may lead to a reduction in testosterone and contribute to bone loss seen in males. In females the adrenal synthesis of androgens may also be reduced due to suppression of endogenous corticosteroid synthesis.

Hyperparathyroidism

Primary hyperparathyroidism is a relatively common disorder with a prevalence in adults estimated at 0.1 to 0.2 per cent. The disease is most common between the fifth and seventh decades of life, and is three times more common in women than men, the female predominance arising entirely from an increased incidence after the menopause (Bhalla 1986). Most patients with mild hyperparathyroidism are asymptomatic, the diagnosis being made during a biochemical screening when a high serum calcium is detected. PTH stimulates osteoclastic bone resorption by an effect primarily on osteoblasts, which in turn release a factor or factors that cause bone resorption or enhance osteoclastic activity. In young individuals, increased bone resorption may be compensated for by increased bone formation, but in older individuals, and in particular postmenopausal females, the compensatory mechanism is less efficient, leading to reduced bone mass in the axial and the appendicular skeleton (Devogelar *et al*. 1984; Martin *et al*. 1986), which, however, does not affect all patients. There is conflicting evidence for recovery of bone mass after parathyroidectomy and further studies are necessary (Martin *et al*. 1986). Since reduced bone mass is a predictor of fractures, parathyroid-mediated bone loss should result in an increased risk of fracture but this possibility remains controversial. Some studies have reported a higher prevalence of spinal fracture (Dauphine *et al*. 1975) that others have not been able to confirm (Wilson *et al*. 1988).

Hypogonadism

All causes of hypogonadism lead to bone loss and may be associated with an increased risk in osteoporotic fractures. Amenorrhoea, whether primary or secondary, is associated with decreased bone mass and increased risk of fracture. Women with endometriosis who are treated with antagonists to luteinizing hormone-releasing hormone develop reversible hypo-oestrogen states in which trabecular bone loss occurs but recovers once treatment is discontinued (Matta *et al*. 1981). Similarly, women with hyperprolactinaemia and anorexia nervosa accompanied by amenorrhoea develop spinal osteopenia that in some cases leads to vertebral fractures. Amenorrhoea is observed in elite female athletes, who develop a reduced spinal bone destiny compared to that of eumenorrhoeic runners (Drinkwater *et al*. 1984), indicating that exercise alone cannot protect the skeleton in the face of oestrogen deficiency. The duration of amenorrhoea, from whatever cause, that will lead to irreversible bone loss is unknown, but is thought to be more than 6 months. Testosterone deficiency is present in approx. 30 per cent of men with spinal osteoporosis; they usually present in the sixth decade but most have symptoms of hypogonadism in the preceding 20 to 30 years (Jackson and Kleerekoper 1990). In males, hypogonadism from any cause may be associated with osteoporosis, including those with Klinefelter syndrome, hypogonadotrophic hypogonadism, hyperprolactinaemia, anorexia nervosa, and mumps orchitis.

Hyperthyroidism

In thyrotoxicosis, bone turnover may be significantly increased due to enhanced recruitment and increased resorption. Patient with a past history of thyrotoxicosis when compared with those who have not had thyrotoxicosis have an increased relative risk of hip fractures of 2.4 (Cummings *et al*. 1992). Thyrotoxicosis causing high-turnover osteoporosis should be thought of in elderly individuals in whom the diagnosis may not be clinically obvious and will be dependent on laboratory testing. Excessive thyroxine replacement in the hyperthyrotoxic individual may also increase bone loss, although the evidence for this is conflicting.

Nutrition

Calcium

While bone mass can may be transiently increased by calcium supplementation in adolescents, there is no evidence to suggest that above-normal intakes of calcium influence the peak bone mass (see 'Management of osteoporosis' below).

Drugs

The effects of glucocorticoids and thyroxine have been discussed above. Chronic heparin therapy can lead to decreased bone mass by enhancing PTH-mediated bone resorption, but warfarin is not associated with this complication despite a low circulating osteocalcin and impaired carboxylation of osteocalcin in warfarin-treated patients. Chronic administration of anticonvulsants lead to disturbance in calcium and vitamin D metabolism. These drugs increase the activity of hepatic microsomal mixed oxidases, leading to increased catabolism of steroid hormones (including vitamin D) and resulting in reduced concentrations of 25-hydroxyvitamin D [**25(OH)D**]. In addition, anticonvulsants directly suppress calcium absorption and antagonize the effects of PTH and vitamin D metabolites on bone. The end result is osteopenia and osteomalacia, particularly if there are other risk factors such as immobility, although not all individuals on chronic therapy are so affected. Methotrexate and other chemotherapeutic agents may affect skeletal metabolism and mass by a direct toxic effect on bone cells or by inducing hypogonadism. The use of antagonists or agonists to luteinizing hormone-releasing hormone to induce ovarian failure in premenstrual tension syndrome and endometriosis results in bone loss that is reversible if treatment is discontinued within 6 months.

Malignancy

Malignant infiltration and replacement of marrow tissue occur in multiple myeloma, lymphoma, leukaemia, systemic mastocytosis, and diffuse bony metastases. Myeloma is usually characterized by focal lytic lesions associated with the accumulation of myeloma cells, but in a small number of patients a diffuse loss of bone that mimics primary osteoporosis occurs. The increased bone resorption is due to the production of osteoclast-activating cytokines such as IL-1 and TNF-a and -b by plasma cells (Garrett *et al*. 1987). In lymphoma a generalized osteoporosis and hypercalcaemia may be the result of local synthesis of 1,25(OH) $_2D_3$ by malignant tissue. Mast cells secrete various products (heparin, histamine, prostaglandin) that could lead to increased osteoclastic bone resorption, and increased numbers of mast cells have been reported in the bone marrow in primary osteoporosis (Fallon *et al*. 1983).

Immobilization

Prolonged immobilization leads to disuse osteoporosis affecting trabecular and cortical bone. The rate of bone loss may be as high as 5 per cent per month for the

first 6 months, after which it slows down to a new steady-state and some recovery is possible on resuming weight-bearing activities. Immobilization leads to increased bone resorption and decreased bone formation associated with hypercalciuria, and to an increase in plasma calcium, increased risk of renal stone formation, and potential suppression of PTH, 1,25(OH)$_2$D and calcium absorption.

Inflammatory disorders

Rheumatoid arthritis

In rheumatoid arthritis, bone loss can be periarticular, a hallmark of the disease, or a more generalized form, as demonstrated in histological, computed tomographic, and dual-photon absorptiometric studies (Dequeker and Geusens 1990). Increased rates of bone loss have been demonstrated by neutron-activated analysis of total body calcium. The pathogenesis of localized osteoporosis involves the production of numerous cytokines such as IL-1, IL-6, transforming growth factor-b, TNF, other growth factors, prostaglandins, and mast cell products (Dequeker and Geusens 1990). At present it is unclear whether all or only some patients with rheumatoid arthritis are affected by generalized osteoporosis, and when present its clinical significance is debatable. While it is not surprising that patients with long-standing destructive and disabling disease have low bone mass (Verstraeten and Dequeker 1990), population-based epidemiological studies show an increased risk of fracture in patients with rheumatoid arthritis, related to age, impaired ambulation, body mass, and corticosteroid use (Hooyman *et al*. 1984; Cooper and Wickham 1990). Factors likely to lead to osteoporosis in rheumatoid arthritis are shown in Table 5. Usually, local factors such as cytokines exert their effect within the local microenvironment. Some cytokines, such as IL-1 and IL-6, may have systemic actions such as the induction of fever and the acute-phase response during inflammation. In rheumatoid arthritis, large amounts of local factors are produced within the joint and have the potential for systemic absorption affecting bone turnover at distant sites. Elevated serum IL-1 that correlates with disease activity has been demonstrated in patients with rheumatoid arthritis (Eastgate *et al*. 1988). However, various inhibitors that may antagonize cytokines have also been found in both serum and synovial fluid (Arend and Dayer 1990) and the net systemic effect on bone turnover of cytokine produced during joint inflammation is unclear. Some studies have demonstrated an association between disease activity and decreased bone mineral content (Sambrook *et al*. 1986) whilst others have not (Rosenspire *et al*. 1980). The difficulty in identifying a correlation between disease activity and bone mineral content is not surprising as almost all the studies involve cross-sectional measurements of the latter, a reflection of long-term accumulation of acute changes in bone turnover, with measurements of disease activity that varies relatively rapidly and may be associated with short-term changes in bone turnover best made by serological and urinary markers of bone metabolism.

Systemic actions of inflammatory products associated with disease activity
Alterations in circulating sex and calcitropic hormones
Altered calcium metabolism
Changes in load bearing of the skeleton
Effect of drugs used in treatment of arthritis

Table 5 Potential factors causing generalized osteoporosis in rheumatoid arthritis

During inflammation an alteration in the circulating levels of hormones has the potential for directly affecting bone turnover or modulating the effects of local factors. Various hormonal disturbances have been described in rheumatoid arthritis, such as changes in circulating sex hormone levels, but their significance in terms of disease activity and bone turnover is unclear. Sambrook *et al*. (1988) found a decreased oestrone, dehydroepiandrosterone sulphate (**DHEAS**), and testosterone; they suggested that the reduction in oestrone and testosterone, but not DHEAS, was secondary to the use of prednisolone. They also noted that reduced DHEAS in postmenopausal women with rheumatoid arthritis correlated with reduced bone mineral density of the femoral neck, suggesting that the disease itself may somehow reduce production of adrenal androgens. A decreased free and total testosterone have been found in some males with rheumatoid arthritis but the significance is as yet unclear. There are conflicting results on the effects of rheumatoid arthritis on calcitropic hormones. In most studies, the circulating levels of 25(OH)D, calcitonin, and PTH are reportedly normal but some have shown a low 25(OH)D (Dequeker and Geusens 1990).

Although calcium malabsorption has been described in patients with rheumatoid arthritis, the majority of workers have reported normal serum calcium and phosphate as well as urinary calcium excretion.

Decreased physical activity leads to decreased functional loading of the skeleton resulting in increased bone resorption. Decreased bone mineral content at the femoral neck and lumbar spine of patients with rheumatoid arthritis is significantly correlated with functional activity and is probably one of the major factors responsible for the bone loss associated with rheumatoid arthritis.

Drugs used in the treatment of rheumatoid arthritis may affect bone loss adversely by affecting bone cell activity or favourably by reducing inflammation and indirectly its effect on bone loss. Corticosteroids in high doses result in generalized bone loss and an increased rate of fracture (Hooyman *et al*. 1984; Cooper *et al*. 1995). This adverse effect is not as obvious at lower doses and there is conflicting evidence as to whether bone loss occurs with doses of less than 7.5 mg of prednisolone per day (Sambrook *et al*. 1986; Butler *et al*. 1991). The effect of intermittent intravenous corticosteroids is unclear, although they are likely to result in an acute suppression of bone metabolism that may be cumulative with frequent treatment. Non-steroidal anti-inflammatory drugs *in vitro* affect osteoblast and osteoclast function and prostaglandin and cytokine production (Gowen 1991), but an adverse independent effect on bone loss has not been demonstrated. The effect of disease-modifying antirheumatic drugs on generalized bone mineral content has not been studied extensively. The rate of generalized bone loss in patients with rheumatoid arthritis on disease-modifying antirheumatic drugs was not different from those not taking them in one study (Reid *et al*. 1981), but others have shown that the drugs prevent bone loss or increase bone mineral content (Kalla *et al*. 1991), probably by suppressing disease activity and thereby improving mobility.

Ankylosing spondylitis

Patients with severe ankylosing spondylitis often develop a dorsal kyphosis, and radiological findings of anterior vertebral wedging in the dorsal spine and of biconcave vertebrae suggest moderate osteoporosis (Ralston *et al*. 1990). Early reports on the prevalence of fractures in ankylosing spondylitis found a low prevalence of vertebral fractures (Wilkinson and Bywater 1958; Hanson *et al*. 1971; Hunter and Dubo 1978; Hunter and Dubo 1983). These studies also indicated that fractures were related to spinal trauma and correlated directly with the duration of the disease. In contrast, other studies such as by Ralston *et al*. (1990) found a much higher prevalence of fractures. Ralston *et al*. (1990) also observed that the development of fractures in their patients correlated directly with disease duration, and noted that vertebral fractures occurred spontaneously and were not related to trauma Cooper *et al*. (1994) found a significantly increased risk of clinically diagnosed vertebral compression fractures in patients with ankylosing spondylitis, but no increase in the risk of limb fractures.

Reid *et al*. (1986), using absorptiometry, reported high bone mineral density of the lumbar spine in patients with ankylosing spondylitis, but Will *et al*. (1989), also using dual X-ray absorptiometry, reported reduced bone mineral density in the lumbar spine and femoral neck of patients with early ankylosing spondylitis when spinal mobility and radiographs were still normal. In patients with more advanced disease there was a further reduction of bone density at the femoral neck and the carpus, suggesting that in ankylosing spondylitis bone loss is progressive and initially involves trabecular bone but later also cortical bone. The same group studied patients with early disease and matched them with same-sex siblings of similar physical activity. The bone density of the femoral neck was reduced in the group with ankylosing spondylitis, indicating that the reduced bone mass is a result of bone loss related to disease rather than failure to achieve adequate bone mass within families with ankylosing spondylitis (Will *et al*. 1990). The reduction of bone density in early ankylosing spondylitis was confirmed by Devogelar *et al*. (1992) and Lanyi *et al*. (1993); in patients with well-established ankylosing spondylitis, both groups of investigators noted increased bone density that was the direct result of syndesmophyte formation. Decreased density of hip bone has also been noted in ankylosing spondylitis (Will *et al*. 1989; Mullaji *et al*. 1994). The pathogenesis of reduced bone density in ankylosing spondylitis is unclear. Since reduced bone density has been found in patients with early disease who have normal spinal mobility, it is unlikely that immobilization of the spine is the primary event. Will *et al*. (1989) suggest that osteoporosis in ankylosing spondylitis is a primary pathological event due to the inflammatory nature of the disease. However, the precise mechanism by which inflammatory mediators are responsible for bone loss remains to be

elucidated.

Clinical features

Typically a patient with osteoporosis is a thin Caucasian female aged between 50 and 70 years who presents with back pain resulting from vertebral fractures. Pain is usually of spontaneous onset, well localized, aggravated by movement, and radiates anteriorly to the lower rib cage and around the flank into the abdomen and thigh. Compression of nerve roots is uncommon and the painful episode usually subsides after 4 to 6 weeks. Some patients, particularly after multiple vertebral fractures, develop a constant dull pain in the lower thoracic or lumbar area that shows many of the features associated with mechanical disorders of the spine. Compression of vertebrae, particularly in the lumbar area, is associated with loss of height, which may be the sole complaint in some patients in whom vertebral compression has been painless. Anterior wedging of the vertebrae leads to a dorsal kyphosis (also known as 'dowager's hump'), with downward angulation of the ribs and a reduction in the gap between the ribs and the iliac crest of the pelvis. Contact between the lower ribs and the iliac crest may cause severe pain in the abdomen and inguinal area. Abdominal protuberance results from downward pressure on the abdominal cavity and may be associated with reflux oesophagitis and a feeling of fullness at meal times (Fig. 7).

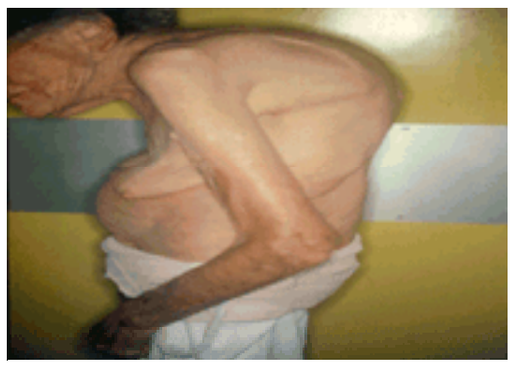


Fig. 7 Severe dorsal kyphosis and protuberant abdomen in a 75-year-old woman with osteoporosis.

Patients with appendicular fractures, usually of the wrist and femur, are often seen in emergency departments, although in the elderly individual a diagnosis of a hip fracture may be easily missed as pain may be minimal.

Pathogenesis

Bone growth and peak bone mass

During childhood and adolescence, growth and modelling leads to an increase in the size, shape, strength, and composition of bone. With the closure of the growth plate (also called the epiphyseal cartilage), growth ceases, but the need for mineral homeostasis and skeletal self-repair by removing and replacing effete bone is served by remodelling, a process that continues throughout life. In remodelling, unlike modelling, the sequence of bone formation and resorption is anatomically and quantitatively coupled. The resorption of a given amount of bone is followed by the deposition of an equivalent amount of new bone, but after the age of 35 to 40 years, presumably due to decreased osteoblastic activity, the amount of bone laid down is less than the amount resorbed during each remodelling sequence and leads to the bone loss seen with ageing.

A rapid gain in bone mass occurs in infancy and adolescence in both sexes. In early childhood males and females have similar bone mass, but after puberty males have a greater amount of cortical bone than females. In girls there is a dramatic increase in bone density during puberty and those with low oestrogen concentrations and a history of irregular menses attain a lower bone mass. Bone density in both white and black girls is similar during the early stages of puberty and an increase is observed only in the later stages of puberty, but at a greater rate in black than white girls. Using computed tomography to measure axial bone density, Gilsanz *et al*. (1991) observed that black and white girls have similar bone density in the early stages of puberty; at Tanner stage 4 and 5 of puberty, bone density increased substantially in both races, with black girls achieving a greater increase (39 per cent) than white girls (11 per cent) (Fig. 8).

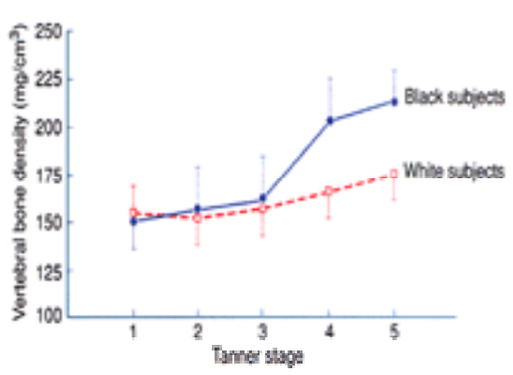


Fig. 8 Vertebral bone density in normal black (●) and white females (□) during different stages of sexual development (adapted from Gilsanz *et al*. 1991).

Peak bone mass, or maximal bone density, is usually achieved in the third decade. Between 85 and 90 per cent of peak bone mass is accumulated during longitudinal growth (i.e. before age 20), but following the closure of the growth plate a further increase in bone mass in the axial and appendicular skeleton occurs between the ages of 20 and 35. Studies by Davies *et al*. (1990) suggest that a gain in bone mass of 3 to 6 per cent per decade may occur from the third decade. Further increases may be possible by changing lifestyle and behavioural patterns (e.g. increased exercise and calcium intake) in adolescence, and it has been suggested, but not yet proven, that increasing the peak bone mass may be an effective means of preventing the consequences of age-related and postmenopausal bone loss. The amount of peak bone mass present at the age of 30 to 35 years is determined by genetic and environmental factors, with the former accounting for 80 per cent. Females achieve a lower peak bone mass, which may explain their increased risk of skeletal fractures in later life. The variance in bone mass between monozygotic twins is less than with dizygotic pairs (Pocock *et al*. 1987) (Fig. 9), and daughters of women with spinal crush fractures achieve a lower peak bone mass than age-matched daughters of non-osteoporotic females, indicating the genetic factors behind peak bone mass. Bone density at several sites (spine, hip, and forearm) is more similar between monozygotic twins than dizygotic twins (Pocock *et al*. 1987) (Fig. 10). The intraclass correlation, a measure of the proportion of the total variance in bone density due to variability among twin pairs, is approx. 0.8 in monozygotic twins and 0.3 in dizygotic twins. If bone mass was entirely genetically determined, the intraclass correlation would be 1.0 and 0.5 in monozygotic and dizygotic twins, respectively (Pocock *et al*. 1987; Slemenda *et al*. 1990*a*). An alternative way of examining the contribution of the environment to bone mass is to look at the maximum within-pair differences in bone density. At all sites, the mean differences between dizygotic twins are almost twice as great as the mean differences between monozygotic twins (Fig. 10). In monozygotic twins the within-pair difference is due to environmental and behaviour factors, and defines the amount of bone mass that may be changed by altering environmental factors. Further evidence for genetic factors is provided by the finding of greater bone mass in black than white individuals, yet the former may have a less suitable environment for skeletal development. The reason for the greater bone mass in black individuals is unclear, but, at least in girls, the difference is apparent at puberty. The mechanism of this effect remains to be elucidated, with increased resistance to the bone-resorptive effects of the calcitropic hormones, namely PTH and 1,25-dihydroxyvitamin D_3, postulated by some (Bell *et al*. 1985) but not confirmed by others.

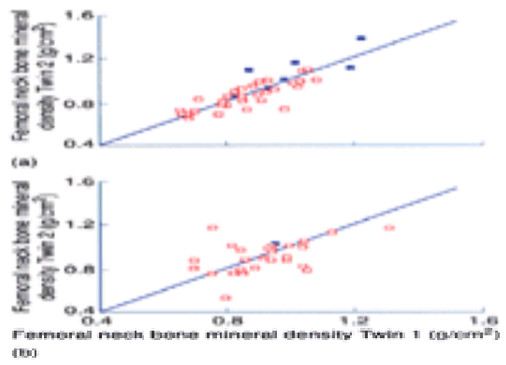


Fig. 9 Correlation of femoral-neck bone mineral density between (a) monozygotic (MZ) twins and (b) dizygotic (DZ) twins showing male (■) and female (☐) twin pairs (reproduced from Pocock *et al*. 1987, with permission).

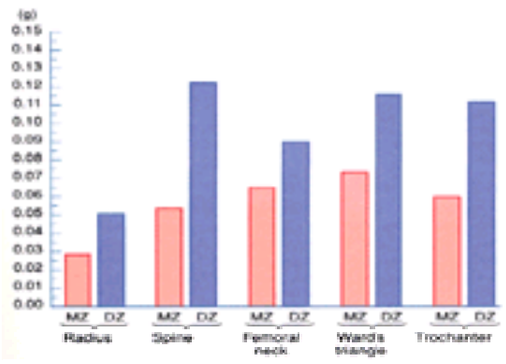


Fig. 10 The mean absolute value of within-pair bone mineral density difference between monozygotic (MZ) and dizygotic (DZ) twins for all skeletal sites (reproduced from Slemenda *et al*. 1991, with permission).

Environmental factors have an important influence on peak bone mass, not only during adolescence but also during the consolidation phase when statural growth has ceased. The influence of nutritional factors, particularly calcium, has attracted considerable attention but remains a matter of debate. Matkovic *et al*. (1979) found that metacarpal cortical bone mass was greater and the incidence of hip fractures lower in regions of the former Yugoslavia where calcium intake was higher. However, increased calcium intake may be associated with an improved socioeconomic status and nutrition, so that it is difficult to be certain that calcium intake was the sole factor for improved metacarpal bone mass. Results from a 3-year double-blind, placebo-controlled trial involving 6- to 14-year-old monozygotic twin pairs, in which one twin received 1 g of calcium supplements daily while the other received placebo, suggest that increased calcium intake leads to a more rapid increase in bone mineral density in growing children (Johnston *et al*. 1992). It is too soon to know whether this early increase leads to a higher peak bone mass or to an earlier establishment of peak bone mass that later does not differ from that of the control twins. Nevertheless, these studies indicate that a population-based strategy for management of osteoporosis may be through improving dietary calcium intake in childhood.

Physical inactivity leads to bone loss and increased physical fitness leads to a greater bone mass in the axial and appendicular skeleton (Pocock *et al*. 1986). In children, weight-bearing physical activity is associated with a moderate increase in bone mass at several sites, which may translate into a greater peak bone mass and a reduced incidence of osteoporotic fractures (Slemenda *et al*. 1991). Excess physical activity, however, is to be avoided, since it may lead to hypothalamic amenorrhoea and bone loss (Drinkwater *et al*. 1984). Although initially the lower bone mass in oligoamenorrhoeic and amenorrhoeic athletes, compared to eumenorrhoeic athletes, was only seen in the lumbar spine, recent studies using improved measurement techniques suggest that bone loss occurs at several sites (Rencken *et al*. 1996). Furthermore, following the return of normal menses, bone mass in these females remains below normal (Drinkwater *et al*. 1990).

Vitamin D receptor gene polymorphism and osteoporosis

Twin studies suggest a strong genetic effect on bone mass, with heritability estimates of 0.5 to 0.9. Higher concentrations of osteocalcin, a marker of increased bone turnover, were reported in twins with a particular vitamin D receptor (*VDR*) gene polymorphism (Morrison *et al*. 1992). Subsequently the same group, using monozygotic and dizygotic twins, noted that 75 per cent of the genetic component of bone density ould be attributed to polymorphism of the *VDR* gene (Morrison *et al*. 1994). Similar findings, but with a weaker correlation, were noted in a British population of postmenopausal females (Spector *et al*. 1995). Ferrari *et al*. (1995) suggested that *VDR* gene polymorphism may predict the rate of changes in spinal bone density in elderly people. However, the association between allelic variations in the *VDR* gene and bone mineral density has not been confirmed by others; both sides of the controversy have been discussed recently (Eisman 1995; Peacock 1995). Thus, Hustmyer *et al*. (1994) found no evidence of linkage between *VDR* alleles and bone density at the spine, hip or wrist in white female twins in the United States. Similarly, in French premenopausal women no association was found between *VDR* alleles and bone density or bone turnover (Garnero *et al*. 1995). At present the difference between these studies is difficult to explain. It is possible that *VDR* gene alleles make a much smaller difference than expected to peak bone mass, which might have been missed in some studies. Alternatively, Parfitt (1994) suggests that *VDR* polymorphism may be linked to another gene polymorphism that regulates bone turnover and that this linkage is influenced by environmental factor(s), which in turn may be confined to one geographical area.

Age-related bone loss

Until the age of 35 to 40 years the amount of bone formed equals the amount removed at each remodelling unit. Thereafter there is a slight mismatch, with less bone being formed, and this results in an age-related loss of bone that is universal in humans and affects all skeletal sites. In both sexes, trabecular and cortical bone mass decline by 6 to 8 and 2 to 4 per cent per decade, respectively. In postmenopausal females, further bone loss occurs associated with declining ovarian function, and in the vertebrae the trabecular loss may be as high as 5 to 10 per cent per year (Stevenson 1988), while cortical bone loss may be 2 per cent per year. The rapid rate of bone loss after ovarian failure, a consequence of the decline in circulating oestrogen level, decreases exponentially over 5 to 8 years eventually to match the age-related bone loss (Krolner and Nielsen 1982).

The mechanism of age-related bone loss is unknown but several possibilities exist. Calcium absorption declines with age in both sexes, as does the ability of the intestine to adapt to a reduced calcium intake by improving absorption. Furthermore, the reduced ability to synthesize 1,25-dihydroxyvitamin D, due to reduced activity of the renal 1a-hydroxylase, will affect intestinal calcium absorption. Reduced calcium absorption will lead to chronic stimulation of parathyroid glands and the increased PTH will accelerate bone loss by stimulating osteoclasts and increasing the number of remodelling units at which uncoupling has occurred. Other factors of importance in age-related bone loss include defective osteoblast function, the conversion of haemopoietic marrow into fatty marrow with loss of precursor cells and locally generated growth factors, and impaired reserves and secretion of calcitonin (Stevenson 1988).

Diagnosis

Risk factors

Since it is easier to prevent than to treat established osteoporosis, it has been suggested that assessment for risk factors may be used to select individuals at sufficient risk and offer them further investigation (Riggs and Melton 1986). These risk factors are detailed in Table 6.

Race (white or Asian)
Premature loss of ovarian function
Postmenopausal
Short stature and low body-mass index
Positive family history
Nulliparity
Low calcium intake
Inactivity
Excessive alcohol intake
Smoking
Long-term use of corticosteroids

Table 6 Risk factors for osteoporosis

The major risk factors are race, loss of ovarian function, and peak bone mass. As mentioned previously, peak bone mass is predominantly genetically determined and racial origin is important in that it is likely that black individuals achieve a greater peak bone mass than white. Loss of ovarian function is associated with bone loss, and premature ovarian failure and amenorrhoea lead to reduced bone mass and are associated with increased risk of osteoporosis. Cigarette smoking may indirectly be harmful to the skeleton because it increases the metabolism of oestrogen, and smokers tend to have a slim build and lower body weight. Obesity may be protective to the skeleton of postmenopausal females by increasing the forces transmitted through it and by increasing conversion of androstenedione to oestrone in fatty tissue (Schindler *et al*. 1972). Females of small build tend to have reduced bone mass. Daughters of women with vertebral osteoporosis tend to have reduced bone mass and a positive family history may be regarded as a risk factor. While risk-factor assessment has been useful for epidemiological studies, its ability to predict bone mass and risk of fracture is poor and it cannot be relied upon to estimate the presence or absence of osteoporosis in individual patients (Slemenda *et al*. 1990*b*; Cooper *et al*. 1991).

Bone mass measurement

A high correlation exists between bone mineral density and bone strength and therefore the risk of fragility fracture. Several prospective studies suggest that the best means of assessing risk of fracture is through measurement of bone mineral density (Johnston *et al*. 1991).

Conventional radiology is a relatively insensitive method of assessing bone mass. Morphological changes, readily assessed semiquantitatively by such methods as the metacarpal index, vertebral body index, and the Singh index for the proximal femur, have been and are still used as a basis for bone mass assay where other techniques are unavailable, but these methods cannot detect early bone loss and also lack sensitivity. Reliable measurements of bone mass can be accurately and precisely obtained by a number of techniques (Table 7). In general these techniques calculate bone mass by measuring tissue absorption of photons obtained from a radioactive or X-ray tube source (Cohn 1991).

Technique	Sites	Radiation dose (μSv)	Units of measurement
Radiogrammetry	Carpus	(Skin dose)	Metacarpal index
	Femoral neck	Carpus 50–500	Singh index
	Vertebrae	Lumbar vertebrae 10 000–240 000	Vertebral body index
Single photon absorptiometry	Forearm	50–100	g/cm g/cm^2
Dual-photon absorptiometry	Lumbar spine Femoral neck Forearm Whole body	10–40	g/cm^2
Dual X-ray absorptiometry	Lumbar spine Femoral neck Forearm Whole body	10–40	g/cm^2
Quantitative computed tomography	Lumbar spine Femur Tibia Forearm	100–10 000	g/cm^2 g/cm^3
Ultrasound	Calcaneus Patella	No ionizing radiation	AVU (apparent velocity of ultrasound) (m/s) BUA (broadband ultrasonic attenuation) (dB/MHz)

Table 7 Techniques available for the measurement of bone mass *in vivo*

Single-photon absorptiometry, available for over 25 years and using low-energy photon beams, measures appendicular bone mass at the forearm and heel. The sites are usually immersed in water to make the soft-tissue thickness constant. The photon beam is directed through the forearm of the patient and detected by a scintillation counter on the other side of the arm. The intensity of the transmitted beam is measured at intervals along the radius and ulna to determine the total bone mineral content. The technique is limited because it can only measure peripheral sites and the position of the beam is critical to avoid errors when doing repeated scans.

Dual-photon absorptiometry uses two photon energies, usually derived from gadolinium, and permits distinction between bone and soft tissue; this allows measurement at sites of fragility fractures in the lumbar spine and proximal femur. The radiation dose is small (50–150 mSv) and precision good (1–4 per cent). The disadvantages are a long scanning time and the need to replace the isotope sources. In many places, dual-photon absorptiometry has been superseded by dual-energy X-ray absorptiometry, which uses X-rays as the source for protons. Its advantages are a reduced scanning time, improved resolution, precision, and accuracy.

Quantitative computed tomography allows volume measurements and also permits distinction between cortical and trabecular bone in the vertebrae. Its main disadvantages are cost and high radiation exposure. Ultrasonography is a non-invasive technique currently under investigation; its use and correlation with bone mineral density measured by dual-energy X-ray absorptiometry is currently being assessed and it cannot as yet be recommended for clinical use.

Bone biopsy

Since the advent of non-invasive techniques to measure bone mass, transiliac bone biopsy is no longer essential to establish the diagnosis of osteoporosis. It may be indicated where osteomalacia is suspected on the basis of abnormal biochemical or radiological features. When undertaken, however, the result often indicates osteoporosis rather than osteomalacia. Bone histomorphometry is of use in quantitating rates of bone formation and resorption, but few centres can undertake this and it is usually done as part of clinical research (Malluche and Faugere 1991).

Biochemical markers

Routine biochemical tests are usually normal in osteoporosis. Biochemical markers of skeletal metabolism are available, but their precise role has not been established. Markers of bone formation include serum alkaline phosphatase (particularly the bone isoenzyme), osteocalcin (also known as bone gla protein), and type 1 procollagen propeptides. Fasting urinary calcium and hydroxyproline, and serum tartrate-resistant acid phosphatase, are markers of bone resorption that lack sensitivity. The urinary excretion of the collagen cross-links, pyridinoline and deoxypyridinoline, reflects breakdown of bone and cartilage collagens and appears to be raised in disorders characterized by high bone turnover such as after the menopause and in primary hyperparathyroidism. The biochemical markers may be of help in investigating bone turnover in established osteoporosis by separating individuals with high-turnover from those with low-turnover osteoporosis, and in monitoring the effects of appropriate treatments.

Management

Prevention

Prevention of bone loss appears to be the most effective means of reducing the incidence of fragility fracture. A reduced bone mass in an individual, and therefore the increased risk of fractures, is related to the peak bone mass achieved by that individual, the rate of bone loss with ageing, and, in females, the rate and duration of postmenopausal bone loss. The determinants of peak bone mass include genetic factors, which cannot be manipulated, and nutritional and environmental factors, which may be manipulated to the individual's advantage. For pharmacological intervention, particularly hormone replacement therapy, the identification of the individual at risk will help restrict therapy to those who need it and avoid treating those whose risk for fragility fracture is low. At present the only means of achieving this is through bone mass measurement, and prospective studies suggest that a reduction in bone density by 1 SD is associated with a relative risk for fracture that varies from 1.5 to 3.0 with site-specificity. In other words, bone mineral density in the spine and hip is a better predictor for fracture of the respective site than that in the forearm is for fracture of the spine or hip.

Calcium

The role of calcium supplementation in preventing bone loss remains controversial. Matkovic's study (Matkovic *et al*. 1979) is often cited as an important piece of evidence appearing to indicate that the preservation of cortical bone was greater in women living in a community with a high-calcium diet than in one with a low-calcium diet. The study has been criticized in that the association was apparently related to other differences between the two communities: the community with a high-calcium diet was physically more active than the low-calcium (Kanis 1991). In their comprehensive review on calcium, Kanis and Passmore (1989) suggest that the inverse relation between calcium intake and hip fracture rate around the world indicates that calcium intake cannot account for the marked variation in fracture rates and argues against a role for calcium in preventing fractures. Studies by Heaney *et al*. (1978) indicated that increased dietary calcium intake was capable of restoring calcium balance in pre- and postmenopausal women, but others have failed to record the advantages of calcium supplementation in preventing bone loss or acting as an effective substitute for hormone replacement (Nilas *et al*. 1984; Riis *et al*. 1987). There may be a role for increasing calcium intake in the elderly osteoporotic female in preventing cortical bone loss and reducing the incidence of further vertebral fractures. A meta-analysis (Cummings 1990) indicated that the beneficial effects of calcium were greatest when calcium intake is low; the advantages in other situations were small and require further study. Calcium supplementation of prepubertal children enhances the rate of increase in bone mineral density but it is unclear if this gain translates into increased peak bone mass (Johnston *et al*. 1992). It is likely that the benefit may be lost once the supplements are stopped. Recent studies over 4 years have shown that, in postmenopausal women, calcium supplements compared to placebo lead to a reduced rate of loss of total-body bone mineral density (Reid *et al*. 1993; Reid *et al*. 1995). In the lumbar spine and femoral neck the difference between the two groups occurred predominantly in the first year; in the subsequent years the calcium-treated group lost bone but at a lower rate than controls. Although there were significantly fewer symptomatic fractures in the calcium group, the study was not large enough to provide a definite answer, which can only be given by a much larger, prospective study. At present, therefore, calcium supplementation should be restricted to those with definite osteoporosis and with a poor calcium diet (daily calcium intake of less than 400 mg), as supplementary therapy in any anabolic treatment, malabsorption states, and in elderly people. Prophylactic use cannot be justified and would have considerable cost implications.

Exercise

It is generally accepted that skeletal mass at any site is dependent on the stress to which it is subjected (Rubin and Lanyon 1984) and that strenuous physical activity in adults is associated with a greater bone mass (Nilsson and Westlin 1971). There is also some evidence that physical activity may reduce the rate of bone lost due to menopause, although the exact type of exercise needed is unclear (Smith *et al*. 1984). In general, however, the exercise must be weight-bearing in type. Physical activity may also be of importance in children and leads to moderate increases in bone mass at several sites, including the hip. The resultant increase in peak bone mass could lead to a significant decrease in the incidence of osteoporotic fractures in later years (Slemenda *et al*. 1991).

Hormone replacement

The most effective treatment in preventing postmenopausal bone loss is oestrogen replacement therapy, which has been shown to reduce the risk of bone loss and incidence of fractures, particularly those of the hip. Loss of ovarian function is associated with a rapid rate of bone loss over a 5- to 10-year period, after which it gradually slows and returns to the age-related rate of bone loss (Lindsay *et al*. 1987). There is histological evidence of increased osteoclastic bone resorption associated with a relative reduction in osteoblastic activity after the menopause. The administration of oestrogen is associated with a reduced rate of bone resorption and therefore a decreased rate of bone loss (Kanis *et al*. 1990). How oestrogens achieve this effect is unclear but they appear to have a direct action on osteoblast activity (Eriksen *et al*. 1988) with a secondary effect on osteoclast activity mediated by cytokines (McSheety and Chambers 1986). Unfortunately, replacement of oestrogen alone is associated with an increased risk of endometrial cancer, which may persist even after discontinuing oestrogen replacement (Paganini-Hill *et al*. 1989). The endometrial cancers are not aggressive and the 5-year survival was near to 100 per cent. The addition of progestogen allows the shedding of the endometrium, and minimizes the risk of hyperplasia and subsequent neoplasia. In general, 12 days of progestogen with continuous oestrogen use carries a negligible risk of hyperplasia (Whitehead and Fraser 1987) and a relative risk of endometrial cancer of 0.9 compared with 1.4 in women using oestrogen alone (Persson *et al*. 1989). Thus, in women who have a regular scheduled bleed, there is no need for a routine endometrial examination to detect hyperplasia unless the bleed occurs prematurely.

Use of the oral contraceptive pill, early menarche, and a late menopause are risk factors for breast cancer and suggest an oestrogen-related effect. Most studies suggest that oestrogen used in hormone replacement therapy (**HRT**) does not increase the overall risk of breast cancer (Brinton *et al*. 1986). There is, however, some concern over very long-term treatment with oestrogens, especially over 10 years, when there may be a small increase in risk (Bergkvist *et al*. 1989). One study from Sweden has suggested a relative risk of 1.7 in women using HRT for more than 7 years (Bergkvist *et al*. 1989), but a recent large meta-analysis suggests a relative risk of 1.3 (Steinberg *et al*. 1991). The Royal College of Obstetrics and Gynaecology Study Group concluded that there may be a duration-dependent risk of breast cancer with HRT and that there was no evidence that progestogens protect against that risk. More recent evidence from the Nurses Health Study suggests that the risk of breast cancer is significantly increased amongst women taking HRT, the relative risk being 1.32, which is not appreciably different if progestogen is also added to the treatment. The use of HRT for 5 to 9 years led to a relative risk of breast cancer of 1.46, which was greatest amongst women aged 60 to 64 years. The relative risk of death due to breast cancer was also increased, at 1.45 in women who had taken medication for five or more years (Colditz *et al*. 1995). Women who are concerned about the breast cancer risk should be encouraged to participate in the United Kingdom national mammographic screening programme, which offers 3-yearly screening from 50 years of age.

With regard to hypertension, there is no evidence that HRT elevates blood pressure and in some instances it has been found to lower the diastolic pressure. Although oral contraceptives increase blood clotting factors and risk of pulmonary embolism, there is currently no good evidence that postmenopausal oestrogen replacement therapy increases the risk of thromboembolic disease. An increased risk of deep-vein thrombosis and pulmonary emboli has recently been reported in women currently taking HRT (Daly *et al*. 1996; Grodstein *et al*. 1996*a*; Jick *et al*. 1996). In all three studies, although the relative risk of venous thrombosis was raised to between 2.1 and 3.5, the absolute risks remained low. For example, Daly *et al*. (1996) found an adjusted relative risk of 3.5, but this would yield only one extra case in 5000 users per year.

HRT is associated with decreased mortality from ischaemic heart disease (by 40–50 per cent). A recent prospective study by Stampfer *et al*. (1991) showed a 50 per cent reduction in the risk of major coronary arterial disease or fatal cardiovascular disease, and no increase in the risk of stroke, in women who had ever used oestrogen replacement therapy compared with women who had never used oestrogen. Part of this effect is probably due to the beneficial effects of oestrogen on lipid levels. The addition of a progestogen does not alter the cardioprotective effects of postmenopausal oestrogen (Grodstein *et al*. 1996*b*). After the menopause, levels of high-density lipoprotein (associated with removing precholesterol from the blood) decline while total cholesterol and low-density lipoprotein increase. Oral oestrogen is associated with a rise in the level of the protective high-density lipoprotein–cholesterol fraction and a decreasing low-density fraction as documented by the Lipid Research Clinic's follow-up study (Bush *et al*. 1983). The effects of combined therapy on ischaemic heart disease are unknown. It is possible, but not proven, that some of the beneficial effects of oestrogen may be reduced by the administration of progestogens (Goldman and Tosteson 1991).

Recently, evidence has also been presented that HRT is associated with an increased risk of deep vein thrombosis and/or pulmonary emboli. The relative risk of venous thrombosis appears to be increased two- to threefold in current users of HRT. The absolute risk in current users remains small, with estimates of 16 and 23 excess cases per 100 000 women per year for all venous thromboembolisms and 6 per 100 000 women per year for pulmonary embolism only. In all three studies presented the risk of venous embolism disappeared on cessation of HRT (Daly *et al*. 1996; Grodstein *et al*. 1996*a*; Jick *et al*. 1996).

Treatment with exogenous oestrogen after the menopause results in a significant reduction in bone resorption as reflected in a fall in the fasting urinary calcium:creatinine and hydroxyproline:creatinine ratios (Lindsay *et al*. 1976). The mechanism of the oestrogen effect is unclear, but is likely to be mediated through the oestrogen-sensitive osteoblast releasing growth factors that inhibit the recruitment and differentiation of osteoclasts from their bone marrow precursors. The minimum oral dose of oestradiol required to prevent bone loss is 2.0 mg/ day, and of conjugated oestrogen 0.625 mg /day. For percutaneous delivery, 3 mg of oestradiol gel, or 50 mg oestradiol implant, or at least 50 mg/24 h of transdermal oestrogen will also be effective. If HRT is discontinued the rate of bone loss is similar

to that seen in early years of the menopause. The optimal duration of treatment is unknown, but it should be long-term (from 5 to 10 years) to maximize the benefit to the cardiovascular system and bone without significant risk. Melton (1987) suggests that treatment for 5 years would shift the hip fracture:age curve to the right. Although life expectancy is increased by about 2 years in HRT users, he argued that the shift in the fracture:age curve would still mean that many women would not live beyond the age where the shifted curve is placed (Melton 1987). From findings by Lindsay *et al*. (1987) it is evidently appropriate to start treatment early in the menopause, since it is unlikely that any existing bone loss can be recovered. Although it remains unclear how long the benefits of HRT on bone mass persist after its discontinuation, a Swedish cohort study suggests that the more potent oestrogens gave a 0.37 relative risk of trochanteric fracture in women who had started treatment before the age of 60 years; in contrast the weaker oestrogen did not confer any benefit (Naessen *et al*. 1990). Other data suggest that the use of HRT 14 years after the menopause may still result in reduction of further bone loss and a slight increase in trabecular bone mass in the axial skeleton (Quigley *et al*. 1987; Lindsay and Tomme 1990).

Calcitonin

Calcitonin is a 32-residue peptide hormone produced by the parafollicular cells of the thyroid gland. The hormone binds to high-affinity receptors on osteoclasts, but its precise physiological role in man is uncertain. Preliminary studies suggest that calcitonin may be as effective as oestrogen in abolishing postmenopausal bone loss (Reginster *et al*. 1987). Further studies are necessary to determine optimal dose, route of administration, and long-term effect. If confirmed it may prove an alternative to HRT in individuals in whom oestrogens are contraindicated or unacceptable. At present, calcitonin is given subcutaneously but the development of new formulations means that an effective nasal spray and suppository delivery systems may soon become available.

Treatment of established osteoporosis

For the very elderly individual with osteoporosis-related fractures there is no effective treatment available. However, for postmenopausal women and men aged 50 to 70 years with established disease, available treatments can be grouped into class I agents, which either impair bone resorption and/or reduce activation frequency, or class II agents, which increase bone formation (Table 8).

Class I—impair bone resorption and/or reduce activation frequencies
Hormone replacement therapy
Calcitonin
Bisphosphonates
Anabolic steroids
Calcium
Vitamin D and metabolites
Class II—stimulate bone formation
Sodium fluoride
Intermittent parathyroid injections

Table 8 Potential treatments for established osteoporosis

HRT

As a class I agent, HRT leads to reduced bone turnover and is effective in preventing postmenopausal bone loss. But HRT may also retard bone loss in established osteoporosis: over a 2-year period, a small increase in bone mass may be seen in trabecular rather than cortical bone as formation continues and refilling of the remodelled space occurs. After about 2 years a plateau is achieved and further increases in bone mass are unlikely. A similar plateau for bone mass may occur with other antiresorptive agents such as calcitonin and bisphosphonates. One problem with HRT in older patients is the reluctance to accept the resumption of menstruation. Long-term studies examining the effectiveness of HRT in preventing further vertebral fractures in patients with spinal osteoporosis are not available.

Calcitonin

In patients with vertebral or wrist fractures, treatment with calcitonin leads to a small increase in bone mass over the first 1 to 2 years. The usefulness of this increase in bone mass will become clearer when results become available from prospective studies in progress on the rates of new vertebral fractures. It may be possible to obtain continued increases in bone mass if calcitonin is administered cyclically. At present, synthetic salmon calcitonin is approved for the treatment of postmenopausal osteoporosis but is required to be given parenterally. It is an expensive drug and similar benefits may be achieved by giving lower doses every 2 to 3 days, although this must be proved by clinical trials.

The analgesic effect of calcitonin, probably mediated through b-endorphin release, is sometimes useful in controlling pain resulting from vertebral crush fractures and enables early mobilization of the patient.

Bisphosphonates

The bisphosphonates are potent inhibitors of bone resorption and have been successfully used to inhibit osteoclastic activity in Paget's disease of bone and malignant hypercalcaemia. Disodium etidronate has recently been used in the treatment of established vertebral osteoporosis. In two double-blind, placebo-controlled studies, etidronate, given cyclically in a dose of 400 mg daily for 2 weeks every 3 months for 3 years, modestly but significantly increased spinal bone density and reduced the incidence of further vertebral deformity and fractures (Storm *et al*. 1990; Watts *et al*. 1990). There was no significant change in cortical bone mass. In an open-label follow-up of 37 postmenopausal females with osteoporosis treated for 5 years, a further increase in bone density of 1.4 per cent was observed during the last 2 years of study; the reduction in vertebral fracture rate observed in the earlier part of the study was also maintained (Storm *et al*. 1996). Etidronate has been approved for use in established vertebral osteoporosis in the United Kingdom and most European countries, but not in the United States,

The aminobisphosphonate, alendronate sodium (alendronate), increases bone mineral density in the spine, hip and total body, and reduces the incidence of new morphometric vertebral fractures (Liberman *et al*. 1993) (Fig. 11). The results of a randomized trial involving 2027 women with at least one vertebral fracture at baseline and low bone-mass density of the femoral neck, treated with alendronate or placebo, were recently reported; they showed that the relative risk of sustaining a new morphometric vertebral fracture was 0.33 in the treated compared to the control group (Black *et al*. 1996). In addition, there was a reduction in clinically apparent vertebral fractures as well as of any clinical fracture in the hip and wrist. The reduced risk of hip fracture noted in this study of postmenopausal women dwelling in the community suggests that bisphosphonates may be of value in preventing such fractures in women with low bone-mass density of the hip. However, until a formal analysis of cost-effectiveness is made, the expense of using alendronate to prevent such fractures may appear prohibitive.

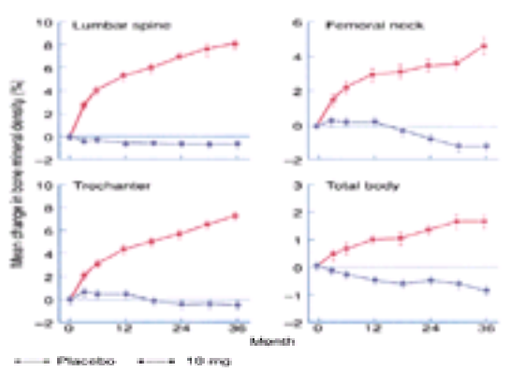


Fig. 11 Changes in bone mineral density from baseline values in women with postmenopausal osteoporosis receiving alendronate 10 mg/day or placebo for 3 years

(after Liberman *et al*. 1993)

In general, bisphosphonates are well tolerated but should be used with caution in patients with impaired renal function. Their absorption is impaired by calcium, iron, magnesium, and aluminium, to which they bind. Some unwanted side-effects of etidronate include hypersensitivity reactions, urticaria, and angio-oedema, but these are uncommon. Etidronate should be used cyclically and not daily as this will lead to impaired normal bone mineralization. In contrast, alendronate, because of its high potency, can be given daily. Although the frequency of upper gastrointestinal side-effects with alendronate was similar to placebo in the randomized trial discussed above, the clinical use of alendronate has been associated with oesophagitis, at times severe enough to warrant admission to hospital, and other upper gastrointestinal symptoms. These symptoms may be more common in those with previous upper gastrointestinal problems and in those taking non-steroidal anti-inflammatory drugs. To some extent, these complications can be avoided by asking the patients to take alendronate with at least 100 ml of water and to avoid lying down during the next 30 min.

Testosterone and anabolic steroids

Hypogonadism in males is a well-established cause of osteoporosis and can be prevented by rendering the individual eugonadal with replacement therapy using testosterone, which increases bone mass by stimulating bone formation. The testosterone ester is usually given intramuscularly every 2 weeks. The newer oral testosterone preparations may also be used, but their effects on bone mass have not been fully evaluated. The use of testosterone in osteoporotic male patients with normal gonadal function cannot be recommended until further data become available showing benefit.

The anabolic steroids, stanozolol and nandrolone denoate, appear to reduce bone resorption as determined by hydroxyproline excretion, and lead to an increase in bone mass in both males and females. The reported increases in bone mass with anabolic steroids may in part be due to measurement artefact: these drugs decrease fat mass and increase lean body mass, which leads to a falsely elevated measure of bone mass. Anabolic steroids also cause mild degrees of masculinization, which is unacceptable to many women, and, in addition, they have adverse effects on plasma lipids, raising doubts about their long-term safety.

Vitamin D and metabolites

It has been proposed that circulating 1,25(OH)D (calcitriol) and calcium absorption are decreased in some, especially elderly, osteoporotic patients. A potential therapeutic role for calcitriol is suggested by its ability to stimulate intestinal calcium absorption, reduce secondary hyperparathyroidism, and abolish PTH-mediated bone resorption. Unfortunately, trials of calcitriol treatment for women with osteoporosis have yielded conflicting results. Some found that the bone density of the spine increased with calcitriol (Gallagher and Goldgar 1990) while others found that it decreased by similar amounts in patients and controls (Ott and Chesnut 1989). In all studies the cohorts have not been large enough to evaluate the effects of calcitriol on the rate of new vertebral fractures. More recently, Tilyard *et al*. (1992) have reported their findings on the occurrence of new vertebral fractures in 622 postmenopausal women with spinal fractures entered into a prospective study of calcitriol on calcium. A threefold reduction in the occurrence of new vertebral fractures or deformity was observed over 3 years in women treated with calcitriol. This effect was evident only in women who had five or fewer vertebral fractures on entry into the study. A reduction in the number of peripheral fractures was also noted in those treated with calcitriol. The dose of calcitriol was 0.25 mg twice a day: it was not associated with any major adverse effects, but concerns about the effect of calcitriol on urinary calcium excretion and renal function remain, since these were not estimated in all patients. In elderly women the addition of 800 units of vitamin D_3 and calcium reduced the number of fractures in comparison to a control group; a significant difference in number was apparent only after 12 to 18 months of the 3-year study (Chapuy *et al*. 1992). The incidence of hip fractures showed a marked increase in the placebo group but remained unchanged in the group treated with vitamin D_3, suggesting that vitamin D_3 prevented the exponential increase in the incidence of hip fractures that takes place in this age group.

Vitamin D deficiency is common in elderly persons, owing to reduced exposure to sunlight, decreased synthesis of vitamin D_3 in the ageing skin, and low dietary intake of vitamin D. Ooms *et al*. (1995) noted that supplementation of the elderly with 400 units of vitamin D_3 daily leads to a slight decrease in secondary hyperparathyroidism and an increase in bone density at the femoral neck. Although Chapuy *et al*. (1992) showed a reduction in the incidence of hip fracture following supplementation with 800 units of vitamin D_3, it is not known whether the benefits were due to the supplemental vitamin D or calcium. Supplementation with 400 units of vitamin D_3 administered to elderly men and women dwelling in the community in Amsterdam did not lead to a decrease in the incidence of hip or peripheral fractures (Lips *et al*. 1996). However, this Dutch population differed in many respects to that studied by Chapuy *et al*. (1992): they were younger, fewer had a raised PTH, and evidence for secondary hyperparathyroidism and vitamin D deficiency was less common. In addition, the dose of vitamin D_3 was 400 units not 800 units as in the French study.

The parenteral administration of vitamin D (150 000 to 300 000 units annually) to elderly individuals did not lead to any reduction in fractures of the leg, but fractures of the arms were reduced by half (Heikinheimo *et al*. 1992).

Sodium fluoride

Sodium fluoride stimulates the proliferation and activity of osteoblasts, leading to an increase in osteoid production and trabecular bone volume. It should be given with substantial calcium supplements to allow mineralization to take place, otherwise osteomalacia will develop.

The drug is rapidly absorbed from the gastrointestinal tract if taken in the fasting state; food will impair absorption. Approximately 60 per cent of the drug accumulates in bone and the rest is excreted by the kidneys. Fluoride attaches to apatite in bone, making it more resistant to resorption. In spinal osteoporosis, moderate doses of sodium fluoride (50 mg/day) reduced the incidence of spinal fracture by 25 per cent after 2 years of treatment (Mamelle *et al*. 1988), but others have found it to be less effective. In recent, randomized, placebo-controlled studies of high-dose sodium fluoride (75 mg daily), increases in axial bone-mineral density were found without any reduction in the incidence of fractures, suggesting that the mechanical strength of the newly formed bones may be reduced (Riggs *et al*. 1990; Kleerekoper and Nelson 1992). These studies have been criticized for using excessive doses of fluoride that may have led to the accumulation of unmineralized osteoid. This issue will only be resolved by further, carefully conducted clinical trials, so that at present fluoride remains an investigational drug.

Although slow-release sodium fluoride reportedly prevented new vertebral fractures and increased bone mineral density in the spine and femoral neck (Pak *et al*. 1995), those trials have been criticized for their design and inability to detect any osteotoxic effects of fluoride (Kleerekoper 1996)

Parathyroid hormones

Animals given daily or alternate-day injections of PTH develop increased amounts of trabecular bone that is biologically normal. In man, daily injections of the 1–34 fragments of PTH, when given alone, increased trabecular bone volume but decreased cortical bone mass; the treatment was abandoned because of the potential disadvantage of increasing the risk of limb fractures. Recent studies using PTH 1–34 in combination with HRT (Reeve *et al*. 1990), calcitriol, or calcitonin demonstrated impressive increases in spinal trabecular bone mass, comparable to that seen with fluoride, but achieved within 12 months as compared to 2 years with sodium fluoride. Side-effects were few and the results suggest the need for a randomized control study of PTH in combination with another antiresorptive agent. At present, therefore, PTH, like sodium fluoride, should be regarded as an investigational drug only.

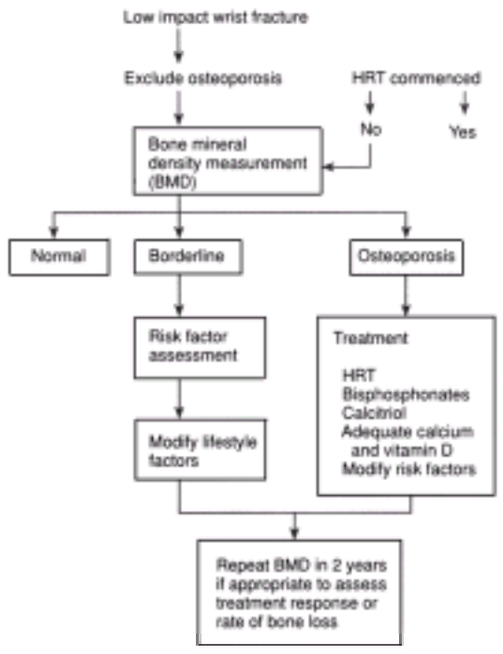
The investigation and management of low-trauma fractures

A low-trauma fracture (i.e., one that follows a fall from standing height or less) should always be investigated to exclude the possibility of osteoporosis. The three most common sites of fracture are the wrist, vertebrae, and hip, and an approach to their management is summarized below.

Wrist fracture

These occur more frequently in younger postmenopausal females than males. The presence of underlying osteoporosis can be confirmed by measuring bone mineral density, although this may not be necessary if the woman elects to go on HRT. If the bone mineral density is consistent with the diagnosis of osteoporosis, treatment suggested in Box 1 should be considered.

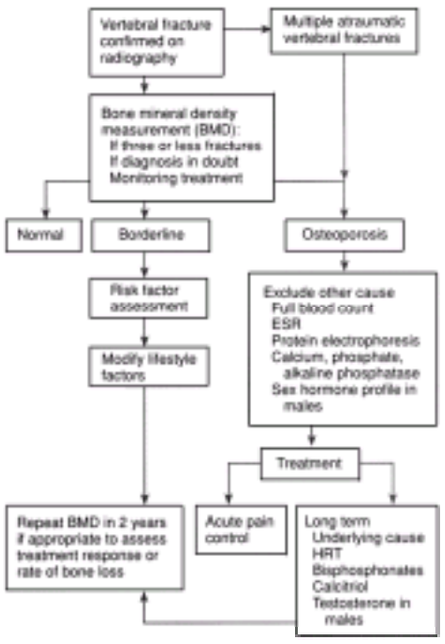
Box 1



Vertebral fracture

Secondary causes such as multiple myeloma and metastases should be excluded where clinically indicated. In those with multiple, atraumatic vertebral fractures, measurement of bone mineral density may be unnecessary, but it should be considered in those with fewer than three vertebral fractures. A repeat measurement of bone mineral density should be considered in those whose baseline density is borderline and where a response to treatment is needed (e.g. when using a bisphosphonate, calcitriol, or calcium and vitamin D) (see Box 2).

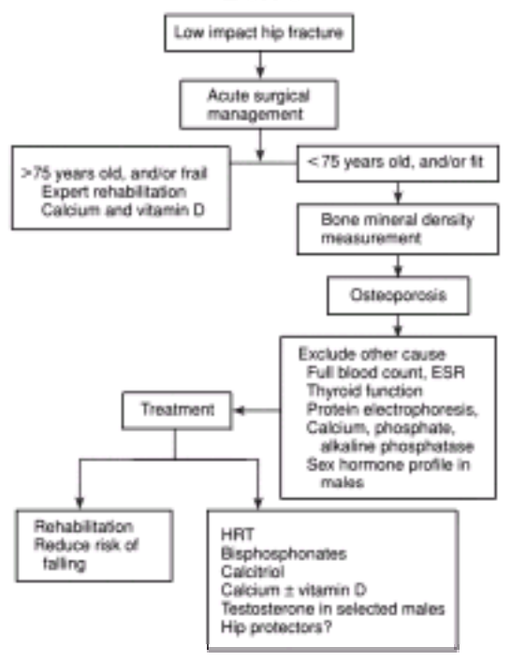
Box 2



Hip fracture

The acute treatment of the hip fracture should follow the Royal College of Physicians guidelines. In the frail elderly person, the future management should consist of adequate vitamin D and calcium. In the younger patient and the independently living elderly person, management with other drugs should be considered as outlined in Box 3. The use of hip protectors remains experimental until their benefit in preventing hip fracture can be demonstrated in the community-dwelling elderly person who falls frequently.

Box 3



Osteomalacia and rickets

Osteomalacia and rickets are characterized by defective mineralization of bone and cartilage leading to an accumulation of unmineralized bone matrix called osteoid. The normal lag between osteoid synthesis and mineralization is prolonged, leading to a decrease in the rate of bone formation. Rickets represents the occurrence of this defect in growing children before the closure of the epiphyses. Unlike osteoporosis, osteomalacia is characterized by a reduction in the ratio of mineralized bone to matrix.

Although osteomalacia can be caused by many disorders, it mostly results either from a primary disorder of vitamin D metabolism or a primary, and non-PTH-related, defect in renal tubular resorption of phosphate (Table 9).

Abnormal vitamin D metabolism
- Reduced bioavailability
  - Nutritional vitamin D deficiency (rare)
  - Inadequate exposure to sunlight
  - Malabsorption states (gastrectomy, coeliac disease, Crohn's disease, pancreatic insufficiency, hepatobiliary disease)
- Defective metabolism
  - Hepatobiliary disease
  - Chronic renal failure
  - Anticonvulsant drugs
  - Vitamin D-dependent rickets type I
  - Oncogenic hypophosphataemic osteomalacia
  - X-linked hypophosphataemia
- Receptor defects
  - Vitamin D-dependent rickets type II

Altered phosphate homeostasis
- Chronic phosphate malabsorption
  - Antacid abuse, malabsorption
- Renal phosphate loss
  - X-linked hypophosphataemia (familial hypophosphatemic rickets)
  - Fanconi syndromes; hereditary forms (Wilson's disease, cystinosis, oculocerebral (Lowe) syndrome, tyrosinaemia), and acquired renal tubular defects (paraproteinaemias, Sjögren's syndrome, systemic lupus erythematosus, monoclonal gammopathies, heavy metal poisoning)
  - Oncogenic hypophosphataemic osteomalacia
- Defective mineralization
  - Aluminium toxicity, bisphosphonates
  - Fibrogenesis imperfecta ossium
  - Hypophosphatasia

Table 9 Classification of osteomalacia

Clinical and laboratory features

The classic symptoms of osteomalacia are generalized bone pain and tenderness, weakness of proximal muscles, and difficulty in walking. However, symptoms may be vague and of insidious onset, making the diagnosis difficult. The bone pain of osteomalacia is dull, poorly localized, and made worse by weight-bearing. It is usually symmetrical and may involve the back, pelvis, thigh, and ribs, often because of the presence of pseudofractures. Compression of the sternum and rib results in pain. Some patients have no pain, particularly in X-linked hypophosphataemic rickets where, in spite of life-long osteomalacia, pain may occur only in middle age.

Muscle weakness, usually proximal, may be of varying severity with only minimal atrophy and no fasciculation. The patient has a waddling gait, and difficulty in climbing stairs and getting out of bed and chair. Muscle biopsy shows no evidence of primary muscle disorder. Proximal weakness is not a feature of X-linked hypophosphataemic rickets.

The features of rickets depend on the age of the child and the underlying syndrome. The child may be hypotonic and apathetic, with growth retardation and delayed walking. Bowing deformities of the long bones will be present when weight-bearing occurs. The epiphyseal plate is widened due to the accumulation of disorganized, unmineralized cartilage, leading to cupping and irregularity at the metaphyseal–epiphyseal junction, usually at the wrist; at the costochondral junction this change gives rise to the 'rachitic rosary'. An indentation may develop at the attachment of the diaphragm to the softened ribs (Harrison's groove). In the neonate and young child the rapid growth of the skull leads to softening of the calvarium (craniotabes), parietal flattening, and frontal bossing. There is delayed dental eruption and enamel defects. Unfortunately, skeletal deformities often persist, even when the histological defect has been corrected, and may require surgical correction.

In osteomalacia and rickets, hypocalcaemia often causes no symptoms, but it sometimes results in paraesthesias, tetanus, cramps, carpopedal spasm, dysarthria, and cardiac arrhythmias. Hypocalcaemia is infrequently severe enough to lead to depression, psychoses, and convulsions.

Radiological features

The radiological features of rickets are most pronounced at the growth plate, where the epiphyses are widened and the calcification border of the metaphysis becomes cup-shaped. The size and density of the ossification centres are reduced, the cortex is indistinct, and there may be bowing of the long bones. In milder forms there may be no diagnostic radiological features.

The pathognomonic radiological feature of osteomalacia is pseudofracture (also known as Looser's zones and Milkman's fracture). Initially the fracture is incomplete, and extends perpendicularly from the cortex with poor callus formation; radiologically it is seen as a radiolucent line. Pseudofractures, often multiple and bilateral, are mainly found in the pubic rami, femoral neck, ribs, clavicles, the outer border of the scapulae, and the metatarsals. They occur in bones subjected to mechanical stress and may have an anatomical relation to entry sites of blood vessels into bone.

Biochemical features

Vitamin D-deficiency osteomalacia is characterized by a low serum calcium and phosphate with elevated serum alkaline phosphatase, increased urinary phosphate excretion, decreased urinary calcium excretion, low circulating levels of 25(OH)D, and a mild elevation of PTH (secondary hyperparathyroidism). Levels of $1,25(OH)_2D$ may be normal or elevated, and therefore unhelpful in the diagnosis of osteomalacia. If the serum calcium and 25(OH)D are normal but the phosphate low, a renal phosphate leak syndrome or end-organ resistance should be suspected. A mild hyperchloraemic acidosis, due to renal bicarbonate loss, may accompany vitamin D deficiency and is a consequence of secondary hyperparathyroidism. Severe hyperchloraemic acidosis suggests a renal tubular disorder, which may be confirmed by failure to acidify the urine after loading with ammonium chloride.

When there is doubt about the occurrence of osteomalacia, the diagnosis can be made by bone biopsy. Specific features of osteomalacia are the presence of excess unmineralized matrix (osteoid) over bone trabeculae, and increased thickness of osteoid. If doubt remains, careful histomorphometry may be needed (Malluche and Faugere 1991). In vitamin D deficiency the biopsy will also show features of hyperparathyroidism, which are usually absent in hypophosphataemic disorders.

Metabolism of vitamin D

The two forms of vitamin D, vitamin D_2 (ergocalciferol) and vitamin D_3 (cholecalciferol), derived from plant and animal tissues respectively, undergo identical metabolism in humans and have similar effects on target tissues. The absence of a subscript often implies a reference to both forms of vitamin D (Fig. 12).

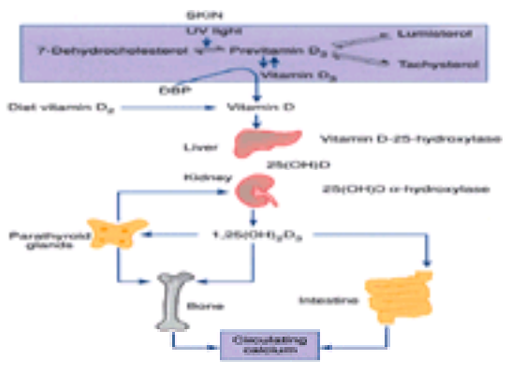


Fig. 12 Schematic representation for metabolism of vitamin D. Previtamin D_3 is formed in the skin and isomerizes to vitamin D_3 or to other biologically inert isomers. Vitamin D-binding protein (DBP) has affinity only for vitamin D_3, which is translocated to the circulation. Vitamin D is then hydroxylated in the liver and kidney to the active metabolite $1,25(OH)_2D$.

Ultraviolet light converts 7-dehydrocholesterol to previtamin D_3. Continued stimulation leads to photoisomerization of previtamin D_3 to two biologically inert products (lumisterol and tachysterol). Previtamin D_3 undergoes thermal isomerization to vitamin D_3, and by binding to vitamin D-binding protein (**DBP**), to which it has a high affinity, it is translocated into the circulation. Dietary vitamin D is absorbed in the jejunum after incorporation into chylomicrons, a process that requires the presence of

bile salts, fatty acids, and monoglycerides.

All forms of vitamin D and their metabolites are transported in the circulation through binding to DBP. The protein has a higher affinity for 25(OH)D than vitamin D or $1,25(OH)_2D_3$. The circulating levels of DBP exceed those of vitamin D and its metabolites, so that all but 1 per cent of these metabolites are present in the free, non-bound form to enter into target cells.

Vitamin D is initially transported to the liver where it is hydroxylated on carbon-25, by vitamin D 25-hydroxylase enzyme, to produce 25(OH)D. The production of 25(OH)D is dependent on the supply of substrate; circulating levels of 25(OH)D reflect an individual's vitamin D status. Only in severe liver disorder is the hepatic production of 25(OH)D impaired, and even then the associated intestinal malabsorption is a contributory factor. From the liver, 25(OH)D, bound to DBP, is transported to the kidney, where it is hydroxylated either to $1a,25(OH)_2D$ (calcitriol) by the renal 25(OH)D 1a-hydroxylase or to $24,25(OH)_2D$. In mammals the principal source of $1a,25(OH)_2D$ is the kidney, except during pregnancy when the placenta can also synthesize it. It is unclear whether other tissues can produce $1a,25(OH)_2D_3$ in normal humans. Granulomatous tissue may possess 1a-hydroxylase activity and this may be the source of the inappropriately high levels of 1,25(OH)D observed in these disorders (Adams *et al*. 1983). Indeed, macrophages activated by lipopolysaccharide or interferon-a, and human T-cell leukaemia-1 virus-infected cord lymphocytes, can also convert 25(OH)D to 1,25(OH)D. The renal 1a-hydroxylase is tightly regulated by PTH, calcium, phosphate, and calcitriol. A high PTH and low calcium stimulate $1,25(OH)_2D$ production, while a raised 1,25(OH)D inhibits calcitriol production but increases the production of $24,25(OH)_2D$.

Calcitriol acts by binding to specific intracellular receptors that occur not only in the target tissues of bone, intestine, kidney, and parathyroid glands but also in mononuclear cells, activated T and B lymphocytes, and skin. Calcitriol stimulates calcium and phosphate absorption in the intestine and increases bone resorption to maintain mineral homeostasis. In bone, however, calcitriol also stimulates the osteoblast and is essential for bone formation, an activity that may also be stimulated by $24,25(OH)_2D$. As mentioned above, receptors for calcitriol are found in tissues that do not have a direct role in mineral homeostasis, including breast cancer cells (Eisman *et al*. 1979), skin keratinocytes and dermal fibroblasts, and haemopoietic cells (Bar-Shavit *et al*. 1983; Bhalla *et al*. 1983). Calcitriol inhibits proliferation and induces differentiation of myeloid cells towards the monocytic lineage (Amento *et al*. 1984; Bhalla *et al*. 1989), perhaps through an early regulation of the c-*myc* gene (Reitsma *et al*. 1983; Kamali *et al*. 1989). *In vivo*, calcitriol, or its 1a analogue, prolong survival of mice inoculated with mouse myeloid leukaemia cells (M1). Calcitriol affects monocyte and T-cell function by influencing the production of cytokines such as IL-1, IL-2, and TNF (Bhalla *et al*. 1986). Since calcitriol can also be synthesized by immune cells it may have an autocrine or paracrine role in the local regulation of immune cell function, but in humans the physiological significance of this needs further elucidated.

Abnormal vitamin D metabolism

Reduced bioavailability

Vitamin D deficiency

Vitamin D deficiency through poor dietary intake is rare unless combined with reduced exposure to sunlight, which occurs especially in elderly people and in the immigrant Asian population of the United Kingdom. Osteomalacia is present in 25 per cent of bone biopsies from elderly patients in the United Kingdom and Scandinavia who have sustained hip fractures, and is likely to be due to insufficient intake of vitamin D and reduced exposure to sunlight. In some countries, dairy products are fortified to prevent nutritional vitamin D deficiency. However, it may still occur in individuals who avoid dairy products and in children who are exclusively breast-fed for prolonged periods.

It is important not to miss a diagnosis of vitamin D-deficiency rickets or osteomalacia, since effective therapy is available and, especially in rickets, limb deformity may be prevented. Bone pain and muscle weakness improve rapidly with replacement therapy, but biochemical and radiological improvements take longer.

The prevention of nutritional vitamin D deficiency is a public-health issue and the deficiency can be eradicated by fortification of certain foods with vitamin D. Simple vitamin D deficiency responds to therapy with oral vitamin D (ergocalciferol) in doses of 2000 to 4000 units (50–100 mg) daily and can be prevented by physiological replacement doses of 200 to 400 units (5–10 mg) daily. In severe vitamin D deficiency it may be better to give a loading dose of 50 000 units (1.25 mg) daily for 1 month followed by physiological replacement. Calcifediol (25(OH)D) can be used instead of vitamin D, where available, and may have fewer problems with accumulation than does vitamin D itself.

Gastrointestinal disorders

Intestinal fat malabsorption due to gastrointestinal, hepatic, or pancreatic disease leads to vitamin D deficiency without the need for other risk factors such as reduced exposure to sunlight. It leads to three distinct types of bone lesions (Rao 1990): (a) increased surface and volume, but not thickness, of osteoid, with normal serum calcium and phosphate but with evidence of secondary hyperparathyroidism—these patients are asymptomatic unless a fracture occurs; (b) increased thickness of osteoid and a prolonged mineralization lag time with reduced concentrations of calcium, phosphate, and elevated alkaline phosphatase—these individuals are symptomatic and have the classical radiological features of osteomalacia; (c) a third group, probably the most common, has low-turnover osteoporosis and is symptomatic only if a fracture occurs. In these disorders and other states of hypovitaminosis it is important to remember that cortical bone loss is irreversible; observe individuals with gastrointestinal disorders regularly to detect vitamin D deficiency. Prevention is best achieved by being aware of the risk and measuring the serum 25(OH)D annually. Treatment of hypovitaminosis D in these situations can be with vitamin D, although larger doses will be needed, or with 25(OH)D or calcitriol. The aim is to keep serum 25(OH)D in the high normal range. In treating established disease, give calcium supplements as well as magnesium if hypomagnesaemia is present.

Defective metabolism

Chronic renal failure

Chronic renal failure leads to reduced production of $1,25(OH)_2D$ and $24,25(OH)_2D$ from 25(OH)D by the kidney. Since a major effect of $1,25(OH)_2D$ on the intestine is to facilitate calcium absorption, in renal failure the intestinal calcium absorption falls, and, in conjunction with skeletal resistance to the calcium-mobilizing action of PTH, leads to hypocalcaemia. The inability of the diseased kidney to excrete phosphate leads to hyperphosphataemia, which aggravates the hypocalcaemia either by decreasing the renal 1a-hydroxylase activity or by impairing the release of calcium from the skeleton. Hypocalcaemia is a large stimulus to the parathyroid glands and leads to parathyroid hyperplasia and secondary hyperparathyroidism. The net result is a complex effect on bone leading to various skeletal disorders described under the term 'renal osteodystrophy'. Patients may have osteitis fibrosa, due to high levels of PTH, osteomalacia, due in part to vitamin D deficiency, or a combination of both. Retention of aluminium, and its deposition at the calcification front, may also play an important part in the development of osteomalacia, which may be refractory to treatment with $1,25(OH)_2D$ (Fig. 13).

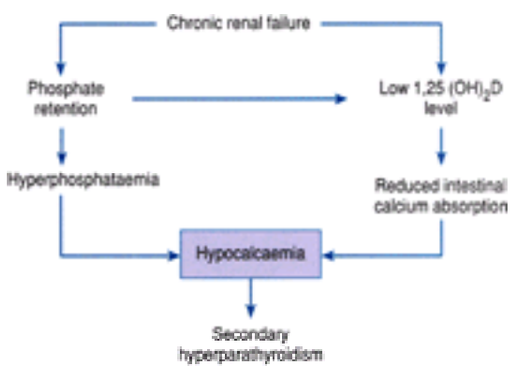


Fig. 13 Pathogenesis of renal bone disease.

Renal osteodystrophy is best managed and prevented by reducing the dietary intake of phosphate, avoiding aluminium-containing phosphate binders, increasing the dietary intake of calcium except when there is marked hyperphosphataemia (because of the risk of extraskeletal calcification), and by supplementation with either $1,25(OH)_2D$ or $1a\text{-}OHD_2$ (alfacalcidol). These treatments usually improve mineral homeostasis, improve clinical symptoms of bone disease, and suppress parathyroid secretion. Occasionally, however, secondary hyperparathyroidism is so severe, with hypercalcaemia, extraskeletal calcification, fractures, and bone pain, that parathyroid surgery may be necessary.

Anticonvulsants

Chronic anticonvulsant therapy induces hepatic microsomal mixed-function oxidases that degrade steroid hormones such as oestrogen and cortisol. Since the hypothalamic–pituitary axis is intact, feedback regulation helps to maintain normal circulating levels of these hormones. Similarly, the chronic use of anticonvulsants leads to increased metabolic conversion of vitamin D and 25(OH)D to biologically inactive metabolites, which are excreted in the urine and bile. Unlike other steroid hormones, feedback regulation is limited and dependent on dietary intake of vitamin D and sunlight exposure. If dietary intake is defective, vitamin D deficiency may result. Anticonvulsants, however, may also affect calcium homeostasis by inhibiting intestinal calcium transport and the bone-resorbing effects of PTH and $1,25(OH)_2D$, and by suppressing osteoblastic and osteoclastic activity. There are significant differences between the various anticonvulsant drugs. Phenobarbitone, a more potent inducer of hepatic microsomal enzymes than phenytoin, does not lead to a low 25(OH)D because increased catabolism is matched by increased formation. Phenytoin does not directly affect vitamin D metabolism but causes bone disease more often than phenobarbitone because it impairs intestinal calcium absorption and reduces the release of calcium from bone, leading to hypocalcaemia and secondary hyperparathyroidism. The net effect is that vitamin D requirements are increased by chronic use of anticonvulsants, especially in those on multiple drugs and in institutional care. Such individuals should be given additional vitamin D, but it is unclear whether all patients taking anticonvulsants should receive prophylactic vitamin treatment.

Vitamin D-dependent rickets type I

This is a rare, autosomal-recessive disease, in which a low or undetectable circulating $1,25\text{-}(OH)_2D$ is the result of defective renal 25(OH)D 1a-hydroxylase activity, and which manifests itself before the age of 2 years with classical features of rickets. Affected children fail to respond to conventional doses of replacement therapy with vitamin D but do so to large doses of vitamin D and 25(OH)D, or to physiological doses of 1,25(OH)D or 1a(OH)D.

Receptor defects

Vitamin D-dependent rickets type II

This is a rare disorder with rickets and/or osteomalacia but with no biochemical evidence of vitamin D deficiency, hypocalcaemia, secondary hyperparathyroidism, high levels of $1,25(OH)_2D$, and failure to respond to large doses of vitamin D or its metabolites. About 70 per cent of the cases reported have alopecia (often complete) and absent eyelashes yet the number and morphology of scalp hair follicles is normal. Alopecia is worse in those resistant to treatment with 1,25(OH)D. Affected patients come from consanguineous marriages and the inheritance is most likely to be autosomal-recessive. The disorder is due to defects in the intracellular interaction of 1,25(OH)D with its receptor. Defects include reduced receptor affinity, abnormal numbers of receptors, and impaired ability of the receptor–hormone complex to bind the DNA. In patients with normal hair a clinical remission can be induced with high doses of vitamin D analogues. A 10-fold higher dose is needed in those with alopecia, but even then only one-half will respond.

Altered phosphate homeostasis

Phosphate depletion

Osteomalacia from chronic phosphate depletion is uncommon and is seen in individuals who consume large amounts of phosphate-binding antacids over a number of years (Dent and Winter 1974). The osteomalacia is clinically, radiologically, and histologically similar to that seen in vitamin D deficiency. Biochemically, however, there are differences in that the serum calcium is normal or even increased, serum phosphate may be low or normal, but urinary phosphate excretion is reduced while urinary calcium excretion may be increased. The concentration of calcitriol is increased and may be responsible for increased absorption of calcium. Treatment consists of phosphate supplementation and avoidance of antacids.

Renal phosphate loss

Hypophosphataemic rickets/osteomalacia may result from hereditary or acquired disorders.

X-linked hypophosphataemic rickets

Also known as familial hypophosphataemic rickets, X-linked hypophosphataemic rickets was first described by Albright *et al*. (1937), who noted that in some patients rickets did not improve with normal doses of vitamin D but did when pharmacological doses of vitamin D were used. The disorder was called vitamin D-resistant rickets and later became known as hypophosphataemic (vitamin D-resistant) rickets when it was noted that it occurred in association with a low serum phosphate. The inheritance is usually X-linked dominant, although autosomal forms also occur. The mutant gene in the disease is located in the distal part of the short arm of the X chromosome (Thakker and O'Riordan 1988).

The disease is fully expressed as rickets and short stature in homozygous males, but the expression of the disease in girls is variable, from none to severe rickets and stunted growth (Fig. 14). A low serum phosphate is found soon after birth, but the disease is only clinically manifest soon after the child begins to walk. Bow legs appear first and there is failure of normal growth. There may be a waddling gait. Skull deformities develop from premature closure of the sutures. Unlike vitamin D-deficiency rickets, in X-linked hypophosphataemic rickets there is no proximal myopathy or dental hypoplasia, but tooth eruption is delayed and tooth development poor, leading to dental abscesses and early dental caries.

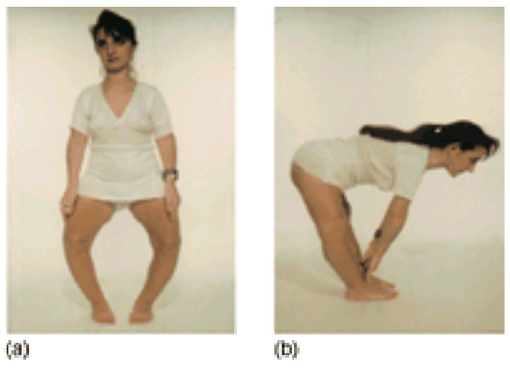


Fig. 14 Twenty-one-year-old girl with X-linked hypophosphataemic rickets. (a) Bowing of the femur and tibia due to previous rickets; (b) lumbar flexion is minimal due to spondylitic changes in the spine.

Radiologically, bone mass is not diminished and may even be above the age-matched normal since there is no secondary hyperparathyroidism. In adults there may be reduced joint mobility, especially at the elbows, shoulders, hip, and spine, due to a widespread enthesopathy (Davies and Stanbury 1981; Polission *et al*. 1985). Spinal enthesopathy may occasionally lead to cord compression. The rigid spine, held straight and with loss of lumbar lordosis, may be confused with ankylosing spondylitis. Indeed, radiographs of the spine may show calcification of the outer fibres of the annulus fibrosus, giving a 'squared-off' appearance to the vertebral bodies, and later complete bridging between vertebrae may lead to a 'bamboo spine'. Ligamentous calcification may obscure the sacroiliac joints, often only in the

upper part (Fig. 15). In the long bones, the cortex appears thickened and the trabeculae are coarse and thickened. Pseudofractures may occur. Degenerative changes develop in weight-bearing joints, probably secondary to the limb deformity.

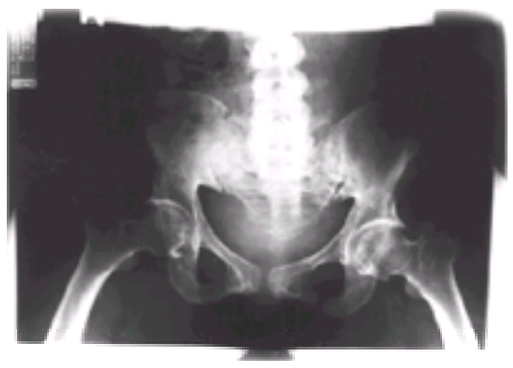


Fig. 15 X-linked hypophosphataemic rickets. Ligamentous calcification of the upper part of the sacroiliac joints and around the hip joints. There is also bridging of the vertebral interspaces.

Biochemical tests show a low serum phosphate, a normal calcium and PTH, and normal or low levels of $1,25(OH)_2D$. Hypophosphataemia is due to a renal tubular defect, with increased renal loss of phosphate due to lowered renal threshold for phosphate. Renal function is otherwise normal and there are no acidification defects, glycosuria, or aminoaciduria.

It is important to make a correct diagnosis of this disorder as early as possible since proper treatment of the infant will prevent skeletal deformities and allow for normal growth. Although hypophosphataemia is central to the disorder, treatment with phosphate alone leads to secondary hyperparathyroidism and with vitamin D alone to the risk of hypercalcaemia and renal failure. A combination of $1,25(OH)_2D$ and phosphate is most effective, resulting in a return to normal of serum phosphate and healing of rickets. Hypercalciuria should be avoided; serum phosphate, calcium, and urinary calcium should be monitored regularly (Petersen *et al*. 1992).

Metabolic acidosis

A chronic metabolic acidosis may result from Fanconi syndrome or renal tubular acidosis, and may be associated with rickets and osteomalacia. Fanconi syndrome embraces a number of acquired and hereditary disorders in which there are varying combinations of defects of the renal proximal tubule. The proximal tubular abnormality leads to glycosuria, aminoaciduria, hyperphosphaturia and hypophosphataemia, and a systemic acidosis from failure of bicarbonate reabsorption. The hereditary forms are seen most commonly in cystinosis, Wilson's disease, and Lowe syndrome, while the acquired forms are seen in association with multiple myeloma, light-chain nephropathy, amyloidosis, Sjögren syndrome, heavy metal poisoning, and malignancy. The hereditary form can present with rickets, while rickets or osteomalacia may occur in acquired forms. The mechanism for the rickets/osteomalacia is not fully known. The failure of mineralization is a consequence of hypophosphataemia, hypocalcaemia, and metabolic acidosis. In some forms, e.g. Wilson's disease, hepatic damage may lead to low levels of 25(OH)D, while in other forms levels of $1,25(OH)_2D$ are reduced due to the proximal renal tubular defect. Treatment of the bone disease is with phosphate supplements and vitamin D metabolites, especially the 1a(OH)D analogue.

Renal tubular acidosis describes disorders in which there is renal loss of bicarbonate resulting in reduced plasma bicarbonate and a systemic acidosis. The pH of the urine is inappropriately high compared to the metabolic acidosis. Renal tubular acidosis exists as two major forms. In the distal tubular form, known as type I, there is failure to secrete protons so that a large pH gradient between blood and the tubular fluid cannot form. In the proximal tubular type (type II) there is bicarbonate wasting but the ability of the distal tubule to excrete protons is intact. Renal tubular acidosis type II is also often associated with Fanconi syndrome and other proximal tubular defects. The development of rickets and osteomalacia in both forms is due to hypophosphataemia and perhaps failure to synthesize $1,25(OH)_2D$, a consequence of the systemic acidosis. The treatment consists of correcting the acidosis and using vitamin D if the rickets fails to heal.

Oncogenous hypophosphataemic osteomalacia

Some tumours (usually mesenchymal) lead to rickets or osteomalacia characterized by hypophosphataemia with phosphaturia. The bone disease regresses after removal of the tumour. The biochemical defect is characterized by a low maximum tubular reabsorption of phosphate per volume of glomerular filtrate, suggesting renal phosphate wasting, and, in some patients, by low or inappropriately low levels of $1,25(OH)_2D$. While the biochemical abnormalities are similar to those seen in X-linked hypophosphataemic rickets, there are more severe skeletal deformities. Clinical features are predominantly those of bone pain, proximal muscle weakness, sometimes fractures, and, in children, growth retardation (Ryan and Reiss 1984). A similar type of hypophosphataemic osteomalacia may occur in association with fibrous dysplasia and neurofibromatosis. The mechanism responsible for the development of hypophosphataemia is unknown, but the tumour may produce a humoral factor that affects proximal tubular function and leads to phosphate wasting and suppressed production of $1,25(OH)_2D$.

The tumour may be small and difficult to locate, but if found and completely removed the bone disorder reverses. Clinical features may return if there is a recurrence of the tumour. If it is not possible to resect the tumour completely, treatment is aimed at correcting the hypophosphataemic osteomalacia by phosphate and $1,25(OH)_2D_3$ supplements.

Defective mineralization

Aluminium

As discussed above, in chronic renal failure aluminium is deposited in bone and impairs mineralization.

Fibrogenesis imperfecta ossium

This is a rare disorder that affects middle-aged men, who develop generalized bone pain and fractures. Approximately 50 per cent of reported cases have an associated monoclonal gammopathy. Radiologically the bones appear dense with loss of normal trabecular pattern; histologically bone shows an increased amount of thick osteoid, reduced mineralization, and decreased birefringence due to the lack of normal lamellar arrangement of collagen. It has been suggested that the disorder is due to an acquired defect of collagen that impairs normal mineralization in lamellar bone, but collagen synthesis elsewhere is normal. Some clinical but not histological improvement may occur with vitamin D. Melphalan reportedly induced a clinical remission, with some improvement in the histological appearances of bone, in a patient with coexistent monoclonal gammopathy (Stamp *et al*. 1985).

Hypophosphatasia

This is a rare, heritable cause of severe rickets in childhood and recurrent stress fractures in adults, transmitted in an autosomal-recessive manner with an incidence of 1 in 100 000 live births for the severe form. There is a reduction in serum and tissue alkaline phosphatase in association with increased urinary excretion of phosphoethanolamine and increased circulating inorganic pyrophosphate and pyridoxal-5'-phosphate. Infantile forms present before the age of 6 months with growth retardation, rachitic deformities, hypercalcaemia, hypercalciuria, and renal failure. Childhood hypophosphatasia may present with premature loss of deciduous teeth alone or in association with rickets. It may improve spontaneously. Adult hypophosphatasia presents with recurrent stress fractures that heal poorly. Chondrocalcinosis is common and attacks of pyrophosphate arthropathy may occur. The cause for the mineralization defect is unknown but it has been suggested that the accumulation of pyrophosphate, due to failure of alkaline phosphatase to cleave it, leads to deficient skeletal mineralization (Whyte 1989). There is no treatment available for this disorder. Since there is no hypocalcaemia, vitamin D and its metabolites should be avoided; their use may lead to hypercalcaemia and hypercalciuria (Whyte 1989).

Chapter References

Adams, J. S., Sharma, O. P., Gacad, M. A., and Singer, F. R. (1983). Metabolism of 25-hydroxyvitamin D_3 by cultured pulmonary alveolar macrophages in sarcoidosis. *Journal of Clinical Investigation*, **72**, 1856–60.

Adinoff, A. D. and Hollister, J. R. (1983). Steroid-induced fractures and bone loss in patients with asthma. *New England Journal of Medicine*, **309**, 265–68.

Advisory Group on Osteoporosis. (1994). *Report*. pp. 5–34. Department of Health, London.

Aitken, M. (1987). Relationship between mortality and femoral neck fracture J and osteoporosis. In *Osteoporosis 1987*, Vol. 1, (eds. C. Christiansen, J. S. Johnsen, and B. J. Riis), pp. 45–8. Norhaven Press, Viborg, Denmark.

Albright, F., Butler, A. M., and Bloomberg, E. (1937). Rickets resistant to vitamin D therapy. *American Journal of Diseases of Children*, **54**, 529–47.

Albright, F., Bloomberg, F., and Smith, P. H. (1976). Postmenopausal osteoporosis. *Transactions of the Association of American Physicians*, **55**, 298–305.

Amento, E. P. *et al*. (1984). 1a, 25-dihydroxyvitamin D_3 induces the maturation of the human monocyte cell line U937, and, in association with a factor from human T lymphocytes, augments production of the monokine, mononuclear cell factor. *Journal of Clinical Investigation*, **73**, 731–9.

Arend, M. P. and Dayer, J. M. (1990). Cytokines and cytokine inhibitors or antagonists in rheumatoid arthritis. *Arthritis and Rheumatism*, **33**, 305–15.

Bar-Shavit, Z. *et al*. (1983). Induction of monocyte differentiation and bone resorption by 1,25-dihydroxyvitamin D_3. *Proceedings of the National Academy of Sciences (USA)*, **80**, 5907–11.

Bell, N. M., Greene, A., Epstein, S., Oexmann, M. J., Shaw, S., and Shary, J. (1985). Evidence for alteration of the vitamin D–endocrine system in blacks. *Journal of Clinical Investigation*, **76**, 470–3.

Benger, V., Johnell, O., and Redlund-Johnell, I. (1988). Changes in incidence and prevalence of vertebral fractures during 30 years. *Calcified Tissue International*, **42**, 293–6.

Bergkvist, L., Adami, H. O., Persson, I., Hoover, R., and Schaiver, C. (1989). The risk of breast cancer after estrogen and estrogen–progestin replacement. *New England Journal of Medicine*, **321**, 293–7.

Bhalla, A. K. (1986). Musculoskeletal manifestations of primary hyperparathyroidism. *Clinics in Rheumatic Diseases*, **12**, 691–705.

Bhalla, A. K., Amento, E. P., Clemens, T. L., Holick, M. F., and Krane, S. M. (1983). Specific high-affinity receptors for 1,25-dihydroxyvitamin D_3 in human peripheral blood mononuclear cells: presence in monocytes and induction in T lymphocytes following activation. *Journal of Clinical Endocrinology and Metabolism*, **57**, 1308–10.

Bhalla, A. K., Amento, E. P., and Krane, S. M. (1986). Differential effects of 1,25-dihydroxyvitamin D_3 on human lymphocytes and monocyte/macrophages: inhibition of interleukin 2 and augmentation of interleukin 1 production. *Cellular Immunology*, **98**, 311–22.

Bhalla, A. K., Williams, M. M., Lal, S., and Lydyard, P. M. (1989). 1,25 Dihydroxyvitamin D_3 but not retinoic acid, induces the differentiation of U937 cells. *Clinical and Experimental Immunology*, **76**, 274–7.

Black, D. M. *et al*. for the Fracture Intervention Trial Research Group (1996). Randomized trial of effect of alendronate on the risk of fracture in women with existing vertebral fractures. *Lancet*, **348**, 1535–41.

Boyce, W. J. and Vessey, M. P. (1985). Rising incidence of fractures of the proximal femur. *Lancet*, **i**, 150–1.

Braidman, I. P., Davenport, L. K., Carter, D. H., Selpy, P. L., Mawer, E. B., and Freemont, A. J. (1995), Preliminary *in situ* identification of oestrogen target cells in bone. *Journal of Bone and Mineral Research*, **10**, 74–80.

Brinton, L. A., Hoover, R., and Fraumeni, J. F. J. (1986). Menopausal oestrogens and breast cancer risk: an expanded case-control study. *British Journal of Cancer*, **54**, 825–32.

Bush, T. L. *et al*. (1983). Estrogen uses and all-cause mortality. Preliminary results from the Lipid Research Clinics Program Follow-up Study. *Journal of the American Medical Association*, **249**, 903–6.

Butler, R. C., Pauce, M. W. J., Worsfield, M., and Sharp, C. A. (1991). Bone mineral content in patients with rheumatoid arthritis: relationship to low dose steroid therapy. *British Journal of Rheumatology*, **30**, 86–90.

Chapuy, M. C. *et al*. (1992). Vitamin D_3 and calcium to prevent hip fractures in elderly women. *New England Journal of Medicine*, **327**, 1637–42.

Cohn, S. H. (1991). Noninvasive measurement of bone mass. In *Metabolic bone disease and clinically related disorders* (eds. L. V. Avioli and S. M. Krane), pp. 264–82. Saunders, Philadelphia.

Colditz, G. A. *et al*. (1995). The use of oestrogen and progestins and the risk of breast cancer in post-menopausal women. *New England Journal of Medicine*, **332**, 1589–93.

Consensus Development Conference (1991). Prophylaxis and treatment of osteoporosis. *American Journal of Medicine*, **90**, 107–10.

Cooper, C. and Wickham, C. (1990). Rheumatoid arthritis, corticosteroid therapy and hip fracture. In *Osteoporosis 1990* (eds. C. Christiansen and K. Overgaard), pp. 1578–9. Handelstrykkeriet, Aalborg, Denmark.

Cooper, C., Barker, D. J., and Wickham, C. (1988). Physical activity, muscle strength, and calcium intake in fracture of the proximal femur in Britain. *British Medical Journal*, **297**, 1443–6.

Cooper, C. *et al*. (1991). Screening for vertebral osteoporosis using individual risk factors. *Osteoporosis International*, **2**, 48–53.

Cooper, C., Champion, G., and Melton, L. J. (1992). Hip fractures in the elderly: a worldwide projection. *Osteoporosis International*, **2**, 285–9.

Cooper, C., Carbone, L., Michet, C. J., Atkinson, E. J., O'Fallow, W. M., and Melton, L. J., III (1994). Fracture risk in patients with ankylosing spondylitis. A population based study. *Journal of Rheumatology*, **21**, 1877–84.

Cooper, C., Copland, M., and Mitchell, M. (1995). Rheumatoid arthritis, corticosteroid therapy in hip fractures. *Annals of the Rheumatic Diseases*, **54**, 49–52.

Cummings, R. G. (1990). Calcium intake and bone mass: a quantitative review of the evidence. *Calcified Tissue International*, **47**, 194–201.

Cummings, S. R., Tao, J. L., and Browner, W. S. (1992). Hyperthyroidism increases the risk of hip fractures: a perspective study. *Journal of Bone and Mineral Research*, **7** (Suppl. 1), 1–21.

Cummings, S. R. *et al*.(1995). Risk factors for hip fracture in white women. *New England Journal of Medicine*, **332**, 767–73.

Daly, E., Vessey, M. P., Hawkins, M. M., Carson, J. L., Gough, P., and Marsh, S. (1996). Risk of venous embolism in users of HRT. *Lancet*, **348**, 977–80.

Dauphine, R. T., Riggs, B. L., and Scholz, D. A. (1975). Back pain and skeletal crush fractures: an unemphasized mode of presentation for primary hyperparathyroidism. *Annals of Internal Medicine*, **83**, 365–7.

Davies, M. and Stanbury, S. W. (1981). The rheumatic manifestations of metabolic bone disease. *Clinical Rheumatic Diseases*, **5**, 596–646.

Davies, K. M., Recker, R. R., Stegman, M. R., Heaney, R. P., Kimmel, D. B., and Leist, J. (1990). Third decade bone gain in women. In *Calcium regulation and bone metabolism* (ed. D. V. Cohn, F. H. Glorieux, and T. J. Martin). Elsevier Science, Amsterdam.

Dent, C. E. and Winter, C. S. (1974). Osteomalacia due to phosphate depletion from excessive aluminium hydroxide ingestion. *British Medical Journal*, **1**, 551–2.

Dequeker, J. and Geusens, P. (1990). Osteoporosis and arthritis. *Annals of the Rheumatic Diseases*, **49**, 276–80.

Devogelar, J. P., Haux, J. P., and Nagart de Deuxchaisnes, C. (1984). Does mild, asymptomatic, primary hyperparathyroidism require surgery to avoid bone loss in postmenopausal females? In *Osteoporosis*(eds. C. Christiansen *et al*.), p. 365. Aallorg Stiftsbocytykken, Copenhagen.

Devogelar, J. P., Maldague, R., Malghem, J., and de Deuxchaisnes, C. N. (1992) Appendicular and vertebral bone mass in patients with ankylosing spondylitis. *Arthritis and Rheumatism*, **35**, 1062–7.

Drinkwater, B. L., Nilson, K., Chestnut, C. H., Bremner, W. J., Shainholtz, S., and Southwood, M. B. (1984). Bone mineral content of amenorrheic and eumenorrheic athletes. *New England Journal of Medicine*, **311**, 277–81.

Drinkwater, B. L., Bremner, B., and Chesnut, C. H., III (1990). Menstrual history as a determinant of current bone density in young athletes. *Journal of the American Medical Association*, **263**, 545–48.

Eastgate, J. A., Symons, J. A., Woods, N. R., Grinlinton, F. M., di Givoine, F., and Duff, G. W. (1988). Correlation of plasma interleukin 1 levels with disease activity in rheumatoid arthritis. *Lancet*, **ii**, 706–9.

Eisman, J. A. (1995). Vitamin D receptor gene alleles and osteoporosis: an affirmative view. *Journal of Bone and Mineral Research*, **10**, 1289–93.

Eisman, J. A., Martin, T. J., McIntyre, I., and Mosley, J. M. (1979). 1,25-dihydroxyvitamin D receptors in breast cancer cells. *Lancet*, **ii**, 1335–6.

Eriksen, E. F. *et al*. (1988). Evidence of estrogen receptors in normal human osteoblast cells. *Science*, **241**, 84–6.

Fallon, M. D., Whyte, M. P., Craig, R. B., Jr, and Teitelbaum, S. L. (1983). Mast-cell proliferation in postmenopausal osteoporosis. *Calcified Tissue International*, **35**, 29–31.

Faulkner, K. G., Cummings, S. R., Black, D., Palermo, L., Glüer, C. C., and Gerant, H. K. (1993). Simple measurement of femoral geometry predicts hip fracture: the study of osteoporotic fracture. *Journal of Bone and Mineral Research*, **8**, 1211–17.

Ferrari, S., Rizzoli, R., Chevalley, D., Slosman, D., Eisman, J. A., and Bonjour, J-P. (1995). Vitamin D receptor gene polymorphism and change in lumbar spine mineral density. *Lancet*, **345**, 423–4.

Gallagher, J. C. and Goldgar, D. (1990). Treatment of postmenopausal osteoporosis with high doses of synthetic calcitriol: a randomized controlled trial. *Annals of Internal Medicine*, **113**, 649–55.

Garnero, P., Borel, O., Sornay-Rendu, E., and Delman, P. D. (1995). *Journal of Bone and Mineral Research*, **10**, 1283–8.

Garrett, I. R. *et al*. (1987). Production of lymphotoxin, a bone resorbing cytokine by cultured human myeloma cells. *New England Journal of Medicine*, **317**, 526–32.

Gilsanz, V., Roe, T. F., Mora, S., Costin, G., and Goodman, A. G. (1991). Changes in vertebral bone density in black girls and white girls during childhood and puberty. *New England Journal of Medicine*, **325**, 1597–600.

Girasole, G., Passeri, G., Jilka, R. L., and Manolagas, S. C. (1994). Interleukin-11: a new cytokine critical for osteoclast development. *Journal of Clinical Investigation*, **93**, 1516–24.

Goldman, L. and Tosteson, A. N. A. (1991). Uncertainty about postmenopausal estrogen. Time for action not debate. Editorial. *New England Journal of Medicine*, **325**, 800–2.

Gowen, M. (1991). Cytokines regular bone cell function. *Rheumatology Review*, **1**, 43–50.

Grimley Evans, J. (1990). The significance of osteoporosis. In *Osteoporosis 1990* (ed. R. Smith), pp. 1–8. Royal College of Physicians, London.

Grodstein, F. *et al*. (1996*a*). Prospective study of exogenous hormones and risk of pulmonary embolism in women. *Lancet*, **348**, 983–7.

Grodstein, F. *et al*. (1996*b*). Postmenopausal estrogen and progestin use and the risk of cardiovascular disease. *New England Journal of Medicine*, **335**, 453–61.

Hanson, C. A., Shagrin, J. W., and Duncan, H. (1971). Vertebral osteoporosis in ankylosing spondylitis. *Clinical Orthopedics*, **74**, 59–64.

Heaney, R. P., Recker, R. R., and Saville, P. D. (1978). Menopausal changes in calcium balance performance. *Journal of Laboratory and Clinical Medicine*, **92**, 953–63.

Heikinheimo, R. J. *et al*. (1992). Annual injection of vitamin D and fractures of aged bones. *Calcified Tissue International*, **51**, 105–10.

Hooyman, J. R., Melton, J. R., Nelson, A. M., O'Fallon, N. M., and Riggs, B. L. (1984). Fractures of rheumatoid arthritis. A population based study. *Arthritis and Rheumatism*, **27**, 1353–61.

Hunter, T. and Dubo, M. I. C. (1978). Spinal fracture complicating ankylosing spondylitis. *Annals of Internal Medicine*, **88**, 546–9.

Hunter, T. and Dubo, M. I. C. (1983). Spinal fracture complicating ankylosing spondylitis: a long term study. *Arthritis and Rheumatism*, **26**, 751–9.

Hustmyer, F. G., Peacock, M., Hui, S., Johnston, C. C., and Christian, J. (1994). Bone mineral density in relation to polymorphism at the vitamin D receptor gene locus. *Journal of Clinical Investigation*, **94**, 2130–4.

Jackson, J. A. and Kleerekoper, M. D. (1990). Osteoporosis in men: diagnosis, pathophysiology and prevention. *Medicine*, **69**, 137–52.

Jackson, J. A., Kleerekoper, M., Parfitt, A. M., Rao, D. S., Villanoeva, A. R., and Frame, B. (1987). Bone histomorphometry in hypogonadal and eugonadal men with spinal osteoporosis. *Journal of Clinical Endocrinology and Metabolism*, **65**, 53–8.

Jick, H., Derby, L. E., Myers, M. W., Vasilakis, C., and Newton, K. M. (1996). Risk of hospital admission for idiopathic venous thromboembolism among users of postmenopausal oestrogen. *Lancet*, **348**, 981–3.

Jilka, R. L. *et al*. (1992). Increased osteoclast development after oestrogen loss: mediation by interleukin-6. *Science*, **257**, 88–91.

Johnell, O. *et al*. (1995). Risk factors for hip fracture in European women: the Medos Study. *Journal of Bone and Mineral Research*, **10**, 1802–15.

Johnston, C. C., Slemenda, C. W., and Melton, J. L. (1991). Clinical use of bone densitometry. *New England Journal of Medicine*, **324**, 1105–9.

Johnston, C. C., Jr *et al*. (1992). Calcium supplementation and increases in bone mineral density in children. *New England Journal of Medicine*, **327**, 82–7.

Kalla, A. A., Meyers, O.L., Chalton, D., *et al*. (1991). Increased metacarpal bone mass following 18 months of slow acting anti-rheumatic drugs for rheumatoid arthritis. *British Journal of Rheumatology*, **30**, 91–100.

Kamali, R., Bhalla, A.K., Farrow, S.M., *et al*. (1989). Early regulation of c-*myc* mRNA by 1,25-dihydroxyvitamin D_3 in human myelomonocytic U937 cells. *Journal of Molecular Endocrinology*, **3**, 43–8.

Kanis, J. A. (1991). Requirement of calcium for optimal skeletal health in women. *Calcified Tissue International*, **49** (Suppl.), 533–41.

Kanis, J. A. and Passmore, R. (1989). Calcium supplementation of the diet. *British Medical Journal*, **296**, 137–40; 205–8.

Kanis, J. A. *et al*. (1990). Osteoporosis: causes and therapeutic implications. In *Osteoporosis 1990* (ed. R. Smith), pp. 46–56. Royal College of Physicians, London.

Kanis, J. A., Melton J, Christiansen, C., Johnston, C.C., and Khaltaev, N. (1994 *a*). The diagnosis of osteoporosis. *Journal of Bone and Mineral Research*, **9**, 1137–41.

Kanis, J. A. *et al*. (1994*b*). *Assessment of fracture risk and its application to screening for postmenopausal osteoporosis: report of a WHO study group* (WHO Technical Report Series 843). WHO, Geneva

Kleerekoper, M. (1992). Vertebral fracture or vertebral deformity? *Calcified Tissue International*, **50**, 50–6.

Kleerekoper, M. (1996). Fluoride: the verdict is in, but the controversy lingers. *Journal of Bone and Mineral Research*, **11**, 565–7.

Kleerekoper, M. and Nelson, D.A. (1992). A randomized trial of sodium fluoride as a treatment for postmenopausal osteoporosis. *Osteoporosis International*, **1**, 155–61.

Krolner, B. and Nielsen, S. P. (1982). Bone mineral content of the lumbar spine in normal and osteoporotic women: cross-sectional and longitudinal studies. *Clinical Science*, **62**, 329–36.

Laan, R. E. J. M., van Riel, P. L. C. M., and van Erning, L. J. T. O. (1992). Vertebral osteoporosis in rheumatoid arthritis patients: effects of low dose prednisolone therapy. *British Journal of Rheumatology*, **31**, 91–6.

Laan, R. E. J. M. *et al*. (1993*a*). Differential effects of glucocorticoids on cortical, appendicular and cortical vertebral bone mineral content. *Calcified Tissue International*, **52**, 5–9.

Laan, R. E. J. M. *et al*. (1993*b*). Low dose prednisolone induces rapid reversible axial bone loss in patients with rheumatoid arthritis. *Annals of Internal Medicine*, **119**, 963–8.

Lanyi, E., Ratko, I., and Gomor, B. (1993). On the diagnostic value of DEXA for osteoporosis in case of patients with ankylosing spondylitis. In *Osteoporosis* (Proceedings 4th International Symposium on Osteoporosis and Consensus Development Conference), pp. 192–4. Bath Institute for Rheumatic Diseases, Bath, UK.

Liberman, V. A. *et al*. (1993). Effects of oral alendronate on bone mineral density and the incidence of fractures in postmenopausal osteoporosis. *New England Journal of Medicine*, **333**, 1437–8.

Lindsay, R. and Tomme, J. F. (1990). Estrogen treatment of patients with established postmenopausal osteoporosis. *Obstetrics and Gynecology*, **76**, 290–5.

Lindsay, R., Hart, D. M., Aitken, J. M., MacDonald, E. B., Anderson, J. B., and Clark, A. C. (1976). Long term prevention of postmenopausal osteoporosis by oestrogen. *Lancet*, **i**, 1038–41.

Lindsay, R., Hart, D. M., Abdalla, H., and Al-Azzawi, F. (1987). Inter-relationship of bone loss and its prevention, and fracture expression. In *Osteoporosis 1987* (eds. C. Christiansen, J. S. Johansen, and B. J. Riis), pp. 508–9. Osteoporosis, Copenhagen.

Lips, P., Graafman, W. C., Ooms, M. E., Bezemer, D., and Bouter, L. M. (1996). Vitamin D supplementation and fracture incidence in elderly persons. A randomised, placebo-controlled clinical trial. *Annals of Internal Medicine*, **124**, 400–6.

Luengo, M., Picado, C., Rio, L. D., Guanbons, N., Monserrat, J. M., and Setoain, J. (1991). Vertebral fractures in steroid dependent asthma and involutional osteoporosis. *Thorax*, **46**, 803–6.

McSheety, P. M. J. and Chambers, T. J. (1986). Osteoblastic cells mediate osteoclastic responsiveness to parathyroid hormones. *Endocrinology*, **118**, 824–8.

Malluche, H. H. and Faugere, M-C. (1991). Bone biopsies: histology and histomorphometry of bone. In *Metabolic bone disease and clinically related disorders* (eds. L. V. Avioli and S. M. Krane), pp. 283–328. Saunders, Philadelphia.

Mamelle, N. *et al*. (1988). Risk–benefit ratio of sodium fluoride treatment in primary vertebral osteoporosis. *Lancet*, **ii**, 361–5.

Manolagas, S. C. and Jilka, R. L. (1995). Bone marrow, cytokines, and bone remodelling. Emerging insights into the pathophysiology of osteoporosis. *New England Journal of Medicine*, **332**, 305–11.

Martin, P., Bergmann, P., Gillet, C., *et al*. (1986). Partial reversible osteopenia after surgery for primary hyperparathyroidism. *Archives of Internal Medicine*, **146**, 689–91.

Matkovic, V., Kostial, K., Simonovic, I., Buzina, R., Brodarec, A., and Nordin, B. E. C. (1979). Bone status and fracture rates in two regions of Yugoslavia. *American Journal of Clinical Nutrition*, **32**, 540–9.

Matta, W. M., Shaw, R. W., Hesp, R., and Katz, D. (1981). Hypogonadism induced by luteinising hormone releasing hormone against analogues: effects on bone density in premenopausal women. *British Medical Journal*, **294**, 1523–4.

Melton, L. J. (1987). Postmenopausal bone loss and osteoporosis. In *The climacteric and beyond* (eds. L. Zochella, M. Whitehead, and P. A. van Keep), pp. 125–35. Parthenon, Carnforth.

Melton, L. J., Chrischilles, E. A., Cooper, C., Lane, A. W., and Riggs, B. L. (1992). How many women have osteoporosis. *Journal of Bone and Mineral Research*, **7**, 1005–10.

Meyer, H. E., Tverdal, A., and Falch, J. A. (1993). Risk factors for hip factors for hip fracture in middle-aged Norwegian women and men. *American Journal of Epidemiology*, **137**, 1203–11.

Michel, B. A, Bloch, D. A., and Fries, J. F. (1991). Predictors of fracture in early rheumatoid arthritis. *Journal of Rheumatology*, **18**, 804–8.

Morrison, N. A., Yeoman, R., Kelley, P. J., and Eisman, J. A. (1992). Contribution to trans-acting factor alleles to normal physiological variability: vitamin D receptor gene polymorphism and circulating osteocalcin. *Proceedings of the National Academy of Sciences (USA)*, **89**, 6665–9.

Morrison, N. A. *et al*. (1994). Prediction of bone density from vitamin D receptor alleles. *Nature*, **367**, 284–7.

Mullaji, A. B., Upadhay, S. S., and Ho, E. K. W. (1994). Bone mineral density in ankylosing spondylitis. *Journal of Bone and Joint Surgery*(B), **76**, 660–5.

Naessen, T., Persson, I., Adami, H. O., Bergstrom, R., and Bergkvist, L. (1990). Hormone replacement therapy and the risk for hip fracture. A prospective, population-based cohort study. *Annals of Internal Medicine*, **113**, 95–103.

Nilas, L., Christiansen, C., and Rodbro, P. (1984). Calcium supplementation and postmenopausal bone loss. *British Medical Journal*, **7**, 1103–6.

Nilsson, B. E. and Westlin, N. E. (1971). Bone density in athletes. *Clinical Orthopaedics*, **77**, 177–82.

Obront, K. J., Bengner, U., Johnell, O., Nilsson, B. E., and Sernbo, I. (1989). Increasing age-adjusted risk of fragility fractures: a sign of increasing osteoporosis in successive generations? *Calcified Tissue International*, **44**, 157–67.

O'Neill, T. W., Felsenberg, D., Varlow, J., Cooper, C., and Kanis, J. A. (1996). The prevalence of vertebral deformity in European men and women: The European Vertebral Osteoporosis Study. *Journal of Bone and Mineral Research*, **11**, 1010–17.

Ooms, M. E., Roos, J. C., Bezemer, D., van der Vijgh, W. J., Bouter, L. M., and Lips, P. (1995). Prevention of bone loss by vitamin D supplementation in elderly women: a randomised double-blind trial. *Journal of Clinical Endocrinology and Metabolism*, **80**, 1052–8.

Ott, S. M. and Chesnut, C. H. (1989). Calcitriol treatment is not effective in postmenopausal osteoporosis. *Annals of Internal Medicine*, **110**, 267–74.

Pacifici, R. *et al*. (1987). Spontaneous release of interleukin 1 from human blood monocytes reflects bone formation in idiopathic osteoporosis. *Proceedings of the National Academy of Sciences (USA)*, **84**, 4616–20.

Pacifici, R. *et al*. (1989). Ovarian steroid treatment blocks a postmenopausal increase in blood monocyte interleukin-1 release. *Proceedings of the National Academy of Sciences (USA)*, **86**, 2398–402.

Pacifici, R. *et al*. (1991). Effect of surgical menopause and estrogen replacement on cytokine release from human mononuclear cells. *Proceedings of the National Academy of Sciences (USA)*, **88**, 5134–8.

Paganini-Hill, A., Ross, R. K., and Henderson, B. E. (1989). Endometrial cancer and pattern of use of oestrogen replacement therapy: a cohort study. *British Journal of Cancer*, **59**, 445–7.

Pak, C. Y. C., Sakmaee, K., Adams-Huct, B., Pziak, V., Peterson, R. O., and Poindexter, J. R. (1995). Treatment of postmenopausal osteoporosis with slow release sodium fluoride. *Annals of Internal Medicine*, **112**, 401–8.

Parfitt, A.M. (1994). Vitamin D receptor genotypes in osteoporosis. *Lancet*, **334**, 1580.

Peacock, M. (1995). Vitamin D receptor gene alleles and osteoporosis: a contrasting view. *Journal of Bone and Mineral Research*, **10**, 1294–7.

Perry, H. M., Fallon, M. D., Bergfeld, M., Teitelbaun, S. L., and Avioli, L. V. (1982). Osteoporosis in young men. A syndrome of hypercalciuria and accelerated bone turnover. *Archives of Internal Medicine*, **142**, 1295–8.

Persson, I. *et al*. (1989). Risk of endometrial cancer after treatment with oestrogen alone or in conjunction with prostagens: results of a prospective study. *British Medical Journal*, **298**, 147–51.

Petersen, D. J., Boniface, A. M., Schranck, F. W., Rupich, R. C., and White, M. P. (1992). X-linked hypophosphatemic rickets: a study (with literature review) of linear growth response to calcitriol and phosphonate therapy. *Journal of Bone and Mineral Research*, **7**, 583–97.

Pocock, N. A., Eisman, J. A., Yeates, M. G., Sambrook, P. N., and Eberl, S. (1986). Physical fitness is a major determinant of femoral neck and lumbar spine bone mineral density. *Journal of Clinical Investigation*, **78**, 618–21.

Pocock, N. A., Eisman, J. A., Hooper, J. L., Yeates, M. G., Sambrook, P. N., and Eberl, S. (1987). Genetic determinants of bone mass in adults. *Journal of Clinical Investigation*, **80**, 706–10.

Polission, R. P., Martinez, S., and Khoury, M. (1985). Calcification of entheses associated with X-linked hypophosphataemic osteomalacia. *New England Journal of Medicine*, **313**, 1–6.

Quigley, M. E., Martin, P. L., Burnier, A. M., and Brooks, P. (1987). Estrogen therapy arrests bone loss in elderly women. *American Journal of Obstetrics and Gynecology*, *156*, 1516–23.

Ralston, S. M., Urquhart, G. D., Brezeski, M., and Sturrock, R. D. (1990). Prevalence of vertebral compression fractures due to osteoporosis in ankylosing spondylitis. *British Medical Journal*, **285**,

563–5.

Rao, D. S. (1990). Metabolic bone disease in gastrointestinal and biliary disorders. In *Primer on the metabolic bone diseases and disorders of mineral metabolism* (ed. M. J. Favus). American Society for Bone and Mineral Research, California.

Reeve, J., Davies, U. M., Hesp, R., McNally, E., and Katz, D. (1990). Treatment of osteoporosis with human parathyroid peptide and observations on effect of sodium fluoride. *British Medical Journal*, **301**, 314–18.

Reginster, J. Y. *et al*. (1987). 1-year controlled randomized trial of prevention of early postmenopausal bone loss by intranasal calcitonin. *Lancet*, **ii**, 1481–3.

Reid, D. M. (1990). Corticosteroid-induced osteoporosis. In *Osteoporosis 1990* (ed. R. Smith), pp. 99–117. Royal College of Physicians, London.

Reid, D. M. *et al*. (1981). Bone loss in rheumatoid arthritis and primary generalized osteoarthritis. Effects of corticosteroids, suppressive antirheumatic drugs and calcium supplements. *British Journal of Rheumatology*, **25**, 253–9.

Reid, D.M., Nicoll, J.J., Kennedy, N.S., Smith, M.A., Tothill, P., and Nuki, G. (1986). Bone mass in ankylosing spondylitis. *Journal of Rheumatology*, **13**, 932–5.

Reid, I. R., Ames, R. W., Evans, M. C., Gamble, G. D., and Sharpe, S. J. (1993). Effect of calcium supplementation on bone loss in postmenopausal women. *New England Journal of Medicine*, **328**, 460–4.

Reid, I. R., Ames, R. W., Evans, M. C., Gamble, G. D., and Sharpe, S. J. (1995). Long term effects of calcium supplementation on bone loss and fractures in postmenopausal women: a randomised controlled trial. *American Journal of Medicine*, **98**, 331–5.

Reitsma, P. H. *et al*. (1983). Regulation of *myc* gene expression in HL-60 leukaemic cells by a vitamin D metabolite. *Nature*, **306**, 492–4.

Rencken, M. L., Chesnut, C. H., III, and Drinkwater, B. L. (1996). Bone density at multiple skeletal sites in amenorreic athletes. *Journal of the American Medical Association*, **276**, 238–40.

Riggs, B. L. and Melton, I. J. (1986). Involutional osteoporosis. *New England Journal of Medicine*, **314**, 1676–86.

Riggs, B. L. *et al*. (1990). Effect of fluoride treatment on the fracture rate in postmenopausal women with osteoporosis. *New England Journal of Medicine*, **322**, 802–9.

Riis, B., Thomsen, K., and Christiansen, C. (1987). Does calcium supplementation prevent postmenopausal bone loss? *New England Journal of Medicine*, **316**, 173–7.

Rosenspire, K. C., Kennedy, A. C., Steinbeck, J., Blau, M., and Green, F. A. (1980). Investigation of the metabolic activity of bone in rheumatoid arthritis. *Journal of Rheumatology*, **7**, 469–73.

Ross, E. J. and Linch, D. C. (1982). Cushing's syndrome-killing disease: discriminatory value of signs and symptoms aiding early diagnosis. *Lancet*, **ii**, 646–9.

Royal College of Physicians (1989). *Fractured neck of femur: prevention and management*. Royal College of Physicians, London.

Rubin, C. T. and Lanyon, L. E. (1984). Regulation of bone formation by applied dynamic levels. *Journal of Bone and Joint Surgery*, **66**, 397–402.

Ryan, E. A. and Reiss, E. (1984). Oncogenous osteomalacia: review of the world literature of 42 cases and report of two new cases. *American Journal of Medicine*, **77**, 501–12.

Sambrook, P. N., Eisman, J. A., Yeates, M. G., Pocock, N. A., Eberl, S., and Champion, G. D. (1986). Osteoporosis in rheumatoid arthritis, safety of low dose corticosteroids. *Annals of the Rheumatic Diseases*, **45**, 950–3.

Sambrook, P. N., Eisman, J. A., Champion, G. D., and Pocock, N. A. (1988). Sex hormone status and osteoporosis in postmenopausal women with rheumatoid arthritis. *Arthritis and Rheumatism*, **31**, 973–8.

Schindler, A. E., Ebert, A., and Friedrich, E. (1972). Conversion of androstanedione to estrone by human fat tissue. *Journal of Clinical Endocrinology and Metabolism*, **35**, 627–30.

Slemenda, C., Hui, S. L., and Johnston, C. C., Jr (1990 *a*). Patterns of bone loss and physiologic growing: prospects for prevention of osteoporosis by attainment of greater peak bone mass. In *Osteoporosis 1990* (eds. C. Christiansen and K. Overgaard), pp. 948–58. Handelstrykkenel, Aalborg, Denmark.

Slemenda, C. W., Hui, S. L., Longcope, D., Wellman, H., and Johnston, C. C., Jr (1990 *b*). Predictors of bone mass in perimenopausal women: a prospective study of clinical data using photon absorptiometry. *Annals of Internal Medicine*, **112**, 96–101.

Slemenda, C. W., Miller, J. Z., Hui, S. L., Reister, T. K., and Johnston, C. C. (1991). Role of physical activity in the development of skeletal mass in children. *Journal of Bone and Mineral Research*, **11**, 1227–33.

Smith, E. L., Smith, P. E., Ensign, C. J., and Shea, M. M. (1984). Bone involution decreases in exercising middle aged women. *Calcified Tissue International*, **36**, 129–38.

Spector, T. D., Hall, G. M., McCloskey, E. V., and Kanis, J. A. (1993). Risk of vertebral fractures in women with rheumatoid arthritis. *British Medical Journal*, **1**, 558.

Spector, T. D. *et al*. (1995). Influence of vitamin D receptor genotype on bone mineral density in postmenopausal women: a twin study in Britain. *British Medical Journal*, **310**, 1357–60.

Stamp, T. C. B., Byers, P. D., Ali, Y., and Jenkins, M. V. (1985). Fibrogenesis imperfecta ossium: remission with melphan. *Lancet*, **i**, 582–3.

Stampfer, M. J. *et al*. (1991). Postmenopausal estrogen therapy and cardiovascular disease. Ten-year follow-up from the Nurses' Health Study. *New England Journal of Medicine*, **325**, 756–62.

Steinberg, K. K. *et al*. (1991). A meta-analysis of the effect of estrogen replacement therapy and the risk of breast cancer. *Journal of the American Medical Association*, **264**, 2648–53.

Stevenson, J. C. (1988). Osteoporosis. Pathogenesis and risk factors. In *Bailliere's clinical endocrinology and metabolism*, Vol. 2, (ed. T. J. Martin), pp. 87–101. Saunders, London.

Storm, T., Thamsborg, G., Steiniche, T., Gerant, H. K., and Sorenson, O. H. (1990). Effect of intermittent cyclical etidronate therapy on bone mass and fracture rate in women with postmenopausal osteoporosis. *New England Journal of Medicine*, **322**, 1265–71.

Storm, T., Kollerup G, Thamsborg, G., Gerant, H. K., and Sorenson, O. H. (1996). Five years of clinical experience with intermittent cyclical etidronate for postmenopausal osteoporosis. *Journal of Rheumatology*, **23**, 1560–4.

Svensson, O., Stromberg, L., Ohlen, G., and Lindgren, U. (1996). Prediction of the outcome after hip fracture in elderly patients. *Journal of Bone and Joint Surgery*(B), **78**, 115–18.

Thakker, R. V. and O'Riordan, J. L. H. (1988). Inherited forms of rickets and osteomalacia. In *Bailliere's clinical endocrinology and metabolism*, Vol. 2, (ed. T. J. Martin), pp. 157–92. Saunders, London.

Tilyard, M. W., Spears, G. F. S., Thompson, J., and Dovey, S. (1992). Treatment of postmenopausal osteoporosis with calcitriol or calcium. *New England Journal of Medicine*, **326**, 357–62.

Verstraeten, A. and Dequecker, J. (1990). Vertebral and peripheral bone mineral content and fracture incidence in postmenopausal patients with rheumatoid arthritis: effect of low dose corticosteroids. *Annals of the Rheumatic Diseases*, **45**, 852–7.

Watts, N. B. *et al*. (1990). Intermittent cyclical etidronate in postmenopausal osteoporosis. *New England Journal of Medicine*, **323**, 73–9.

Whitehead, M. I. and Fraser, D. (1987). Controversies concerning the safety of estrogen replacement therapy. *American Journal of Obstetrics and Gynecology*, **156**, 1315–22.

Whyte, M. P. (1989). Hypophosphatasia. In *The metabolic basis of inherited disease* (6th edn) (eds. C. R. Scriver, A. L. Beaudet, W. S. Sly, and D. Valle), pp. 2843–56. McGraw-Hill, New York.

Wilkinson, M. and Bywaters, E. (1958). Clinical features and course of ankylosing spondylitis. *Annals of the Rheumatic Diseases*, **17**, 209–28.

Will, R., Palmer, R., Bhalla, A. K., Ring, F., and Calin, A. (1989). Osteoporosis in early ankylosing spondylitis: a primary pathological event? *Lancet*, **ii**, 1483–5.

Will, R., Palmer, R., Elvins, D., Ring, F., and Bhalla, A. K. (1990). A lower femoral neck BMD occurs in patients with ankylosing spondylitis (AS) compared with same six siblings. In *Osteoporosis 1990* (eds. C. Christiansen and K. Overgaard), pp. 1672–4. Handelstrykkenel, Aalborg, Denmark.

Wilson, R. J., Rao, D. S., Ellis, B., Kleerekoper, M., and Parfitt, M. (1988). Mild asymptomatic primary hyperparathyroidism is not a risk factor for vertebral fractures. *Annals of Internal Medicine*, **109**, 959–62.

Winner, S. J., Morgan, C. A., and Grimley Evans, J. (1989). Perimenopausal risk of falling and the incidence of distal forearm fractures. *British Medical Journal*, **298**, 1486–8.

5.17.2 Paget's disease of bone

Adrian J. Crisp

Introduction and epidemiology

On 14 November 1876 Sir James Paget gave the first comprehensive account of 'osteitis deformans' to the Medical and Chirurgical Society (Paget 1877). Paget's disease is a chronic disorder characterized by accelerated resorption and production of bone, resulting in deformity and fragility, and, in Paget's phrase, by 'change of size, shape and direction'. A radiological survey in Britain of hospital patients of 55 years and over in 31 towns concluded that 4 to 5 per cent had Paget's disease, with a focus in Lancashire of 7 to 8 per cent (Barker 1984). The male:female ratio is 3:2. About 10 to 30 per cent of patients have a single pagetic lesion (monostotic disease) and the rest polyostotic disease. There have been many reports of the disease affecting more than one family member and of its recurrence in successive generations. The disease is most common in Britain, North America, Australia, and New Zealand but rare in Eire, Scandinavia, Asia, and in non-white races (Barker 1984). This remarkable geographical distribution compels speculation about a genetically determined disorder disseminated by the migration of populations. Paget's disease is not only a cause of bone and joint pain but also of chronic disability; in one study, 42 per cent of middle-aged patients were forced into premature retirement from work by the disease (Harinck *et al*. 1986).

Pathophysiology

Histopathology

The striking histological feature of the disease is the increased number and size of both osteoclasts and osteoclast nuclei, with active bone resorption. There is a compensatory increase in the number and size of osteoblasts, leading to exuberant new bone formation. This considerable increase in bone turnover disrupts the normal lamellar structure of the matrix collagen, with increased intraosseous fibrosis, hypervascularity, and enlarged haversian canals. This bone consists of a variable mixture of immature ('woven') bone and irregular mature lamellar bone (without haversian systems) separated by deeply staining cement lines in the characteristic mosaic pattern of the disease (Fig. 1).

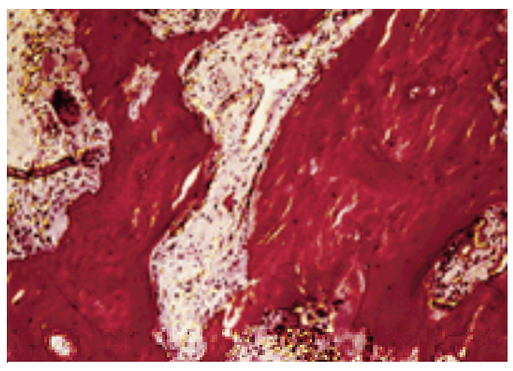


Fig. 1 A section of pagetic bone showing the anarchic architecture and the exaggerated resorption lacunae occupied by multinucleated osteoclasts.

Aetiology

The weight of evidence argues that Paget's disease is caused by a slow-virus infection of bone cells. Polyclonal antibodies reveal paramyxovirus antigens in pagetic osteoclasts compatible with measles and respiratory syncytial virus infections. Studies with monoclonal antibodies implicate measles, simian virus, and human parainfluenza. *In situ* hybridization with DNA probes specific for measles nucleocapsid protein detects measles sequences in osteoclasts, but probes also hybridize with osteoblasts, fibroblasts, and lymphocytes (Basle *et al*. 1987).

Impressive evidence has also come from the demonstration of virus-like particles in osteoclast nuclei (Fig. 2) with a similarity to the measles virus inclusions observed in the brains of patients with subacute sclerosing panencephalitis, a known outcome of measles (Basle *et al*. 1987).

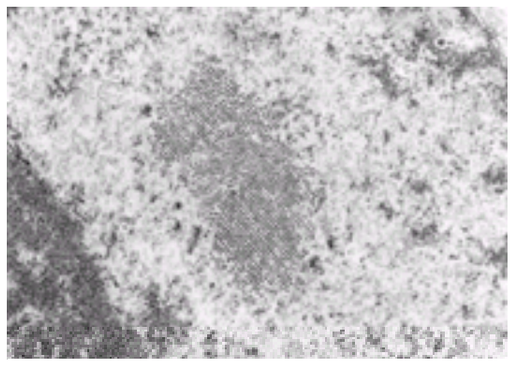


Fig. 2 An electron micrograph of bone tissue from a patient with Paget's disease showing viral microcylindrical elements. (Reproduced with the kind permission of Professor Andre Rebel of Angers, France and of *The Lancet*.)

There is some evidence, based on history of exposure of patients with Paget's disease to dogs in early life, implicating canine distemper, a closely related paramyxovirus (O'Driscoll and Anderson 1985). The same Manchester group has also reported the presence of RNA from canine distemper virus in pagetic bone (Gordon *et al*. 1991). The pattern of disease spread also points to infection, perhaps in early life. Paget's disease is characterized by the concurrent appearance of lesions in one or several bones rather than by the sequential recruitment or spread of new bone involvement in each patient.

However, not all studies have found paramyxovirus particles or RNA in pagetic bone (Ralston *et al*. 1991). There has been no convincing recovery of virus from pagetic cells or pagetic osteosarcoma so far, and it has not been possible to passage the disease to uninfected cell cultures and animals. In addition the antiviral agent, inosiplex, which has successfully treated subacute sclerosing panencephalitis, has been ineffective in Paget's disease in preliminary studies.

Kahn (1990) has suggested that the target cell for viral infection may not be the osteoclast but the osteoblast, which can also contain viral RNA. The structure of the osteoblast is also more deranged in Paget's disease than in any other high-turnover states such as renal hyperparathyroidism. Common progenitor cells give rise to osteoclasts, granulocytes, and monocytes and distribute their progeny widely throughout the circulation; one might then expect that new bone involvement would be 'metastatic' in each patient, which is incorrect. Osteoblasts, which are modified fibroblasts, proliferate locally, and this would be more compatible with the pattern of bone involvement in Paget's disease. We know very little of the molecular basis of Paget's disease. Growth factors are likely to influence the anarchic processes in pagetic bone, which are characterized by bone destruction without respect for the existing architecture and with generous formation of new vessels. Non-transforming growth factors, as well as transforming growth factors secreted by some osteosarcoma cells, are convincing candidates for control of the cellular processes in Paget's disease, which has often been considered a benign tumour of the bone (Krane 1986).

Clinical features (Table 1, Table 2)

	Percentage
Pelvis	72–76
Lumbar spine	33–58
Thoracic spine	24–45
Femur	25–55
Sacrum	29–44
Skull	26–42
Tibia	22–35
Radius	16
Feet	10
Hands	9
Ribs	3

Table 1 Distribution of Paget's disease in the skeleton

Pain
Bone expansion and deformity
Fractures
Heat
Neurological syndromes:
Deafness
Tinnitus
Headache
Spinal cord and root compression
Brainstem/cerebellar compression
Blindness
Other cranial nerve involvement
Malignant osteosarcoma and benign giant-cell tumour
Immobilization hypercalcaemia and hypercalciuria
High-output cardiac failure
Hyperuricaemia and gout
Angioid streaks of retina

Table 2 Clinical features of Paget's disease

Many patients with Paget's disease are unaware of it. Infection may occur in early life and require many years before bone lesions are detectable. The diagnosis is often made incidentally during radiological or biochemical investigations of other systems. The pagetic vertebra seen on a plain abdominal radiograph or a raised serum alkaline phosphatase are very common presenting features. Although long-term prospective studies of outcome in asymptomatic patients with Paget's disease have not been made, perhaps only a minority of patients will eventually develop clinical symptoms or signs clearly attributable to the disease.

Pain

Bone pain

Typically the bone pain of Paget's disease is constant, deep, and boring, sometimes worse at night and at rest. The cause of the pain is not well understood but is likely to be related to increased internal and periosteal blood flow in metabolically active bone increasing intraosseous pressure and stimulating bone pain fibres, which may lie in canaliculi. Long before there is any biochemical evidence of disease control there is often a correlation between pain relief and reduction of bone blood flow, as indicated by thermography, when rapidly acting drugs such as mithramycin (plicamycin) are used (Crisp *et al*. 1989). Compression fractures of verebrae also occur in weakened pagetic bone.

Adjacent joint pain

Bone deformity commonly alters force transmission through the adjacent joint causing premature loss of articular cartilage and secondary osteoarthritis, especially at hip and knee (Fig. 3). Patients can also develop osteoarthritis in an unaffected knee as a result of favouring the contralateral, non-pagetic knee. Metabolically active, hyperaemic, subchondral bone may also be toxic to articular cartilage. Osteoarthritic pain is usually associated with exercise and relieved by rest and sleep. In practice it is often difficult to dissect pagetic bone pain from osteoarthritic pain. Protrusio acetabuli occurs in about one-quarter of hip joints involved by Paget's

disease and sacroiliac joint obliteration occurs rarely (Guyer 1980).

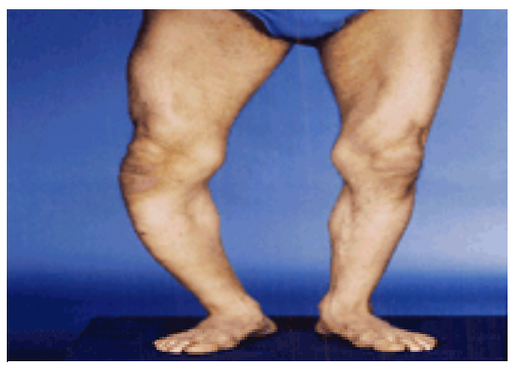


Fig. 3 A patient with Paget's disease of the right femur and tibia causing external rotation of the right leg and secondary osteoarthritis of the knee. A psoriatic patch is also noted on the right knee.

Bone expansion and deformity

Bowing long-bone deformity causes effective inequality of leg length, leading to lumbar scoliosis and accelerated degenerative lumbar spondylosis as well as secondary osteoarthritis. Adaptation to slow, insidious deformity may cause remarkably few symptoms.

Fractures

Pain may arise from incomplete fissure fractures involving only cortex on usually the convex side of a bowed long bone. The pain may settle with rest and medical treatment, but the fissure fracture can extend into a completed fracture. Mild injuries may cause pathological fractures in weakened pagetic bone. In a large series of 180 patients, 15 per cent suffered fractures (Harinck *et al*. 1986). Avulsion fractures, for example, of the patella, can occur after sudden muscular contraction. Non-union may be more common at sites of fracture in pagetic bone.

Heat

Palpation of affected bones near to skin surfaces is often helpful, although bone tenderness is unusual. Auscultation of tibia or skull can sometimes reveal bruits. Temperature may be correlated with both metabolic activity of bone and with bone pain.

Neurological syndromes

Common symptoms of cranial Paget's disease includes non-specific headaches, possibly due to vascular changes in bone and meninges, impaired hearing, and tinnitus. About half the patients with skull involvement will develop deafness, which may be sensineural, conductive, or mixed. Nerve entrapment by the expanded petrous temporal bone is partly responsible, but subtle toxic effects on the inner ear mediated by hypervascularity are also likely. Less commonly the optic nerve and other cranial nerves may be affected with predictable results.

Skull expansion may cause frontal or occipital prominence but if the base of the skull is involved, softened bone may lead to platybasia—the descent of the cranium on to the cervical spine—which may cause dizziness and syncope, owing to kinking of the vertebrobasilar blood vessels, or even cerebellar or brainstem compressive syndromes.

Dysfunction of the spinal cord may rarely result from vertebral expansion and compression but both skull-base and cord dysfunction may follow ill-understood, 'steal' phenomena: increased blood flow in bone 'steals' blood away from neural structures causing ischaemic damage. Hence incipient paraparesis may improve rapidly with intensive medical therapy reducing bone blood flow.

Neoplastic complications

Almost all cases of osteogenic sarcoma in adults arise from pagetic bone but conversely osteosarcoma is a very rare complication of Paget's disease occurring in less than 0.1 per cent of affected patients. The sarcoma is highly malignant and usually rapidly fatal. This diagnosis is worth considering in a patient who develops intense pain in an affected bone, with a progressive lytic lesion and a rising alkaline phosphatase against a background of stable symptoms (Fig. 4). There may be genetic factors: three brothers, who all developed Paget's disease in early life, later died from pagetic osteosarcomas (Brenton *et al*. 1980).

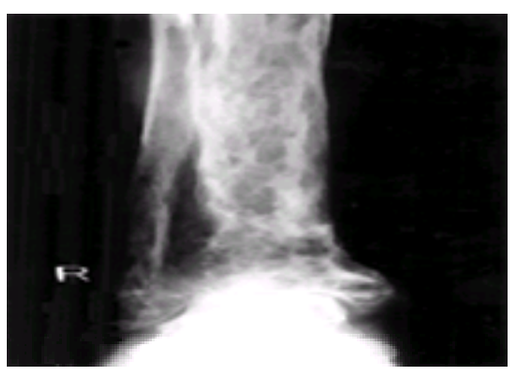


Fig. 4 A radiograph of the right lower leg of a patient with Paget's disease. Just above the ankle joint the markedly lytic region contains an osteosarcoma.

Over 100 cases of the benign giant-cell tumour arising from pagetic bone have been reported. These commonly involve facial bones and mandible, and some cases have been linked by a common family background in Avellino, a small town in Italy (Jacobs *et al*. 1979; Bhambhani *et al*. 1992).

Immobilization, hypercalcaemia, and hypercalciuria

The serum calcium is usually normal but during immobilization for an unrelated illness or orthopaedic surgery, hypercalcaemia and hypercalciuria can occur. Bone formation is inhibited but resorption continues unchecked. Rarely both abnormalities can occur without immobilization, but this can often be traced to coincidental hyperparathyroidism. Hypercalciuria alone is more common, occurring in 21 out of 180 patients, nine of whom described a history of renal calculi (Harinck *et al*. 1986).

Hyperuricaemia

Increased bone-cell turnover might be expected to increase urate synthesis: 13 out of 101 (12.9 per cent) patients without an independent cause had raised urate values. Only three patients suffered from clinical gout (Harinck *et al*. 1986).

High-output cardiac failure

Cardiac index (cardiac output corrected for surface area) is increased in the majority of pagetic patients, especially those with widespread disease. High-output cardiac failure has only rarely been reported when more than one-third of the skeleton is involved.

Other clinical associations

Angioid streaks in the retina caused by disruption of Bruch's membrane may be associated but are also more firmly linked with another condition of abnormal connective tissue, pseudoxanthoma elasticum. Associations with Hashimoto's thyroiditis, Dupuytren's contracture, and chondrocalcinosis have all been proposed but none is firmly based.

Investigations and assessment

Biochemical

Serum alkaline phosphatase

This osteoblastic enzyme, a measure of new bone formation and not resorption, is the most useful and widely available marker and may be elevated by as much as 30 times above the upper limit of the reference range. Analysis of isoenzymes or measurement of alternative liver enzymes help to exclude a significant hepatic contribution to the total alkaline phosphatase. Occasionally, this enzyme is normal with limited Paget's disease but one should not be deterred from treating a painful pagetic lesion by a normal result.

Urine hydroxyproline

Increased osteoclastic bone resorption parallels increased bone formation and may be assessed by assay of urine hydroxyproline excretion, a breakdown product of collagen. Only 20 to 30 per cent of the total hydroxyproline is derived from bone resorption. Ideally the patient should have followed a gelatin (denatured collagen)-free diet for 48 h before the urine collection to exclude a dietary contribution but this is often impractical. The urinary hydroxyproline/creatinine ratio may be determined on a 2-h collection after an overnight fast or on a 24-h specimen.

Urinary pyridinoline collagen cross-links

The recent development of an assay to measure the excretion of bone-specific collagen cross-links offers the prospect of accurate measurement of current osteclastic bone resorption. This is likely to replace measurement of urinary hydroxproline when more widely available (Beardsworth *et al*. 1990).

Serum osteocalcin

Osteocalcin, which is specifically produced by osteoblasts, might in theory be of benefit in the assessment of Paget's disease. However, only one-half of patients with raised alkaline phosphatase have raised osteocalcin and values can be increased by treatment via interactions between parathyroid hormone and 1,25-dihydroxyvitamin D_3. Thus, osteocalcin is unlikely to be a valuable marker of disease activity in Paget's disease.

Serum total acid phosphatase

This osteoclastic enzyme can be raised in active Paget's disease but is of little value. Occasionally metastatic prostatic carcinoma is incorrectly diagnosed in the presence of sclerotic vertebrae and a raised concentration. The two conditions can coexist.

Serum a_2HS glycoprotein

This liver-derived protein is taken up by active bone and levels tend to be lower than normal in active Paget's disease.

Radiological

Plain radiographs

There is a wide variety of appearances in Paget's disease but the key features are lytic areas, especially early in the disease, sclerotic areas, and bone expansion with coarse and disorganized trabecular structures (Fig. 5 and Fig. 6).

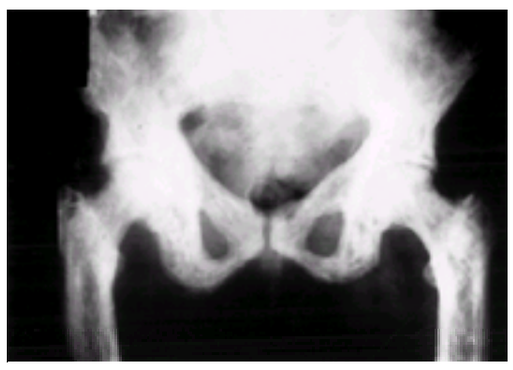


Fig. 5 A radiograph of the pelvis showing typical widespread Paget's disease.

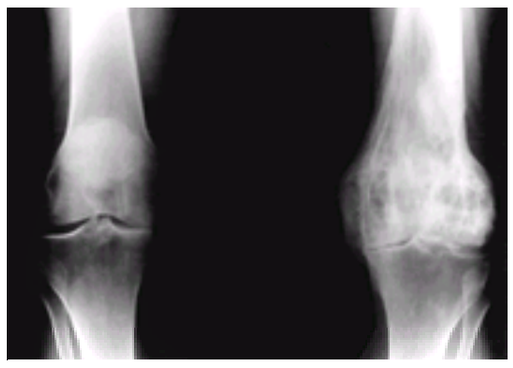


Fig. 6 A radiograph of both knees showing typical pagetic involvement of the left distal femur and worse osteoarthritis of the affected knee when compared with the knee uninvolved by Paget's disease.

Isotope bone scanning

In one study (Meunier *et al*. 1987), scanning demonstrated 8.3 per cent more pagetic sites than did plain radiography and is therefore the most sensitive investigation for defining the extent of lesions (Fig. 7). A scan and then plain radiographs of affected sites should be performed at first diagnosis for future comparisons. The percentage retention of the isotope after 24 h—less than 40 per cent in normal subjects in our laboratory—provides an index of total pagetic bone activity. During aggressive advance of the resorption front, scanning may underestimate disease activity, suggesting active osteoclastic activity with a poor osteoblastic response—a situation comparable with many lytic lesions in multiple myeloma. Quantitative bone scintigraphy comparing the area of increased uptake with the same area on the contralateral (unaffected) side can be very useful in the assessment of the monostotic lesion when alkaline phosphatase is often normal.

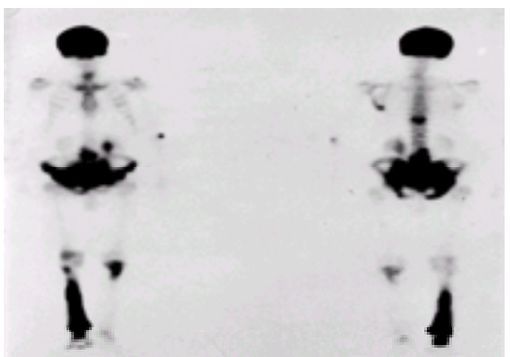


Fig. 7 A scintigram showing a typical pagetic distribution of lesions in skull, vertebrae, pelvis, and both legs.

Thermography

Increased bone and periosseous blood flow is demonstrated effectively in superficial bones and is reduced after effective treatment (Fig. 8). Thermographic improvement and pain reduction seem to be linked (Crisp *et al*. 1989).

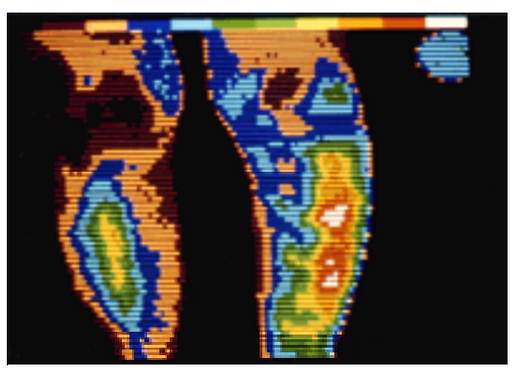


Fig. 8 A thermogram of lower legs in a patient with Paget's disease showing a focus of increased temperature (white and red areas) at the site of active Paget's disease (by kind permission of Dr Brian Hazleman).

Differential diagnosis of Paget's disease

The combination of physical signs, including a warm deformed bone with characteristic radiographs and a raised serum alkaline phosphatase, usually leaves no doubt about the diagnosis of Paget's disease. Consider also:

1. Other causes of raised serum bone alkaline phosphatase:
   a. metastatic bone disease;
   b. osteomalacia;
   c. hyperparathyroidism, with osteitis fibrosa cystica: this may accompany Paget's disease and may be more than coincidental;
   d. idiopathic hyperphosphatasia: characterized by bone deformity in childhood and a probably recessive inheritance—its relationship with adult Paget's disease is not known.
2. Other causes with similar radiographic appearances:
   a. metastatic bone disease; osteolytic and osteoblastic disease, e.g. prostatic secondaries;
   b. fibrous dysplasia;
   c. chronic osteomyelitis;
   d. metaphyseal dysplasia (Engelmann's disease);
   e. sternocostoclavicular hyperostosis.

Treatment

Attitudes towards treating Paget's disease vary considerably from the therapeutic nihilist, who attributes most of the pain to secondary osteoarthritis and considers specific treatment as an expensive superfluity to the tyro who chooses to treat all patients. The criteria of the present author are listed in Table 3. With the advent of safe effective drugs for Paget's disease there is a firm trend towards early, aggressive treatment with the objective of maintaining serum alkaline phosphatase within the normal range. Although remission after treatment is related to the degree of biochemical control (Harinck *et al*. 1987; Kanis and Gray 1987), and there is strong histological evidence that effective suppression promotes more normal bone remodelling, we lack evidence that intensive treatment inhibits the development of long-term complications.

Pain arising from a site of known Paget's disease
Early, potentially deforming disease
Osteolytic lesions especially in weight-bearing bones
Skull disease
Complications:
Progressive neurological syndromes
Fissure fractures (avoid etidronate)
Immobilization hypercalcaemia
High-output cardiac failure
Disease in patients aged under 55 years
Serum alkaline phosphatase and/or urine hydroxyproline concentration more than twice upper limit of normal
Patients likely to undergo joint replacement at involved sites within 6 months

Table 3 Indications for treatment of Paget's disease

Pagetic and related osteoarthritic pain may be reduced by simple analgesics and non-steroidal anti-inflammatory drugs but pure pagetic bone pain often responds poorly to these. Physical treatment to improve muscle function and correct any inequality in leg length with shoe raises may be very helpful. More effective control will be achieved by drugs that suppress bone turnover. All primarily inhibit osteoclast activity, as reflected by an early decrease in urine hydroxyproline excretion followed by a later fall in serum alkaline phosphatase. Intravenous mithramycin (plicamycin), 15 µg/kg per day for 7 to 10 days is highly effective, giving rapid pain relief within 3 days and biochemical improvement, but it has an unacceptable toxicity to bone marrow, liver, and kidneys and is on the verge of obsolescence since the availability of later generation bisphosphonates. It may still be of value for the urgent treatment of acute neurological deterioration, most commonly spinal cord compression, when the rapid control of bone blood flow and perhaps periosteal oedema can lead to impressive improvement.

In most patients the choice of treatment now falls between the bisphosphonates (previously called diphosphonates) and the calcitonins.

The bisphosphonates (Fig. 9)

This family of pyrophosphate analogues share a common P–C–P backbone replacing the P–O–P of pyrophosphate. Side-chain variations confer differing properties but they are all effective for pagetic pain and disease activity. Oral etidronate and more recently oral tiludronate and intravenous pamidronate are licensed for use in Paget's disease in the United Kingdom. There has been additional experience with oral and intravenous clodronate, which is at present licensed only for use in malignant hypercalcaemia.

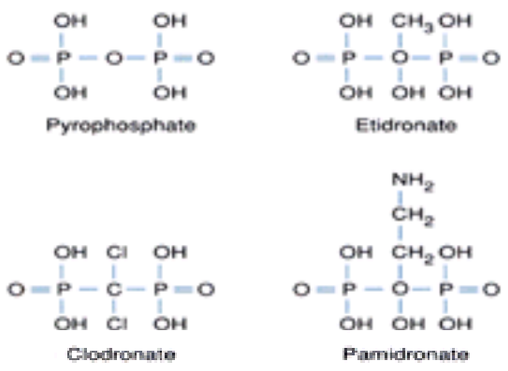


Fig. 9 The structures of pyrophosphate and bisphosphonates in common clinical use.

Calcitonins only partially suppress the disease while treatment continues. Bisphosphonates can achieve prolonged remissions following a course of treatment and are therefore now first choice.

Etidronate disodium

The recommended regimen for oral etidronate is 5 mg/kg per day for up to 6 months, although a month's course of 20 mg/kg per day is probably as effective, without added toxicity (Preston *et al*. 1986; Heath 1987). To achieve maximal gut absorption all oral bisphosphonates should be taken at least 2 h after food. Mild gastrointestinal symptoms and transient increased bone pain are occasional side-effects. Etidronate even at low doses can cause a reversible mineralization defect, which may be minimized by vitamin D supplements, and may increase the risk of fracture. This bisphosphonate is contraindicated for lytic lesions in weight-bearing long bones (Krane 1982). Courses of etidronate should always be followed by a minimum treatment-free period of 3 months to permit recovery of the mineralization defect.

Newer bisphosphonates

The newer bisphosphonates are not completely free of the risk of causing a mineralization defect but the safe therapeutic window is much wider. Clodronate is about 10 times less potent than pamidronate and there has been far less clinical experience with it (Yates *et al*. 1985). Oral tiludronate 400 mg daily for 3 months has been recently licensed for the treatment of Paget's disease in the United Kingdom. Very poor absorption and a high incidence of mild gastrointestinal side-effects with pamidronate effectively limits its use to the intravenous route but absolute compliance with a precise dose is achieved.

The optimal dosage regimen for pamidronate in Paget's disease remains to be determined. Anderson *et al*. (1994) have induced a full biochemical remission (serum alkaline phosphatase falling to normal) in 90 per cent of patients in whom this enzyme was initially raised, with one or more complex courses. Such courses typically induce a remission of about 2 years and Anderson *et al*. claim that a permanent remission can be achieved in 10 to 15 per cent of patients. If the alkaline phosphatase level is below 500 i.u./l they advocate a 30 mg infusion over 2 h followed by three infusions of 60 mg each over 4 h. If the alkaline phosphatase level exceeds 500 i.u./l, six 60 mg infusions are recommended administered at 2-week intervals.

Other groups have argued that a single infusion of a dose large enough to saturate all of the binding sites of bone affected by Paget's disease should achieve as much as complex recurrent regimens (Fig. 10). Watts *et al*. (1993) studied the effect of a single infusion of 105 mg of pamidronate in 14 patients with mean levels of serum alkaline phosphatase about three times the normal upper limit. The alkaline phosphatase level returned to normal (the goal of successful treatment and likely to produce prolonged remission) in 71 per cent of patients with a mean nadir of the enzyme at 5.9 months. Excellent symptomatic control was achieved for 1.5 to 2 years. Four patients received a second infusion of 105 mg of pamidronate, a mean interval of 19 months after the first dose. This resulted in a similar clinical and biochemical response to the first. One patient had a 4-year remission before treatment was required.

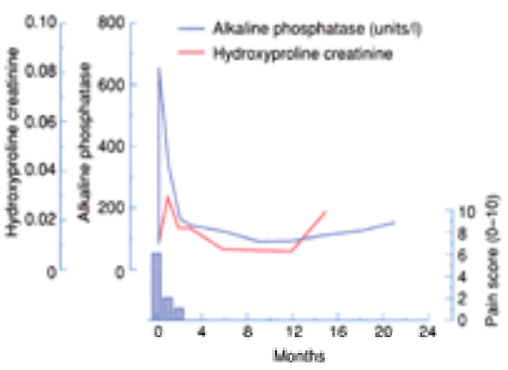


Fig. 10 The effect of a single intravenous infusion of pamidronate (APD) at a dose of 105 mg in one patient with Paget's disease. An early rapid fall in urine hydroxyproline excretion is followed by a slower fall in serum alkaline phosphatase. Pain is abolished after 3 months and the patient is still pain free after 21 months.

Further support for a single dose regimen was reported by Chakravarty *et al*. (1994) who treated 36 patients, with mean levels of serum alkaline phosphatase about four times the normal upper limit, with 60 mg of pamidronate. Alkaline phosphatase levels were returned to normal in 78 per cent of patients at 6 months and 77 per cent of patients at 1 year. The efficacy of the two single-dose regimens using pamidronate at 60 to 105 mg therefore is similar, but is marginally inferior to the complex regimen advocated by Anderson *et al*. (1994) which achieves a normal level in 90 per cent of patients.

The author's current practice is to treat pagetic patients with a single infusion of pamidronate 90 mg over 6 hours. Patients are followed up every 3 months with a clinical assessment and measurement of alkaline phosphatase. If this is supranormal at 6 months, a second infusion of 90 mg is administered. Only a small minority of patients do not achieve normal alkaline phosphatase levels with this regimen. This 'resistant' subgroup is treated with 90 mg infusions every 3 months until remission occurs.

Intravenous pamidronate is usually well tolerated but the following side-effects have been reported: transient pyrexia after the first but not usually after later infusions; mild leucopenia and mild hypocalcaemia; myalgia; transient increased bone pain; acute anterior uveitis, which may recur if the patient is rechallenged (unpublished observation).

The calcitonins

Subcutaneous or intramuscular calcitonin (salmon, porcine, or human) is safe and effective but has many disadvantages. It is so far only widely available by injection. It commonly causes mild side-effects such as nausea, flushing, and diarrhoea. After successful initial treatment, patients may relapse while continuing treatment, because of down-regulation of calcitonin receptors. Early relapse after ceasing treatment is also common, and calcitonin is very expensive.

If salmon calcitonin (salcatonin) is chosen, the starting dose after a 10 i.u. test dose is 50 to 100 i.u. daily reducing to 50 i.u. two or three times weekly once the symptomatic response has been achieved. Patients usually notice decreasing pain after 1 to 2 months but if there has been no response by 3 months, then calcitonin treatment should cease. Effective treatment has been claimed to reverse disease progression and, until the advent of the newer bisphosphonates, calcitonin was the treatment of choice for lytic disease of long bones and fissure fractures. Many patients learn to give their own injections but to condemn them to many years of injections 2 to 3 times weekly is becoming increasingly unacceptable when alternative short courses of bisphosphonates can achieve prolonged clinical and biochemical remissions.

It is possible that new formulations of calcitonin given by nasal spray, for example salmon calcitonin at 200 to 400 i.u. daily, when available might be useful but the typical therapeutic effect is weaker than that of the newer bisphosphonates (Reginster *et al*. 1988).

Surgery

There is still no consensus as to whether patients requiring total hip or knee replacement for joints ravaged by the combination of Paget's disease and osteoarthritis should receive specific medical treatment before surgery. There is some evidence that effective treatment—a newer bisphosphonate or calcitonin rather than the demineralizing etidronate—reduces bone blood flow, improves the quality of bone in which the prosthesis will be sited, and facilitates the adhesion of cement to bone. Certainly orthopaedic surgeons should not be deterred from replacement arthroplasty by the presence of Paget's disease, as the risks of prosthetic loosening are only slightly increased. It would seem logical to time surgery to coincide with a phase of clinical and biochemical pagetic remission induced by medical treatment. Prospective studies are awaited.

Chapter References

Anderson, D.C. *et al*. (1994). Intravenous pamidronate: evolution of an effective treatment strategy. *Seminars in Arthritis and Rheumatism*, **23 (4)**, 273–5.

Barker, D.J.P. (1984). The epidemiology of Paget's disease of bone. *British Medical Bulletin*, **40**, 396–400.

Basle, M.F., Rebel, A., Fournier, J.G., Russell, W.C., and Malkani, K. (1987). On the trail of paramyxoviruses in Paget's disease of bone. *Clinical Orthopaedics*, **217**, 9–15.

Beardsworth, L.J., Eyre, D.R., and Dickson, I.R. (1990). Changes with age in the urinary excretion of lysyl- and hydroxy-lysylpyridinoline, two new markers of bone collagen turnover. *Journal of Bone and Mineral Research*, **5**, 671–6.

Bhambhani, M.M., Lamberty, B.G.H., Clements, M.R., Skingle, S.J., and Crisp, A.J. (1992). Giant cell tumours in mandible and spine—a rare complication of Paget's disease of bone. *Annals of the Rheumatic Diseases*, **51**, 1335–7.

Brenton, D., Isenberg, D.A., and Bertram, J. (1980). Osteosarcoma complicating familial Paget's disease. *Postgraduate Medical Journal*, **56**, 238–43.

Chakravarty, K., Merry, P., and Scott, D.G.I. (1994). A single infusion of pamidronate AH Pr BP in the treatment of Paget's disease of bone. *Journal of Rheumatology*, **21**, 2118–21.

Crisp, A.J., Smith, M.L., Skingle, S.J., Smith, M., Page Thomas, D.P., and Hazleman, B.L. (1989). The localisation of the bone lesions of Paget's disease by radiographs, scintigraphy and thermography: pain may be related to bone blood flow. *British Journal of Rheumatology*, **28**, 266–8.

Gordon, M.T., Anderson, D.C., and Sharpe, P.T. (1991). Canine distemper virus localised in bone cells of patients with Paget's disease. *Bone*, **12**, 195–201.

Guyer, P.B. (1980). Radiology in Paget's disease: clinical and aetiological significance. *Hospital Update*, **6**, 1079–91.

Harinck, H.I.J., Bijvoet, O.L.M., Vellenga, C.J.L.R., Blanksma, H.J., and Frijlink, W.B. (1986). Relations between signs and symptoms in Paget's disease of bone. *Quarterly Journal of Medicine*, **226**, 133–51.

Harinck, H.I.J., Papapoulos, S.E., Blanksma, H.J., Moolenaar, A.J., Vermeij, P., and Bijvoet, O.L.M. (1987). Paget's disease of bone: early and late responses to three different modes of treatment with APD. *British Medical Journal*, **295**, 1301–5.

Heath, D.A. (1987). Treating Paget's disease. *British Medical Journal*, **294**, 1048–50.

Jacobs, T.P., Michelsen, J., Polay, J.S., D'Adamo, A.C., and Canfield, R.E. (1979). Giant cell tumor in Paget's disease of bone. *Cancer*, **44**, 742–7.

Kahn, A.J. (1990). The viral etiology of Paget's disease of bone: a new perspective. *Calcified Tissue International*, **47**, 127–9.

Kanis, J.A. and Gray, R.E.S. (1987). Long term follow up observations on treatment in Paget's disease. *Clinical Orthopaedics*, **217**, 99–125.

Krane, S.M. (1982). Etidronate disodium in the treatment of Paget's disease of bone. *Annals of Internal Medicine*, **96**, 619–25.

Krane, S.M. (1986). Paget's disease of bone. *Calcified Tissue International*, **38**, 309–17.

Meunier, P.J. *et al*. (1987). Skeletal distribution and biochemical parameters of Paget's disease. *Clinical Orthopaedics*, **217**, 37–44.

O'Driscoll, J.B. and Anderson, D.C. (1985). Past pets and Paget's disease. *Lancet*, **ii**, 919–21.

Paget, J. (1887). On a form of chronic inflammation of bones (osteitis deformans). *Transactions of the Medical and Chirurgical Society*, **60**, 37–63.

Preston, C.J. *et al*. (1986). Effective short term treatment of Paget's disease with oral etidronate. *British Medical Journal*, **292**, 79–80.

Ralston, S.H., Digiovine, F.S., Gallacher, S.J., Boyle, I.T., and Duff, G.W. (1991). Failure to detect paramyxovirus sequences in Paget's disease of bone using the polymerase chain reaction. *Journal of Bone and Mineral Research*, **6**, 1243–8.

Reginster, J.Y., Jeugmans-Huynen, A.M., Albert, A., Denis, D., and Franchimont, P. (1988). One year's treatment of Paget's disease of bone by synthetic salmon calcitonin as a nasal spray. *Journal of Bone and Mineral Research*, **3**, 249–52.

Watts, R.A., Skingle, S.J., Bhambhani, M.M., Pountain, G., and Crisp, A.J. (1993). Treatment of Paget's disease of bone with single dose intravenous pamidronate. *Annals of Rheumatic Diseases*, **52**,

616–18.

Yates, A.J.P. *et al*. (1985). Intravenous clodronate in the treatment and retreatment of Paget's disease of bone. *Lancet*, **i**, 1474–7.

5.17.3 Diseases of bone, cartilage, and synovium

P. J. Maddison

In this chapter, a miscellaneous group of disorders affecting components of the joint and periarticular structures is described. They can present a challenge to the clinician in diagnosis and management. There is often a delay in diagnosis causing pain and disability. Frequently this is due to insufficient clinical suspicion, partly because of the rarity of most of these conditions. Their diagnosis generally depends on recognizing a combination of clinical, laboratory, and histological features, but the advent of techniques such as magnetic resonance imaging (**MRI**) facilitates earlier diagnosis. This is particularly important for conditions such as osteonecrosis where the institution of treatment at an earlier stage in the course of the process may prevent joint destruction.

Osteoid osteoma

Skeletal neoplasms are discussed in Chapter 1.2.1.2 and Chapter 5.19.1. Only rarely do they cause problems for the experienced clinician in being mistaken for systemic rheumatic diseases. An exception is hypertrophic osteoarthropathy (Chapter 5.19.1); another is osteoid osteoma. This is a benign osteoid-forming tumour, which can be an elusive cause of bone pain, 'radiculopathy', or 'arthritis' in children and young adults, depending on its site. It is uncommon and accounts for 10 to 12 per cent of benign bone neoplasms (Dahlin and Unni 1986).

The maximum age incidence is in the second and third decades, but this tumour can occur in all age groups, and is two or three times more common in boys than girls. More than two-thirds of the lesions occur in long bones, mostly in the lower extremity and especially involving the femur and tibia. The neck of the femur and the intertrochanteric region are a particularly characteristic location in the femur (Resnick and Niwayama 1988).

The lesion consists of a small core or nidus of cellular, highly vascularized tissue, with an interlacing network of immature bone and osteoid in varying proportions. The nidus is surrounded by a zone of reactive bone, especially in cortical bone. Very rarely, the osteoid osteoma is multifocal with more than one nidus. High levels of prostaglandins are produced within the lesions (Makley and Dunn 1982). Intra-articular lesions, which are rare and arise at the end of a long bone within the insertion of the joint capsule, are accompanied by a synovitis characterized by a hyperplastic synovium and a prominent lymphocytic infiltrate.

Clinical features

Almost without exception the initial symptom is pain. This may be vague and intermittent at first but becomes increasingly intense. Often, though not invariably, it is worse at night, and typically it is relieved by aspirin and other non-steroidal anti-inflammatory drugs. Although pain is usually felt in the region of the bone lesion, the presentation can be much less characteristic and the diagnosis is often delayed for many months. For example pain can sometimes be referred or radicular, accompanied by muscle atrophy and diminished or absent tendon reflexes in the affected limb, thus mimicking a spinal lesion (Kiers *et al*. 1990). Intra-articular lesions can also present a confusing picture, with joint pain, stiffness, effusion, muscle atrophy, and loss of function. Osteoid osteomas that arise in the posterior elements of the spine can present with scoliosis or torticollis.

Diagnosis

A typical lesion (Fig. 1(a)) in the cortex of a long bone is seen on a plain radiograph as a well-defined area of sclerosis, 0.5 to 1.0 cm in diameter, surrounding a radiolucent nidus, which may contain speckled areas of calcification. However, lesions in cancellous bone and neural arch and intra-articular lesions are often difficult to locate with plain radiographs. A bone scan using $^{99}Tc^{m}$ hydroxymethylene diphosphonate is a highly sensitive technique to screen for an osteoid osteoma, as these lesions avidly accumulate isotope (Fig. 1(b)). Helms *et al*. have described the 'double density sign' as characteristic for osteoid osteoma (Helms *et al*. 1984) but it is often not seen and the bone scan is then non-specific for differentiating an osteoid osteoma from lesions such as a stress fracture, synovitis, and a Brodie's abscess. Computed tomography (**CT**) is also valuable for imaging lesions difficult to locate on plain radiographs, especially those in the spine and proximal femur, and for precisely locating the nidus before surgical resection (Swee *et al*. 1979). MRI effectively demonstrates the associated intramedullary and soft tissue changes but the resulting image can lead to diagnostic confusion and CT is generally considered to be the better imaging modality for the diagnosis of osteoid osteoma (Assoun *et al*. 1994).

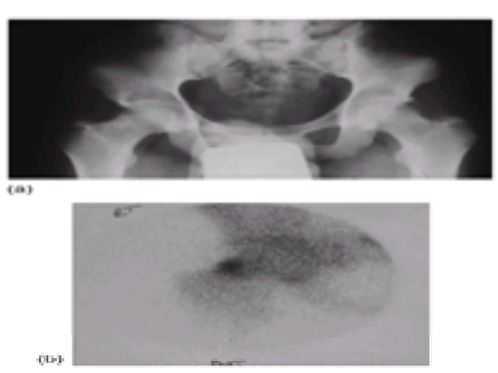


Fig. 1 (a) Plain pelvic radiograph showing an osteoid osteoma of the femoral neck in an 18-year-old male who presented with a 3-month history of progressively worsening pain. (b) Radionuclide bone scan showing increased uptake of isotope by the lesion.

Treatment

Surgery is curative, provided the nidus is excised completely. If this is not achieved, the osteoid osteoma may recur, up to10 years afterwards. Many surgeons use an *en bloc* excision but some use CT- or radioisotope-guided excision of osteomas, mainly of those located in the extremities (Ward *et al*. 1993; Musculo *et al*. 1995). Radiofrequency ablation (Rosenthal *et al*. 1995) or thermocoagulation (De Berg *et al*. 1995) have been reported as promising and less invasive alternatives to surgery in selected patients.

Synovial chondromatosis

This is a disorder of unknown aetiology that affects synovium-lined joints, tendon sheaths, and bursas. It is characterized by chondrometaplasia of the subsynovial connective-tissue cells. The joint cavity becomes filled with a thickened synovium containing pearly-white to blue nodules. Some of these lie free in the joint as loose bodies. The histological appearance is of double-nucleated chondrocytes exhibiting moderate hyperchromasia, arranged in clusters with abundant intervening matrix.

The disorder is uncommon and generally occurs in middle-aged men. It has never been reported before puberty. Most commonly it affects the knee (one-half of cases) or hip, but can also involve other large joints such as the elbow, shoulder, and wrist. Extra-articular involvement of tendons occurs, predominantly in the hands and feet.

Clinically, it resembles pigmented villonodular synovitis (see below). The presenting symptoms are mild, but progressive, pain and locking. Aspirated joint fluid is always yellow or straw-coloured. Plain radiographs (Fig. 2) show small, punctate calcifications outlining the joint margin and sometimes bony erosions due to raised intra-articular pressure. Large osteochondral fragments can be seen but this is not a specific feature. Loss of articular cartilage does not occur. If the cartilaginous nodules are not calcified or ossified, the diagnosis can only be confirmed by arthroscopy.

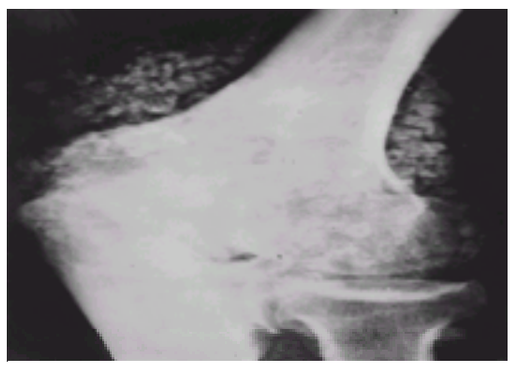


Fig. 2 Anteroposterior view of the elbow demonstrating multiple osteochondral bodies, regularly shaped and uniform in size. This is typical of synovial chondromatosis. (By courtesy of Dr P. Renton, University College London.)

This is a benign condition. During the active phase it is very slowly progressive but it becomes self-limiting and regression can be observed. In rare cases there is transformation to chondrosarcoma.

Treatment is surgical and most authorities suggest that this should consist of removal of loose bodies and excision of the synovial membrane (Christensen and Poulsen 1975), which can be done via the arthroscope (Coolican and Dandy 1989). Removal of loose bodies without an extensive synovectomy has been reported to give similar results (Shpitzer *et al*. 1990).

Pigmented villonodular synovitis

The term pigmented villonodular synovitis was first coined by Jaffe *et al*. in 1941 to encompass a group of conditions, previously known by a variety of other terms, characterized by exuberant proliferation of synovial cells and mesenchymal supporting tissue affecting joints, tendons, and bursas (Jaffe *et al*. 1941; Flandry and Hughston 1987; Goldman and DiCarlo 1988). In their seminal paper, Jaffe and colleagues emphasized the villous and nodular proliferation, the deposition of iron and fat, and the non-malignant nature of these conditions. The terminology subsequently was expanded to distinguish localized lesions (sharply localized or pedunculated lesions involving tendon sheaths or part of the joint lining) from diffuse lesions of the joint synovial membrane, which, although similar histologically, are more progressive and tend to recur after treatment.

Aetiological factors

The aetiology is unknown and there are still various hypotheses based on different interpretations of the histological changes. Although once considered to be neoplastic, pigmented villonodular synovitis does not behave like a malignancy and does not metastasize. In contrast, a benign neoplastic process of synovial, vascular, or fibrohistiocytic origin is proposed by some (Rao and Vigorita 1984). This view is supported by some cytogenetic evidence of clonality (Ray *et al*. 1990). Most favour a non-neoplastic aetiology; suggestions include a response to blood and blood products from trauma. An epidemiological study (Myers *et al*. 1980), while showing no evidence of a genetic basis for pigmented villonodular synovitis, demonstrated a history of chronic repetitive trauma and repeated haemarthroses in approximately 50 per cent of cases. However, there has been a general failure to reproduce typical histological changes in experimental models with injection of various substances from whole blood to colloidal iron. Furthermore, clinical experience from haemophiliacs and others with bleeding disorders points away from a reaction to trauma and haemorrhage *per se*. A commonly held, but rather unhelpful, view is that pigmented villonodular synovitis is an inflammatory response to an as yet unknown trigger.

Pathology

The striking pathological feature is the proliferation of synovial lining cells at the surface and invading the subsynovial stroma. The result in diffuse pigmented villonodular synovitis is a greatly thickened synovial membrane bristling with big, long villi, which may be interspersed with nodules of various size. The tissue is often stained red-brown from repeated haemorrhage, with mottled areas of yellow-orange from lipid deposition. Histologically (Schumacher *et al*. 1982), there is marked proliferation of surface lining cells, together with an invasion of the subsynovial stroma by large epithelioid histiocytic cells that electron-microscopic studies show to be a mixture of proliferating synovial fibroblasts and type B synovial lining cells. The stroma also contains multinucleate giant cells and lipid-laden macrophages. In addition, there are a few lymphocytes and plasma cells. Capillary hyperplasia resulting in numerous thin-walled vascular channels is another characteristic feature. Haemosiderin is deposited in the lining and epithelioid cells, and in the stroma. Another typical feature is the ability of pigmented villonodular synovitis to invade subchondral bone. This is thought to be through erosion into the osteocartilaginous junction, by extension into vascular foramina, and from the effect of increased intra-articular pressure causing focal osteopenia in the subchondral bone.

Clinical features

Pigmented villonodular synovitis is rare, with an estimated annual incidence of around 1.8 cases/million population (Myers *et al*. 1980). Typically it affects adults of both sexes in their third or fourth decade, but there is a wide age range and cases have been reported in children (Docken 1979) and even infants (Curtin *et al*. 1993). Classically, it presents as a monoarthritis. In 80 per cent of cases the knee is involved, followed in order by the hip, ankle, and shoulder; very rarely, multiple joints are involved and there are case reports of involvement at sites such as the spine (Clark *et al*. 1993) and temporomandibular joint (Franchi *et al*. 1994). The onset is usually insidious, with pain the most common complaint (Flandry *et al*. 1994a); this is mild at first but progressive. Rarely, there may be sudden exacerbation of pain

due to torsion or infarction of a nodule of abnormal tissue. Sometimes there are features of internal derangement, such as locking, especially with localized forms of pigmented villonodular synovitis. Occasionally, an affected knee joint becomes unstable. On examination, there is often swelling and sometimes one or more palpable masses of synovium. There may be local warmth and points of tenderness.

Joint aspiration gives fluid that ranges in colour from yellow or straw, with deep xanthochromia from previous haemorrhage, to brown-stained or frankly bloody. Reports of synovial fluid analysis are sparse but findings point to inflammation and include a slight elevation of protein, reduced glucose, and a low to moderate leucocyte count. The results of other laboratory tests are otherwise normal.

Plain radiographs may be normal or only show the soft tissue outline of synovial swelling, which is made more radiodense by haemosiderin deposition. However, calcification, which occurs in malignant lesions such as synovial sarcoma, is absent. In approximately 50 per cent there are osteoarticular changes corresponding to invasion of bone by the lesion. These include multiple subchondral cysts, which can occur on non-weight-bearing surfaces, and well-demarcated erosions due to increased intra-articular pressure. These changes occur earlier in joints with a tight articular capsule, such as the hip, than in the knee in which a more distensible capsule accommodates a greater degree of soft tissue proliferation. In the knee, bony lesions often develop first in the patellofemoral compartment, where intra-articular soft tissue is more likely to be entrapped (Smith and Pugh 1962). Preservation of joint space is reported to be a typical feature of pigmented villonodular synovitis, but, in fact, loss of joint space can occur as a late feature. Juxta-articular osteoporosis and osteophyte formation are not seen. MRI is very characteristic if there is sufficient haemosiderin and fat deposition in the lesion. Haemosiderin deposition produces a low signal intensity on T_1-weighted images, which decreases even further on T_2-weighted images. In contrast, areas with high fat content have high signal intensity. Consequently, an MRI study demonstrating a multinodular intra-articular lesion with patchy areas having characteristics of fat and haemosiderin deposition is highly suggestive of pigmented villonodular synovitis. However, in practice both false positives and false negatives occur. Techniques such as arthrography and arteriography have been used, but the results are rather non-specific.

Diagnosis

The diagnosis is based on a combination of clinical, radiological, and histological findings and the gross appearance of the lesion. The presence of serosanguinous synovial fluid in a young adult in the absence of a history of recent trauma is highly suggestive of the diagnosis of pigmented villonodular synovitis. The definitive diagnosis, however, often rests on the interpretation of a synovial biopsy. Histological criteria for the diagnosis have not been defined. Although features such as epithelioid cells, multinucleate giant cells, and deposition of fat and haemosiderin are highly characteristic, interpretation of the histological features by a specialist in osteoarticular pathology is required.

Conditions to be considered in the differential diagnosis of pigmented villonodular synovitis include:

1. malignant synovioma;
2. synovial haemangioma;
3. synovial chondromatosis;
4. tuberculous arthritis;
5. amyloidosis;
6. haemophilia;
7. lipoma arborescens.

Treatment

Localized forms of pigmented villonodular synovitis are treated by marginal excision of the lesion and have a good prognosis. The diffuse form, however, tends to be progressive and recurrence is not uncommon. Treatment suggestions are largely based on anecdotal experience, the published series are small, and post-treatment follow-up is limited. A range of techniques, which include radiation, wide synovectomy, synovectomy combined with radiation, arthrodesis, bone grafting, primary arthroplasty, and radiation synovectomy, has been used. No single method has a uniformly high proportion of good results.

The most commonly reported treatment is surgical synovectomy. There are recent reports of good results (Flandry *et al*. 1994b), although in previous series there have been a high percentage of recurrences, with persistent joint pain and stiffness as common sequelae (Johansson *et al*. 1982). The use of radiation therapy as an adjunct does not improve the outcome. Radiation synovectomy using intra-articular yttrium-90 silicate has been reported, with promising results (Franssen *et al*. 1989). There are several advantages over surgical synovectomy including technical simplicity and fewer complications. There is limited experience, however, and not much long-term follow-up. In advanced cases with joint destruction (especially in the hip), it may be necessary to resort to a total arthroplasty.

Chronic focal osteomyelitis (Brodie's abscess)

This usually follows an acute haematogenous infection in an adolescent male (Boriani 1980). Three-quarters of Brodie's abscesses develop in the lower extremity, most often in the tibia. The acute episode of infection may have been successfully treated with antibiotics from a clinical point of view, only to be followed by a focus of chronic infection that can persist for years if not surgically drained.

Typically, the infection occurs in the metaphyseal side of the growth plate, where it remains localized. In rare cases, the infection extends through the growth plate into the epiphysis. Usually it is unifocal but multiple abscesses have been reported. Approximately 80 per cent involve *Staphylococcus aureus*, although almost any organism can be implicated.

Pain is the main complaint, described as aching or boring, and may have been present for months or even years. This is accompanied by localized tenderness and sometimes swelling. Occasionally, the abscess dissects through spongy bone and erodes through the cortex to drain into a joint, or through a sinus to the skin surface. The symptoms and signs of systemic illness that generally accompany acute osteomyelitis are conspicuously absent. Laboratory tests may also be normal, although there can be slight leucocytosis or elevation of the erythrocyte sedimentation rate. Radiographs show a sharply demarcated and irregular area of bone destruction surrounded by sclerosis (Fig. 3). Cultures from the abscess are positive in 50 to 80 per cent of cases.

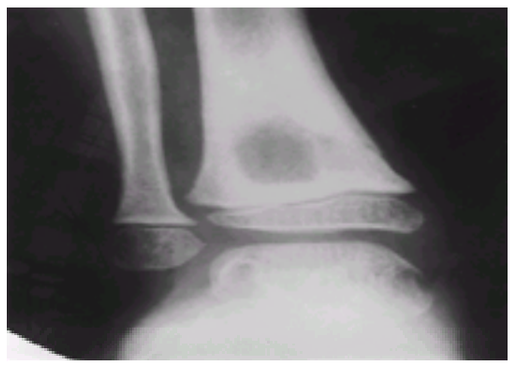


Fig. 3 Anteroposterior view showing a sharply demarcated Brodie's abscess surrounded by reactive sclerosis in the distral diaphysis of the tibia of a child presenting with chronic pain. (By courtesy of Dr P. Renton, University College London.)

Treatment is by surgical drainage and appropriate antibiotics. The prognosis is good.

Osteonecrosis

Osteonecrosis is a major reason for orthopaedic surgery to the hip, particularly in younger patients. There are several synonyms: avascular necrosis, aseptic

necrosis, ischaemic necrosis, steroid necrosis, segmental subchondral infarction. Osteonecrosis at certain sites is associated with eponyms such as Legg–Perthes' disease (hip), Freiberg's disease (metatarsal), and Kienboch's disease (lunate).

In 1860, James Paget described the gross appearance of osteonecrosis in his lectures on surgical pathology (Bullough and DiCarlo 1990). Necrosis of the femoral head was associated with Caisson's disease in 1888, with corticosteroids in 1957 (Pietrogrande and Mastromarino 1957), and the association with systemic lupus erythematosus was reported in 1960 (Dubois and Cozen 1960). Before the 1960s, most reported cases were associated with fracture. The current prevalence of osteonecrosis is difficult to assess because many cases are clinically silent. Different forms of osteonecrosis affect children, young adults, or elderly people. However, most of those with atraumatic osteonecrosis are relatively young, with a peak incidence in the fifth decade of life. One group reports that about 18 per cent of femoral heads removed in total hip-replacement procedures for non-traumatic causes show evidence of osteonecrosis (Bullough and DiCarlo 1990). Approximately 60 per cent are bilateral, women are slightly more affected (1.2:1), and the mean age of presentation is 55 years compared with 67 years for patients with primary osteoarthritis.

Osteonecrosis, like infarction anywhere, results from a reduction or the obliteration of the blood supply to the affected area. Subchondral bone has a limited collateral circulation and the perfusion pressure and blood flow of epiphyses and fatty marrow is low compared with red diaphyseal marrow. Therefore, ends of bones such as the femoral head are more susceptible to ischaemia. Various mechanisms have been implicated (Mankin 1992):

1. interruption of extraosseous arterial blood supply (e.g. trauma, vasculitis);
2. interruption of intraosseous sinusoidal circulation (e.g. nitrogen bubbles, sickled erythrocytes, thrombi, fat emboli);
3. extravascular compression of sinusoidal circulation (e.g. nitrogen bubbles, intramedullary lipocyte hypertrophy, accumulation of Gaucher's or malignant cells).

Often a combination of factors is involved. A final common pathway appears to be increased bone-marrow pressure, which can be demonstrated in the very earliest stages of the process (Zizic *et al*. 1986) and which further contributes to the impaired intraosseous microcirculation and progression of necrosis.

The pathology of osteonecrosis is well defined. The first phase is necrosis of bone and bone marrow. A cut section of bone shows a wedge-shaped necrotic zone in the subchondral region, which is demarcated from the normal bone marrow by a hyperaemic border. Granulation tissue develops and then advances from the margin of the infarct and necrotic bone is resorbed. Behind this, a second front of osteoblasts lays down new bone. If articular stress exceeds the structural integrity of the altered bone, there will be collapse of the articular surface and disruption of the joint.

Aetiological factors (Zizic 1990; Mankin 1992; Chang et al. 1993)

Some of the causes of osteonecrosis are:

1. trauma;
2. sepsis;
3. radiation, thermal, and electrical injury;
4. Caisson's disease;
5. haemoglobinopathies;
6. haemophilia;
7. coagulopathies;
8. Gaucher's disease;
9. alcoholism;
10. Cushing's syndrome;
11. corticosteroid usage;
12. systemic lupus erythematosus;
13. rheumatoid arthritis;
14. systemic sclerosis;
15. vasculitis;
16. organ transplantation (kidney, heart, marrow);
17. chronic dialysis;
18. human immunodeficiency virus infection;
19. pancreatitis;
20. chronic liver disease;
21. hypertriglyceridaemia;
22. pregnancy (especially in the third trimester);
23. (idiopathic).

The major cause is a fracture that interferes with the blood supply to areas such as the femoral or humeral head, talus, and scaphoid. Osteonecrosis is a major complication of intracapsular fracture of the femoral neck, which is accompanied by disruption of the circulation in approximately 20 per cent of cases. Osteonecrosis occurs in 18 per cent of compressed-air workers and 4 per cent of divers; it appears to be the result of the development of intravascular and extravascular nitrogen gas bubbles during decompression, effectively occluding the circulation to sites such as the femoral and humeral heads. Lesions are frequently multiple and bilateral; involvement of the humeral heads is particularly characteristic. In Caisson's disease, for some reason, the subchondral bone of the knee and the ankle are not involved.

Depending on the genotype, sickle cell haemoglobinopathies have a 5 to 14 per cent prevalence of radiographically detectable osteonecrosis. As in Caisson's disease, the femoral and humeral heads are involved.

The two most common causes of non-traumatic osteonecrosis are alcoholism and hypercortisolism, which account for two-thirds of the cases. A feature they have in common is alteration in systemic fat metabolism. Osteonecrosis may be produced as a result of fatty emboli, or of intramedullary lipocyte hypertrophy and consequent intraosseous sinusoidal compression. Osteonecrosis in rheumatoid arthritis and systemic lupus is mostly associated with corticosteroid treatment. However, as in systemic lupus (see below), osteonecrosis is described as a complication of rheumatoid arthritis in the absence of steroids (Wollheim 1984), although there is little information about the prevalence of this. Osteonecrosis has been reported after repeated intra-articular injections of long-acting corticosteroids (Laroche *et al*. 1990).

Osteonecrosis in systemic lupus

This is a relatively common and disabling complication of lupus (Cronin 1988). The prevalence of the symptomatic, radiographic lesion in adults with systemic lupus is reported to be between 5 and 10 per cent. However, asymptomatic lesions detected by MRI are more common and present in up to 35 per cent (Nagasawa *et al*. 1994). Although some become apparent on plain radiographs, present in up to 25 per cent of adult systemic lupus erythematosus patients (Klippel *et al*. 1979) and possibly in even more children (Bergstein *et al*. 1974), in only a minority of lesions is the necrosis extensive enough to cause clinical problems. The major risk factor for osteonecrosis is corticosteroid therapy, but there are reports of the complication before the steroid era (Leventhal and Dorfman 1974). Suggested additional risk factors present in the analysis of some studies, but not in others, include younger age, vasculitis, Raynaud's phenomenon, and leucopenia. Osteonecrosis has been reported in association with antiphospholipid antibodies in patients with the 'primary antiphospholipid syndrome' but others (Alarcon-Segovia *et al*. 1989; Migliaresi *et al*. 1994) have found no correlation between anticardiolipin antibodies and osteonecrosis in their lupus population. It appears that it is the use of large doses of corticosteroids (e.g. prednisolone, 60 mg daily), sustained over a period of months, that is important for the development of osteonecrosis, rather than the duration of corticosteroids or the accumulative dose *per se*. Giving pulses of megadose corticosteroids either once or repeated after intervals of several weeks does not independently predispose to osteonecrosis. A meta-analysis reported by Felson and Anderson suggests that this relationship to dose and duration of corticosteroids applies to patients with a variety of clinical conditions (Felson and Anderson 1987). It is worth remembering that a number of malpractice suits have been brought against physicians for failing to inform patients of this complication of corticosteroids.

The onset of localized pain, particularly in a weight-bearing joint, should give rise to a high index of suspicion. The femoral head is most commonly involved, but other sites include the femoral condyle, tibial plateau, humeral head, talus, scaphoid, and lunate. Simultaneous involvement of multiple sites is not infrequent.

Diagnosis

Pain is the principal symptom. In about two-thirds of the patients it occurs at rest and may be troublesome at night (due to increased intraosseous pressure). It is well established that once radiographic changes occur in osteonecrosis the natural history is generally subchondral bone collapse and severe disability. The Arlet and Ficat classification of osteonecrosis is based on the plain radiographic appearance (Ficat and Arlet 1980), as follows:

Stage I: normal appearance;

Stage II: early changes consisting of diffuse osteoporosis, sclerosis, and cyst formation producing a mottled appearance (Fig. 4(a));

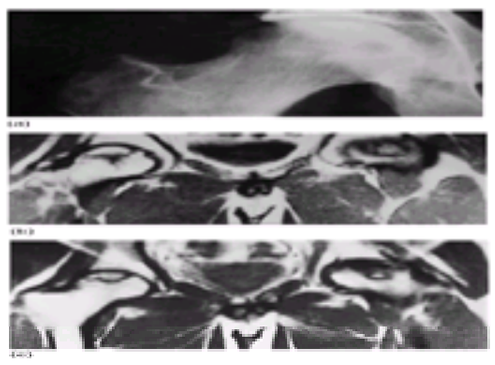


Fig. 4 (a) Early changes of osteonecrosis of the femoral head shown on plain radiograph (Arlet and Ficat stage II). (b) MRI clearly detects osteonecrosis at an early stage in T_1-weighted and (c) T_2-weighted images. (By courtesy of Dr P. Renton, University College London.)

Stage III: subchondral bone collapse (crescent sign) with normal joint space;

Stage IV: abnormal contour of bone with joint-space loss.

This has been extended by Steinberg *et al*. to provide a more sensitive method of following the progress of osteonecrosis in the femoral and humeral heads (Steinberg *et al*. 1984).

In early stages the plain radiograph is normal. The imaging technique currently combining the greatest sensitivity and specificity for the diagnosis of early cases is MRI (Zizic 1990) (Fig. 4(b) and (c)). The presence of a low intensity band on T_1-weighted images is an early specific finding of osteonecrosis and MRI has the advantage of identifying early changes in necrotic bone marrow before other changes in bone have taken place (Halland *et al*. 1993; Sugano *et al*. 1994). If MRI is not available, an isotope bone scan with $^{99}Tc^{m}$ diphosphonate is indicated (Tawn and Watt 1989). This is also helpful in early diagnosis but it is less specific, anatomical resolution is often poor, and false negatives commonly occur, especially when there is bilateral disease. Other methods used for early diagnosis have included CT with multiplanar reconstruction (Sartoris *et al*. 1986). This is reported to be more sensitive than bone scan but the significant radiation exposure makes MRI a more attractive option. Measurement of intramedullary bone pressure and intraosseus venography proposed by Zizic *et al*. is probably also redundant as a routine procedure since the advent of MRI (Zizic *et al*. 1986). However, bone biopsy of the affected site is sometimes necessary, as local infection can be associated with osteonecrosis (Habermann and Friedenthal 1978).

Osteonecrosis at specific sites

Vertebral

This is uncommon but compared with compression fractures of osteoporosis is more often associated with neurological complications (Feldmann *et al*. 1988). This mainly happens in elderly people; the lesions are usually single, and at the thoracolumbar junction. It is not usually associated with malignancy and MRI is helpful in early diagnosis. The intravertebral vacuum cleft is the characteristic feature on the plain radiography.

Keinbock's disease

This describes osteonecrosis of the lunate. It may occur after a fracture but in most cases there is no specific history of this. Often it happens in the dominant hand and is thought to be the result of chronic repetitive trauma. It is associated with a short ulna, which probably reinforces the effects of repeated trauma.

Preiser's disease

This is spontaneous osteonecrosis of the scaphoid, usually affecting the proximal pole. Involvement of the scaphoid is usually post-traumatic and in idiopathic cases the trauma may have been minor and unrecognized.

Hegemann's disease

This describes osteonecrosis of the humeral trochlear. This is rare and happens mainly in preadolescent and adolescent boys, who present with a swollen elbow; this has reduced movement but is not particularly painful. Usually it resolves spontaneously.

Legg–Perthes' disease

This is osteonecrosis of the femoral head, which occurs in children between the age of 4 and 12 years and affects one or both hips. The aetiology is unknown but the condition is associated with increased intraosseous pressure and venous hypertension (Liu and Ho 1991). There may be spontaneous resolution, especially in younger patients, in whom conservative management is indicated.

Osteonecrosis of the femoral condyle

This occurs spontaneously in older people, predominantly affecting the medial condyle.

Kohler's disease

This is the rare involvement of the tarsal navicular, which primarily occurs in male children between the ages of 4 and 10 years. It is usually self-limiting.

Freiberg's disease

This is osteonecrosis of the metatarsal head, usually the second. It mainly affects adolescent females and is usually self-limiting.

Treatment

Advanced osteonecrosis of weight-bearing surfaces, such as the femoral head, leads to secondary osteoarthritis and severe disability. Total joint replacement may be the only solution in such cases. The results are satisfactory, giving a long period of good function. However, one still hesitates to recommend total joint replacement in young people and the failure rate may be higher in operations for osteonecrosis, especially when associated with corticosteroids, than in arthroplasties for other conditions (Cornell *et al*. 1985). Attempts have therefore been made to treat the condition at an earlier stage in order to preserve the integrity of affected bone.

Osteonecrosis of the femoral condyle in elderly people can often be treated conservatively with initial limitation of weight-bearing, non-steroidal anti-inflammatory drugs, hydrotherapy, and muscle strengthening exercises (Motohashi *et al*. 1991).

There is little evidence from retrospective studies that bed rest, modified weight bearing, analgesics, or non-steroidal anti-inflammatory drugs are of much benefit in other cases, at least for osteonecrosis of the hip in adults. Because increased intraosseous pressure is a common feature in early stages, core decompression has been recommended. Encouraging results with prevention of progression in early stages of osteonecrosis of the hip (Hungerford 1989) and the shoulder (Mont *et al*. 1993) have been reported by some groups. However, this is a controversial topic and not all studies have been so encouraging (Learmonth *et al*. 1990). The use of MRI to assess the extent of osteonecrosis may be a way of selecting those most likely to respond to this procedure (Holman *et al*. 1995).

Other ways of preserving the femoral head include bone grafting using a variety of procedures (Rosenwasser *et al*. 1994; Urbaniak *et al*. 1995), the use of electrical stimulation, and sometimes a combination of these (Steinberg *et al*. 1985). In addition, various techniques of osteotomy, such as the transtrochanteric osteotomy of the femoral head, have been used to alter the weight-bearing surface away from the involved area. These techniques are still being refined and evaluated.

Osteochondritis dissecans

Osteochondritis dissecans is usually a solitary lesion of the medial femoral condyle in which a fragment composed of articular cartilage and subchondral bone becomes demarcated from the surrounding bone and cartilage and may form an intra-articular loose body (Green and Banks 1990). Occasionally, it may involve the elbow, hip, and talus. The aetiology is unknown but anomalies of ossification and low-grade trauma appear to be important.

It predominantly affects adolescent males and should be suspected in a child or teenager who, after minor trauma, develops a relatively sudden onset of knee pain followed by mechanical dysfunction. There may be a hereditary component and familial occurrence has been reported (Paes 1989). This is characterized by multiple articular lesions, particularly affecting the hips and knees, and an autosomal dominant inheritance. Sometimes there is associated dwarfism and a generalized epiphyseal abnormality.

Symptoms are mainly pain, reduced joint movement, effusion, and limp. A plain radiograph shows a well-circumscribed, sclerotic lesion demarcated by a radiolucent line from surrounding bone (Fig. 5).

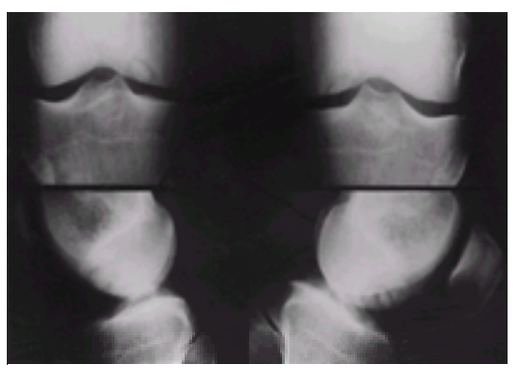


Fig. 5 Anteroposterior and lateral views showing typical lesions of osteochondritis dissecans in both medial femoral condyles. A radiolucent line separates the oval-shaped, *in situ* body from the femoral condyle. (By courtesy of Dr P. Renton, University College London.)

In young patients before skeletal maturity there is a good chance of healing, and treatment is consequently conservative. Once the epiphyses have closed, osteochondritis dissecans is more likely to cause intra-articular loose bodies and subsequent symptoms of internal derangement, and eventually secondary osteoarthritis. Arthroscopy can be helpful in assessing the extent of the lesion. A variety of surgical procedures have been recommended, from drilling of the lesion to promote ingrowth of fibrocartilage to replacement of a large deficit with an osteochondral allograft.

Relapsing polychondritis

This is an uncommon, multisystem disorder of unknown aetiology characterized by episodic and sometimes progressive inflammation of cartilaginous structures and tissues rich in glycosoaminoglycans. The characteristic clinical syndrome, with involvement of the pinna of the ears, nose, larynx and upper airways, joints, heart, blood vessels, inner ear, cornea, and sclera, was first described by Jaksch-Wartenhorse (Jaksch-Wartenhorse 1923). In about 30 per cent of cases it is associated with other systemic rheumatic or autoimmune diseases such as rheumatoid arthritis, systemic lupus, Sjögren's syndrome, thyroiditis, ulcerative colitis, psoriasis, and Behçet's syndrome (Kitridou *et al*. 1987; Tishler *et al*. 1987; Orme *et al*. 1990; Harisdangkul and Johnson 1994).

Aetiology and pathogenesis

The most specific lesion is inflammation of cartilage. This leads to cartilage destruction and fibrosis. The lesion is characterized by a dense inflammatory infiltrate of neutrophils, lymphocytes, macrophages, and plasma cells. Initially this involves the perichondral region. There is loss of proteoglycans, destruction of the collagen matrix, and chondrocyte death. Destroyed cartilage is replaced by granulation tissue and there is subsequent fibrosis (Fig. 6). Cartilage matrix proteins, such as cartilage oligomeric matrix protein (COMP), are released and raised levels are present during disease activity (Saxne and Heinegard 1995).

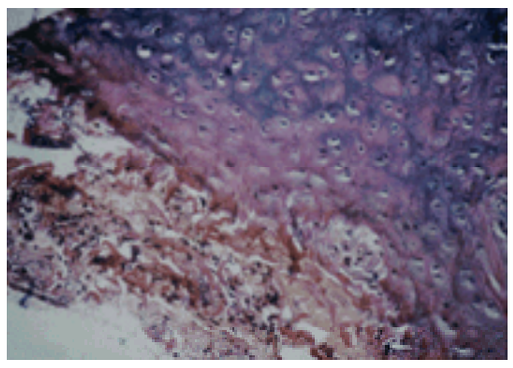


Fig. 6 Acute inflammation of the perichondral region of the external ear with infiltration of neutrophils, mononuclear cells, and plasma cells with destruction of underlying cartilage (haematoxylin and eosin). (By courtesy of Dr A. Balsa, La Paz Hospital, Madrid.)

The cause is unknown but, on the basis of circumstantial evidence, it is thought that there is an immunological pathogenesis. Components of cartilage, such as

collagen, elastin and proteoglycan, express multiple antigenic determinants but these are normally sequestered from the immune system. A breach of the integrity of cartilaginous structures could potentially stimulate an immune response to these constituents. The ubiquitous nature of matrix components such as elastin and proteoglycans could explain the pattern of involvement, which includes respiratory tract cartilage, structures of the eye, and the cardiovascular system.

Antibodies, predominantly IgG, which react specifically with collagens type II, IX, and XI (forming the major fibrillar scaffold in cartilage), can be detected in some but not all patients with relapsing polychondritis (Yang *et al*. 1993). The highest titres are found in the early phase of the disease and the titre may relate to disease activity (Ebringer *et al*. 1981). These antibodies are not disease specific and occur, for example, in rheumatoid arthritis, although possibly they are directed to different epitopes on the collagen molecule (Terato *et al*. 1990). The possibility that humoral factors are involved in the pathogenesis is suggested by the report of polychondritis occurring in the newborn infant of an affected mother and the subsequent recovery of the baby (Arundell and Haserick 1960). Granular deposits of immunoglobulin and complement have been observed at the chondrofibral junction in biopsies from affected ears (Valenzuela *et al*. 1980), suggesting the involvement of immune complexes. Collections of fluid in the middle ear have been reported to be hypocomplementaemic, suggesting complement consumption (McKenna *et al*. 1976). Thus, there is evidence for humoral immune mechanisms being involved in cartilage injury. In addition, earlier reports of cellular immune reactions to proteoglycan and other matrix components (Herman and Dennis 1973) have been supported by the more recent demonstration of cell-mediated immunity to collagens type II, IX, and XI paralleling the humoral response to these structural proteins (Alsalameh *et al*. 1993).

A significant increase in DR4 antigen frequency has been found (Lang *et al*. 1993) but, in contrast to rheumatoid arthritis, there is no predominance of any DR4 subtype.

Clinical features

The disease predominantly affects middle-aged white subjects, with a peak incidence in the fourth to fifth decades, although it has been reported in all races and age groups. The patient typically presents with recurrent swelling and pain of the external ear and/or nose, or with uveitis, or with an arthropathy. The cumulative involvement of various organ systems is summarized in Table 1.

Organ	Percentage involvement
External ear	85
Arthritis	75
Nose	50
Eye	50
Respiratory tract	50
Internal ear	40
Skin	25
Kidney	20
Heart	10
Blood vessels	8
Central nervous system	Rare

Table 1 Extent of organ involvement in relapsing polychondritis

Episodes of inflammation of the cartilaginous portion of one or both ears and the nose are often sudden and last several days. Repeated episodes of protracted inflammation, leading to cartilage destruction, produce deformity such as saddle nose (Fig. 7(a) and Fig. 7 (b)).

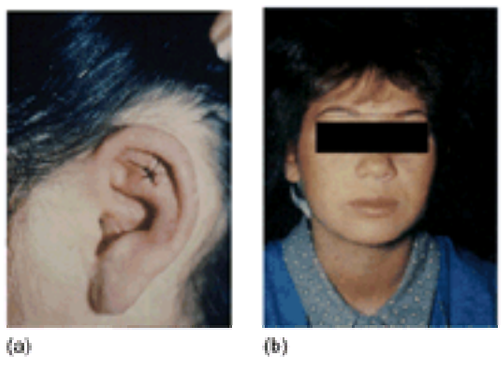


Fig. 7 Clinical features of relapsing polychondritis. (a) Acute inflammation of the external ear. (b) Saddle nose deformity. (By courtesy of Dr A. Balsa, La Paz Hospital, Madrid.)

Joint involvement is also common and occurs independently of other manifestations. The typical clinical picture is episodic, asymmetrical inflammation of both large and small joints, including the parasternal and sacroiliac joints, lasting several days to weeks. Generally it is non-deforming, non-erosive, and seronegative for rheumatoid factor (O'Hanlan *et al*. 1976; Balsa *et al*. 1995). Radiographs have demonstrated narrowing of the joint space without erosion, presumably reflecting the pathological process of loss of hyaline cartilage only (Booth *et al*. 1989). Aspirated synovial fluid is non-inflammatory. Mitchell and Shepard reported that, in addition to non-specific proliferation of lining cells, a unique feature was the presence in the middle and deeper layers of the synovium of large, clear spaces with projecting multinucleate bodies surrounding chondrocytes (Mitchell and Shepard 1972). In contrast to the 'pure' polyarthritis of relapsing polychondritis, the condition can develop against the background of a well-defined, erosive polyarthritis such as rheumatoid arthritis.

A wide range of rather non-specific ocular manifestations occur in relapsing polychondritis. The most common is episcleritis but more severe involvement includes scleritis and peripheral corneal thinning, both of which can lead to perforation, uveitis, retinal vasculitis, and optic neuritis, any of which can lead to blindness (Hoang-Xuan *et al*. 1989). Also reported are palsy of ocular muscles, orbital inflammation, and papilloedema.

The disease affects the respiratory tract in at least one-half of the patients. This can lead to the breakdown of tracheal and bronchial cartilage, with resulting airway collapse during the respiratory cycle (Crockford and Kerr 1988). The larynx and upper trachea are frequently involved first, and symmetrical subglottic narrowing is a common finding. Eventually the process can involve the distal trachea and main bronchi. Symptoms of respiratory tract involvement include dysphonia, cough, stridor, and dyspnoea; this involvement can dominate the clinical picture. During the active phase there may be tenderness over the thyroid cartilage and trachea. Young patients presenting with involvement of the upper respiratory tract early in the course of the disease tend to be rather resistant to treatment and have a poor prognosis (Neilly *et al*. 1985).

Involvement of the inner ear tends to occur later in the course of the disease and leads to vestibular and auditory dysfunction. A wide range of non-specific skin lesions has been reported, including erythema nodosum and leucocytoclastic vasculitis. The MAGIC (mouth and genital ulcers with inflamed cartilage) syndrome describes an overlap with Behçet's syndrome characterized by prominent orogenital ulceration (Orme *et al*. 1990). Renal manifestations have been reported in as many as 20 per cent of cases (Chang-Miller *et al*. 1987). On renal biopsy the predominant lesions are reported to be mesangial proliferation and segmental, necrotizing glomerulonephritis with crescents. Immunofluorescence and electron-microscopic studies show faint deposition of immunoglobulin and complement, and small amounts of electron-dense deposits, respectively, mainly in the mesangium. Patients with renal involvement tend to have more severe disease with extrarenal vasculitis and worse prognosis.

The heart is involved in about 10 per cent of patients. Aortic incompetence, either from dilatation of the aortic root or valvular destruction, is the most common manifestation. It is often severe, progressing even when other aspects of the disease are in remission (Buckley and Ades 1992), and requires valve replacement in at

least a third (Manna *et al*. 1985). Even when valve replacement is successful, there is the possibility of future dehiscence of the prosthetic valve because of continuing inflammation of perivalvular tissues (Lang-Lazdunski *et al*. 1995). Other cardiovascular manifestations include pericarditis, myocarditis, heart block, coronary vasculitis leading to myocardial infarction, and aneurysms of the aorta and other large arteries. Vasculitis involving large and medium-sized arteries may occur independently of cardiac involvement and often carries a poor prognosis. Meningioencephalitis has been reported in association with polychondritis but central nervous system involvement is rare (Hanslik *et al*. 1994)

Laboratory abnormalities are non-specific and for the most part reflect chronic inflammation. Common features are an acute-phase response, anaemia of chronic disease, and thrombocytosis. There may be a moderate leucocytosis. The development of a myelodysplastic syndrome has been reported (Diebold *et al*. 1995). Serological tests demonstrate anti-type II collagen antibodies in up to one-half of patients, circulating immune complexes in the majority, and antinuclear antibodies in approximately 20 per cent. ANCA have been reported in 24 per cent of patients with polychondritis, mainly during the acute phase of the disease, and in some patients in association with vasculitis (Papo *et al*. 1993; Handrock and Gross 1993). Various imaging techniques are useful for assessing complications of the disease. For example plain radiographs of the respiratory tract should include a soft tissue exposure of the neck in the lateral projection to demonstrate the larynx and upper trachea, together with penetrated frontal views of the trachea and major bronchi in the chest (Crockford and Kerr 1988). CT scanning is an accurate method for assessing the upper and lower airways (Mendelson *et al*. 1985; Davis *et al*. 1989) (Fig. 8). The value of MRI in assessing disease involvement, especially of the upper respiratory tract, is also reported (Fornadely *et al*. 1995).

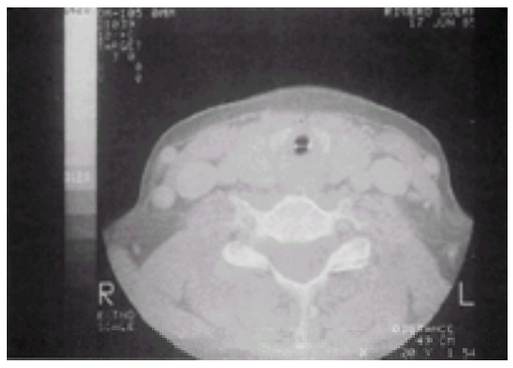


Fig. 8 A CT scan showing cartilage destruction and marked narrowing of the trachea. (By courtesy of Dr A. Balsa, La Paz Hospital, Madrid.)

The clinical course of this disease is highly variable. It is probable that the literature emphasizes the worst end of the spectrum and that many milder cases go unrecognized. It has been suggested that the 5-year and 10-year survival after diagnosis are 74 and 55 per cent (Michet *et al*. 1986). The most common causes of death are infection, cardiac and respiratory involvement, and systemic vasculitis.

Diagnosis

Early diagnosis is important in an attempt to prevent potentially life-threatening complications. There are no universally accepted diagnostic criteria but McAdam *et al*. proposed that the presence of three of the six following clinical features should be diagnostic (McAdam *et al*. 1976):

1. recurrent chondritis of both auricles;
2. non-erosive polyarthritis;
3. chondritis of the nasal cartilage;
4. ocular inflammation including conjunctivitis, keratitis, scleritis/episcleritis, uveitis;
5. involvement of laryngeal and/or tracheal cartilage;
6. cochlear and/or vestibular involvement.

These were modified slightly by Damiani and Levine (Damiani and Levine 1979) to include:

1. three or more of the above criteria;
2. at least one clinical criterion plus histological confirmation;
3. chondritis in two or more separate anatomical locations with a response to treatment.

Treatment

The rarity of relapsing polychondritis, the diversity of its presentation, and the unpredictability of recurrences make it difficult to recommend a particular treatment protocol. There are no controlled trials therefore the following is based on anecdotal experience.

Mild symptoms of joint involvement and inflammation of ear and nose cartilages can sometimes be controlled by non-steroidal anti-inflammatory drugs alone. Dapsone has been reported to be effective for systemic manifestations not controlled with symptomatic treatment alone (Barranco *et al*. 1976). Corticosteroids can be effective in suppressing acute, severe manifestations. High doses (at least 1 mg/kg per day of prednisolone or equivalent) are required. Alternate-day regimens are not recommended but there are reports of bolus parenteral methylprednisolone being effective for manifestations such as acute airways' obstruction (Lipnick and Fink 1991). It appears that corticosteroids are more effective in treating disease in cartilage of the respiratory tract than, for example, involvement of the joints and eyes. Also, there is no evidence that corticosteroids influence long-term outcome, and in some patients the disease undoubtedly progresses despite giving corticosteroids (Kilman 1978).

In severe cases, the addition of an immunosuppressive agent is required to control disease activity and to be steroid sparing. Successful responses have been reported using agents such as azathioprine and cyclophosphamide. The latter is the drug of choice for features such as necrotizing scleritis and systemic vasculitis. In refractory cases there have been anecdotal reports of success with cyclosporin A (Svenson *et al*. 1984) and monoclonal antibodies to CD4 (Van der Lubbe *et al*. 1991).

Additional supportive measures may be needed for those with severe respiratory, cardiac, and renal complications. For example whereas only 14 per cent of patients with this disease present initially with respiratory tract involvement, 80 per cent of these will require tracheostomy. Initially, this is usually for glottic, laryngeal, and subglottic inflammation and oedema producing airways' obstruction (McAdam *et al*. 1976). Later, the indication for tracheostomy is often the collapse of laryngeal or tracheal cartilages. The recent introduction of expandable metal stents represents an important development in managing patients with major airway stenosis and collapse (Shah *et al*. 1995).

Chapter References

Alarcon-Segovia, D. *et al*. (1989). Antiphospholipid antibodies and the antiphospholipid syndrome in systemic lupus erythematosus: a prospective study of 500 consecutive patients. *Medicine*, **68**, 353–65.

Alsalameh, S., Mollenhauer, J., Scheuplein, F., Stoss, H., Kalden, J.R., Burkhardt, H., and Burmester, G.R. (1993). Preferential cellular and humoral immune reactivities to native and denatured collagen types IX and XI in a patient with fatal relapsing polychondritis. *Journal of Rheumatology*, **20**, 1419–24.

Arundell, F.W. and Haserick, J.R. (1960). Familial chronic atrophic polychondritis. *Archives of Dermatology*, **82**, 439–41.

Assoun, J., Richardi, G., Railhac, J.J., Baunin, C., Fajadet, P., Giron, J., Maquin, P., Haddad, J., and Bonnevialle, P. (1994). Osteoid osteoma: MR imaging versus CT. *Radiology*, **191**, 217–23.

Balsa, A., Espinosa, A., Cuesta, M., MacLeod, T.I., Gijon-Banos, J., and Maddison, P.J. (1995). Joint symptoms in relapsing polychondritis. *Clinical and Experimental Rheumatology*, **13**, 425–30.

Barranco, V.P., Minor, D.P., and Solomon, H. (1976). Treatment of relapsing polychondritis with dapsone. *Archives of Dermatology*, **112**, 1286–8.

Bergstein, J.M., Wiens, C., Fish, A.J., Vernier, R., and Michael, A. (1974). Avascular necrosis of bone in systemic lupus erythematosus. *Journal of Pediatrics*, **85**, 31–6.

Booth, A., Dieppe, P.A., Goddard, P.L., and Watt, I. (1989). The radiological manifestations of relapsing polychondritis. *Clinical Radiology*, **40**, 147–9.

Boriani, S. (1980). Brodie's abscess: a study of 181 cases with special reference to radiographic diagnostic criteria. *Italian Journal of Orthopaedic Traumatology*, **6**, 373–83.

Buckley, L.M., and Ades, P.A. (1992). Progressive aortic valve inflammation occurring despite apparent remission of relapsing polychondritis. *Arthritis and Rheumatism*, **35**, 812–14.

Bullough, P.G. and DiCarlo, E.F. (1990). Subchondral avascular necrosis: a common cause of arthritis. *Annals of the Rheumatic Diseases*, **49**, 412–20.

Chang, C.C., Greenspan, A., and Gershwin, M.E. (1993). Osteonecrosis: current perspectives on pathogenesis and treatment. *Seminars in Arthritis and Rheumatism*, **23**, 47–69.

Chang-Miller, A. *et al*. (1987). Renal involvement in relapsing polychondritis. *Medicine (Baltimore)*, **66**, 202–17.

Christensen, J.H. and Poulsen, J.O. (1975). Synovial chondromatosis. *Acta Orthopaedica Scandinavica*, **46**, 919–25.

Clark, L.J., McCormick, P.W., Domenico, D.R., and Savory, L. (1993). Pigmented villonodular synovitis of the spine. *Journal of Neurosurgery*, **79**, 456–9.

Coolican, M.R. and Dandy, D.J. (1989). Arthroscopic management of synovial chondromatosis of the knee. Findings and results in 18 cases. *Journal of Bone and Joint Surgery*, **71B**, 498–500.

Cornell, C.N., Salvati, E., and Pincus, P.M. (1985). Longterm followup of total hip replacement in patients with osteonecrosis. *Orthopaedic Clinics of North America*, **16**, 757–69.

Crockford, M.P. and Kerr, I.H. (1988). Relapsing polychondritis. *Clinical Radiology*, **39**, 386–90.

Cronin, M.E. (1988). Musculoskeletal manifestations of systemic lupus erythematosus. *Rhuematic Disease Clinics of North America*, **14**, 99–116.

Curtin, W.A., Lahoti, O.P., Fogarty, E.E., Dowling, F.E., and Regan, B.F. (1993). Pigmented villonodular synovitis arising from the sheath of the extensor hallucis longus in an eight-month-old infant. *Clinical Orthopaedics and Related Research*, **292**, 282–4.

Dahlin, D.C. and Unni, K.K. (1986). *Bone tumours: general aspects and data on 8542 cases* (4th edn), pp. 88–102. Thomas, Springfield, IL.

Damiani, J.M. and Levine, H.L. (1979). Relapsing polychondritis: report of ten cases. *Laryngoscope*, **89**, 929–44.

Davis, S.D., Berkmen, Y.M., and King, T. (1989). Peripheral bronchial involvement in relapsing polychondritis: demonstration by thin-section CT. *American Journal of Roetgenology*, **153**, 953–4.

De Berg, J.C., Pattynama, P.M., Obermann, W.R., Bode, P.J., Vielvoye, G.J., and Taminiau, A.H. (1995). Percutaneous computed-tomography-guided thermocoagulation for osteoid osteomas. *Lancet*, **346**, 350–1.

Diebold, L., Rauh, G., Jager, K., and Lohrs, U. (1995). Bone marrow pathology in relapsing polychondritis: high frequency of myelodysplastic syndromes. *British Journal of Haematology*, **89**, 820–30.

Docken, W.P. (1979). Pigmented villonodular synovitis: a review with illustrative case reports. *Seminars in Arthritis and Rheumatism*, **9**, 1–22.

Dubois, E.L. and Cozen, L. (1960). Avascular (aseptic) bone necrosis associated with systemic lupus erythematosus. *Journal of the American Medical Association*, **174**, 966–84.

Ebringer, R., Rook, G., Swana, G.T., Botazzo, G.F., and Doniach, D. (1981). Autoantibodies to cartilage and type II collagen in relapsing polychondritis and other rheumatic diseases. *Annals of the Rheumatic Diseases*, **40**, 473–9.

Feldmann, J.L., Alcalay, M., Queinnec, J.Y., and De Bray, J.M. (1988). Spinal cord compression related to vertebral osteonecrosis. *Clinical Experimental Rheumatology*, **6**, 297–300.

Felson, D.T. and Anderson, J.J. (1987). A cross-study evaluation of association between steroid dose and bolus steroids and avascular necrosis of bone. *Lancet*, **i**, 902–5.

Ficat, R.P. and Arlet, J. (1980). *Ischaemia and necrosis of bone*. Williams and Wilkins, Baltimore.

Flandry, F. and Hughston, J.C. (1987). Pigmented villonodular synovitis. *Journal of Bone and Joint Surgery*, **69A**, 942–9.

Flandry, F.C., Hughston, J.C., McCann, S.B., and Kurtz, D.M. (1994 *a*). Diagnostic features of diffuse pigmented villonodular synovitis of the knee. *Clinical Orthopaedics and Related Research*, **298**, 212–20.

Flandry, F.C., Hughston, J.C., Jacobson, K.E., Barrack, R.L., McCann, S.B., and Kurtz, D.M. (1994 *b*). Surgical treatment of diffuse pigmented villonodular synovitis of the knee. *Clinical Orthopaedics and Related Research*, **300**, 183–92.

Fornadely, J.A., Seibert, D.J., Ostrov, B.E., and Warren, W.S. (1995). The role of MRI when relapsing polychondritis is suspected but not proven. *International Journal of Pediatric Otorhinlaryngology*, **31**, 101–7.

Franchi, A., Frosini, P., and Santoro, R. (1994). Pigmented villonodular synovitis of the temporomandibular joint. *Journal of Laryngology and Otology*, **108**, 166–7.

Franssen, M.J.A.M., Boerbooms, A.M. Th., Karthaus, R.P., Buijs, W.C.A.M., and van de Putte, L.B.A. (1989). Treatment of pigmented villonodular synovitis of the knee with yttrium-90 silicate: prospective evaluations by arthroscopy, histology, and ^{99m}Tc pertechnetate uptake measurements. *Annals of the Rheumatic Diseases*, **48**, 1007–13.

Goldman, A.B. and DiCarlo, E.F. (1988). Pigmented villonodular synovitis. Diagnosis and differential diagnosis. *Radiologic Clinics of North America*, **26**, 1327–47.

Green, W.T. and Banks, H.H., (1990). The classic osteochondritis dissecans in children. *Clinical Orthopaedics and Related Research*, **255**, 3–12.

Habermann, E.T. and Friedenthal, R.B. (1978). Septic arthritis associated with avascular necrosis of the femoral head. *Clinical Orthopaedics*, **134**, 325–31.

Halland, A.M., Klemp, P., Botes, D., Van Heerden, B.B., Loxton, A., and Scher, A.T. (1993). Avascular necrosis of the hip in systemic lupus erythematosus: the role of magnetic resonance imaging. *British Journal of Rheumatology*, **32**, 972–6.

Handrock, K. and Gross, W.L. (1993). Relapsing polychondritis as a secondary phenomenon of primary systemic vasculitis. *Annals of the Rheumatic Diseases*, **52**, 895–6.

Hanslik, T., Wechsler, B., Piette, J.C., Vidailhet, M., Robin, P.M., and Godeau, P. (1994). Central nervous system involvement in relapsing polychondritis. *Clinical and Experimental Rheumatology*, **12**, 539–41.

Harisdangkul, V. and Johnson, W.W. (1994). Association between relapsing polychondritis and systemic lupus erythematosus. *Southern Medical Journal*, **87**, 753–7.

Helms, C.A., Hattner, R.S., and Vogler, J.B. (1984). Osteoid osteoma: radionuclide diagnosis. *Radiology*, **151**, 779–84.

Herman, J.H. and Dennis, M.V. (1973). Immunopathological studies in relapsing polychondritis. *Journal of Clinical Investigation*, **52**, 549–58.

Hoang-Xuan, T., Foster, C.S., and Rice, B.A. (1989). Scleritis in relapsing polychondritis: response to therapy. *Ophthalmology*, **97**, 892–8.

Holman, A.J., Gardner, G.C., Richardson, M.L., and Simkin, P.A. (1995). Quantitative magnetic resonance imaging predicts clinical outcome of core decompression for osteonecrosis of the femoral head. *Journal of Rheumatology*, **22**, 1929–33.

Hungerford, D.S. (1989). Response: the role of core decompression in the treatment of ischaemic necrosis of the femoral head. *Arthritis and Rheumatism*, **32**, 801–6.

Jaffe, H.L., Lichtenstein, L., and Sutro, C.J. (1941). Pigmented villonodular synovitis, bursitis and tenosynovitis. *Archives of Pathology*, **31**, 731–65.

Jaksch-Wartenhorse, R. (1923). Polychondropathia. *Wiener Archiv für Innere Medizin*, **6**, 93–4.

Johansson, J.E., Ajjoub, S., Coughlin, L.P., Wener, J.A., and Cruess, R.L. (1982). Pigmented villonodular synovitis of joints. *Clinical Orthopaedics*, **163**, 159–66.

Kiers, L., Shield, L.K., and Cole, W.G. (1990). Neurological manifestations of osteoid osteoma. *Archives of Disease in Childhood*, **65**, 851–5.

Kilman, W.J. (1978). Narrowing of the airway in relapsing polychondritis. *Radiology*, **126**, 373–6.

Kitridou, R.C. Wittmann, A-L., and Quismorio, F.P. (1987). Chrondritis in systemic lupus erythematosus: clinical and immunopathological studies. *Clinical and Experimental Rheumatology*, **5**, 349–53.

Klippel, J.H., Gerber, L.H., and Pollak, I. (1979). Avascular necrosis in systemic lupus erythematosus: silent symmetric osteonecrosis. *American Journal of Medicine*, **67**, 83–8.

Lang, B., Rothenfusser, A., Lanchbury, J.S., Rauh, G., Breedveld, F.C., Urlacher, A., Albert, E.D., Peter, H.H., and Melchers, I. (1993). Susceptibility to relapsing polychondritis is associated with HLA-DR4. *Arthritis and Rheumatism*, **36**, 660–4.

Lang-Lazdunski, L., Hvass, U., Paillole, C., Pansard, Y., and Langlois, J. (1995). Cardiac valve replacement in relapsing polychondritis. A review. *Journal of Heart Valve Disease*, **4**, 227–35.

Laroche, M., Arlet, J., and Mazieres, B. (1990). Osteonecrosis of the femoral and humeral heads after intraarticular corticosteroid injection. *Journal of Rheumatology*, **17**, 549–51.

Learmonth, I.D., Maloon, S., and Dall, G. (1990), Core decompression for early atraumatic osteonecrosis of the femoral head. *Journal of Bone and Joint Surgery (British volume)*, **72**, 387–90.

Leventhal, G.H. and Dorfman, H.D. (1974). Aseptic necrosis of bone in systemic lupus erythematosus. *Seminars in Arthritis and Rheumatism*, **4**, 73–93.

Lipnick, R.N. and Fink, C.W. (1991). Acute airway obstruction in relapsing polychondritis: treatment with pulse methylprednisolone. *Journal of Rheumatology*, **18**, 98–9.

Liu, S-L. and Ho, T-C. (1991). The role of venous hypertension in the pathogenesis of Legg-Perthes disease. *Journal of Bone and Joint Surgery*, **73A**, 194–200.

McAdam, L.P., O'Hanlan, M.A., Bluestone, R., and Pearson, C.M. (1976). Relapsing polychondritis: prospective review of 23 patients and review of the literature. *Medicine (Baltimore)*, **55**, 193–215.

McKenna, C.H., Luthra, M.S., and Jordan, R.E. (1976). Hypocomplementaemic ear effusion in relapsing polychondritis. *Mayo Clinic Proceedings*, **51**, 495–8.

Makley, J.T. and Dunn, M.G. (1982). Prostaglandin synthesis by osteoid osteoma. *Lancet*, **ii**, 42–5.

Mankin, H.J. (1992). Nontraumatic necrosis of bone (osteonecrosis). *New England Journal of Medicine*, **326**, 1473–9.

Manna, R. *et al*. (1985). Relapsing polychondritis with severe aortic insufficiency. *Clinical Rheumatology*, **4**, 474–80.

Mendelson, D.S., Son, P.M., Crane, R., Cohen, B.A., and Spiera, H. (1985). Relapsing polychondritis studied by computed tomography. *Radiology*, **157**, 489–90.

Michet, C.J. Jr., McKenna, C.H., Luthra, H.S., and O'Fallen, W.M. (1986). Relapsing polychondritis: survival and predictive roles of early disease manifestations. *Annals of Internal Medicine*, **104**, 74–8.

Migliaresi, S., Picillo, U., Ambrosone, L., Di Palma, G., Mallozzi, M., Tesone, E.R., and Tirri, G. (1994) Avascular osteonecrosis in patients with SLE: relation to corticosteroid therapy and anticardiolipin antibodies. *Lupus*, **3**, 37–41.

Mitchell, N. and Shepard, N. (1972). Relapsing polychondritis: an electron microscopic study of synovium and articular cartilage. *Journal of Bone and Joint Surgery*, **54**, 1235–45.

Mont, M.A., Maar, D.C., Urquhart, M.W., Lennox, D., and Hungerford, D.S. (1993). Avascular necrosis of the humeral head treated by core compression. A retrospective review. *Journal of Bone and Joint Surgery (British volume)*, **75**, 785–8.

Motohashi, M., Morii, T., and Koshino, T. (1991). Clinical course of roentgenographic changes of osteonecrosis on the femoral condyle under conservative treatment. *Clinical Orthopedics*, **266**, 156–61.

Musculo, D.L., Velan, O., Pineda-Acero, G., Ayerza, M.A., Calabrese, M.E., and Santini Araujo, E. (1995). Osteoid osteoma of the hip. Percutaneous resection guided by computed tomography. *Clinical Orthopaedics and Related Research*, **310**, 170–5.

Myers B.W., Masi, A.T., and Feigenbaum, S.L. (1980). Pigmented villonodular synovitis and tenosynovitis. A clinical epidemiologic study of 166 cases and literature reivew. *Medicine*, **59**, 223–38.

Nagasawa, K., Tsukamoto, H., Tada, Y., Mayumi, T., Satoh, H., Onitsuka, H., Kuwabara, Y., and Niho, Y. (1994). Imaging study on the mode of development and changes in avascular necrosis of the femoral head in systemic lupus erythematosus: long-term observations. *British Journal of Rheumatology*, **33**, 343–7.

Neilly, J.B., Wister, J.H., and Stevenson, R.D. (1985). Progressive tracheo-bronchial polychondritis: need for early diagnosis. *Thorax*, **40**, 78–9.

O'Hanlan, M., McAdam, L.P., Bluestone, R., and Pearson, C.M. (1976). The arthropathy of relapsing polychondritis. *Arthritis and Rheumatism*, **19**, 191–4.

Orme, R.L. Nordlund, J.J., Barich, L., and Brown, T. (1990). The MAGIC syndrome (mouth and genital ulcers with inflamed cartilage). *Archives of Dermatology*, **126**, 940–4.

Paes, R.A. (1989). Familial osteochondritis dissecans. *Clinical Radiology*, **40**, 501–4.

Papo, T., Piette, J.C., Du, L.T.H., Godeau, P., Meyer, O., Kahn, M.F., and Bougeous, P. (1993). Antineutrophil cytoplasmic antibodies in polychondritis. *Annals of the Rheumatic Diseases*, **52**, 384–6.

Pietrogrande, V. and Mastromarino, R. (1957). Osteopat da prolongata trattamento cortisonico. *Ortopedia e Traumatologia dell' Apparato Matore*, **25**, 791–810.

Rao, A.S. and Vigorita, V.J. (1984). Pigmented villonodular synovitis (giant cell tumour of tendon sheath and synovial membrane): a review of 81 cases. *Journal of Bone and Joint Surgery*, **66A**, 76–94.

Ray, R.A., Morton, C.A., Lipinski, K.K., Corson, J.M., and Fletcher, J.A. (1990). Cytogenetic evidence of clonality in a case of pigmented villonodular synovitis. *Cancer*, **67**, 121–5.

Resnick, D. and Niwayama, G. (1988). *Diagnosis of bone and joint disorders* (2nd edn), Vol. 6. W. B. Saunders, Philadelphia.

Rosenthal, D.I., Springfield, D.S., Gebhardt, M.C., Rosenberg, A.E., and Mankin, H.J. (1995). Osteoid osteoma: percutaneous radio-frequency ablation. *Radiology*, **197**, 451–4.

Rosenwasser, M.P., Garino, J.P., Kiernan, H.A., and Michelsen, C.B. (1994). Long term followup of thorough debridement and cancellous bone grafting of the femoral head for avascular necrosis. *Clinical Orthopaedics and Related Research*, **306**, 17–27.

Sartoris, D.J. (1986). Computed tomography with multiplanar reformation and 3-dimensional image analysis in the preoperative evaluation of ischaemic necrosis of the femoral head. *Journal of Rheumatology*, **13**, 153–60.

Saxne, T., and Heinegard, D. (1995). Serum concentrations of two cartilage matrix proteins reflecting different aspects of cartilage turnover in relapsing polychondritis. *Arthritis and Rheumatism*, **38**, 294–6.

Schumacher, H.R., Lotke, P., Athreya, B., and Rothfuss, S. (1982). Pigmented villonodular synovitis: light and electron microscopic studies. *Seminars in Arthritis and Rheumatism*, **12**, 32–43.

Shah, R., Sabanathan, S., Mearns, A.J., and Featherstone, H. (1995). Self-expanding tracheobronchial stents in the management of major airway problems. *Journal of Cardiovascular Surgery*, **36**, 343–8.

Shpitzer, T., Ganel, A., and Engelberg, S. (1990). Surgery for synovial chondromatosis: 26 cases followed up for 6 years. *Acta Orthopaedica Scandinavica*, **61**, 567–9.

Smith, J.H. and Pugh, D.G. (1962). Roentgenographic aspects of articular pigmented villonodular synovitis. *American Journal of Roentgenology*, **87**, 1146–56.

Steinberg, M.E., Hayken, G.D., and Steinberg, D.R. (1984). A new method for evaluation and staging of avascular necrosis of the femoral head. In *Bone circulation*, (ed. J. Arlet, R.P. Ficat, and D.S. Hungerford), pp. 398–403. Williams and Wilkins, Baltimore.

Steinberg, M.E., Brighton, C.T., Hayken, G.D., Tooze, S.E., and Steinberg, D.R. (1985). Electrical stimulation in the treatment of osteonecrosis of the femoral head. *Orthopedic Clinics of North America*, **16**, 747–56.

Sugano, N., Ohzono, K., Masuhara, K., Takaoka, K., and Ono, K. (1994). Prognostication of osteonecrosis of the femoral head in patients with systemic lupus erythematosus by magnetic resonance

imaging. *Clinical Orthopaedics and Related Research*, **305**, 190–9.

Svenson, K.L.G. *et al*. (1984). Cyclosporin A treatment in a case of relapsing polychondritis. *Scandinavian Journal of Rheumatology*, **13**, 329–31.

Swee, R.G., McLeod, R.A., and Beabout, J.W. (1979). Osteoma detection, diagnosis, and localization. *Radiologic Diagnosis*, **13**, 117–23.

Tawn, D.J. and Watt, I. (1989). Bone marrow scintigraphy in the diagnosis of post traumatic avascular necrosis of bone. *British Journal of Radiology*, **62**, 790–5.

Terato, K. *et al*. (1990). Specificity of antibodies to type II collagen in rheumatoid arthritis. *Arthritis and Rheumatism*, **33**, 1493–1500.

Tishler, M., Caspi, D., and Yaron, M. (1987). Classical rheumatoid arthritis associated with nondeforming relapsing polychondritis. *Journal of Rheumatology*, **14**, 367–8.

Urbaniak, J.R., Coogan, P.G., Gunneson, E.B., and Nunley, J.A. (1995). Treatment of osteonecrosis of the femoral head with free vascularized fibular grafting. A long-term follow-up study of one hundred and three hips. *Journal of Bone and Joint Surgery (American volume)*, **77**, 681–94.

Valenzuela, R., Cooperrider, P.A., Gogate, P., Deodhar, S.D., and Berfeld, W.F. (1980). Relapsing polychondritis: immunomicroscopic findings in cartilage of ear biopsy specimens. *Human Pathology*, **11**, 19–24.

Van der Lubbe, P.A., Multenberg, A.M., and Breedveld, F.C. (1991). Anti-CD4 monoclonal antibody for relapsing polychondritis. *Lancet*, **337**, 1349.

Ward, W.G., Eckardt, J.J., Shayestehfar, S., Mirra, J., Grogan, T., and Oppenheim, W. (1993). Osteoid osteoma diagnosis and management with low morbidity. *Clinical Orthopaedics and Related Research*, **291**, 229–35.

Wollheim, F.A. (1984). Acute and long term complications of corticosteroid pulse therapy. *Scandinavian Journal of Rheumatology*, **54**, (Suppl.), 27–32.

Yang, C.L., Brinckmann, J., Rui, H.F., Vehring, K.H., Lehmann, H., Kekow, J., Wolff, H.H., Gross, W.L., and Muller, P.K. (1993). Autoantibodies to cartilage collagens in relapsing polychondritis. *Archives of Dermatological Research*, **285**, 245–9.

Zizic, T.M. (1990). Avascular necrosis of bone. *Current Opinion in Rheumatology*, **2**, 26–37.

Zizic, T.M., Marcoux, C., Hungerford, D.S., and Stevens, M.B. (1986). The early diagnosis of ischaemic necrosis of bone. *Arthritis and Rheumatism*, **29**, 1177–86.

5.17.4 Diseases of bone and cartilage in children

Barbara M. Ansell

The developmental skeletal disorders are difficult to classify and depend largely on radiological examination (Poznanski 1974; Wynne-Davies *et al*. 1985; Spranger 1992). Their importance lies in that they can present with musculoskeletal pain, alterations in joint movement and growth disturbances, so it is necessary for paediatricians and rheumatologists to be able to recognize such conditions.

The osteochondrodysplasias

There are more than 160 osteochondrodysplasias, divided into three groups: (i) defects of tubular and flat bone, (ii) disorganized development of cartilagenous and fibrous components of the skeleton, and (iii) idiopathic osteolyses. Features of some of the more common ones which give rise to problems follow.

Achondroplasia

This is the most common form of short-limbed dwarfism. Although inherited as an autosomal dominant trait, many cases represent new mutations. Arms and legs are short and such children usually have hypotonia, and with the head normal in size, may have difficulty in supporting it. There is an increased lumbar lordosis; spinal stenosis may result in neurological complications in adolescence or adult life. Hypochondroplasia resembles achondroplasia but lacks the facial characteristics.

Diatrophic dwarfism

This is characterized not only by short stature but also by cleft palate, thickening of the ear pinnas, severe club feet and thumb deformation, as well as joint dysplasia and scoliosis. It is thought to be inherited as an autosomal recessive trait.

Metatrophic dysplasias

In this, not only are short limbs present at birth and increasing loss of joint mobility, but there is also a kyphoscoliosis which is severe and progressive and cord compression is not uncommon. This disorder is an autosomal dominant one.

Epiphyseal dysplasias

Multiple epiphyseal dysplasia

One of the most common skeletal dysplasias which is inherited as an autosomal dominant trait (Maudsley 1955; Amir *et al*. 1985). It presents in childhood with pain and stiffness in affected limbs and is characterized by shortening and contractures and occasionally scoliosis. It can easily be confused with chronic arthritis (Patroni and Kredich 1985), but absence of inflammatory signs is important in its differentiation. Radiologically there are progressive irregularities of the end plates of the mid-thoracic vertebral bodies, short metacarpals and terminal phalanges and flattening, sclerosis, and fragmentation of the epiphyses of the hips, (Fig. 1) knees and other joints.

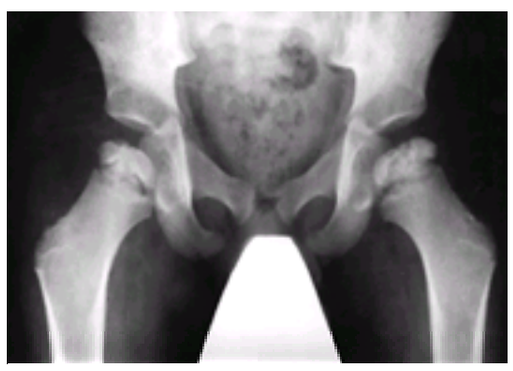


Fig. 1 A 10-year-old boy presenting with recurrent hip problems over 1 year; he had mild epiphyseal dysplasia, but fragmentation of the femoral heads is obvious.

Spondyloepiphyseal dysplasia

This is characterized by short stature with a disproportionately short trunk. Epiphyses of the hips and shoulders are the most severely affected early, and there may be platyspondyly.

Three forms are presently recognized—congenita, pseudoachondroplasia and tarda. The congenita form is inherited as an autosomal dominant; it is present at birth and is associated with myopia and retinal detachment (Anderson *et al*. 1990a). The pseudoachondroplasia may be inherited as an autosomal dominant or recessive and presents in early childhood (Hall *et al*. 1987). The patient's stature is short but with a normal face. There is marked irregularity of the epiphyses of the long bones which are late in maturing, and progressive scoliosis (Maroteaux *et al*. 1980). The tarda form may be inherited as an X-linked recessive (Iceton and Horne 1986) or be dominant, but it can be seen as a new mutation in two or three sibs (Szpiro-Tapio *et al*. 1988). There is little difference between these forms clinically as most look the same, however radiological changes may differ and a tentative classification has been suggested by Maroteaux and Spranger (1991).

There are a growing number of mutations in the gene for type II collagen (Col 2 A1), which include point mutations, deletions, insertions and splicing defects (Williams *et al*. 1993; Reginato *et al*. 1994). Unrelated families have been reported with Ang ^{75}Cys or Ang^{519}Cys mutation (Williams *et al*. 1993). Late-onset spondyloepiphyseal dysplasia is associated with precocious osteoarthritis, but in tall patients, has also been shown to be associated with type II procollagen gene (Col 2 A1) mutation in exon 11; recurrent mutations at a few specific sites of Col 2 A1 suggest there may be susceptibility 'hot spots' (Bleasel *et al*. 1995). However, not all cases are associated with primary defects of type II collagen (Anderson *et al*. 1990b).

Late-onset spondoepiphyseal dysplasia usually presents between the age of 3 and 10 years, but can be later, with pain in the knees and fingers followed by deformity and can simulate juvenile arthritis (Lewkonia and Beck-Hansen 1992); such children are noted to tire easily on activity and from an early age tend to find stairs difficult, presumably due to the poor development of the hips (Fig. 2). From time to time there may be acute episodes of pain, sometimes associated with swelling. It is not uncommon to have a family history of 'similar arthritis'. Clinically such children are usually of short stature; initially there is a suggestion of bony enlargement of the finger joints which becomes more obvious as deformity progresses (Fig. 3). This is followed by enlargement of the knee joints, particularly the patellas, with gradual loss of function of knees and hips (Wynne-Davies *et al*. 1982). All the acute phase reactants are normal as is the blood count, there are no immune complexes present; rheumatoid factor, antinuclear antibodies (ANA) and DNA, are negative. Radiological changes are characteristic and diagnostic (Fig. 4 and Fig. 5). Spranger (1983) described this condition as a progressive 'pseudojuvenile rheumatoid arthritis'. Pseudogout has also been noted (Bradley 1987). Spinal stenosis as a sequel is not uncommon, while total replacement arthroplasty of the hips for secondary degenerative changes may be required at an early age (Fig. 6). Hypoplasia of the odontoid process can also predispose the patient to cervical cord injury as a result of minor trauma.

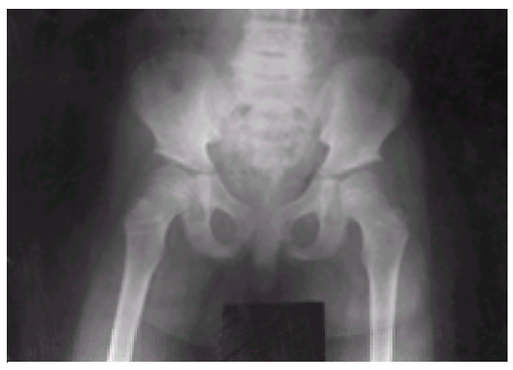


Fig. 2 This girl was described as normal at birth, but at about the age of 3 was noticed to have difficulty climbing up and down stairs; the hip radiographs revealed these poorly developed acetabula and persistence of anteversion of the hips: shortly after this, hand function became impaired and she was diagnosed as spondyloepiphyseal dysplasia.

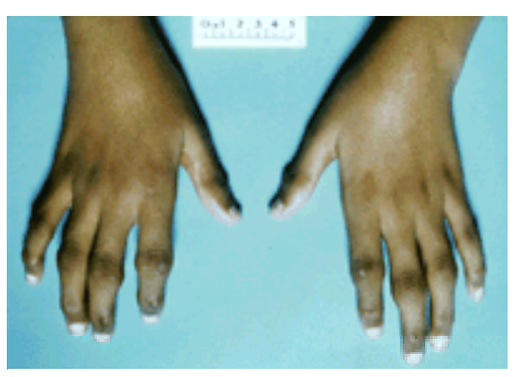


Fig. 3 This shows the clinical appearance of the hands in another case of spondyloepiphyseal dysplasia commencing at the age of 4. By the age of 12, there is obvious bony enlargement of the proximal interphalangeal joints associated with some loss of movement and slight bony enlargement at terminal interphalangeal joints.

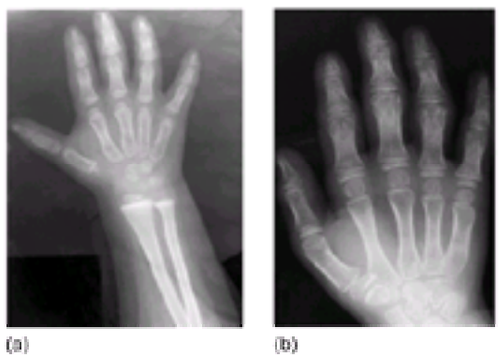


Fig. 4 (a) The initial hand radiograph of the patient in Fig. 2 was regarded as normal, although bony changes are starting in the phalanges. By the time bony enlargement was obvious (b) the radiograph shows marked widening of the phalanges at the proximal and distal interphalangeal joints, as well as alteration in texture and epiphyseal changes.

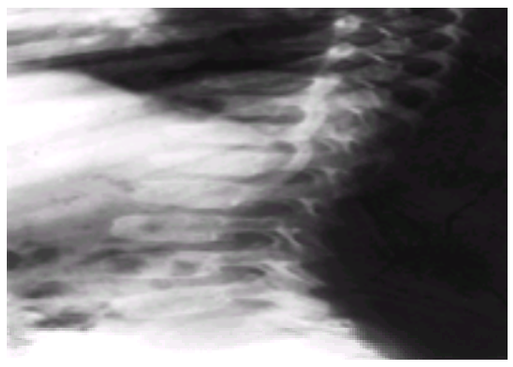


Fig. 5 Lateral radiograph of the spine of a 6-year-old presenting with lumbar lordosis, difficulty in using the hands and limited walking ability. Note the characteristic beaking in vertebral bodies, particularly in the lower thoracic spine, in this child with spondyloepiphyseal dysplasia, which had also affected his father.

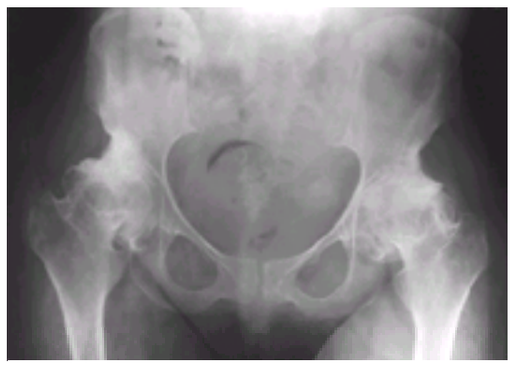


Fig. 6 The shape of the femoral heads, and the avascular necrosis and secondary degenerative changes that have occurred by 16 years of age in the patient illustrated in Fig. 2; this was just prior to bilateral total replacement arthroplasties.

Cartilage–hair hypoplasia

This is inherited as an autosomal recessive trait and can usually be diagnosed at birth as these children have achondroplastic features, the hands are short and fat, there is hyperextensibility of the joints with the exception of the elbows and the hair is fine and sparse. Radiographs show widened metaphyses with scalloping, irregular sclerosis and cystic defects (McKusick 1964). These children tend to have a T-cell immunodeficiency which predisposes them to infection (Harris *et al*. 1981; Polmar and Pierce 1986).

Dyggve–Melchior–Clausen syndrome

This appears late in an initially normal child and manifests as dwarfism, claw hands, sternal problems, lumbar lordosis and progressive mental retardation (Dyggve *et al*. 1962; Beighton 1990). Clinically it mimics Morquio's disease, but in addition there is mental retardation. This condition is probably heterogeneous since both autosomal recessive and X-linked recessive forms have been described.

Acro/acromesomelic dysplasia

Trichorhinophalangeal dysplasia

This is characterized by enlargement of the interphalangeal joints in particular (Fig. 7) which may cause confusion with juvenile chronic arthritis (Noltorp *et al*. 1986). Clinically the bulbous nose with hyperplastic nares, large ears, short brittle hair and short stature suggest the correct diagnosis (Gieodron *et al*. 1973). All acute phase reactants are normal and radiographs show cone-shaped epiphyses with short metacarpals and metatarsals and there may be fragmentation of femoral epiphyses. Many families have shown autosomal dominant inheritance. The chromosome 8q deletion is subtle in that only the very narrow 9-positive band 24.12 is missing (Fryns and Van den Berghe 1986; Bühler *et al*. 1987). There are considerable similarities to the Langer–Gieodron syndrome (TRP II) in which exostoses occur as well as mental retardation and in these children 8q 24.13 is deleted as well.

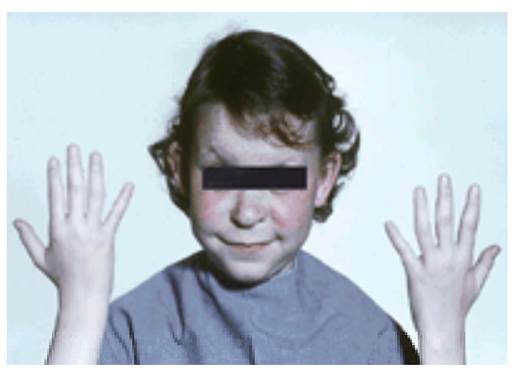


Fig. 7 Trichorhinophalangeal dysplasia; note the bony enlargement of proximal interphalangeal joints together with facies. (By courtesy of Dr Alan Craft, Newcastle.)

Storage diseases

The deficiency of a lysosomal degradative enzyme causes an accumulation of its substrate within the lysosomes of the cell; the tissue distribution of the enzyme deficiency determines its expression. Classification is still not entirely satisfactory.

Mucopolysaccharidoses

These are due to deficiency of enzymes involved in the metabolism of glycosaminoglycans (Table 1) (Beck *et al*. 1986; Whitley 1993). All are autosomal recessive disorders (Stanbury *et al*. 1989). A prominent feature is a progressive skeletal dysplasia affecting particularly the hands, hips and vertebrae. It is important to be aware that all these syndromes can present as joint problems, and the first clue to diagnosis may be a claw hand (Fisher *et al*. 1974), but other problems include stiffness of the shoulders, general stiffening and clumsiness in using joints and occasionally dislocation of the hips. In the more severe types such as Hurler's syndrome, there is a gradual coarsening of the facial features associated with dwarfism, mental retardation and clouding of the cornea. There is no difficulty in diagnosis when the fully developed features of Hurler's syndrome are present, but it may take up to 2 years for these to become obvious; flexion deformities of the fingers allow suspicion early, so that prenatal diagnosis of future sibs can be made.

Type		Urinary excretion	Enzyme deficiency	Gene	Locus
I	Hurler	Dermatan sulphate Heparan sulphate	α-L-Iduronidase	4p	16.3
I	Scheie	Dermatan sulphate	α-L-Iduronidase		
I	Hurler/Scheie	Dermatan sulphate Heparan sulphate			
II	Hunter	Dermatan sulphate Heparan sulphate	Iduronate-, sulphatase	Xq	27.3
II	Hunter/Scheie	Dermatan sulphate Heparan sulphate			
III	Sanfilippo A	Heparan sulphate	Heparan		
	B		α-(N-Acetyl)glucosaminidase		
	C	Heparan sulphate	Acetyl-CoA- α-glucosaminide		
	D	Heparan sulphate	N-Acetyl-glucosamine-6-sulphatase	12q	14
IV	Morquio-Brailsford A	Keratan sulphate Chondroitin-6 sulphate	Galactose-6-sulphatase		
	B		β-Galactosidase		
VI	Maroteaux-Lamy	Dermatan sulphate	N-Acetyl-galactosamine-4-sulphatase		
VII	Sly	Dermatan sulphate Heparan sulphate Chondroitin sulphate	β-Glucuronidase		

Table 1 Mucopolysaccharidoses

The two conditions that warrant special mention are Scheie and the Morquio–Brailsford syndrome, as these children are of normal intelligence and develop symptoms after the age of 2 or 3 years. In Scheie syndrome, where the disease is caused by a deficiency of the enzyme a-L-iduronidase (Scott *et al*. 1990), the stature is well preserved, but there is progressive stiffening of the joints of the hands (Fig. 8), elbows and knees without swelling and pain and with no evidence of inflammatory indices. Urinary excretion of dermatan sulphate is increased. Clouding of the cornea may mimic iridocyclitis. Later carpal tunnel compression and atlantoaxial subluxation can cause problems; cardiac difficulties involving the aortic and mitral valves are also seen (Butman *et al*. 1989).

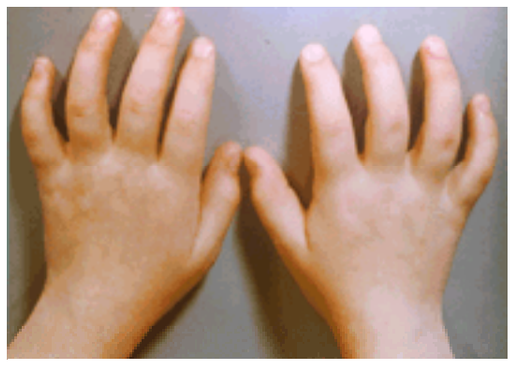


Fig. 8 Hands of child with Hunter–Scheie syndrome which had progressively stiffened from the age of 1 year.

A phenotype described by Roubicek *et al*. (1985), intermediate between Hurler and Scheie, has the coarse facial features and corneal clouding, skeletal changes are mild to severe, but mental retardation is mild.

In the Hunter–Scheie syndrome, stiffening of the shoulders associated with some degree of dwarfism occurs early as well as moderate skeletal involvement, but intellectual impairment of varying degree does not develop until later.

In the Morquio–Brailsford syndrome, two different enzyme deficiencies have been noted (Table 1), but symptoms and signs are similar. There is progressive musculoskeletal stiffening which usually begins about the age of 3 or 4 years and is associated with dwarfing. The hands tend to enlarge, valgus deformity of the knees gradually develops and the gait becomes stiff and waddling with some kyphosis and protrusion of the sternum. Initially as the joints enlarge they are hypermobile, but later become stiff. Characteristic radiological findings include platyspondyly and odontoid hypoplasia, and dysplastic hips with poorly developed acetabula. Radiologically they are distinct from the various forms of spondyloepiphyseal dysplasia. Progressive cervical spinal cord damage can occur (Lipson 1977). Clouding of the cornea is mild initially, but deafness may be an early problem. Rarely echocardiographic changes occur and aortic regurgitation develops in a proportion (John *et al*. 1990).

All these children need, not only to be protected against unnecessary therapy with long-acting drugs such as gold, penicillamine, or methotrexate, but also to be watched for complications such as carpal tunnel compression (Wraith and Alani 1990), atlantoaxial subluxation (Lipson 1977), and spinal cord compression (Blaw and Langer 1969).

Mucolipidoses

The term mucolipidosis is applied to a group of disorders that are characterized by the intracellular accumulation of both glycosaminoglycans and sphingolipids, but without excess urinary excretion of glycosaminoglycans. Progressive ocular and neurological abnormalities are common to all these disorders, which are autosomal recessive.

Mucolipidosis type I

Isolated neuraminidase deficiency (sialidase) produces a Hurler-like syndrome. The urinary excretion of sialated urinary oligosaccharides is markedly elevated.

Mucolipdosis type II

Sometimes called 'I cell disease', also causes a Hurler-like syndrome with progressive loss of joint movement. There are prominent intracytoplasmic (I) inclusions in lymphocytes and skin fibroblasts. Biochemically there is a deficiency of Glc- *N*-acetylglucosaminyl-phosphotransferase in fibroblasts and leucocytes (Cipolloni *et al*. 1980).

Mucolipidosis type III

This is sometimes known as pseudo-Hurler polydystrophy and is characterized by restriction of joint mobility, which does not become apparent until the second or third year of life. Such patients have presented to us as juvenile arthritis starting with difficulty in raising the shoulders (Fig. 9) and problems with grip (Fig. 10), and atypical carpal tunnel syndrome (Starreveld 1975), or with thickening in the hands and some tightening of the skin mimicking scleroderma. No inflammatory arthritis is present and routine blood tests are all normal. Mentality may be normal or slightly below standard. Radiographic findings are those of a dysotosis multiplex. The condition tends to stabilize in the teens. Ocular features include cloudy cornea, astigmatism, abnormalities of the retina and optic nerve, and visual field defects (Traboulsi and Maumence 1986). Aortic incompetence is not uncommon. The basic defect is in the enzyme that specifically phosphorylates mannose residues of lysosomal glycoproteins. Characteristic inclusions are found in cultured fibroblasts.

Fig. 9 This boy, who proved to have mucolipidosis type III, presented with difficulty in using his hands; he was found to have very limited shoulder movement, contractures at the elbow and the beginning of contractures at the knees, hips and feet.

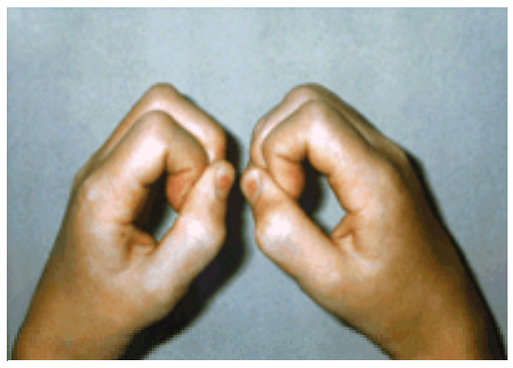


Fig. 10 Hands of the child from Fig. 9; note the inability to pinch with finger and thumb, and the thickening of the tissues due to deposition.

Sphingolipidoses

Here there is an accumulation of lipid in the cell as a result of specific enzyme deficiencies; only three of the many different sphingolipidoses have musculoskeletal signs and symptoms, notably Farber's lipogranulomatosis, Gaucher's disease, and Fabry's disease.

Farber's lipogranulamatosis

This is due to deficiency of acid ceramidase and is an autosomal recessive sphingolipidosis. Usually it begins in the neonatal period with a hoarse cry and irritability. Painful red masses develop along tendon sheaths and around joints, followed by the development of contractures. Delayed motor development and mental retardation are prominent. Epiglottal and laryngeal swelling cause repeated pulmonary infections, usually leading to death early in life. It has been suggested that there are milder types of this defect, and recent descriptions of older children presenting with joint deformity and some mental retardation are appearing (Jameson *et al.* 1987). These may confuse the rheumatologist.

Gaucher's disease

This results from a deficiency in the enzyme acid glucosidase and is characterized by an accumulation of glucosyl ceramide in reticuloendothelial cells in the bone marrow, spleen, liver, lymph nodes and other internal organs. Three clinical forms are now recognized. The most common, type 1, is the chronic non-neuropathic adult form, more prevalent in Ashkenazi Jews. Type 2 is the infantile or active neuropathic form, and death occurs before the age of 2 years. Type 3 is the juvenile or subacute neuropathic form, which occurs particularly among the Swedish, and is thought to be caused by a single mutation in exon 10 of the glucocerebrosidase gene (Dahl *et al*. 1990). Osteoarticular complaints are an important feature, particularly in type 3, where they may be the earliest manifestation of the disease and result from infiltration of the marrow of the subchondral bone. A common complaint is polyarthralgia affecting the large peripheral joints, pathological fracture of a long bone or compression of a vertebra which gives rise to pain. Severe degenerative hip disease as a result of avascular necrosis and collapse of the femoral head is not uncommon. The most frequent early radiological finding is widening of the distal portion of the femur just above the medial condyle, the Erlenmeyer flask deformity. Similar flaring may be present in the tibia and humerus (Peters *et al*. 1977; Stowens *et al*. 1985).

Fabry's disease (angiokeratoma corporis diffusum)

This is an X-linked recessive disorder characterized by the progressive accumulation of birefringent deposits of triglycosylceramide in the endothelial and smooth muscle cells of blood vessels and in ganglions and perineural cells of the autonomic nervous system. It results from a deficiency of ceramide trihexoside-a-galactosidase and ceramide trihexosidase, the gene for which has been localized to the Xq 22–23 region of the X chromosome (Sakuraba *et al*. 1990). The affected boys have recurrent attacks of fever and a severe arthritis with a burning and tingling pain in the extremities. This is aggravated by hot weather or exercise. The fingers and elbows may become swollen with difficulty in extending the fingers. The typical rash consists of purple papules—angiokeratoma diffusum universale. Ocular signs include opacification of the cornea in a whirl-like configuration. Secondary osteonecrosis may become increasingly important in weight-bearing joints such as the hips. Female heterozygotes may have a milder form of the disease, as did the mother in a case presenting with fever and lymphadenopathy (Mayou *et al*. 1989). Renal, cardiac and cerebral disease tend to cause death in the mid-adult years (Kramer *et al*. 1985; Sakuraba *et al*. 1986; Morgan *et al*. 1990).

Other rare disorders

Multicentric reticulohistiocytosis (see Chapter 5.13.7)

A disease sometimes referred to as 'lipoid dermatoarthritis'. As yet no biochemical abnormality has been defined, but foamy giant cells and histiocytes are found on biopsy of the skin, synovium and mucous membranes. It causes a mutilating, destructive arthropathy with cutaneous nodules often on the face as well as other sites (Zayid and Farroj 1973).

Winchester syndrome

This begins after the age of a few weeks up to one year with swelling of the proximal interphalangeal joints and enlargement of the wrists; later corneal clouding, coarsening of the face, and joint contractures appear (Winchester *et al*. 1969). Patchy thickening of skin with pigmentation are found over the back, flanks and lateral aspects of the arms. Other features are corneal opacities in mid-childhood, retarded growth, carpal-tarsal osteolysis and destruction of the small joints. Increased urinary oligosaccharide excretion is found (Lambert *et al*. 1989; Winter 1989).

Kniest syndrome

This is associated with congenital short limbs, and with a large head, round face and depressed nasal bridge. It is inherited usually as an autosomal dominant disorder, but sporadic cases have been seen (Kim *et al*. 1975). Stiffness of the fingers, dislocation of the hips and kyphoscoliosis develop; later enlargement of the joints and severe contractures occur (Fraya *et al*. 1979). The cartilage has hypertrophic chondrocytes surrounded by a loose matrix that contains large holes

resembling Swiss cheese, hence the name. Other abnormalities, notably cleft palate, vitreo-retinal degeneration and retinal detachment, deafness and hernia are characteristic. Other chondrodysplasias may have a similar appearance (Sconyers *et al*. 1985).

Moore–Federman disease

Small hands with brachydactyly are noted early, followed by contractures of the hand joints (Fig. 11) and dwarfism (Moore and Federman 1965). This may be the same condition as acromicric dysplasia (Winter *et al*. 1989).

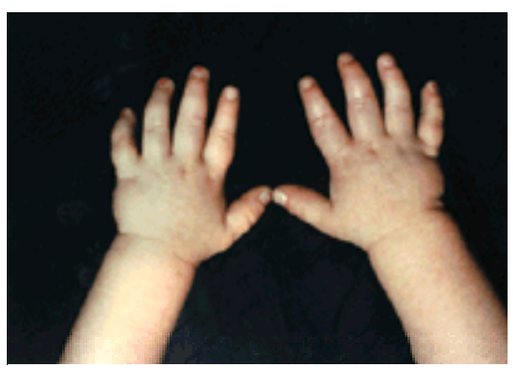


Fig. 11 This 7-year-old presented with some difficulty in gripping and possible swelling of the fingers. She was noted to be short. Her grandfather had very similar hands.

Thiemann's disease

This tends to commence at about the age of 10 or 11 years and is characterized by progressive enlargement of the proximal interphalangeal joints of the hands, the interphalangeal joints of the great toe and occasionally other toes, followed by slight flexion of the enlarged joints. Clinically there is bony swelling but no evidence of soft-tissue swelling and the erythrocyte sedimentation rate is normal. Radiologically there is irregularity of the epiphyses of the phalanges. This condition can be familial (Molloy and Hamilton 1978).

Idiopathic acro-osteolysis

Although destruction and disappearance of bone as a primary condition is rare, a number of different types have been described (Brown *et al*. 1976). Hereditary osteolysis is an autosomal dominant; it usually begins about the age of 3 years and affects children of both sexes (Naranjo *et al*. 1992). The bones of the carpus and tarsus (Urlus *et al*. 1993) are particularly affected, and to a lesser extent the hands and feet as well as elbows and knees. It presents with tender, swollen, limited wrists and ankles closely mimicking juvenile chronic arthritis, but there is no obvious synovitis and the erthrocyte sedimentation rate is usually normal (Beals and Bird 1975). Initially the radiographs are normal, but after a time there is porosis of the carpal bones followed by localized destruction (Fig. 12) and finally complete disappearance of the bone. This condition tends to stabilize spontaneously in early adult life. The Haydu–Cheney acro-osteolysis is characterized by facial as well as skeletal hanges. The eyes slope downwards, the philtrum is long and the nostrils anteverted. The skull radiograph shows wormian bones, an elongated pituitary fossa and a thickened cranium with persistent sutures; the classical acro-osteolysis only develops later in childhood. Inheritance is likely to be autosomal dominant, although most cases are sporadic (Macpherson and Pai 1989). Multicentric osteolysis has also been reported as an autosomal recessive trait with symptoms developing between the age of 2 and 5 years (Torg *et al*. 1969).

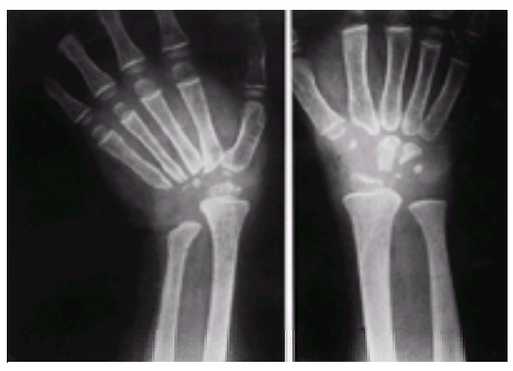


Fig. 12 This child presented some 3 years earlier with pain and swelling in the wrists and feet; note the tapering of the bases of the metacarpals, and absorption of carpal bones at a different rate on the two sides, together with the early absorption of the ulnar head on the left, due to idiopathic acro-osteolysis.

The Thieffry–Kohler form is similar to idiopathic multicentric osteolysis, but with a facial appearance characterized by frontal bossing, micrognathia, a small mouth and protruding eyes. It is usually inherited as an autosomal dominant trait. The joint manifestations begin in the wrists and ankles, leading to progressive deformity. It is not certain whether the form associated with progressive nephropathy resulting in hypertension and renal failure in early adult life is the same as Thieffry–Kohler (Shurtleff *et al*. 1964; Carnevale *et al*. 1987).

Phantom bone disease (Gorham's disease)

This is a non-hereditary form which occurs usually between the ages of 5 and 10 years (Gorham and Stout 1955). Any bone can be affected and there are usually multiple sites which are asymmetrical in distribution; histologically there is proliferation of thin-walled blood vessels.

Distal osteolysis

Thought to be inherited as an autosomal dominant trait, it is characterized by progressive osteolysis of the phalanges, metatarsals and metacarpals. It tends to occur in children of 8 years and upwards and often heals spontaneously, although sometimes part of a finger or toe may be lost (Elias *et al*. 1978).

Other oddities

Not all children who present with joint deformities can be characterized, for example the patient shown in Fig. 13 and Fig. 14, and all skeletal disorders mimicking arthritis have not been covered here, so cross-reference to other chapters is essential. Readers are also referred to Wynne-Davies *et al*. (1985) and Spranger (1992).

Fig. 13 This boy was first noted to have a left elbow contracture as a baby; this progressed and deformities were noted in the other elbow, wrists, knees, ankles and feet. The only relevant history was that his mother had had mumps during the pregnancy.

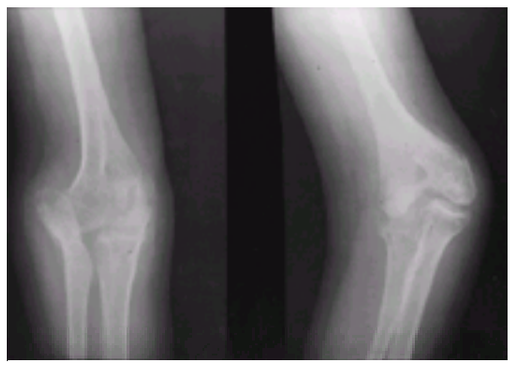


Fig. 14 Unusual bony abnormalities in the elbow radiograph from the patient shown in Fig. 13. The radial head on the left in particular is grossly abnormal.

Arthrogryposis

This is a symptom complex characterized by stiffness and contractures of the joints at birth. It can affect both the upper and lower limbs in about 50 per cent of cases, while in 40 per cent it is only the lower limbs and 10 per cent the upper limbs. Distal joints are often severely affected with the feet showing talipes equinovarus and the wrists severe flexion deformities. These are frequently associated with flexion contractures of the knees and fixed extended elbows (Lloyd-Roberts and Lettin 1970). The normal contours of the joints are lost as are the skin creases over them; there is absence or underdevelopment of surrounding muscles. The rigidity of joints is thought to be due to fibrosis. In addition, there may be congenital defects of the skeleton. The patients that give rise to problems in differentiation from juvenile arthritis are those with isolated limitation of movement, particularly at the hips, knees, or elbows (Fig. 15). It is important to recognize them early as improvement can be achieved by manipulation and splinting, while soft-tissue release of ligaments and tendons around the joints will help fixed deformities, but should be carried out in relatively young patients.

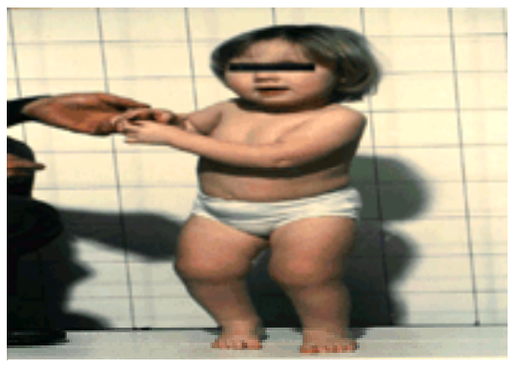


Fig. 15 Although well covered, it was considered that this baby, who had presented with contractures of the hips, knees and ankles at birth, was suffering from a limited form of arthrogryposis.

Kashin–Beck disease

This endemic arthritis with systemic or visceral manifestations has been reported from Eastern Russia, Siberia, Northern China, and Korea and appears to result from eating bread baked from fungus-infected grain (Nesterov 1964). Unknown toxic products cause an epiphyseal and metaphyseal dysplasia of the bones of the interphalangeal, wrist, knee, and ankle joints. This isorder continues to become more severe as long as the child lives in the endemic area eating such products, and will ultimately cause symmetric progressive limitation of movement involving multiple joints. The initial symptoms are in school-aged children and consist of aching and muscle weakness. Laboratory indices of inflammation are absent. The eventual dwarfing, bony dysplasia, and short digits resemble those seen in the lysosomal stores diseases. Experimental animals fed grain infected with *Fusarium* species have been shown to develop a similar form of dysplasia.

Mseleni disease

The population of the Mseleni area of Northern Zululand has a high incidence of a chronic polyarthritis, commencing in childhood or adolescence and characterized by a restriction of movement and limitation of mobility, with the hips, knees, and ankles as the predominant sites of involvement; mild stunting of growth is common (Lockitch *et al*. 1973). The characteristic radiograph abnormalities include irregularities on the surface of the epiphyses, change in shape and, ultimately, progress to osteoarthritis in the hips, often with protrusio acetabulae; short metacarpals and deformity at the distal end of the ulna may also occur. The aetiology is thought to involve a nutritional deficiency or toxin similar to the Kashin–Beck disease.

Chapter References

Amir, D., Mogle, P., and Weinberg, H. (1985). Multiple epiphyseal dysplasia in one family; a further review of seven generations. *Journal of Bone and Joint Surgery*, **67B**, 809–13.

Anderson, I.J. *et al*. (1990*a*). Spondyloepiphyseal dysplasia congenita: genetic linkage to type II collagen (COL2 A1). *American Journal of Human Genetics*, **46**, 896–901.

Anderson, I.J. *et al*. (1990*b*). Spondyloepiphyseal dysplasia, mild autosomal is not due to primary defects of type II collagen. *American Journal of Medical Genetics*, **37**, 272–6.

Beals, R.F. and Bird, G.B. (1975). Carpal and tarsal osteolysis: a case report and review of the literature. *Journal of Bone and Joint Surgery*, **57A**, 681–6.

Beck, M., Glossi, J., Grubisic, A., and Spranger, J. (1986). Heterogeneity of Morquio disease. *Clinical Genetics*, **29**, 325–31.

Beighton, P. (1990). Syndrome of the mouth. Dyggve–Melchior–Clausen syndrome. *Journal of Medical Genetics*, **27**, 512–15.

Blaw, M. and Langer, L.O. (1969). Spinal cord compresssion Morquio–Brailsford disease. *Journal of Pediatrics*, **74**, 593–600.

Bleasel, J.F. *et al*. (1995). Type II procollagen gene (COL 2A1) mutation in exon 11 associated with spondylo-epiphyseal dysplasia tarda. Tall stature and precocious osteoarthritis. *Journal of Rheumatology*, **22**, 255–61.

Bradley, J.D. (1987). Pseudo gout in progressive pseudo-rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **46**, 709–12.

Brown, D.M. *et al*. (1976). The acro-osteolysis syndrome, morphologic and biochemical studies. *Journal of Pediatrics*, **88**, 573–80.

Bühler, E.M., Bühler, U.K., Beutler, L., and Fessler, R. (1987). A final word on the trichorhino-phalangeal syndromes. *Clinical Genetics*, **31**, 273–5.

Butman, S.M., Karl, L., and Copeland, J.S. (1989). Combined aortic and mitral valve replacement in an adult with Scheie's disease. *Chest*, **96**, 209–10.

Carnevale, A. *et al*. (1987). Idiopathic multicentral osteolysis with facial anomalies and nephropathy. *American Journal of Medical Genetics*, **26**, 877–86.

Cipolloni, C. *et al*. (1980). Neonatal mucolipidosis II (I cell disease) clinical, radiological and biochemical studies in a case. *Helvetica Paediatrica Acta*, **35**, 85–95.

Dahl, N., Lagerstrom, M., Erikson, A., and Petterson, U. (1990). Gaucher disease type III (Norbottnian type) is caused by a single mutation in exon 10 of the glucocerebrosidase gene. *American Journal of Human Genetics*, **47**, 275–8.

Dyggve, H.V., Melchoir, J.C., and Clausen, J. (1962). Morquio–Ullrich's disease. *Archives of Disease in Childhood*, **37**, 525–34.

Elias, A.N. *et al*. (1978). Hereditary osteodysplasia and acro-osteolysis. *American Journal of Medicine*, **65**, 627–36.

Fisher, R.C., Horner, R.L., and Wood, V.E. (1974). The hand in mucopolysaccharide disorders. *Clinical Orthopaedics and Related Research*, **104**, 191–9.

Fraya, R., Malheim, R., and Idriss H. (1979). The Kniest (Swiss cheese cartilage) syndrome; description of a distinct arthropathy. *Arthritis and Rheumatism*, **22**, 286–9.

Fryns, J.P. and Van den Berghe, H. (1986). Bq 24.12 interstitial deletion in tricho-phalangeal syndrome type 1. *Human Genetics*, **74**, 188–9.

Gieodron, A. *et al*. (1973). Autosomal dominant transmission of the tricho-rhino-phalangeal syndrome. *Helvetica Paediatrica Acta*, **28**, 249–59.

Gorham, L.V. and Stout, A.P. (1955). Massive osteolysis (acute spontaneous absorption of bone, phantom bone, disappearing bone) in relationship to haemangiomatosis. *Journal of Bone and Joint Surgery*, **37A**, 985–1004.

Hall, J.G., Dorst, J.P., Rotta, J., and McKusick, Y.A. (1987). Gonadal mosaicisus in pseudo-achondroplasia. *American Journal of Medical Genetics*, **28**, 143–52.

Harris, R.E. *et al*. (1981). Cartilage–hair hypoplasia, defective T cell function and Diamond–Blackfan anaemia in an Amish child. *American Journal of Medical Genetics*, **8**, 291–7.

Iceton, J.A. and Horne, G. (1986). Spondylo-epiphyseal dysplasia tarda—the X-linked variety in three brothers. *Journal of Bone and Joint Surgery*, **68B**, 616–19.

Jameson, R.A., Holt, P.J.L., and Keen, J.H. (1987). Farber's disease (lysosomal acid ceramidase deficiency). *Annals of the Rheumatic Diseases*, **46**, 559–61.

John, R.M., Hunter, D., and Swanton, R.H. (1990). Echocardiographic abnormalities in type IV mucopolysaccharidoses. *Archives of Diseases in Childhood*, **65**, 746–9.

Kim, H.J. *et al*. (1975). Kniest syndrome with dominant inheritance and mucopolysaccharidosis. *American Journal of Human Genetics*, **27**, 755–64.

Kramer, W., Thormann, J., Meuller, K., and Frenzel, H. (1985). Progressive cardiac involvement by Fabry's disease despite successful renal allotransplantation. *International Journal of Cardiology*, **7**, 72–5.

Lambert, J.C. *et al*. (1989). Biochemical and ultrastructural studies of two familial cases of Winchester syndrome. *Journal de Genetique Humaine*, **37**, 231–6.

Lewkonia, R.M. and Beck-Hansen, N.T. (1992). Spondylo-epiphyseal dysplasia tarda simulating juvenile arthritis: clinical and molecular genetic observations. *Clinical and Experimental Rheumatology*, **10**, 411–14.

Lipson, S.J. (1977). Dysplasia of the odontoid process in Morquio's syndrome causing quadriparesis. *Journal of Bone and Joint Surgery*, **59A**, 340–44.

Lloyd-Roberts, G.C. and Lettin, A.W.F. (1970). Arthrogryposis multiplex congenita. *Journal of Bone and Joint Surgery*, **52B**, 494–508.

Lockitch, G. *et al*. (1973). Mseleni joint disease: a pilot study. *South African Medical Journal*, **1**, 2283–93.

Macpherson, R.J. and Pai, G.S. (1989). The Haydu–Cheney syndrome. Case report and review of the literature. *Dysmorphic Clinical Genetics*, **3**, 70–78.

Maroteaux, R. and Spranger J. (1991). The spondylometaphyseal dysplasias: a tentative classification. *Pediatric Radiology*, **21**, 293–7.

Maroteaux, R. *et al*. (1980). The mild form of pseudo-achondroplasia. *European Journal of Pediatrics*, **133**, 227–31.

Maudsley, R.H. (1955). Dysplasia epiphysialis multiplex: a report of 14 cases in 3 familes. *Journal of Bone and Joint Surgery*, **37B**, 228–40.

Mayou, S.C., Kirby, J.D., and Morgan, S.H. (1989). Anderson–Fabry disease: an unusual presentation with lymphadenopathy. *Journal of the Royal Society of Medicine*, **82**, 555–6.

McKusick, V.A. (1964). Metaphyseal dysostosis and thin hair; a new recessively inherited syndrome? *Lancet*, **i**, 832–3.

Molloy, M.G. and Hamilton E.B.D. (1978). Thiemanns disease. *Rheumatology and Rehabilitation*, **17**, 179–80.

Moore, W.T. and Federman, D.D. (1965). Familial dwarfism and stiff joints. *Archives of Internal Medicine*, **115**, 398–404.

Morgan, S.H. *et al*. (1990). The neurological complication of Anderson–Fabry disease (a-galactosidse A deficiency) investigation of symptomatic and presymptomatic patients. *Quarterly Journal of Medicine*, **75**, 491–507.

Naranjo, A. *et al*. (1992). Primary idiopathic osteolysis: description of a family. *Annals of the Rheumatic Diseases*, **51**, 1074–8.

Nesterov, A.L. (1964). The clinical course of Kashin–Beck disease. *Arthritis and Rheumatism*, **7**, 29–40.

Noltorp, S., Kristoffersson, U., and Mardahl N. (1986). Trichorhinophalangeal syndrome type I: symptoms and signs, radiology and genetics. *Annals of the Rheumatic Diseases*, **45**, 31–6.

Patroni, N.A. and Kredich, D.W. (1985). Arthritis in children with multiple epiphyseal dysplasia. *Journal of Rheumatology*, **12**, 145–9.

Peters, S.P., Lee, R.E., and Glue, R.H. (1977). Gaucher's disease—a review. *Medicine*, **56**, 425–42.

Polmar, S.H. and Pierce, G.F. (1986). Cartilage hair hypoplasia: immunological aspects and their clinical implications. *Clinical Immunology and Immunopathology*, **40**, 87–93.

Poznanski, A.K. (1974). *The hand in radiologic diagnosis*. W.K. Saunders, Philadelphia.

Reginato, A.J. *et al*. (1994). Familial spondyloepiphyseal dysplasia tarda, brachydactyly and precocious osteoarthritis associated with an argenine 75®cysteine mutation in the procollagen type II gene in a kindred of Chilue Iolander. *Arthritis and Rheumatism*, **37**, 1078–86.

Roubicek, M., Gehler, J., and Spranger, J. (1985). The clinical spectrum of a-L-iduronidase deficiency. *American Journal of Medical Genetics*, **20**, 471–8.

Sakuraba, H. *et al*. (1986). Cardiovascular manifestations in Fabry's disease. A high incidence of mitral valve prolapse in hemizygotes and heterozygotes. *Clinical Genetics*, **29**, 276–83.

Sakuraba, H. *et al*. (1990). Identification of point mutations in the a-galactosidase A gene in classical and atypical hemizygotes with Fabry disease. *American Journal of Human Genetics*, **47**, 784–9.

Sconyers, S.M. *et al*. (1985). A distal chondrodysplasia resembling Kniest dysplasia. Clinical, roengenographic, histologic and ultrastructural findings. *Journal of Pediatrics*, **103**, 898–904.

Scott, H.A. *et al*. (1990). Chromosomal localisation of the human a-L-iduronidase gene (IUDA) to 4p 236.3. *American Journal of Human Genetics*, **47**, 802–7.

Shurtleff, D.B. *et al*. (1964). Hereditary osteolysis with hypertension and neuropathy. *Journal of the American Medical Association*, **188**, 363–8.

Spranger, E.J.P. (1992). International classification of osteochondro dysplasias. (For the International working group on constitutional diseases of bone). *European Journal of Paediatrics*, **151**, 407–15.

Spranger, J. *et al*. (1983). Progressive pseudo-rheumatoid arthropathy of childhood. A hereditary disorder simulating juvenile rheumatoid arthritis. *European Journal of Paediatrics*, **140**, 34–40.

Stanbury, J.B. *et al*. (1989). *The metabolic basis of disease*, (6th edn). McGraw-Hill, New York.

Starreveld, E. (1975). Bilateral carpal tunnel syndrome in childhood. A report of two sisters with mucolipidosis. *Neurology*, **25**, 234–8.

Stowens, D.W. *et al*. (1985). Skeletal complications of Gaucher's disease. *Medicine*, **64**, 310–22.

Szpiro-Tapiaa, S. *et al*. (1988). Spondyloepiphyseal dysplasia tarda: linkage with genetic markers from the distal short arm of the X chromosome. *Human Genetics*, **81**, 61–3.

Torg, J.S. *et al*. (1969). Hereditary multicentre osteolysis with recessive transmission: a new syndrome. *Journal of Pediatrics*, **75**, 243–52.

Traboulsi, E.I. and Maumence, I.H. (1986). Ophthalmologic findings in mucolipidosis III (pseudo Hurler's polystrophy). *American Journal of Ophthalmology*, **102**, 592–7.

Urlus, M. *et al*. (1993). Carpo-tarsal osteolysis. Case report and review of the literature. *Genetic Counseling*, **4**, 25–36.

Whitley, C.B. (1993). The mucopolysaccharidoses. In *McKusick's heritable disorders of connective tissue*, (5th edn) (ed. P. Beighton), pp. 367–499. Mosby, London.

Williams, C.J. *et al*. (1993). Spondyloepiphyseal dysplasia and precocious osteoarthritis in a family with an Arg ^{75}Cys mutation in the precollagen type II gene (COL 2A1). *Human Genetics*, **92**, 499–505.

Winchester, P.H. *et al*. (1969). A new acid mucopolysaccharidosis with skeletal deformities simulating rheumatoid arthritis. *American Journal of Roentgenology*, **106**, 121–8.

Winter, R.M. (1989). Syndrome of the month—Winchester's syndrome. *Journal of Medical Genetics*, **26**, 772–5.

Winter, R.M. *et al*. (1989). Moore–Federman syndrome and acromicric dysplasias: are they the same entity? *Journal of Medical Genetics*, **26**, 320–25.

Wraith, J.E. and Alani, S.M. (1990). Carpal tunnel syndrome in the muco-polysaccharidoses and related disorders. *Archives of Disease in Childhood*, **65**, 962–6.

Wynne-Davies, R., Hall, C., and Ansell, B.M. (1982). Spondylo-epiphyseal dysplasia tarda with progressive arthropathy. *Journal of Bone and Joint Surgery*, **64B**, 442–5.

Wynne-Davies, R., Hall, C., and Apley, A.(ed.) (1985). *Atlas of skeletal dysplasias.* Churchill Livingstone, Edinburgh.

Zayid, I. and Farroj, J. (1973). Familial histiocytic dermato-arthritis. *American Journal of Medicine*, **54**, 793–800.

5.18.1 Intervertebral disc disease and other mechanical disorders of the back

Malcolm I. V. Jayson

Introduction

The primary roles of the human spine are to bear the weight of the upper structures of the body, provide flexibility for movements, and protect the vital structures—in particular the spinal cord and the nerves that emerge through the intervertebral foramina. The stresses associated with a lifetime's use, combined with ageing and degenerative changes and the effects of individual and repeated traumatic episodes, are commonly associated with mechanical problems so that back pain due to structural change of the spine is almost a universal experience. However, although mechanical and degenerative changes of the lumbar spine are common, the correlation with back pain is weak (Lawrence 1977). Some patients have minor evidence of damage to the spine yet experience major problems; others may show degenerative change and yet be symptom free. Moreover, structural changes in the spine are permanent but the symptoms of back pain are commonly transient. Acute episodes of pain are separated by periods in which symptoms are minimal or absent. This complicates our understanding of the pathogenesis of back problems. The remitting course of individual episodes emphasizes the need for controlled studies of various treatment programmes.

Epidemiology

Back pain is an extremely common problem. Recent population studies in Britain show a point prevalence of 14 per cent at any one time (Mason 1994), 39 per cent in the last month (Croft and Jayson 1994), 37 per cent in the last year (Mason 1994), and lifetime prevalence of between 58 per cent (Croft and Jayson 1994) and 80 per cent (Waddell 1987). The peak prevalence is between 45 and 59 years although it is common in the young and in the elderly (Papageorgiou *et al*. 1995). True sciatica appears to be much less frequent. With strict diagnostic criteria, the overall prevalence in studies by Heliovaara *et al*. was estimated to be 5.3 per cent in men and 3.7 per cent in women (Heliovaara *et al*. 1987).

Despite a dramatic increase in recent years in disability due to back problems, there is no clear evidence of an increase in the number of back-related injuries (Troup and Edwards 1985). In a recent survey (Mason 1994), 11 per cent of the population (that is 30 per cent of those with back pain in the last year) reported their activities limited by back pain during the previous 4 weeks; 1.9 per cent of all employed people and 6 per cent of employed people with back pain had lost at least 1 day off in the last month. In Britain in 1993, the total work loss was estimated as approximately 150 000 000 days (Clinical Standards Advisory Group 1994). The prevalence of back disability is increasing rapidly with the number of working days lost now running at some four times the figures for 20 years ago, and similar changes are found in most countries. Cats-Baril and Frymoyer estimated the costs in the United States to be between $25 and 100 billion in 1990 (Cats-Baril and Frymoyer 1991). There is, however, little evidence of increase in the numbers of work-related injuries (Health and Safety Executive 1993). We see a dramatic increase in the morbidity associated with back problems but there does not seem to be any change in the nature of the back injuries. It is likely that social, psychological, and employment problems are playing the major role in this dramatic increase in disability.

There is little difference in the prevalence of back pain and disability between men and women. There is an increased prevalence of low back pain and disability with lower social class; the relationship being stronger in males than females (Walsh *et al*. 1992; Mason 1994). Many patients report the development of back problems in relationship to an accident or injury although commonly this is difficult to evaluate (Mason 1994). Most are work related injuries (Health and Safety Executive 1993) with the highest incidence rate in the construction industry and agricultural sector. There appears to be an association between low back problems and heavy manual work. Videman *et al*. found that vertebral osteophytes were associated with heavy physical work (Videman *et al*. 1990). However, symmetrical disc degeneration was associated with sedentary work. A number of studies have shown an increased lifetime prevalence of back problems in those involved in heavy occupations as compared with light work (Biering-Sorenson 1985; Riihimaki *et al*. 1989) but it is difficult to know whether this is due to the direct effects of overloading the spine or to repeated minor trauma. Mitchell found that people performing heavy manual work had no increase in the number of spells off work with back pain but they were of increased duration (Mitchell 1985). There also appears to be an association between back pain and long-term exposure to driving and vibration (Hulshof and van Zanten 1987).

Epidemiological studies show associations of back pain with neck pain (Porter and Hibbert 1986), chronic musculoskeletal pains (Makela 1993), smoking, psychological distress, and depression (Wright *et al*. 1995).

Structure and function of the lumbosacral spine

The five lumbar and first sacral vertebrae form the principal load carrying structure for the back. Each pair of vertebrae is joined by the intervertebral disc anteriorly and the two apophyseal joints posteriorly. These joints cannot move in isolation and all movements between vertebrae necessarily affects all three. For this reason each pair of vertebrae and the connecting three joints are known as a spinal unit.

The intervertebral disc in the adult consists of a central gelatinous nucleus pulposus surrounded by the tough fibres of annulus fibrosus. The annular fibres spiral obliquely in fascicles between the vertebral rims, lying at about 60° to the spinal axis and interdigitating with each other (Fig. 1). This arrangement makes the intervertebral disc an effective shock absorber which accommodates vertical loads by slight squashing of the disc and bulging of the periphery, and flexion, extension, and lateral flexion movements by alteration of the angles of the crossing fascicles. However, torsion of the lumbar spine is less readily accommodated as twisting of one vertebra on that below means that some collagen fibres of the annulus will be stretched. This is one of the reasons why twisting movements of the lumbar spine are more likely to be associated with annular damage.

Fig. 1 The collagen fibres of the annulus fibrosus spiralling obliquely around the margins of the intervertebral disc. (By courtesy of Dr J.B. Weiss.)

There is a lumbar lordosis that is most marked at L4/5 and L5/S1. The associated increase in pressures in the posterior part of the disc, together with the lower lumbar discs carrying higher loads than the upper discs, account for the greater prevalence of lumbar problems at these sites.

The gel-like nucleus of the normal disc consists of a matrix of water and glycosaminoglycan in a random meshwork of Type II collagen fibres (McDevitt 1988). With ageing and disc degeneration there is alteration of the glycosaminoglycan, with a reduction of its molecular size (Adams and Muir 1976), and the nucleus becomes more fibrous. There is loss of water content, the gel-like character of nucleus is lost, and, as a result, it is no longer able to redistribute pressures in an isotropic fashion. Localized areas of pressure concentration appear and may be associated with focal damage and an increase in proportion of the body load is borne by the annulus fibrosus. The normal annulus primarily consists of Type I collagen fibres at its outer periphery but with a gradual change to Type II fibres at its inner margin with the nucleus (Eyre and Muir 1976). Type I collagen is characteristic of collagens involved in resisting tensile loads, which is the primary function of the outer annulus, whereas the inner annular and nuclear Type II collagen is more typical of cartilage collagen and has to withstand compressive loads. There is clearly an appropriate, functional adaptation of collagen structure within the disc. In degenerative disease this pattern is lost. There is an excess of Type I collagen in both the annulus and nucleus (Herbert *et al*. 1975) interfering with the biomechanical efficiency of the disc. There are alterations in the patterns of degradative enzymes with disc degeneration and herniation (Ng *et al*. 1986) but it is not clear whether these changes play a fundamental role in degeneration of the disc or are secondary to disc damage. Experimental induction of disc prolapse shows that extensive vascular ingrowth is followed by more florid evidence of internal disc disruption (Vernon Roberts 1992). Angiogenic factors associated with the proliferation of new blood vessels are capable of activating degradative enzymes (Weiss and McLaughlin 1993), so providing a possible mechanism for the initiation of disc degeneration.

Measurements of pressures within the intervertebral disc show the increase associated with the upright posture and load bearing (Nachemson 1992). The load on the lumbar spine is increased by contraction of the paravertebral muscles, which stabilize the spine, and reduced by the lumbar lordosis. The design of car seats is largely related to information about the stresses on the disc with varying degrees of lumbar support and back rest inclination.

In contrast to the discs, the apophyseal joints are true synovial joints with fibrous capsules lined by synovium. They lie in an anteroposterior direction although their centres of rotation are behind the disc. Rotation of the lumbar spine is therefore limited and will produce sheer stresses on the disc—again accounting for the risks associated with torsion of the lumbar spine.

Pain is commonly experienced in the back and lower limbs. It may arise from stimulation of the nerve roots or nociceptive receptors (Wyke 1987). These are distributed in:

1. the skin and subcutaneous tissues;
2. the fibrous capsules of the apophyseal and sacroiliac joints;
3. the outer layers of the annulus fibrosus;
4. the longitudinal ligaments of the spine and, in particular, the posterior longitudinal ligaments;
5. the periostium and attached fascias, aponeuroses, and tendons;
6. walls of blood vessels in and around the spine;
7. the dura mata and epidural adipose tissues.

There is a complex anastomotic network of sensory fibres around the vertebral column and, in particular, around the margins of the intervertebral disc, penetrating the outer annulus and encircling blood vessels (Ashton *et al*. 1994). Substance P immunoreactive fibres can be demonstrated in these areas (Coppes *et al*. 1990). As a result, back pain is often diffuse and it can be difficult to locate the source of the problem from the distribution of symptoms.

Afferent impulses pass through the sensory neurones and the dorsal horn ganglion into the spinal cord. Complex neuropharmacological changes involving substance P, vasoactive intestinal peptide, and other neuropeptides are modulated after nociceptive stimulation of the spine (Weinstein 1986). As a result, pain felt in the back and in the lower limbs can arise due to a wide variety of different reasons. These include:

1. radicular pain from direct nerve or nerve root compression;
2. referred pain from damage of nociceptive receptors in the various spinal tissues with referral of symptoms into the lower limb;
3. deafferentation pain resulting from a loss of the afferent connections of the spinal neurones;
4. dysregulatory or reactive pain involving the pathophysiology of the afferent motor system with resulting hypertonus of postural muscles;
5. psychosomatic pain such as pain enhanced by emotion, depression, or social distress.

In addition, recent work suggests central mechanisms by which the perception of pain is modulated. In particular, following nociceptive damage neuromorphological and neurochemical changes occur within the dorsal horn which may be long lasting so the sensation of pain may persist despite the lack of peripheral cause. The cells may be sensitized with increased sensitivity to minor stimulation and the message may spread within the spinal cord so that symptoms are perceived over a wide area (Dubner and Basbaum 1994).

The sympathetic chain runs alongside the vertebral column with multiple anastamosing sympathetic fibres and the peripheral sympathetic system accompanies blood vessels extending into the lower limbs. Many patients with chronic lower limb pain showed evidence of a sympathetic dysfunction syndrome. Most often this follows surgery to the lumbar spine although it can occur in association with mechanical problems (Sachs *et al*. 1993).

Pathological changes in the spine

Herniated intervertebral disc

Nuclear material can burst through the annulus fibrosus displacing and damaging the surrounding structures. Most often this occurs at the level of the L4/5 and L5/S1 intervertebral discs and the prolapse is usually in a posterolateral position, although sometimes it can be much nearer or in the midline. As a result, the prolapse may damage nerve roots in addition to ligaments, periostium, blood vessels, dura, and other tissues. The degrees of herniation are known as protrusion, extrusion, and sequestration according to their extent and whether the herniated material retains contact with the disc. Herniation most often develops after heavy load bearing with the spine in a flexed and twisted position. However, there is evidence that compression of a healthy lumbar disc will usually cause vertebral end plate fracture and that posterior and posterolateral herniation only occur when there is previous disc degeneration and fissure producing weakening of the annulus fibrosus (Jayson *et al*. 1973). The particular stress may have acted solely as the final precipitating cause of the problem.

The herniated material is often surrounded by an area of erythema and oedema. Usually it will gradually heal with fibrosis and shrinkage so that eventually the symptoms will improve. However, the annulus will be damaged and there is considerable risk of recurrent disc herniation and later development of secondary degenerative changes. The herniated material may press directly on nerves or compress the epidural veins producing venous obstruction and ischaemia over the relevant nerve. It may be complicated by thickening of the neural sheaths with perineural and intraneural fibrosis, and in turn by neuronal atrophy (Hoyland *et al*.

1989).

A more central, posterior prolapse can damage the posterior longitudinal ligament, which is innervated by the sinuvertebral nerve, and produce central back pain. A large, central posterior prolapse can compress the cauda equina leading to multiple nerve root damage in the lower limbs and sphincter disturbance.

There seems to be some predisposition towards development of herniation. Pathological studies in cadavers (Jayson and Barks 1973) and magnetic resonance scanning in life (Powell *et al*. 1986) show that when disc degeneration is present there is an increased risk of changes at multiple levels. There may be familial factors predisposing towards the development of this prolapse although no specific immunogenetic pattern has been identified (Postacchini *et al*. 1988). Environmental factors clearly play a part. Moderate amounts of physical activity are good for the back but lack of exercise and excessive spinal loads are both associated with increased risks of back pain, disc herniation, and degenerative disease of the spine (Videman *et al*. 1990).

The classical presentation of a herniated intervertebral disc is the acute onset of back pain followed by radicular pain, numbness, and paraesthesiae in one or other lower limb. The problem may develop after some stress and in particular bending, twisting, and lifting which, when combined, seem to produce the greatest risk. Most patients, however, will give a history of some preceding aching and stiffness in the back over the previous few days.

Clinical pattern

Direct damage to the nerve root will produce pain, numbness, and paraesthesiae as shown in Table 1. A small lateral herniation of the L4/5 disc will commonly affect the L5 nerve and herniation of the L5/S1 disc will affect the S1 root. However, a large prolapse, particularly if it is more central, may affect several nerve roots with widespread neurological symptoms and signs, and perhaps sphincter disturbance. There is considerable overlap in the distributions of the various nerve roots and the symptoms and physical signs only act as approximate guides to which disc has been damaged. In addition, pressure on ligaments and other soft tissues will produce referred pain felt in the buttock and lower limb. This is often confused with radicular symptoms due to nerve root damage. Referred pain is poorly defined and patients commonly have difficulty in describing its distribution. The symptoms of disc herniation are made worse by spinal movements and also by Valsalva manoeuvres such as coughing, sneezing, micturition, and defecation, which raise the cerebrospinal fluid pressure. Examination may show a sciatic scoliosis due to unilateral spasm of the paraspinal muscles. There may be severe limitation of spine movements but often in only one or two directions and in particular movements, including flexion. Palpation of the spine may show local areas of tenderness over the spine or sometimes elsewhere, such as the sacroiliac joints. Straight leg raising is limited with a positive Lasegue's test. The specific neurological signs of nerve root damage should be sought (Table 1).

Root	Superficial paraesthesiae and sensory change	Muscle weakness	Tendon reflex changes
L2	Upper thigh: anterior, medial, and lateral surfaces	Flexion and adduction of hip	None
L3	Anterior surface of lower thigh Anterior and medial surfaces of knee	Adduction of hip; extension of knee	Knee jerk possibly decreased
L4	Anteromedial surface of leg	Extension of knee; dorsiflexion and inversion of foot	Knee jerk decreased
L5	Anterolateral surface of leg, dorsum and medial surface of foot, especially dorsal surface of hallux	Extension and abduction of hip Flexion of knee; dorsiflexion of foot and toes, especially hallux	None
S1	Lateral border and sole of foot, back of heel, and lower calf	Flexion of knee; plantar flexion and eversion of foot	Ankle jerk decreased

Table 1 Principal changes used for identifying the sites of lumbar nerve root lesions (note that this table does not list the total distribution of each root)

Investigations

The blood sedimentation rate, plasma viscosity, and routine haematological and biochemical tests are normal. The cerebrospinal fluid obtained by lumbar puncture is usually normal but when there is a large prolapse with a spinal block the protein may be elevated.

Imaging techniques commonly used include:

Plain radiography

In an acute prolapse the spine is normal except perhaps for a scoliosis. When followed over several years, disc narrowing may develop and later on there may be secondary spondylotic changes.

Radiculography

Water-soluble radio-opaque dye is injected into the subarachnoid space at lumbar puncture. The dye penetrates along the nerve roots and obstruction to flow produced by a prolapse can be clearly identified (Fig. 2).

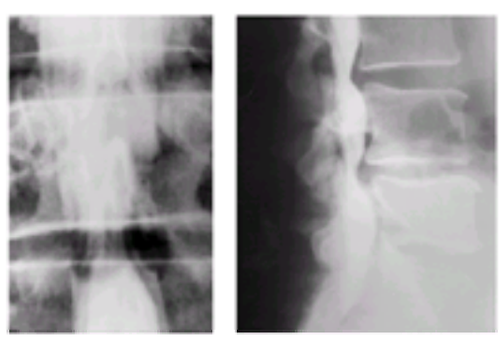


Fig. 2 Radiculogram showing (a) left sided L4/5 disc herniation and (b) smaller disc protrusion at L3/4.

Discography

Direct injections may be made into an intervertebral disc. A normal, healthy disc will accept about 0.5 ml of fluid without producing significant symptoms. Larger volumes will be accepted by herniated discs with reproduction of symptoms. The pain may be relieved by injection of local anaesthetic. The radiographic appearances of disc herniation may be identified (Fig. 3).

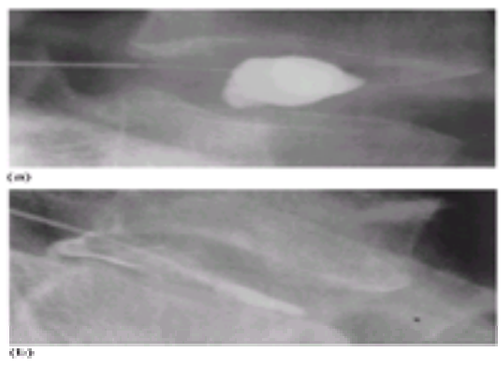


Fig. 3 Discography: (a) normal and (b) disc degeneration with leakage of the dye into the epidural space.

Computed tomography (CT)

Herniation of a disc and the size and extent of nerve root involvement can be determined. This technique is sometimes use to complement radiculography when the details of nerve root damage may be clarified.

Magnetic resonance imaging (MRI)

This technique is becoming the investigation of choice. It has the advantage of avoiding the use of X-rays. It will demonstrate the size and extent of a disc herniation as well as degenerative change in the disc by loss of signal on the T_2 image and reduction of disc height (Fig. 4).

Fig. 4 MRI scan, vertical reconstruction. There is loss of signal at the L3/4, L4/5, and L5/S1 discs indicating disc degeneration with herniation at L4/5 and L5/S1.

Lumbar spondylosis

Degenerative changes in the spine are common. Pathological studies have demonstrated their first appearance to be often around 25 years and they are almost universal in older people. Although patients with severe degrees of spondylosis more commonly develop back pain than those without, the correlation between back pain and radiographic evidence of lumbar spondylosis (Lawrence 1977) and MRI evidence of disc bulging and protrusion (Jensen *et al*. 1994) is poor. Because they are so common it may be better to call these appearances 'ageing change' rather than 'degenerative change' as the latter term conveys a perjorative image to patients and an implication of long-term problems, which is frequently not the case.

There is disorganization of the internal structure of the disc with the loss of the clear distinction between the nucleus pulposus and the annulus fibrosus. Cleft formation within the disc is common. The disc becomes narrowed with osteophytosis around the vertebral margins. Osteoarthritic changes may also develop in the apophyseal joints and it is commonly not clear whether such changes preceded or followed degeneration in the disc. Lumbar spondylosis with osteophytes and often minor degrees of disc herniation can directly impinge upon nerve roots and other structures. More commonly, however, they may contribute towards obstruction of vascular flow and be associated with peri- and intraneural fibrosis and neuronal atrophy (Hoyland *et al*. 1989).

There is no defined clinical pattern associated with lumbar spondylosis. Some patients only suffer postural backache with pain associated with prolonged sitting or standing in a poor posture or aggravated by heavy manual work. Others suffer more severe episodes, often provoked by trivial stress. These acute episodes will usually resolve within a few days but the patient is always at risk of further episodes. The pain may be felt across the back but often spreads into the buttock and the lower limbs. If there is nerve root entrapment there may be the specific radicular symptoms and signs (Table 1). Soft tissue damage can produce referred pain spreading to the limb, which is often not in a well defined distribution.

The radiographic changes include narrowing of disc space, gas in the disc, sclerosis of the vertebral end plates, and marginal osteophytosis (Fig. 5). It is important to remember that these changes are common in the symptom-free population and particularly in older subjects. Their presence, therefore, does not mean they are the source of the symptoms and careful evaluation is always required to ensure there is no other problem in any individual patient.

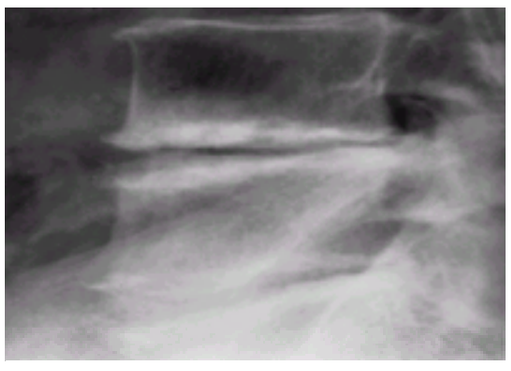


Fig. 5 Lumbar spondylosis at L4/5 and L5/S1.

Internal disc disruption

After trauma, and in particular compressive loads on the spine perhaps associated with the vertebral end plate fracture, there may be degradation of the nucleus pulposus spreading outwards to involve the annulus fibrosus. This may irritate nerve endings in the outer border of the annulus and lead to the development of pain. Loss of disc height may contribution towards venous obstruction and tissue ischaemia (Crock 1992).

The symptoms are felt in the back or referred into the buttock or lower limb. They are aggravated by movement. However, the nerve roots may not be affected in which case there will be no specific neurological signs. In the early stages, plain radiographs may be normal but later there is progressive narrowing of the disc space. Likewise, the myelogram and CT scan may be normal although the MRI scan may show loss of signal in the T_2 image. On discography the radiographs, perhaps supplemented by a CT scan, may show characteristic morphological appearances. Reproduction of the pain by raising the pressure within the individual disc and relief by local anaesthetic is more valuable.

In cases where the disc can be specifically identified as the source of the symptoms, spinal fusion may be helpful (Colhoun *et al*. 1988).

Facet joint syndromes

The apophyseal joints are true synovial joints and commonly develop osteoarthritic changes. Patients may develop pain which arises from osteoarthritic apophyseal joints. The typical picture is of pain in the back spreading into one or other buttock or into the back of the thigh. It is aggravated by extension and on palpation there is marked tenderness over the facet joints. However, each facet joint has a nerve supply arising from several levels and, as a result, the pain often has a widespread distribution making it difficult to identify which specific joint is the source of the symptoms.

Facet joint arthrography can be helpful as it may demonstrate osteoarthritic change and, in particular, if injection of the joint aggravates the symptoms and anaesthetic provides relief, this strongly suggests that the joint is responsible for the development of pain (Fairbank *et al*. 1981). However, the diagnosis is controversial as the correlation between the response to provocative surgical procedures and clinical findings is poor and it remains possible that when symptoms are relieved by anaesthetic it has leaked out from the joint and provided symptomatic relief simply by affecting nerve endings in a widespread distribution.

Spinal stenosis

There is considerable variation in the size and shape of the vertebral canals. In particular, some patients have a canal with a small diameter and with a trefoil shape due to osteoarthritic hypertrophy of the apophyseal joints. The nerve roots may be tightly packed. As a result, there is an increased risk of back problems following any further intrusion into the canal. It is known that patients developing symptoms from a disc protrusion have significantly smaller canal measurements than the normal population (Porter *et al*. 1980). Symptoms due to spinal stenosis may develop in later life as a result of osteophytosis, disc bulging, and ligament thickening intruding into a tight canal space.

There are two principal syndromes associated with spinal stenosis—central stenosis and foraminal stenosis.

Central stenosis—neurogenic claudication

On walking the patients progressively develop discomfort, numbness, paraesthesiae, heaviness, and a dead feeling in the lower limbs that will eventually make them stop. On resting the symptoms gradually ease over a period of 5 to 10 min and the patient may start walking again. The problem is frequently confused with claudication due to a poor arterial blood supply in the lower limbs. As the patients are frequently elderly and may also have an arteriopathy the differential diagnosis may sometimes be difficult. The vertebral canal has slightly greater dimensions when the spine is flexed and many patients describe bending forwards to relieve their pain. Likewise they may find it easier to walk uphill when they lean forwards than downhill when they lean back.

On examination they may stand with the spine, hips, and knees slightly flexed. They can usually bend forwards without problem but extension is absent. Neurological examination is unhelpful.

Plain radiographs may show the evidence of degenerative change but it is difficult to determine the dimensions of the vertebral canal. Myelography can be technically difficult because of the crowding of the nerve roots. CT and MRI scans are most helpful for identifying the extent and severity of the stenosis and measurement of the vertebral canal dimensions may be made (Fig. 6).

Fig. 6 MRI scan showing spinal stenosis, most marked at L3/4 and L4/5, with degenerative disc disease.

Foraminal stenosis—root entrapment syndrome

Narrowing of the intervertebral foramen with nerve root entrapment can give rise to a specific pattern of radicular pain in the lower limb (Porter 1992). This narrowing may be due to bony osteophytosis from the apophyseal joints, posterolateral osteophytes on the vertebral bodies, ossification and hypertrophy of the posterior spinous ligaments, and fibrous proliferation of the nerve root sheaths.

The patient develops a pattern of radicular pain distinct from that of a prolapsed intervertebral disc. In particular, the pain is often severe and unremitting, being present at rest in bed as well as on exercise, although it may be aggravated by physical activity. Back movements may be full and many patients do not show restriction of straight leg raising.

CT (Fig. 7) and MRI scans are most helpful for demonstrating foraminal stenosis. Nerve root sheath injection with local anaesthetic can also be helpful for identifying the site of the pain producing lesion (Dooley *et al*. 1988).

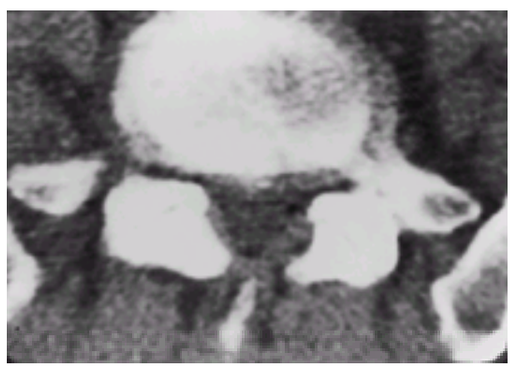


Fig. 7 CT scan showing foraminal stenosis.

Spondylolysis and spondylolisthesis

Defects in the pars interarticularis (spondylolysis) are common and usually not of significance. Spondylolisthesis refers to slipping, usually forwards but occasionally backwards, on the vertebra below. Spondylolisthesis has been classified by Wiltse and Rothman (Wiltse and Rothman 1990) as follows:

1. congential;
2. isthmic:
    a. lytic—stress failure of the pars interarticularis,
    b. an elongated but intact pars secondary to a healed stress fracture,
    c. fracture;
3. degenerative;
4. post-traumatic;
5. pathological—due to destruction of posterior elements;
6. surgical.

Patients may develop pain in the back and in the lower limb aggravated by standing and walking and relieved by rest. Narrowing of the vertebral canal can cause symptoms suggestive of spinal stenosis. The physical findings will depend on the degree of slip and extent of neurological damage. It may be possible to palpate a step on their spine on forward flexion in gross cases.

Postsurgical back pain

Many patients undergo surgery for spinal disorders. A significant proportion fail to respond or develop recurrent back problems. Careful assessment is needed to determine the cause of the problem in each individual case. Problems that may be identified are as follows:

1. the operation was undertaken at the incorrect level;
2. inadequate surgery, perhaps with failure to decompress adequately a nerve root or excise a disc herniation;
3. recurrent disc herniation at the same or another level;
4. secondary or degenerative change, perhaps associated with extensive removal of bone;
5. perineural fibrosis and arachnoiditis (see below);
6. sympathetic dysfunction syndromes (Sachs *et al*. 1993);
7. psychological factors;
8. unrelated pathology.

Perineural fibrosis and arachnoiditis

Abnormal fibrosis may develop in the periradicular tissue (perineural fibrosis) or within the dural sac (arachnoiditis). The former develops most commonly in a florid form in patients who have undergone spine surgery and may be due to retained cotton debris from the surgical swabs and patties (Hoyland *et al*. 1988). However, minor degrees of periradicular fibrosis are common in association with lumbar spondylosis. It is frequently difficult to correlate the presence of epidural fibrosis and the development of back and lower limb symptoms (Cooper *et al*. 1991).

Arachnoiditis was first described following meningitis. It can develop following surgery but is then restricted to the level of the operation. The most widespread forms of arachnoiditis are found in those patients who in the past underwent myelography using oil based radiographic media. There is inflammation around the nerve roots followed by necrosis and adhesions. The pathological changes and clinical features of arachnoiditis have been described by Burton (Burton 1990).

Very severe and widespread pain often is experienced not only in the back and lower limbs but over the whole trunk and even extending into the upper limbs and face. It frequently has a burning and dysaesthetic quality and patients are commonly very distressed. Radiculograms will show a grossly distorted thecal sac (Fig. 8). High resolution MRI scans are most helpful for demonstrating the presence of arachnoiditis (Fig. 9).

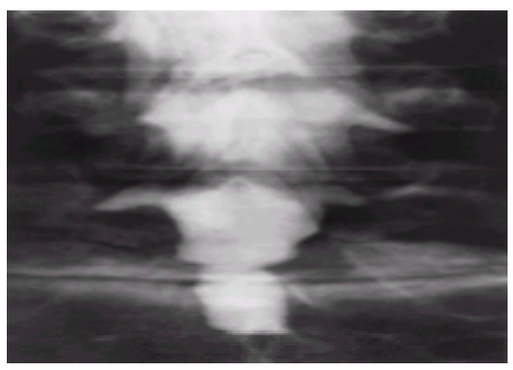


Fig. 8 Radiculogram in a patient with arachnoiditis showing gross distortion of the thecal sac.

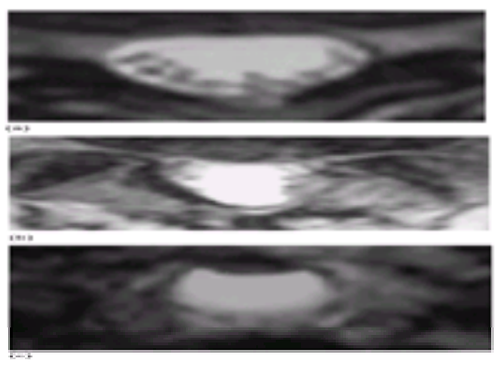


Fig. 9 High resolution MRI scans showing the nerve roots (a) normal, (b) arachnoiditis. The nerve roots are adherent to each other, atrophic, and stuck to the wall of the thecal sac. (c) The nerve roots have atrophied completely and can no longer be seen. Residual Myodil is still present anteriorly.

Non-specific back pain

In many patients it is difficult to define a precise cause for the symptoms. They suffer persistent or recurrent episodes of back pain. Examination may reveal limitation of spine movement but without neurological signs. Radiographs may show varying degrees of lumbar spondylosis but it is difficult to determine.

Within this group many will have localized areas of tenderness over the spine or the sacroiliac joints. Pressure on areas may reproduce the patient's symptoms and

they are sometimes called areas of subluxation or dysfunction by osteopaths and chiropractors although there is no pathological basis for these terms. Collee *et al*. described iliolumbar and greater trochanteric pain syndromes occurring in patients with chronic low back pain (Collee *et al*. 1990). However, no specific pathology has been demonstrated and the cause of the problems remain uncertain.

Hypermobile patients often suffer recurrent back pain. Radiographs may show marked evidence of degenerative change yet these subject retain a remarkably good range of back movements. It is believed that injury of the ligaments and capsule, perhaps associated with small fractures, may be the source of pain in these circumstances.

The fibromyalgia syndrome is also a source of confusion. These patients have multiple tender points, particularly around the shoulder and hip girdles, and this is commonly associated with a poor sleep pattern. Many are depressed and there is an association with migraine and an irritable bowel syndrome. There is no evidence of a specific back pathology, although the fibromyalgia syndrome can complicate other back problems so that patients experience widespread pain (see also Chapter 5.14).

Management

Each patient will require individual assessment. In particular, it is important to distinguish patients with acute, recurrent, and chronic mechanical back pain syndromes.

Acute back pain

Most acute back problems are seen in primary care. Many patients do not seek medical advice and a high proportion will be treated by a physiotherapist, osteopath, or chiropractor. It is generally believed that about 90 per cent of acute episodes settle within 6 weeks (Waddell 1987). However, these estimates may be over optimistic. Croft *et al*. found that of patients consulting their primary care practitioner, 69 per cent had a clear cut new episode of back pain, 20 per cent had an acute exacerbation of a more chronic or persistent complaint, and in 8 per cent the consultation was part of a continuing problem (Croft and Jayson 1994). At 6 months follow-up, 27 per cent had recovered completely, 28 per cent improved, 30 per cent were the same, and 14 per cent had become worse.

In acute back pain the patients should be triaged into those with simple back ache, nerve root pain, and possible serious spinal pathology. Guidelines for classifying these three groups have been published (Clinical Standards Advisory Group 1994).

Simple backache

Clinical criteria

Patients aged 20 to 55 years;

pain in the lumbosacral region, buttock, and thighs;

pain mechanical in nature varying with physical activity and with time;

patient is well.

Blood tests and lumbar spine radiographs are not indicated unless there is doubt about the diagnostic triage.

Such patients should be managed with simple analgesics such as paracetamol (aminocetophen). Occasionally they may require non-steroidal anti-inflammatory drugs. Narcotics should be avoided and never given for more than 2 weeks. Bed rest should only be prescribed if essential and then normally for only 1 to 3 days. This produces less disability than more prolonged periods of bed rest (Deyo *et al*. 1986).

It is always important to provide ergonomic advice on posture while standing, working, sitting, and lifting in order to protect the back against excessive loads.

Early physical activity is encouraged and reduces pain (Malmivaara *et al*. 1995). Indeed, continuing ordinary activities whenever possible seems to produce the best outcome (Malmivaara *et al*. 1995). Patients should be advised that exercise is not harmful and reduces pain, and that physical fitness is beneficial. The patient should be encouraged to return to work as soon as possible. Simple analgesic, or perhaps anti-inflammatory drugs, may be required and narcotics should be avoided if possible. With a regimen such as this, most patients will return to normal function rapidly.

The role of the therapist is restoration of function and encouraging the patient to be mobile. Physical exercise has an important role. The actual type of exercise is relatively unimportant. The benefits seem to lie in the quantity of exercise rather than in its specific nature. A positive attitude to activity and encouragement to return to work as soon as possible is all important. Early identification of psychosocial problems is important as they predict both the development of back pain (Croft, P., Papageorgiou, A.C., Ferry, S., Thomas, P., Jayson, M.I.V., and Silman, A.C., in press) and of acute back pain becoming chronic (Burton *et al*. 1995).

In order to remobilize the patient there is a wide variety of forms of physiotherapy and manual treatments. There is no evidence that heat, cold, ultrasound, or massage provide any benefit other than comfort at the time they are administered. Exercises and physical fitness are important and an incremental aerobic fitness programme of physical reconditioning is advised.

With this programme the vast majority of patients will recover within 6 weeks. If the pain still persists at 6 weeks, the diagnostic triage should be reviewed and investigations such as erythrocyte sedimentaion rate and a radiograph of the lumbar spine should be requested, if specifically indicated. For the patient with persistent back ache an active rehabilitation programme is necessary and at this stage referral may be required for a second opinion, rehabilitation, additional assessment, pain management, and, occasionally, surgery.

Manipulation is practised by physiotherapists, osteopaths, chiropractors, physicians, and surgeons. Although there has been some conflicting evidence from various trials, overall it appears that manipulation accelerates the rate of recovery in recent onset back pain, although not making any long-term difference to the outcome (Twomey and Taylor 1995). There is no evidence distinguishing the effects of manipulation by the various practitioners except that manipulation under general anaesthesia carries significant risk of neurological damage to the cauda equina (Haldeman and Rubinstein 1992).

Nerve root pain

Clinical criteria

Unilateral leg pain greater than back pain with the pain radiating to the feet or toes;

there may be numbness and paraesthesias in the same distribution;

reduced straight leg raising indicating nerve irritation;

motor sensory or reflex changes limited to one nerve root.

The management of such patients is similar to simple backache but undertaken more slowly. The recovery rate is not as good. If they fail to improve within 6 weeks, they may need more detailed assessment. Such patients may require scans and may be considered for surgery.

Possible serious spinal pathology

Clinical criteria

Aged under 20 or over 55 years;

previous trauma;

constant, progressive non-mechanical pain;

gradual onset —morning stiffness, limitation of movements in all directions and peripheral joint involvement, and other features suggesting ankylosing spondylitis or related disorders;

systemically unwell;

widespread neurological signs and cauda equina syndrome;

weight loss;

significant medical history such as previous carcinoma, steroid therapy, etc.

In such circumstances the patient will require intensive and appropriate investigations.

Recurrent back pain

Many patients have recurrent episodes of back pain which may be precipitated by lifting, twisting, and bending. A detailed understanding of the structure and function of the spine and the appropriate ways to protect it in a variety of physical activities combined with simple exercises to increase physical fitness and strengthen the abdominal and paraspinal muscles seem helpful in preventing further recurrences of back problems. The role of the physiotherapist is back education and the training may be formalized as a series of lessons in a 'back school' (Andersson 1992). The results of such training programmes are at least as good and probably better than conventional physiotherapy treatments.

Chronic back pain

Many patients suffer persistent pain in the back which may spread into the lower limbs. They may become very severely and permanently disabled. Many have previously undergone one or more spinal operations. Detailed assessments are required to elucidate the pathogenesis of the pain in individual subjects and to plan a treatment programme.

Chronic back pain may be due to:

1. chronic disc herniation with persistent nerve root damage;
2. severe degenerative change in the intervertebral discs;
3. non-mechanical pathologies such as inflammatory spondylarthropathies, neoplasms, infections, Paget's disease, etc;
4. incomplete discectomy or nerve decompression;
5. recurrent or new disc prolapse;
6. scar tissue forming around the nerves, perhaps as a reaction to surgery. Retained microscopic cotton debris from the swabs and patties used at operations may be of direct relevance here (Hoyland *et al*. 1988);
7. arachnoiditis due to previous oil based myelography (Fig. 8);
8. sympathetic syndromes with referred symptoms felt in the lower limb (Sachs *et al*. 1993);
9. fibromyalgia;
10. psychological factors—these play an important part in many patients, perpetuating the chronic nature of back pain. They include depression, anxiety, and compensation factors. Operant conditioning refers to the psychological reinforcement of pain behaviour and may include not only financial benefits but also sympathy and concern expressed by relatives, friends, medical, and paramedical staff;
11. central nervous system modulation.

Assessment of the chronic back pain patient requires a very careful history and examination. The findings suggestive of a substantial non-organic component to the problem (Waddell *et al*. 1990) include:

1. pain behaviour;
2. pain reproduction on simulated movements of the spine such as pressure on the skull, rotation of the pelvis;
3. restricted straight leg raising on formal testing but unrestricted distracted straight leg raising such as being able to sit up with the lower limbs extended;
4. regional weakness or sensory disturbance in a non-neurological distribution;
5. widespread superficial tenderness over the back.

Adequate management can only be undertaken after a thorough evaluation of each individual patient. This should include a very careful review of the medical history together with physical and psychological assessments. When specific pathologies can be identified they may be amenable to the appropriate therapy. For many patients this is not possible. Forms of treatment which may be beneficial include:

1. The use of appropriate medication—pure analgesics such as paracetamol, or perhaps codeine or its derivatives, or dextropropoxyphene are adequate for most patients. For some, particularly if there is a pain pattern suggestive of a secondary vascular/fibrotic/inflammatory element with pain aggravated by rest, non-steroidal anti-inflammatory drugs could be very helpful. Muscle relaxants, such as chlormezanone and baclofen, may be helpful when there is a major element of muscle spasm. Some patients describe a neuralgic element to the pain with electric shock sensations radiating down the lower limbs. This may be relieved by antiepileptic drugs such as carbamezepine or sodium valproate. Others have widespread, superficial paraesthesiae and tenderness and may respond to tricyclic antidepressant drugs such as amitriptyline. In the spinal stenosis syndrome there is some evidence (Porter and Hibbert 1983) that calcitonin may provide relief, perhaps by altering blood flow dynamics within the vertebral column.
2. Physiotherapy and ergonomic advice—this teaches the patients how to perform tasks within their physical capabilities and helps to give them confidence.
3. Local injections may be helpful for some people. They include:
   a. Trigger point and local injections, usually of steroid and anaesthetic. Although commonly used, the evidence for their value is in doubt. Garvey *et al*. compared lignocaine alone, lignocaine plus steroid, needle insertion with no injection of material, and application of a vapour coolant spray followed by acupressure and found no differences between these treatments (Garvey *et al*. 1989).
   b. Facet joint injections. No study has demonstrated any benefit from this technique.
   c. Epidural injections. There have been a number of controlled studies with conflicting results. My opinion is that the technique is of value for the patient with radicular pain which has failed to resolve completely. Major adverse effects are rare but they are principally associated with dural puncture and intrathecal injection (Bogduk 1995).
   d. Acupuncture. Although some studies suggest that needling has advantages over control treatment, all had major methodological flaws. There do not appear to be any differences in outcome with needling in the Chinese meridians compared with misplaced needling. The TENS machines are an alternative method of providing acupuncture.
4. Acupuncture and transcutaneous electrical nerve stimulation is believed to stimulate large afferent fibres blocking pain transmission through nociceptive small fibres. Despite widespread claims that this is of value, no trial has convincing demonstrated significant advantage to these techniques.
5. Patient education. Failure to provide an explanation of the problems leads to patient dissatisfaction (Deyo and Diehl 1986). The back education programme may be structured as a back school. Although appreciated by the patients, their specific value with regard to outcome is in doubt.
6. Lumbar corsets and belts. The evidence for their efficacy is confused. Many patients come to depend on their corsets and may develop severe restriction of spine movements. My own view is that they should not be in general use.

7. Activity modification. Bed rest should be avoided in chronic back pain. Patients should be encouraged to remain active and undertake a regular exercise programme to improve physical fitness (Frost *et al*. 1995).
8. Further surgery—the results of secondary and subsequent operations tend to be poor. Further surgery should only be contemplated in patients for whom there is a very clearly defined lesion causing symptoms for which there is an adequate surgical solution.
9. Multidisciplinary functional restoration programmes—these programmes are recent developments. Detailed assessment is required. Once the patient accepts that no further specific interventions will help they should enter a programme combining intensive physical activation, counselling for the understanding of pain and related problems, reducing the use of medication and health care, and dealing with depressive symptoms and anxiety together with encouragement to return to normal activities. These programmes are time consuming and expensive but effective in improving function, return to work rates, and work retention (Burke *et al*. 1994). Mayer *et al*. were able to return 87 per cent of patients to work after such a programme in contrast to 41 per cent of controls (Mayer *et al*. 1987).

Conclusion

In recent years, there has been a major increase of interest in the back pain problem. We now have much better understanding of the mechanisms of pathogenesis of various back pain syndromes and the reasons why the problems may persist and become chronic. With this better understanding, targeted therapy appears effective and lends hope of providing better control of this problem in the future.

Chapter References

Adams, P. and Muir, H. (1976). Qualitative changes with age of proteoglycans of human lumbar discs. *Annals of the Rheumatic Diseases*, **35**, 289–96.

Andersson, G.B.G. (1992). Back schools. In *The lumbar spine and back pain* (4th edn) (ed. M.I.V. Jayson), pp. 409–16. Churchill Livingstone, Edinburgh.

Ashton, I.K. *et al*. (1994). Neuropeptides in the human intervertebral disc. *Journal of Orthopaedic Research*, **12**, 186–92.

Biering-Sorenson, F. 1985. Risk of back trouble in individual occupations in Denmark. *Ergonomics*, **28**, 51–60.

Bogduk, N. (1995). Epidural steroids. *Spine*, **20**, 845–8.

Burke, S.A., Harms-Constas, C.K., and Aden, P.S. (1994). Return to work/work retention outcomes of a functional restoration programme. A multi-centre prospective study with a comparison group. *Spine*, **19**, 1880–6.

Burton, A.K., Tillotson, K.M., Main, C.J., and Hollis, S. (1995). Psychosocial predictors of outcome in acute and subacute low back trouble. *Spine*, **20**, 722–8.

Burton, C.V. (1990). Adhesive arachnoiditis. In *Neurological surgery III*. Vol II (ed. Youmans), pp. 2856–63. Saunders, Philadelphia.

Cats-Baril, W.L. and Frymoyer, J.W. (1991). The economics of spinal disorders. In *The adult spine: principles and practice*(ed. W. Frymoyer), pp. 85–105. New York. Raven Press.

Clinical Standards Advisory Group (1994). *Back pain*. HMSO, London.

Colhoun, E., McCall, I.W., Williams, I., and Cassar-Pollicino, V.N. (1988). Provocation discography as a guide to planning operations on the spine. *Journal of Bone and Joint Surgery*, **70B**, 267–71.

Collee, G., Dijkmans, B.A.C., Vanderbrouke, J.P., Rouzing, P.M., and Cats, A. (1990). A clinical epidemiological study in low back pain: description of two clinical syndromes. *British Journal of Rheumatology*, **29**, 354–7.

Cooper, R.G., Mitchell, W.S., Illingworth, K., St Clair Forbes, W., Gillespie, J.E., and Jayson, M.I.V. (1991). Role of epidural fibrosis and defective fibrinolysis in the persistence of post-laminectomy back pain. *Spine*, **16**, 1044–8.

Coppes, M.H., Marani, E., Thomeer, R.T.W.M., Oudega, M., and Groen, G.J. (1990). Innervation of annulus fibrosus in low back pain. *Lancet*, **336**, 189–90.

Crock, H.V. (1992). Isolated disc resorption. In *The lumbar spine and back pain* (4th edn) (ed. M.I.V. Jayson) pp 307–12. Churchill Livingstone, Edinburgh.

Croft, P. and Jayson, M.I.V. (1994). *Low back pain in the community and in hospitals*. A report to the Clinical Standards Advisory Group of the Department of Health. Prepared by the Arthritis and Rheumatism Council, Epidemiology Research Unit, Manchester.

Deyo, R.A. and Diehl, A.K. (1986). Patient satisfaction with medical care for low-back pain. *Spine*, **11**, 28–30.

Deyo, R.A., Diehl, A.K., and Rosenthal, M. (1986). How many day's rest for acute low back pain? A randomised clinical trial. *New England Journal of Medicine*, **315**, 1064–70.

Dooley, J.F., McBroom, R.J., Taguchi, T., and MacNab, I. (1988). Nerve root infiltration in the diagnosis of radicular pain. *Spine*, **13**, 79–83.

Dubner, R. and Basbaum, A.I. (1994). Spinal dorsal horn plasticity following tissue or nerve injury. In *Textbook of pain* (ed. P.D. Wall and R. Melzak), pp. 225–41. Churchill Livingstone, Edinburgh.

Eyre, D.R. and Muir, H., (1976). Types I and II collagens in intervertebral discs. *Biochemical Journal*, **157**, 267–70.

Fairbank, J.L.T., Park, W.M., McCall, I.W., and O'Brien, J.P. (1981). Apophyseal injection of local anaesthetic as a diagnostic aid in low back syndromes. *Spine*, **6**, 598–605.

Frost, H., Klaber Moffett, J.L., Moser, J.S., and Fairbanks, J.C.T. (1995). Randomised controlled clinical trial for patients with chronic low back pain. *British Medical Journal*, **310**, 151–4.

Garvey, T.A., Marks, M.R., and Wiesel, S.E. (1989). A prospective, randomized double-blind evaluation of trigger-point injection therapy for low-back pain. *Spine*, **14**, 962–4.

Haldeman, S. and Rubinstein, S.M. (1992). Cauda equina syndrome in patients undergoing manipulation of the lumbar spine. *Spine*, **17**, 1469–73.

Health and Safety Executive (1993). *Key fact sheet on back injuries to employees between 1987/88 and 90/91*. Health and Safety Executive, Statistical Services Unit.

Heliovaara, M. *et al*. (1987). Incidence and risk factors of herniated lumbar intervertegral disc or sciatica leading to hospitalization. *Journal of Chronic Diseases*, **40**, 251–8.

Herbert, C.M., Lindberg, K.A., Jayson, M.I.V., and Bailey, A.J.(1975). Changes in the collagen of human intervertebral discs during ageing and degenerative disc disease. *Journal of Molecular Medicine*, **1**, 79–91.

Hoyland, J.A., Freemont, A.J., and Jayson, M.I.V. (1988). Retained surgical debris in post laminectomy arachnoiditis. *Journal of Bone and Joint Surgery*, **70B**, 659–62.

Hoyland, J.A., Freemont, A.J., and Jayson, M.I.V. (1989). Intervertebral foramen venous obstruction—a cause of peri-radicular fibrosis. *Spine*, **14**, 558–68.

Hulshof, C. and van Zanten, B.V. (1987). Whole body vibration and back pain. *Archives of Occupational and Environmental Health*, **59**, 205–20.

Jayson, M.I.V. and Barks, J.J. (1973). Structural changes in the intervertebral discs. *Annals of the Rheumatic Diseases*, **32**, 10–15.

Jayson, M.I.V., Herbert, C.M., and Barks, J.S. (1973). Intervertebral discs, nuclear morphology and bursting pressures. *Annals of the Rheumatic Diseases*, **32**, 308–15.

Jensen, M.C., Brant-Zawadzki, M.N., Obuchowski, N., Modic, M.T., Malkaisian, D., and Ross, J.S. (1994). Magnetic resonance imaging of the lumbar spine in people without back pain. *New England Journal of Medicine*, **331**, 69–73.

Lawrence, J.S. (1977). Disc disorders. In *Rheumatism in populations*, pp. 68–97. Heinemann, London.

Makela, M. (1993). *Publications of the Social Insurance Institution*. The Research and Development Unit, Helsinki.

Malmivaara, A., Hakkinen, U., Aro, T., Heinrichs, M-L., Koskenniemi, L., Kuosma, E., Lappi, S., Paloheimo, R., Servo, C., Vaaranen, V., and Hemberg, S. (1995). The treatment of acute low back pain—bed rest, exercises or ordinary activity? *New England Journal of Medicine*, **332**, 351–5.

Mayer, T.G., Gatchel, R.J., Mayer, H., Kishino, N.D., Keeley, J., and Mooney, V. (1987). A prospective two-year study of functional restoration in industrial low back injury. *Journal of the American*

Medical Association, **258**, 1763–7.

McDevitt, C.R. (1988). Proteoglycans of the intervertebral disc. In *Biology of the intervertebral disc* (ed. P. Ghosh), pp. 151–70. CRC Press, New York.

Mason, V. (1994). *The prevalence of back pain in Great Britain*. (A report prepared for the Department of Health by the Office of Population Censuses and Surveys, Social Survey Division based on the Omnibus Survey March, Apr, June 1993.) HMSO, London.

Mitchell, J.N. (1985). Low back pain and the prospects for employment. *Journal of Social and Occupational Medicine*, **35**, 91–4.

Nachemson, A. (1992). Lumbar mechanics as revealed by lumbar intra-discal measurements. In *The Lumbar spine and back pain* (ed. M.I.V. Jayson), pp. 157–71. Churchill Livingstone, Edinburgh.

Ng, S.C.S., Weiss, J.B., Quinnel, R., and Jayson, M.I.V. (1986). Abnormal connective tissue degrading enzyme patterns in prolapsed intervertebral disc. *Spine*, **11**, 695–701.

Papageorgiou, A., Croft, P., Jayson, M.I.V., and Silman, A. (1995). Estimating the prevalence of low back pain in the general population. Evidence from the South Manchester back pain survey. *Spine*, **20**, 1889–94.

Porter, R.W. (1992). Spinal stenosis of the central and root canal. In *The lumbar spine and back pain* (ed. M.I.V. Jayson), pp. 313–32. Churchill Livingstone, Edinburgh.

Porter, R.W. and Hibbert, C. (1983). Calcitonin treatment for neurogenic claudication. *Spine*, **8**, 585–92.

Porter, R.W. and Hibbert, C. (1986). Back pain and neck pain in four general practices. *Clinical Biomechanics*, **1**, 7–10.

Porter, R.W., Hibbert, C., and Wellman, P. (1980). Backache and the lumbar spinal canal. *Spine*, **5**, 99–105.

Postacchini, F., Lami, R., and Pugliese, O. (1988). Familial predisposition to discogenic low back pain. *Spine*, **12**, 1403–6.

Powell, M.C., Wilson, M., Szypryt, P., Symonds, E.M., and Worthington, B.S. (1986). Prevalence of lumbar disc degeneration observed by magnetic resonnance in symptomless women. *Lancet*, **ii**, 1366–7.

Riihimaki, H., Tola, S., Videman, T., and Hannin, K. (1989). Low back pain and occupation. *Spine*, **14**, 204–9.

Sachs, B.T., Zindrick, M.R., and Beasley, R.D. (1993). Reflex sympathetic dystrophy after operative procedures on the lumbar spine. *Journal of Bone and Joint Surgery*, **75A**, 721–5.

Troup, J.D.G. and Edwards, F.C. (1985). *Manual handling and lifting: an information and literature review with special reference to the back*. HMSO, London.

Twomey, L. and Taylor, J. (1995). Exercise and manipulation in the treatment of low back pain. *Spine*, **20**, 615–9.

Vernon Roberts, B. (1992). Age related and degenerative pathology of the intervertebral disc and apophyseal joints. In *The lumbar spine and back pain* (4th edn) (ed. M.I.V. Jayson), pp. 17–41. Churchill Livingstone, Edinburgh.

Videman, T., Nurminen, M., and Troup, J.D.G. (1990). Lumbar spinal pathology in cadaveric material in relation to history of back pain, occupation and physical loading. *Spine*, **15**, 728–40.

Waddell, G. (1987). A new clinical model for the treatment of low back pain. *Spine*, **12**, 632–44.

Waddell, G., Mc Cullough, J.A., Kummel, E., and Venner, R.M. (1980). Non organic physical signs in low back pain. *Spine*, **5**, 117–25.

Walsh, K. *et al*. (1992). Low back pain in eight areas of Britain. *Journal of Epidemiology and Community Health*, **26**, 227–30.

Weinstein, J. (1986). Mechanisms of spinal pain. *Spine*, **11**, 999–1001.

Weiss, J.B. and McLaughlin, B. (1993). Activation of gelatinase-A and reactivation of the gelatinase-A inhibitor complex by endothelial cell stimulating angiogenesis factor. *Journal of Physiology*, **456**, 49.

Wiltse, L.L. and Rothman, S.L.G. (1990). Lumbar and lumbosacral spondylolisthesis: classification, diagnosis and natural history. In *The lumbar spine* (eds J.N. Weinstein and S.W. Weisel), pp. 471–99. Saunders, Philadelphia.

Wright, D., Barrow, S., Fisher, A., Hull, S.D., and Jayson, M.I.V. (1995). Influence of physical, psychological and behavioural factors on consultations for back pain. *British Journal of Rheumatology*, **34**, 156–61.

Wyke, B. (1987). The neurology of low back pain. In *The lumbar spine and back pain* (3rd edn) (ed. M.I.V. Jayson), pp. 56–99. Churchill Livingstone, Edinburgh.

5.18.2 Cervical pain syndromes

Allan I. Binder

Introduction

The vast majority of patients with neck pain have degenerative or mechanical lesions. 'Cervical spondylosis' is often used to describe neck pain of a mechanical nature, but the term is variably applied to include soft tissue, disc, and degenerative bony lesions (Resnick 1985). Furthermore, the boundary between 'normal' ageing and disease is unclear. The pathology of cervical spondylosis is assumed to be identical to lumbar spondylosis, but this similarity has been questioned (Bland and Boushey 1990). Less research has been performed into cervical than lumbar pain, as severe disability is less common with cervical disease.

This chapter will concentrate on the common mechanical and degenerative pain syndromes that predominantly affect the cervical spine and the principles of their treatment. Other conditions which can cause neck pain will be considered in more detail in the relevant chapters devoted to the particular diseases.

Epidemiology

About 10 per cent of the adult population have neck pain at any one time (Lawrence 1969; Hadler 1985), with symptoms often being associated with specific occupations or sporting activities (Holt 1972). While this prevalence is similar to that of low back pain, to lose time from work is unusual, and under 1 per cent of patients develop neurological deficit. As so many patients with neck pain never seek medical care, the true prevalence of chronic disease is uncertain.

Functional anatomy of the cervical spine

The cervical spine is the most mobile and least stable part of the human spine, consisting of seven vertebrae connected by five intervertebral discs. There are 37 separate articulations (Bland and Boushey 1990), and a complex system of ligaments and muscles, which with the varying shapes of individual vertebrae, and different methods of articulation, are responsible for the myriad movements of the head and neck. Any of these structures can be the source of pain.

The vertebral arteries pass close to the zygapophyseal joints, immediately anterior to the emerging cervical nerve roots. The preganglionic sympathetic nerve fibres run closely adherent to the carotid and vertebral vessels, to synapse with the stellate, middle, and superior cervical ganglia. The postganglionic sympathetic fibres then separate into three directions: some fibres go to the upper limbs to provide autonomic control of circulation, sweating, and proprioception; other fibres re-enter the spinal cord via the intervertebral foramina to synapse in the vestibular apparatus, cerebellum, thalamus, and hypothalamus; and some fibres pass upward with the vertebral and carotid arteries to the brain (Bland and Boushey 1990). Involvement of the vertebral arteries and sympathetic nerves in the degenerative process may explain many of the unusual features associated with disease of the cervical spine.

The lack of room in the spinal canal between C4 and T2 is due to the enlargement of the cord in this region. As degenerative changes are also most frequent and severe between C5 and T1 (Hayashi *et al*. 1988), compression of the cord usually develops at this site. Inflammatory arthropathies in contrast, have a predilection for involvement of the atlantoaxial and upper cervical spine. Minor congenital spinal abnormalities increase the risk of degenerative change (Hensinger 1991), and there is a remarkable similarity in the pattern of degenerative disease in monozygous twins (Palmer *et al*. 1984), possibly reflecting the similarity in the shape of their vertebrae.

The anterior and posterior nerve roots from C4 to T1 exit through the dural root sleeves, and traverse the intervertebral foramina. They then merge to form the brachial plexus (Fig. 1), which lies between the clavicle and first rib, in close proximity to the subclavian vessels. The neurovascular bundle is susceptible to compression at various sites in the thoracic outlet, which lies between the neck and axilla. Cervical nerves have a dermatomal representation (Fig. 2), which explains the radicular pattern of symptoms in the upper limbs caused by impingement on nerve roots.

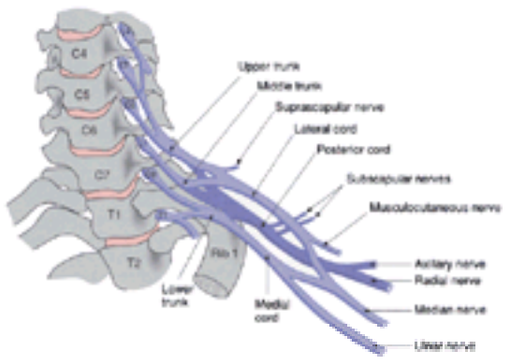


Fig. 1 Diagram showing the nerve roots (C4 to T1) forming the brachial plexus, with the peripheral nerves which arise from the plexus.

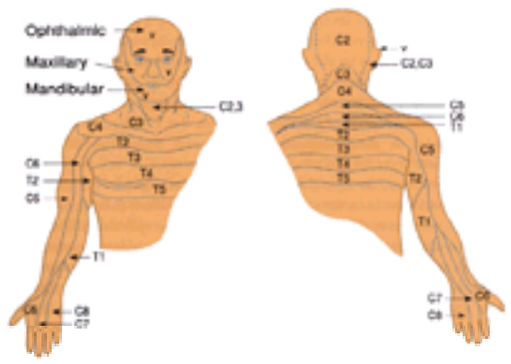


Fig. 2 Dermatomal distribution of the cervical and upper thoracic nerves which reflect the radicular pattern of nerve root lesions.

Aetiology of neck pain

Although 'cervical spondylosis' accounts for most cases of neck pain, there are many other causes of pain which need to be excluded (Table 1). The cervical spine is frequently involved in polymyalgia rheumatica, rheumatoid arthritis, and other arthropathies, and neck pain can result from serious local pathology, such as infection or malignancy. It is also often the site of referral of pain from distant sources (Table 2).

Table 1 Causes of neck pain

Acromioclavicular joint, temporomandibular joint, teeth
Heart — angina pectoris, myocardial infarction
Aorta — aneurysm
Pharynx — infection, tumour
Lung — bronchogenic carcinoma, Pancoast tumour, apical lesion
Abdomen — disease of the gallbladder, stomach, oesophagus (including hiatus hernia) and pancreas
Diaphragm — subphrenic abscess
Central nervous system — migraine, 'tension' headache, tumour, posterior fossa lesion, meningitis, arachnoiditis
Lymph node — cervical lymphadenitis
Shoulder — frozen shoulder, reflex sympathetic dystrophy

Table 2 Sites of referred pain to the neck

Cervical spondylosis

Pathology

With ageing, degenerative change usually develops in the cervical spine (Gore *et al*. 1986). This is readily apparent on radiographs of most adults over the age of 30 years. The term 'cervical spondylosis' refers to this progressive degenerative process, which affects all levels of the cervical spine (Lestini and Wiesel 1989), but with more severe changes at the lower levels (Hayashi *et al*. 1988). There is a sequential change in the intervertebral discs, with osteophytosis of the vertebral bodies and changes in the facet joints and laminal arches. There is a continuum from 'normal' ageing to the overtly pathological state. The correlation between the degree of radiological change and the presence and severity of pain is poor (Van der Donk *et al*. 1991), and hence there is disagreement on the exact definition of cervical spondylosis.

With increase in age, there is a steady decrease in the degree of hydration of the intervertebral discs. Degeneration follows and radiolucent nitrogen-filled spaces (vacuum disc phenomenon) may develop within the disc. When present, these spaces confirm the degenerative nature of the process, and are reassuring, as they exclude a diagnosis of infection (Resnick 1985). With progression of the ageing process, loss of disc height and bony sclerosis follow. Osteophytosis often then develops, with degenerative changes (osteoarthritis) in the nearby zygapophyseal joints and other articulating surfaces. Involvement of the spinal ligaments in the degenerative process can lead to a loss of stability, which is an important factor in causing myelopathy in elderly patients.

Clinical assessment

Detailed history and examination will usually confirm the degenerative nature of the condition or alert one to a need to exclude more serious pathology. Assessment of

the shoulder joints is also necessary to determine if there is coexisting shoulder pathology (Hawkins *et al*. 1990), although cervical pathology itself can cause painful limitation in the range of shoulder flexion and abduction above 90°.

Symptoms

Pain

This is the most common symptom of cervical pathology, and is usually poorly localized to the neck and shoulders when arising from deep structures, such as ligaments, muscles, joints, discs, or bone. The pain can however, be clearly defined and in a dermatomal distribution when caused by irritation of the nerve roots (Fig. 2). Pain arising from structures of the cervical spine is characteristically altered (aggravated or relieved) by their movement. The causes of neck pain are shown in Table 1.

Pain is most often referred to the occiput, nuchal muscles, and superior aspect of the shoulders. Heaviness or aching of the upper limbs also reflects cervical origin, and the pain can closely mimic soft tissue lesions of the shoulder, elbow, and wrist (Gunn and Milbrandt 1976; Murray-Leslie and Wright 1976). Retro-orbital and temporal pain suggest referral from the upper cervical levels (C1 to C3). Temporal pain when associated with tenderness can be misinterpreted as evidence of giant-cell arteritis. Pain can also be referred to the upper thoracic spine and interscapular areas. Some patients, especially with lesions of C6 and C7 complain of anterior chest pain, which closely mimics coronary ischaemia (Brodsky 1985). Pseudoangina of this type is sometimes associated with local tenderness of the chest wall. There is particular diagnostic difficulty in patients with a combination of both coronary insufficiency and cervical spondylosis, as anginal pain is more likely to radiate to the neck in patients with symptomatic cervical spondylosis. Coronary angiography is the key investigation in the assessment of the severity of the cardiac lesion in these cases.

Stiffness

This is a common accompaniment of ageing, degeneration, and many vertebral diseases, and can be reversible or irreversible.

Dizziness

This may occur as a result of involvement of the vertebral arteries, especially in the presence of severe degenerative spinal disease. Atheroma and disturbance of flow in the vertebral vessels may contribute to the development of dizziness in older patients. Vertigo and faintness caused by vertebrobasilar disease is nearly always accompanied by other focal symptoms of transient ischaemia of the brainstem or occipital lobes. In some patients with cervical pathology, tinnitus and gait disturbance (Sudarsky and Ronthal 1983) can occur, as a result of irritation of the sympathetic nerves.

Occipital headache

This is a common manifestation of cervical degenerative disease, especially when the disease affects the upper cervical levels. Occipital neuralgia is another cause of occipital pain in some patients. Edmeads (1988) discussed the controversy over the importance of disease of the cervical spine in the aetiology of non-occipital headache; Wober-Bingol *et al*. (1992) found no link, but Nagasawa *et al*. (1993) found a strong association.

Blurring of vision and diplopia

These symptoms, where associated with neck movement can result from cervical pathology, and have been attributed to irritation of the sympathetic nerve supply to the eye.

Dysphagia

This can be caused by irritation of the cranial or sympathetic nerves, muscular spasm, or compression of the oesophagus by large anterior osteophytes (Sobol and Rigual 1984).

Paraesthesia and sensory loss

Numbness and tingling is usually vague and ill-defined in cervical spondylosis, but can be precise, following the clear segmental dermatomal distribution of nerve entrapment (see Radiculopathy below). The symptoms are often affected by neck movement, or are postural, being worse at night or with specific activities. Lesions of C1 to C3 can cause paraesthesia affecting the face, head, and tongue. Involvement of the C4 root gives symptoms referred to the superior aspect of the shoulder, and C5 to T1 lesions give numbness in the upper limb (see Fig. 2).

Weakness

Mechanical disease of the cervical spine most typically gives a subjective feeling of heaviness or weakness, especially affecting the hands, but without true weakness on formal testing. Objective muscle weakness, wasting, and fasciculation in the absence of systemic upset suggests a radiculopathy, thoracic-outlet syndrome, or neuropathy of the brachial plexus. Abnormality of gait due to spasticity of the lower limbs suggests myelopathy.

Rare manifestations

Unusual features sometimes resulting from pathology of the cervical spine are shown in Table 3. Most of these symptoms result from irritation of the sympathetic nerves, but the clue to their vertebral origin is reproducibility with neck movement.

Visual—blurring, diplopia, retro-orbital pain
Auditory—tinnitus, nerve dealness, earache, poor balance
Intestinal—dysphagia, nausea, vomiting, diarrhoea
Cardiac—'pseudoangina', dyspnoea, palpitations
Respiratory—cough, dyspnoea, sneezing
Central nervous system—syncope, 'drop attacks', vertebro-basilar insufficiency, speech disturbance, migraine

Table 3 Rarer symptoms arising from cervical spine pathology

Signs

Tenderness

When due to degenerative disease, tenderness is poorly localized and of variable severity. It is usually worse in the lower cervical region, and may be associated with muscle spasm. Tender myofascial trigger points are a characteristic feature of the 'fibrositis/fibromyalgia' syndrome (Smythe 1986), but also occur with disease of the facet joints. Exquisite localized tenderness over a vertebral body may suggest osteomyelitis or malignancy, particularly if the patient has features of systemic upset or

abnormality of blood tests such as full blood count, erythrocyte sedimentation rate, C-reactive protein, or protein electrophoresis.

Limitation of movement

This is a feature of ageing and degeneration and may be otherwise asymptomatic or accompanied by pain. Severe irreversible loss of range particularly on lateral flexion and rotation occurs with cervical spondylosis, but is more characteristic of the spondylarthropathies and diffuse idiopathic skeletal hyperostosis (**DISH**). Reversible stiffness, worse in the early morning, is more suggestive of polymyalgia rheumatica or an inflammatory arthropathy than of a degenerative lesion.

Neurological deficit

Neurological abnormalities are characteristic of radiculopathy, but also occurs in thoracic-outlet syndromes and neuropathy of the brachial plexus (see below).

Radiological assessment (see also Chapter 4.9.1)

Routine radiographs of the anteroposterior and lateral spine are sufficient to indicate the severity of bone and disc pathology. A through-mouth view to outline the odontoid peg, when combined with flexion/extension radiographs, will demonstrate existing subluxation. Oblique views of the cervical spine will show the intervertebral foramina, and are useful in patients suspected of having a radiculopathy (Fig. 3). Loss of cervical lordosis on the lateral radiograph in patients with neck pain is usually ascribed to muscle spasm, but this association has been questioned (Helliwell *et al*. 1994).

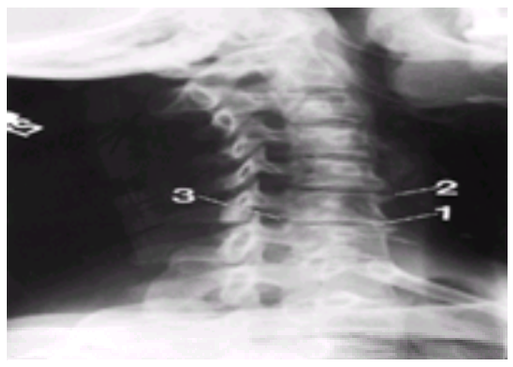


Fig. 3 Oblique radiograph of the cervical spine in a patient with cervical spondylosis, showing the loss of disc height (1), anterior osteophytosis (2), and foraminal narrowing (3).

Myelography, computed tomography (**CT**), and CT-myelography, have greatly improved the radiological assessment of cervical disease, especially at the upper cervical levels, which are hard to visualize on routine radiographs.

Magnetic resonance imaging (**MRI**) is less invasive and equally reliable as CT-myelography for the visualization of cervical abnormalities (Nagata *et al*. 1990), and is the radiological investigation of choice (Kramer *et al*. 1991), giving detailed information about the spinal cord, bones, discs, and soft tissue structures (Fig. 4 and Fig. 5). MRI is particularly valuable in demonstrating congenital abnormalities, cord tumours, demyelination, and disc lesions. Degenerative vertebral changes are well demonstrated, but as they occur normally with ageing, need to be interpreted with care. Asymptomatic people often show important pathological lesions, such as narrowing of the disc space, osteophytosis, or even compression of the spinal cord (Boden *et al*. 1990), and the frequency of abnormal scans in asymptomatic people increases with age. Boden *et al*. (1990) and Lehto *et al*. (1994) have emphasized that decisions about surgery should be based on clinical indications with the support of the MRI findings, and not on radiological data alone. Similar false-positive studies have been described with plain radiography, CT (Wiesel *et al*. 1984), and all other radiological techniques.

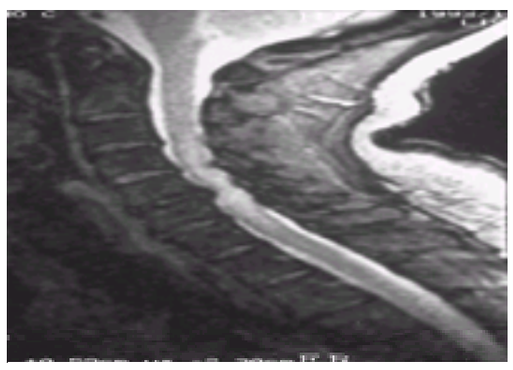


Fig. 4 MRI scan showing loss of height and signal affecting several discs, with multisegmental spondylotic bars. Compression of the cord is noted by protrusion of the C5/6 disc with myelopathic changes in the cord.

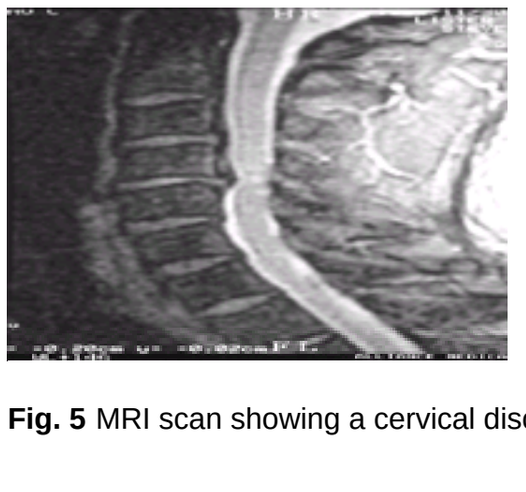


Fig. 5 MRI scan showing a cervical disc prolapse at C4/5 with impingement upon the cervical cord and secondary myelopathic changes in the cord itself.

Electrodiagnostic studies

Electromyography, nerve conduction studies, and somatosensory evoked potentials may help in the elucidation of the diagnosis, especially where primary neurological disease or polymyositis need to be differentiated from radiculopathy or myelopathy secondary to degenerative vertebral disease (Dvorak *et al*. 1990). Specific neurophysiological findings have been described in cervical spondylosis associated with myelopathy or radiculopathy (Yiannikas *et al*. 1986), but there is no consensus as to their reliability or diagnostic value, particularly in uncomplicated cervical spondylosis. Magnetic stimulation of the motor cortex (Di Lazzaro *et al*. 1992) and spinal cord evoked potentials (Baba *et al*. 1993a) have more value in the assessment of myelopathic patients, but at best, the electrodiagnostic techniques

only provide supportive evidence to clinical and radiological findings.

Other tests

Other tests such as full blood count, erythrocyte sedimentation rate, C-reactive protein, protein electrophoresis, liver function tests, urate, rheumatoid factor, and bone scan may be necessary to exclude other causes of neck pain (Table 1 and Table 2).

Complications of cervical spondylosis

Cervical myelopathy

The relatively tight fit of the spinal cord in the lower cervical region, as a result of the natural expansion of the cervical cord, accounts for an increased risk of development of a myelopathy at this site.

Myelopathy can develop *de novo*, or can occur in patients with known cervical spondylosis, particularly in the presence of a congenitally narrow spinal canal (Fukui *et al*. 1990). With progressive degeneration and resulting osteophytosis and instability, pressure on the spinal cord can develop.

In elderly patients, myelopathy can result from compression of the spinal cord, disturbed blood supply, or a combination of both factors. Disc protrusion, posterior osteophyte formation, and retrolisthesis (posterior slide) due to ligamentous laxity, contribute to age-associated spondylotic spinal compression (Hayashi *et al*. 1988). Even when the static spinal canal diameter is adequate, many elderly patients show significant and progressive 'dynamic' spinal stenosis (Hayashi *et al*. 1988; Fukui *et al*. 1990), which can be demonstrated (Jinkins *et al*. 1986) by a comparison of flexion and extension CT or MRI scans.

In young patients, a sudden onset of myelopathy or radiculopathy suggests cervical disc prolapse (see below). Cervical myelopathy can also result from vertebral diseases such as osteomyelitis, malignancy, and crystal arthropathy, and intramedullary disease can present with similar symptoms (Table 4.)

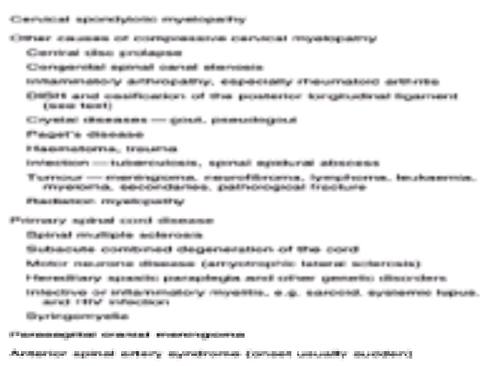


Table 4 Differential diagnosis of cervical myelopathy

Cervical myelopathy typically presents with a slowly progressive disability over a period of weeks or months, although the onset can be sudden, especially when caused by trauma or central disc prolapse. The most common, early manifestations of myelopathy are clumsiness, weakness or dysaesthesia of the hands, or gait disturbance due to upper motor-neurone abnormality in the lower limbs.

Numbness or tingling in the hands is poorly localized, unless radiculopathy is also present. There is usually no gross sensory deficit, but blunting of light touch, pin-prick, and temperature sensation can occur. Stereoanaesthesia, and isolated loss of position sense in the hands, is also sometimes seen, particularly in high cervical lesions. Clumsiness or weakness of the hands may rarely be associated with variable degrees of wasting and evidence of lower motor-neurone abnormality at the level of the lesion. Characteristic features in the legs are of a spastic paraparesis, and in elderly patients, undiagnosed cervical myelopathy is an important case of gait disturbance (Sudarsky and Ronthal 1983).

Tendon reflexes are often reduced in the arms and increased in the legs with up-going plantar responses and ankle clonus, although the balance is dependent on the level of the lesion. Inversion of the supinator jerk (reduced or absent supinator reflex with brisk finger flexion when tapping over the lower radius), with a brisk triceps jerk, suggests involvement of the spinal cord at the C5/6 level. A brisk finger-flexion reflex (on tapping the examiner's own fingers when placed over the flexor surface of the terminal phalanges), and a positive Hoffman's sign (brisk thumb flexion when the distal phalanx of the middle finger is flicked into extension) also suggests hyperreflexia. Sphincter control is usually well maintained in cervical myelopathy complicating cervical spondylosis.

Neck pain may not be a prominent feature of myelopathy, and when present, its mechanism of production is complex and poorly understood. The cervical lesion is often multisegmental (Fig. 4), and in some cases there is a large discrepancy between the sensory level on clinical examination and the true level of cord compression (Simmons *et al*. 1986).

Shooting electric pain down the spine, arms, or legs, or temporary neurological deficit when the head is flexed (Lhermitte phenomenon) may be mentioned by the patient. Great caution (and prior radiography) is necessary when attempting to confirm the presence of this sign, as permanent deterioration in the neurological condition may follow its performance. Lhermitte phenomenon is usually a sign of cervical myelopathy of an inflammatory type (typically multiple sclerosis), but can be due to compressive, or rarely other (e.g. vitamin B_{12} deficiency) causes.

MRI (Fig. 4) is the investigation of choice in patients with suspected myelopathy (Kramer *et al*. 1991; Okada *et al*. 1994), although CT-myelography also has value. The anteroposterior or sagittal diameter of the canal on radiological investigation is a useful indicator of possible myelopathy, and compression of the cord is likely if the diameter at any level is 10 mm or less (Murone 1974). With a diameter above 13 mm, compressive symptoms rarely occur solely on the basis of degenerative disease. Atrophy of the spinal cord on CT-myelography (Jinkins *et al*. 1986), or intrinsic spinal cord damage on MRI (Fig. 4), especially after myelopathic decompression (Nagata *et al*. 1990; Batzdorf and Flannigan 1991; Matsuda *et al*. 1991) are associated with a less favourable outcome.

The most sensitive neurophysiological investigation in the assessment of patients with cervical myelopathy is somatosensory evoked potentials, especially when done in both upper and lower limbs (Yiannikas *et al*. 1986). However, these tests rarely add much to the evaluation, particularly in difficult cases.

Cervical radiculopathy

Nerve root compression as a result of cervical spondylosis usually occurs at the C5 to C7 level with lesions being single or multiple, and uni- or bilateral.

Acute radiculopathy in the young usually follows trauma or lateral prolapse of a disc. In middle-aged and elderly patients, asymptomatic osteophytic encroachment gradually narrows the nerve root foramina until trivial trauma is sufficient to precipitate acute or subacute radiculopathy. Chronic radiculopathy may arise insidiously, or may follow the more acute syndromes. As chronic radicular syndromes are usually due to degenerative disease, they can be associated with myelopathy.

Neurological features follow a segmental distribution in the upper limb (see Fig. 2). Sensory symptoms such as shooting pain, paraesthesia, and hyperaesthesia are more common than motor weakness or a change in the reflexes. Arm weakness, when present, is of a lower motor-neurone type, and may be accompanied by a loss of reflex (biceps and supinator jerk in C5/6 lesions; triceps jerk in C7 lesions).

Radicular pain can be exacerbated by manoeuvres that stretch the affected root, or by an increase in intrathoracic or intra-abdominal pressure, as occurs with coughing, sneezing, or the Valsalva manoeuvre. If sustained shoulder abduction induces a temporary relief in radicular pain, a C6 lesion can be expected (Fast *et al*.

1989). This test has value both diagnostically and as part of a conservative treatment regimen (Farmer and Wisneski 1994).

Oblique cervical radiographs often demonstrate narrowing of the nerve root foramina (Fig. 3) in patients with radiculopathy (Pyhtinen and Laitinen 1993). Electromyography may show abnormalities in patients with clinical signs of root compression, and additional information can sometimes be obtained from studying superficial radial sensory-evoked potentials. However, these techniques are of limited value where symptoms of radiculopathy occur without objective signs (Yiannikas *et al*. 1986).

Cervical radiculopathy (Table 5) can rarely be caused by malignant tumours, such as Pancoast tumour (Vargo and Flood 1990), or nasopharyngeal malignancies, particularly when accompanied by bony infiltration. There is often difficulty in differentiating cervical spondylosis with myelopathy or radiculopathy from primary neurological diseases, such as syringomyelia, motor neurone disease (amyotrophic lateral sclerosis), or tumour, particularly when unrelated osteoarthrosis of the cervical spine coexists (see Table 4 and Table 5).

Cervical spondylotic radiculopathy
Other causes of compressive cervical radiculopathy
Lateral disc prolapse
Inflammatory arthropathy, e.g. rheumatoid arthritis
Trauma, haematoma
Arachnoiditis
Infection
DISH and ossification of the posterior longitudinal ligament (see text)
Tumour—primary, secondary, Pancoast tumour of lung
Brachial plexus and thoracic-outlet syndromes
Peripheral nerve and 'double-crush' syndromes
Complete rotator-cuff tear and reflex sympathetic dystrophy
Motor neurone disease and other primary spinal cord diseases (see Table 4)

Table 5 Differential diagnosis of cervical radiculopathy

Other important mechanical pain syndromes

Cervical disc prolapse

Herniation of a cervical disc is a common cause of neck, shoulder, and arm pain in the younger patient. It typically presents as a sudden onset of neck pain with associated muscle spasm, followed by radicular symptoms and signs in the upper limb. It may follow trivial trauma or awkward movements of the neck, but in many cases there is no obvious cause. Some patients give a history of preceding mechanical cervical pain, and manipulation or chiropractic therapy can precipitate the acute disc prolapse. The lesions almost always affect the lower cervical levels.

Lateral disc protrusion, which is more common, results in radiculopathy affecting the C7 root in 70 per cent of cases, the C6 root in 20 per cent, and the C5 or C8 roots in the remaining 10 per cent of cases (Yoss *et al*. 1957). The site of pain and the neurological features depend on the level of the lesion (Smythe 1994; Nakajima and Hirayama 1995), but the syndromes are often incomplete. Limitation of cervical range is usual, and pain can be aggravated by movement (especially hyperextension), coughing, and sneezing, and relieved by traction on the neck.

Central disc prolapse, if large, causes myelopathy (Fig. 5), and when painless may simulate degenerative syndromes of the spinal cord, such as motor neurone disease or multiple sclerosis (see Table 4). MRI (Vanderburgh and Kelly 1993) or CT-myelography are essential to differentiate disc prolapse or other causes of compressive myelopathy from intrinsic disease of the spinal cord, even when the latter is considered more likely.

Disc prolapse is much less common in the thoracic spine and is difficult to diagnose clinically, unless myelopathy or radiculopathy intervene.

Diffuse idiopathic skeletal hyperostosis (DISH, Forestier's disease, ankylosing hyperostosis)

This common disorder of unknown aetiology is most often diagnosed in middle-aged or elderly white males, who show an ossifying diathesis. Patients develop flowing ossification along the anterior and lateral aspects of the spine, bridging between vertebrae, but with preservation of normal disc heights. The condition affects the thoracic spine most severely, but can also occur in the cervical and lumbar regions. The lesions start at the ligamentous insertions (entheses), with progression of the flowing osteophytes along the ligaments. Similar ossification is also found at entheses along the pelvis and other parts of the skeleton.

Severe limitation of cervical spine movement usually develops without any other symptoms (Boachie-Adjei and Bullough 1987). However, some patients present with pain, myelopathy, and less commonly, radiculopathy. Large anterior osteophytes in the cervical region can also cause dysphagia owing to oesophageal obstruction. If anterior osteophytes due to cervical spondylosis or DISH are suspected as the cause of the dysphagia, barium swallow is the preferred method of initial assessment, as endoscopy carries some risk of inadvertent oesophageal perforation (Kristensen *et al*. 1988). The ossification progresses more rapidly at mobile compared with immobile segments (Suzuki *et al*. 1991).

Ossification of the posterior longitudinal ligament

This is probably a variant of DISH (Hukuda *et al*. 1983), or is closely associated with it (Resnick *et al*. 1978). It is most frequently found in Japanese or other Asian patients, but is occasionally found in all ethnic groups. Like DISH, it occurs most often in middle-aged and elderly males, with a majority of cases producing no symptoms. However, in association with congenital and degenerative vertebral abnormalities, even trivial trauma can precipitate cervical myelopathy or radiculopathy. The ossification can be difficult to detect on plain radiographs, and both MRI (Otake *et al*. 1992; Baba *et al*. 1995) and CT scanning have value where the diagnosis is suspected. Calcification of the ligamentum flavum and posterior longitudinal ligament can also result from calcium pyrophosphate crystal deposition (pseudogout), confirmed on biopsy in patients having surgery for compressive myelopathy (Brown *et al*. 1991; Baba *et al*. 1993b).

Soft tissue syndromes considered to be 'mechanical'

Spasm, postural, and anxiety/tension-related neck pain

Most adults suffer transient neck pain and stiffness, which is attributed to awkward posture of the neck, especially during sleep. The pathology is uncertain and the condition is self-limiting within days. Occupations and hobbies which involve heavy lifting or unusual positions of the neck are a cause of recurrent pain of this type. Anxiety and tension can manifest as episodic or chronic neck pain in susceptible individuals. Poor posture of the cervical and thoracic spine (Griegel-Morris *et al*. 1992), which can cause muscle spasm, and hence traction on the nerve roots, may explain these symptoms. The interaction between cervical spondylosis, tension, and the fibrositis/pain-amplification syndrome needs further clarification.

'Fibrositis/fibromyalgia' and pain-amplification syndrome (see also Chapter 5.14)

'Fibrositis' is a term that has been used for many decades to describe non-specific neck pain. Smythe (1986) defined the diagnosis more precisely, describing a pain-amplification syndrome associated with localized points of myofascial tenderness. This can be confined to the cervical spine (fibrositis), or can be more generalized (fibromyalgia), with the tender points concentrated around the entheses. Patients localize the pain to the posterior and lateral regions of the neck with associated stiffness. Tenderness is noted most specifically anterior to the transverse processes in the lower cervical spine. Sleep disturbance and fatiguability have been described in some cases. There is greater agreement on diagnostic features with the adoption of the American College of Rheumatology Fibromyalgia Classification Criteria (Wolfe *et al*. 1990).

Neck pain associated with tender trigger points is also found as a stress-related phenomenon (Croft *et al*. 1994; Wolfe *et al*. 1995). Poor posture, underlying vertebral degeneration, especially of the facet joints, and pain amplification may be important elements contributing to the severity of pain in these cases. Pain amplification may also explain the disproportionately severe pain in some patients with rheumatoid arthritis and other inflammatory arthropathies. Treatment of fibromyalgia is disappointing (Simms 1994; Wilke 1995), and the level of disability is often similar to rheumatoid arthritis (Martinez *et al*. 1995).

'Whiplash syndrome'

Acute flexion–extension injuries, typically in automobile accidents following rear-end impact, can cause acute neck pain (whiplash injury), even in the absence of bony damage. More severe symptoms are associated with an unprepared occupant, acceleration/deceleration injury, and a rotated or inclined head position (Sturzenegger *et al*. 1994). Although seat belts have reduced the incidence of severe injuries, the frequency of whiplash-type lesions has continued to rise (Bourbeau *et al*. 1993; Galasko *et al*. 1993), often leading to considerable and prolonged disability, complicated by medicolegal difficulties.

Severe cervical pain, spasm, loss of range, and occipital headache, are the most common symptoms, but many patients also complain of widely radiating pain, headache, vertigo, memory loss, poor concentration, fatigue, or neurological symptoms in the upper limbs (Radanov *et al*. 1992). Very occasionally, temporary or permanent quadriplegia may ensue.

The pathology of the condition is unclear, but soft tissue damage, bleeding into the muscles, and zygapophysial joint injury (Barnsley *et al*. 1995) may explain the acute pain and spasm. Traction and microinjury of the cord, and acceleration/deceleration injury to the brain may be responsible for some of the other symptoms. The presence of cervical spondylosis adds to the risks of damage to the spinal cord or emerging nerve roots. Despite the frequency of whiplash injury, there is no objective investigation which if carried out early on, can define the site and severity of the damage and so predict outcome. Technetium isotope scanning (Barton *et al*. 1993), CT (Antinnes *et al*. 1994), and MRI (Jonsson *et al*. 1994; Pettersson *et al*. 1994) are useful in some cases of whiplash, although the correlation with clinical symptoms and signs is often poor.

The management of whiplash immediately after the injury, remains controversial. Rest in a collar is usually advocated, with physiotherapy (see below) being started as soon as the acute spasm has settled. Mealy *et al*. (1986) found significantly better recovery with early active mobilization, and McKinney *et al*. (1989) found similar benefits from a regimen of early exercise at home. However, Pennie and Agambar (1990) were unable to confirm any benefits of early active treatment for whiplash injury in a large series of patients.

Patients need to be warned that recovery may take 18 months or longer. Maimaris *et al*. (1988) and Evans (1992) found a poor outcome to be associated with older age, occipital headache, neck stiffness, referred symptoms, abnormal neurological signs, degenerative changes on radiographs, and symptoms persisting for more than 2 months. Psychosocial factors and personality traits were less important than the physical factors (reflecting severity) mentioned above, in predicting outcome (Radanov *et al*. 1994), and the settlement of litigation was not always associated with rapid recovery (Parmar and Raymakers 1993). Maimaris *et al*. (1988) reported that a third of their patients were symptomatic beyond 2 years, and Watkinson *et al*. (1991) found some symptoms in 86 per cent and intrusive symptoms in 23 per cent of patients reviewed at an average follow-up of 10.8 years. It is uncertain if whiplash injury accelerates the progression of degenerative change in affected patients (Hamer *et al*. 1993), but a point is reached where physical therapy should be stopped, as it is unlikely to result in any further improvement (Hirsch *et al*. 1988).

Treatment of mechanical cervical disease

Most mechanical neck pain is responsive to conservative therapeutic regimens, although surgical intervention is sometimes necessary. Box 1 summarizes the principles of therapy and timing of MRI scanning in patients with uncomplicated acute and chronic mechanical cervical syndromes and those with neurological complications.

Bedrest

This is only necessary in acute lesions, especially with neurological deficit, such as a prolapsed cervical disc, trauma, or when there is a suspicion of infection or malignancy. A hard collar is advisable in these cases, and patients should be watched carefully for a deterioration in neurological status, when urgent neurosurgical intervention may be necessary.

Physiotherapy

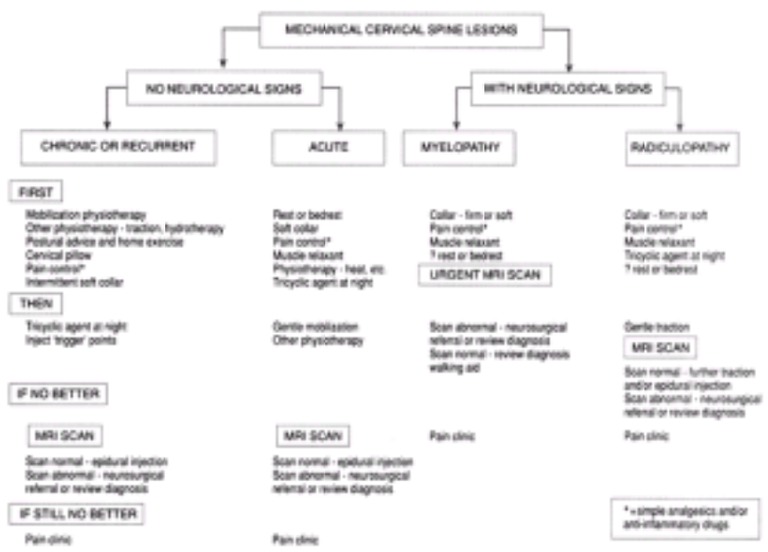
This is the mainstay of treatment of chronic cervical syndromes, and needs to stress the improvement in muscle strength and range of movement, using passive mobilization techniques. Other electrical modalities of treatment may help by giving symptomatic relief to patients, especially where mobilization physiotherapy is contraindicated. Advise on posture and relaxation (such as the Alexander technique), weight reduction, and a home exercise programme are also likely to be beneficial. The avoidance of awkward head positions and lifting heavy weights, especially in occupational or recreational activities, may also prove of value. Dizziness, when due to cervical causes, usually proves unresponsive to physiotherapy and is often a limiting factor to this type of treatment. Transcutaneous electrical nerve stimulation and acupuncture can help reduce pain in some patients. Manipulation by chiropractors is often beneficial (Cassidy *et al*. 1992), but can prove hazardous, especially when radiological examination is not undertaken before treatment (Raskind and North 1990). Dissection of the vertebral arteries and vertebrobasilar strokes (Frisoni and Anzola 1991) are well-recognized complications, for example.

Cervical traction is particularly valuable in patients with radiculopathy, but may also be beneficial in other patients with cervical pathology. It can be used in continuous or intermittent regimens, but whichever regimen is used, great care is necessary, as some patients fail to tolerate this form of treatment or show neurological deterioration as the weight is applied.

Soft collar and cervical or soft pillow

A soft collar can be worn during periods of increased pain, and is especially valuable to facilitate sleep when this is disturbed by pain. Special pillows are also available to provide support for the neck during sleep, although some clinicians prefer to recommend a soft pillow that can be moulded into the appropriate shape to support the neck. Patients should be discouraged from using more than one or two pillows, to prevent undue angulation of the neck which can precipitate symptoms at night and on waking.

Box 1 Principles of treatment of mechanical cervical syndromes and timing of MRI scanning



Drugs

Analgesic tablets should be used during periods of increased pain, but anti-inflammatory and muscle relaxant drugs are only justified in acute situations where pain is particularly severe. Low-dose amitryptiline (10 to 75 mg at night) may alter the pain amplification cycle noted in some patients.

Cervical epidural injection

The injection of depot steroids into the cervical epidural space can maximize pain control in patients in whom serious underlying pathology, such as infection or tumour, or gross neurological deficit have been excluded (Rowlingson and Kirschenbaum 1986). Epidural injections often produce long-lasting improvement in cervical degenerative lesions (Cicala *et al*. 1989), especially in patients with radicular pain (Ferrante *et al*. 1993; Castagnera *et al*. 1994). The injection of local steroid alone, or in combination with local anaesthetic agents, into painful 'trigger' points or near painful facet joints or impinged nerve roots, may also relieve pain and spasm in some cases, although this is less effective than the epidural route (Stav *et al*. 1993).

Surgery

Neurosurgical intervention is only necessary in patients with progressive cervical myelopathy, or rarely with radiculopathy or intractable pain. Even with a prolapsed cervical disc, most patients will recover spontaneously or with conservative measures. Unlike prolapse of a lumbar disc, herniation of a cervical disc is associated with a slower rate of recovery that is more difficult to document with objective measurements. With cervical myelopathy due to degenerative disease, the result of decompressive surgery is often disappointing (Yonenobu *et al*. 1986; Rowland 1992). The rate of progression of the neurological deficit may be slowed by the surgery, but the lost function may not recover. This poor outcome reflects the irreversible nature of the damage to the spinal cord (see Fig. 4 and Fig. 5) and also the compromised vascular supply to the cord in some cases. Some patients show a good initial recovery following surgery, but with relapse some time later (Goto *et al*. 1993). This relapse may reflect recurrence of the original pathology, or scar tissue related to surgery at the site of the initial operation. The likely benefits of further surgery are based on detailed reinvestigation by radiological and other means.

In principle, surgical intervention can be achieved via the anterior and/or posterior routes, with the operative procedure being individually selected on the basis of detailed clinical and radiological assessment. In cases of myelopathy, a multisegmental operation is often necessary to obtain satisfactory results. This multisegmental approach is particularly important where neurological abnormality is associated with DISH or ossification of the posterior longitudinal ligament (Epstein 1993) that have been unresponsive to conservative therapy.

Although many operations using the anterior or posterior approach have been advocated for cervical myelopathy, they are all based on the principles of fusion to prevent excessive motion (especially if only one or two levels are involved), and decompression by laminectomy, laminoplasty, or discectomy.

Surgery is less often necessary for cervical radiculopathy. The principles of treatment are similar to those for myelopathy (Chesnut *et al*. 1992), but decompression can be achieved by foraminotomy and/or partial discectomy.

Pain clinic

Where physical approaches to therapy have been exhausted, the multidisciplinary approach of the pain clinic may assist the patient in learning to cope with the chronic pain. This approach is more often required for low back-pain syndromes, where serious and prolonged disability is more likely.

Neck pain, shoulder pain, and soft tissue lesions in the upper limb

Patients with cervical spondylosis often complain of shoulder pain, and it is important to differentiate neck pain referred to the shoulder (usually the superior aspect) from primary conditions of the shoulder. Many patients have features that suggest both cervical and shoulder lesions, and these can be difficult to separate (Hawkins *et al*. 1990). Treatment may need to be directed at both sites.

The 'double crush' syndrome refers to a combination of both cervical radiculopathy and a peripheral nerve entrapment lesion. The most common peripheral lesion is carpal tunnel syndrome, although the ulnar and radial nerves can also be affected. Elucidation of the peripheral lesion is usually possible using electromyography and nerve conduction studies. Somatosensory evoked potentials and electromyographic sampling of the more proximal muscles may provide evidence of cervical radiculopathy or plexus lesions in addition to the peripheral lesion. Surgical intervention should be directed at the peripheral entrapment lesion before the cervical lesion.

Cervical spondylosis can also mimic epicondylitis and other localized soft tissue lesions in the upper limbs (Gunn and Milbrandt 1976), and there is a well-documented association between cervical spondylosis and these soft tissue lesions (Murray-Leslie and Wright 1976).

Referred pain syndromes

Where appropriate, examination for polymyalgia rheumatica, reflex sympathetic dystrophy, inflammatory arthropathy, infection, malignancy, and even crystal arthropathies may be necessary. Table 2 above shows the sites of origin, and the more common causes of pain referred to the cervical region.

Brachial plexus lesions

Pain of brachial plexus origin is felt in the neck, shoulder, forearm, or hand. Supraclavicular pain is common and may be accompanied by local tenderness, and bony (cervical rib), or pulsatile (aneurysm) swelling at this site. The pain can often be induced by certain manoeuvres or changes in the position of the arm (see below). Table 6 shows the causes of brachial plexus injury.

Trauma — birth and motorcycle injury, surgery, cannulas, or needles
Thoracic-outlet syndrome — e.g. cervical rib
Traction — surgery, e.g. medial sternotomy, rucksack palsy, postanaesthesia
Brachial plexus neuropathy (brachial neuritis):
Infection — viral, bacterial
Toxins — heroin
Injection of foreign serum or vaccine
Systemic illness — lupus, vasculitis, Hodgkin's disease
Familial brachial plexus neuropathy
Idiopathic
Radiotherapy
Tumour — especially Pancoast tumour of lung, breast
Lightning, electric shock

Table 6 Causes of brachial plexus lesions

Thoracic-outlet obstruction

Compression of the distal nerve roots, brachial plexus, subclavian vessels, or the combined neurovascular bundle can occur at various sites between the neck and the axilla (the thoracic outlet). Possible contributing factors are shown in Table 7. 'Cervical ribs', which vary from simple exostoses of the transverse processes of C7 to fully formed extra ribs, are the most common cause of the thoracic-outlet syndrome. As 0.5 per cent of the population have bilateral cervical ribs yet less than 10 per cent ever get symptoms, other factors (Liu *et al*. 1995) must also be important in symptomatic cases. Poor posture and sagging or droopy shoulders (Swift and

Nichols 1984) have been identified as important factors in the development of symptoms, especially in females of early or middle age, in whom the syndrome is most often diagnosed.

Cervical rib
Congenital fibromuscular bands
Congenital abnormality of clavicle or first rib
Interscalenus muscle hypertrophy
Clavipectoral lesions — the 'hyperabduction syndrome'
Old fracture of the first rib or clavicle
Vascular disorders and hyperviscosity syndromes
External compression, e.g. heavy weights, rucksack lesions
Poor posture, sagging or droopy shoulders
Excessive muscle development around shoulders

Table 7 Contributing factors to thoracic-outlet syndrome

The typical symptoms of thoracic-outlet obstruction vary according to the site of compression of the neurovascular bundle (Novak *et al*. 1993). Pain is usually present in the neck, shoulder, or upper arm. Vasomotor abnormalities such as numbness, paraesthesia, coldness, colour change, or Raynaud's phenomenon are characteristic of the lesion and may dominate the picture. Compression of the lower levels of the brachial plexus result in neurological features such as sensory loss in the C8 and T1 dermatomes, with paraesthesia, or weakness in the hand. Some patients show well-defined syndromes due to selective compression of the subclavian vein, artery, or brachial plexus (Wilbourn 1993).

Obliteration of the radial or brachial pulse may be noted, either when the patient takes and holds a full breath with the head tilted back or rotated laterally (Adson Test), or when the arm is abducted and externally rotated while the shoulders are braced (Wright's manoeuvre). Both tests can be noted with thoracic-outlet syndrome, but are not reliable indicators of this condition.

Ten per cent of patients develop serious vascular complications, which can be venous or arterial (Hawkes 1986). Thrombosis of the axillosubclavian vein can present acutely or chronically, sometimes following prolonged exercise. Typical features are pain and swelling of the arm aggravated by exercise, with the diagnosis being confirmed by venography. Compression of the subclavian artery can be followed by post-stenotic dilatation, aneurysm formation, and retrograde thromboembolic phenomena. Claudication, vasomotor phenomena, digital gangrene, and acute limb-threatening ischaemia can occur. Arteriography, and in some cases Doppler ultrasonography are necessary to confirm the arterial abnormality.

Thoracic-outlet syndrome may be difficult to show objectively (Novak *et al*. 1993), and few patients have the diagnosis confirmed (Wilbourn 1993). Radiographs of the cervical spine may demonstrate cervical ribs, congenital bony abnormalities, or old fractures. MRI can also define soft tissue bands and other lesions causing distortion or deviation of nerves and blood vessels in the thoracic-outlet (Panegyres *et al*. 1993). Electromyography and somatosensory evoked potentials may offer some supportive evidence to confirm the diagnosis, or exclude lesions such as cervical radiculopathy, carpal tunnel syndrome, or ulnar neuropathy, which can closely mimic this syndrome. Arteriography and venography are necessary in difficult cases, especially with vascular complications.

Conservative treatment (Walsh 1994) with shoulder-girdle exercises and improved posture often fails to remove all the symptoms, but is all that should be advocated for mild disease, especially where the diagnosis remains in doubt.

Surgical intervention, which is often unsuccessful (Urschel and Razzuk 1986) and carries considerable risks (Marinoni *et al*. 1987), should be avoided (Wilbourn 1991), except in patients with serious vascular or other complications, or in severe cases of proven thoracic-outlet syndrome.

Surgery can be carried out via the supraclavicular or transaxillary routes, or via a combination of both. Operations advocated include resection of the first rib (Martin 1993), resection of the cervical rib, claviculectomy, scalenotomy with soft tissue release, and partial scalenectomy.

Resection of the cervical rib is adequate treatment in patients with predominantly neurological symptoms, but may not be sufficient with serious vascular disease. Therapeutic options for acute venous thrombosis are anticoagulation, fibrinolysis, and thrombectomy, followed by treatment of the thoracic-outlet syndrome, usually by resection of the first rib. For chronic venous obstruction, primary treatment of the actual syndrome may be enough, although some cases also require endovenectomy or venous bypass surgery (Sanders and Haug 1990). Compressive arterial lesions can be treated by rib resection or primary treatment of the syndrome, but once intrinsic arterial damage has developed, aneurysmectomy and arterial reconstruction are necessary. With distal thromboembolic disease, thrombectomy or embolectomy need to be combined with the other measures (Scher *et al*. 1984). Where vasomotor symptoms persist after the initial operation, chemical sympathectomy of the upper thoracic sympathetic chain may be helpful.

Cryptogenic brachial plexus neuropathy (brachial neuritis, neuralgic amyotrophy)

This condition usually develops acutely in healthy adults aged between 25 and 65 years. Some cases follow viral or other infection, injection of serum or vaccine, heroin use, or strenuous exercise. Rare cases are familial. An ache first develops around the shoulder or neck with increasing severity over 1 to 2 weeks. As the acute pain starts to settle, rapid onset of weakness and wasting of the shoulder and upper limb muscles is noted. The pattern of wasting depends on the pattern of injury of the brachial plexus nerves. Sensory loss, paraesthesia, hyperaesthesia, and hyporeflexia may also occur. The pain is exacerbated by use of the affected muscles, which may become totally paralysed. The lesion can be bilateral, and in some cases, is associated with involvement of the phrenic nerve (Walsh *et al*. 1987) or other peripheral and even cranial nerves (Mohanaruban and Fisher 1986). Electromyography of the affected muscles reveals fibrillation potentials and positive waves, in a pattern characteristic of combined damage to nerve roots and peripheral nerves. The electromyographic abnormality is often bilateral, even where this was not clinically apparent.

Although the course is variable, the prognosis for recovery is good. Most recover within 6 months, with 90 per cent showing a complete recovery within 3 years. Recurrences do occur, but are rare.

Brachial plexus neuropathy needs to be differentiated from acute poliomyelitis, rotator cuff tears, extradural malignancy, Pancoast tumour of the lung, thoracic-outlet syndrome, and cervical radiculopathy.

Conclusion

Degenerative lesions in the cervical region cause many varied symptoms, which are not always recognized to be from this source. 'Cervical spondylosis', which is a vague term used to describe a whole range of degenerative and mechanical lesions, causes less disability than similar lesions affecting the lumbar spine, and has therefore been the subject of less study. The more common mechanical syndromes and associated complications and their treatment have been discussed, with consideration being given to other causes of neck pain and neurological deficit in the upper limbs.

Chapter References

Antinnes, J.A., Dvorak, J., Hayek, J., Panjabi, M.M., and Grob, D. (1994). The value of functional computed tomography in the evaluation of soft-tissue injury in the upper cervical spine. *European Spine Journal*, **3**, 98–101.

Baba, H., Kawahara, N., Tomita, K., and Imura, S. (1993 *a*). Spinal cord evoked potentials in cervical and thoracic myelopathy. *International Orthopaedics*, **17**, 82–6.

Baba, H., Maezawa, Y., Kawahara, N., Tomita, K., Furusawa, N., and Imura, S. (1993 *b*). Calcium crystal deposition in the ligamentum flavum of the cervical spine. *Spine*, **18**, 2174–81.

Baba, H., Furusawa, N., Chen, Q., Imura, S., and Tomita, K. (1995). Anterior decompressive surgery for cervical ossified posterior longitudinal ligament causing myeloradiculopathy. *Paraplegia*, **33**, 18–24.

Barnsley, L., Lord, S.M., Wallis, B.J., and Bogduk, N. (1995). The prevalence of chronic cervical zygapophysial joint pain after whiplash. *Spine*, **20**, 20–6.

Barton, D., Allen, M., Finlay, D., and Belton, I. (1993). Evaluation of whiplash injuries by technetium 99m isotope scanning. *Archives of Emergency Medicine*, **10**, 197–202.

Batzdorf, U. and Flannigan, B.D. (1991). Surgical decompressive procedures for cervical spondylotic myelopathy. A study using magnetic resonance imaging. *Spine*, **16**, 123–7.

Bland, J.H. and Boushey, D.R. (1990). Anatomy and physiology of the cervical spine. *Seminars in Arthritis and Rheumatism*, **20**, 1–20.

Boachie-Adjei, O. and Bullough, P.G. (1987). Incidence of ankylosing hyperostosis of the spine (Forestier's disease) at autopsy. *Spine*, **12**, 739–43.

Boden, S.D., McCowin, P.R., Davis, D.O., Dina, T.S., Mark, A.S., and Wiesel, S. (1990). Abnormal magnetic-resonance scans of the cervical spine in asymptomatic subject. A prospective investigation. *Journal of Bone and Joint Surgery*, **72A**, 1178–84.

Bourbeau, R., Desjardins, D., Maag, U., and Laberge-Nadeau, C. (1993). Neck injuries among belted and unbelted occupants of the front seat of cars. *Journal of Trauma*, **35**, 794–9.

Brodsky, A.E. (1985). Cervical angina. A correlative study with emphasis on the use of coronary arteriography. *Spine*, **10**, 699–709.

Brown, T.R., Quinn, S.F., and D'Agostino, A.N. (1991). Deposition of calcium pyrophosphate dihydrate crystals in the ligamentum flavum: evaluation with MR imaging and CT. *Radiology*, **178**, 871–3.

Cassidy, J.D., Lopes, A.A., and Yong-Hing, K. (1992). The immediate effect of manipulation versus mobilization on pain and range of motion in the cervical spine: a randomized controlled trial. *Journal of Manipulative and Physiological Therapeutics*, **15**, 570–5.

Castagnera, L., Maurette, P., Pointillart, V., Vital, J.M., Erny, P., and Senegas, J. (1994). Long-term results of cervical epidural steroid injection with and without morphine in chronic cervical radicular pain. *Pain*, **58**, 239–43.

Chesnut, R.M., Abitbol, J-J., and Garfin, S.R. (1992). Surgical management of cervical radiculopathy. Indication, techniques, and results. *Orthopedic Clinics of North America*, **23**, 461–74.

Cicala, R.S., Thoni, K., and Angel, J.J. (1989). Long-term results of cervical epidural steroid injections. *Clinical Journal of Pain*, **5**, 143–5.

Croft, P., Schollum, J., and Silman, A. (1994). Population study of tender point counts and pain as evidence of fibromyalgia. *British Medical Journal*, **309**, 696–9.

Di Lazzaro, V., Restuccia, D., Colosimo, C., and Tonali, P. (1992). The contribution of magnetic stimulation of the motor cortex to the diagnosis of cervical spondylotic myelopathy. Correlation of central motor conduction to distal and proximal upper limb muscles with clinical and MRI findings. *Electroencephalographic and Clinical Neurophysiology*, **85**, 311–20.

Dvorak, J., Janssen, B., and Grob, D. (1990). The neurologic workup in patients with cervical spine disorders. *Spine*, **15**, 1017–22.

Edmeads, J. (1988). The cervical spine and headache. *Neurology*, **38**, 1874–8.

Epstein, N. (1993). The surgical management of ossification of the posterior longitudinal ligament in 51 patients. *Journal of Spinal Disorders*, **6**, 432–55.

Evans, R.W. (1992). Some observations on whiplash injuries. *Neurological Clinics*, **10**, 975–7.

Farmer, J.C. and Wisneski, R.J. (1994). Cervical spine nerve root compression. An analysis of neuroforaminal pressures with varying head and arm positions. *Spine*, **19**, 1850–5.

Fast, A., Parikh, S., and Marin, E.L.(1989). The shoulder abduction relief sign in cervical radiculopathy. *Archives of Physical Medicine and Rehabilitation*, **70**, 402–3.

Ferrante, F.M., Wilson, S.P., Iacobo, C., Orav, E.J., Rocco, A.G., and Lipson, S. (1993). Clinical classification as a predictor of therapeutic outcome after cervical epidural steroid injection. *Spine*, **18**, 730–6.

Frisoni, G.B. and Anzola, G.P. (1991). Vertebrobasilar ischaemia after neck motion. *Stroke*, **22**, 1452–60.

Fukui, K., Kataoka, O., Sho, T., and Sumi, M. (1990). Pathomechanism, pathogenesis, and results of treatment in cervical spondylotic myelopathy caused by dynamic canal stenosis. *Spine*, **15**, 1148–52.

Galasko, C.S. *et al*. (1993). Neck sprains after road traffic accidents: a modern epidemic. *Injury*, **24**, 155–7.

Gore, D.R., Sepic, S.B., and Gardner, G.M. (1986). Roentgenographic findings of the cervical spine in asymptomatic people. *Spine*, **11**, 521–4.

Goto, S. *et al*. (1993). Anterior surgery in four consecutive technical phases for cervical spondylotic myelopathy. *Spine*, **18**, 1968–73.

Griegel-Morris, P., Larson, K., Mueller-Klaus, K., and Oatis, C.A. (1992). Incidence of common postural abnormalities in the cervical, shoulder, and thoracic regions and their association with pain in two age groups of healthy subjects. *Physical Therapy*, **72**, 425–31.

Gunn, C.C. and Milbrandt, W.E. (1976). Tennis elbow and the cervical spine. *Canadian Medical Association Journal*, **114**, 803–9.

Hadler, N.M. (1985). Illness in the workplace: the challenge of musculoskeletal symptoms. *Journal of Hand Surgery*, **10A**, 451–6.

Hamer, A.J., Gargan, M.F., Bannister, G.C., and Nelson, R.J. (1993). Whiplash injury and surgically treated cervical disc disease. *Injury*, **24**, 549–50.

Hawkes, C.D. (1986). Neurosurgical considerations in thoracic outlet syndrome. *Clinical Orthopaedics*, **207**, 24–8.

Hawkins, R.J., Bilco, T., and Bonutti, P. (1990). Cervical spine and shoulder pain. *Clinical Orthopaedics*, **258**, 142–6.

Hayashi, H., Okada, K., Hashimoto, J., Tada, K., and Ueno, R. (1988). Cervical spondylotic myelopathy in the aged patient. A radiographic evaluation of the aging changes in the cervical spine and etiologic factors of myelopathy. *Spine*, **13**, 618–25.

Helliwell, P.S., Evans, P.F., and Wright, V. (1994). The straight cervical spine: does it indicate muscle spasm? *Journal of Bone and Joint Surgery*, **76B**, 103–6.

Hensinger, R.N. (1991). Congenital anomalies of the cervical spine. *Clinical Orthopaedics*, **264**, 16–38.

Hirsch, S.A., Hirsch, P.J., Hiramoto, H., and Weiss, A. (1988). Whiplash syndrome. Fact or fiction? *Orthopedic Clinics of North America*, **19**, 791–5.

Holt, L. (1972). Frequency of symptoms for different age groups and professions. In *Cervical pain*(ed. C. Hirsch and Y. Zotterman), pp. 17–20. Pergamon, New York.

Hukuda, S., Mochizuki, T., Ogata, M., and Shichikawa, K. (1983). The pattern of spinal and extraspinal hyperostosis in patients with ossification of the posterior longitudinal ligament and the ligamentum flavum causing myelopathy. *Skeletal Radiology*, **10**, 79–85.

Jinkins, J.R., Bashir, R., Al-Mefty, O., Al-Kawi, M.Z., and Fox, J.L. (1986). Cystic necrosis of the spinal cord in compressive cervical myelopathy: demonstration by iopamidol CT-myelography. *American Journal of Roentgenology*, **147**, 767–75.

Jonsson, H., Jr., Cesarini, K., Sahlstedt, B., and Rauschning, W. (1994). Findings and outcome in whiplash-type neck distortions. *Spine*, **19**, 2733–43.

Kramer, J., Rivera, C.A., and Kleefield, J. (1991). Degenerative disorders of the cervical spine. *Rheumatic Disease Clinics of North America*, **17**, 741–55.

Kristensen, S., Sander, K.M., and Pedersen, P.R. (1988). Cervical involvement in diffuse idiopathic skeletal hyperostosis with dysphagia and rhinolalia. *Archives of Otorhinolaryngology*, **245**, 330–4.

Lawrence, J.S. (1969). Disc degeneration. Its frequency and relationship to symptoms. *Annals of the Rheumatic Diseases*, **28**, 121–38.

Lehto, I.J., Tertti, M.O., Komu, M.E., Paajanen, H.E., Tuominen, J., and Kormano, M.J. (1994). Age-related MRI changes at 0.1 T in cervical discs in asymptomatic subjects. *Neuroradiology*, **36**,

49–53.

Lestini, W.F. and Wiesel, S. W. (1989). The pathogenesis of cervical spondylosis. *Clinical Orthopaedics*, **239**, 69–93.

Liu, J.E., Tahmoush, A.J., Roos, D.B., and Schwartzman, R.J. (1995). Shoulder-arm pain from cervical bands and scalene muscle anomalies. *Journal of Neurological Science*, **128**, 175–80.

Maimaris, C., Barnes, M.R., and Allen, M.J. (1988). 'Whiplash injuries' of the neck: a retrospective study. *Injury*, **19**, 393–6.

Marinoni, E.C., Bonfiglio, G., Coletti, M., and Passarelli, O. (1987). Thoracic-outlet syndrome. Proposed protocol for diagnosis and treatment. *Italian Journal of Orthopaedic Traumatology*, **13**, 379–86.

Martin, G.T. (1993). First rib resection for the thoracic-outlet syndrome. *British Journal of Neurosurgery*, **7**, 35–8.

Martinez, J.E., Ferraz, M.B., Sato, E.I., and Atra, E. (1995). Fibromyalgia versus rheumatoid arthritis: a longitudinal comparison of the quality of life. *Journal of Rheumatology*, **22**, 270–4.

Matsuda, Y. *et al*. (1991). Increased MR signal intensity due to cervical myelopathy. Analysis of 29 surgical cases. *Journal of Neurosurgery*, **74**, 887–92.

McKinney, L.A., Dornan, J.O., and Ryan, M. (1989). The role of physiotherapy in the management of acute neck sprains following road traffic accidents. *Archives of Emergency Medicine*, **6**, 27–33.

Mealy, K., Brennan, H., and Fenelon, G.C. (1986). Early mobilization of acute whiplash injuries. *British Medical Journal (Clinical Research)*, **292**, 656–7.

Mohanaruban, K. and Fisher, D.J. (1986). A combination of cranial and peripheral nerve palsies in infectious mononucleosis. *Postgraduate Medical Journal*, **62**, 1129–30.

Murone, I. (1974). The importance of the sagittal diameters of the cervical spine canal in relation to spondylosis and myelopathy. *Journal of Bone and Joint Surgery*, **56B**, 30–6.

Murray-Leslie, C.F. and Wright, V. (1976). Carpal tunnel syndrome, humeral epicondylitis, and the cervical spine: a study of clinical and dimensional relations. *British Medical Journal*, **1**, 1439–42.

Nagasawa, A., Sakakibara, T., and Takahashi, A. (1993). Roentgenographic findings of the cervical spine in tension-type headache. *Headache*, **33**, 90–5.

Nagata, K., Kiyonaga, K., Ohashi, T., Sagara, M., Miyazaki, S., and Inoue, A. (1990). Clinical value of magnetic resonance imaging for cervical myelopathy. *Spine*, **15**, 1088–96.

Nakajima, M. and Hirayama, K. (1995). Midcervical central cord syndrome: numb and clumsy hands due to midline cervical disc protrusion at the C3–4 intervertebral level. *Journal of Neurology, Neurosurgery and Psychiatry*, **58**, 607–13.

Novak, C.B., Mackinnon, S.E., and Patterson, G.A. (1993). Evaluation of patients with thoracic-outlet syndrome. *Journal of Hand Surgery*, **18A**, 292–9.

Okada, Y., Ikata, T., Katoh, S., and Yamada, H. (1994). Morphologic analysis of the cervical spinal cord, dural tube, and spinal canal by magnetic resonance imaging in normal adults and patients with cervical spondylotic myelopathy. *Spine*, **19**, 2331–5.

Otake, S., Matsuo, M., Nishizawa, S., Sano, A., and Kuroda, Y. (1992). Ossification of the posterior longitudinal ligament: MR evaluation. *American Journal of Neuroradiology*, **13**, 1059–70.

Palmer, P.E., Stadalnick, R., and Arnon, S. (1984). The genetic factor in cervical spondylosis. *Skeletal Radiology*, **11**, 178–82.

Panegyres, P.K., Moore, N., Gibson, R., Rushmore, G., and Donaghy, M. (1993). Thoracic outlet syndrome and magnetic resonance imaging. *Brain*, **116**, 823–41.

Parmar, H.V. and Raymakers, R. (1993). Neck injuries from rear impact road traffic accidents: prognosis in persons seeking compensation. *Injury*, **24**, 75–8.

Pennie, B.H. and Agambar, L.J. (1990). Whiplash injuries. A trial of early management. *Journal of Bone and Joint Surgery*, **72B**, 277–9.

Pettersson, K., Hildingsson, C., Toolanen, G., Fagerlund, M., and Bjornebrink, J. (1994). MRI and neurology in acute whiplash trauma: no correlation in prospective examination of 39 cases. *Acta Orthopaedica Scandinavica*, **65**, 525–8.

Pyhtinen, J. and Laitinen, J. (1993). Cervical intervertebral foramen narrowing and myelographic nerve root sleeve deformities. *Neuroradiology*, **35**, 596–7.

Radanov, B.P., Dvorak, J., and Valach, L. (1992). Cognitive deficits in patients after soft tissue injury of the cervical spine. *Spine*, **17**, 127–31.

Radanov, B.P., Sturzenegger, M., De Stefano, G., and Schnidrig, A. (1994). Relationship between early somatic, radiological, cognitive and psychosocial findings and outcome during a one year follow-up of 117 patients suffering from common whiplash. *British Journal of Rheumatology*, **33**, 442–8.

Raskind, R. and North, C.M. (1990). Vertebral artery injuries following chiropractic cervical spine manipulations—case reports. *Angiology*, **41**, 445–52.

Resnick, D. (1985). Degenerative diseases of the vertebral column. *Radiology*, **156**, 3–14.

Resnick, D., Guerra, J., Jr., Robinson, C.A., and Vint, V.C. (1978). Association of diffuse idiopathic skeletal hyperostosis (DISH) and calcification and ossification of the posterior longitudinal ligament. *American Journal of Roentgenology*, **131**, 1049–53.

Rowland, L.P. (1992). Surgical treatment of cervical spondylotic myelopathy: time for a controlled trial. *Neurology*, **42**, 5–13.

Rowlingson, J.C. and Kirschenbaum, L.P. (1986). Epidural analgesic techniques in the management of cervical pain. *Anesthesia and Analgesia*, **65**, 938–42.

Sanders, R.J. and Haug, C. (1990). Subclavian vein obstruction and thoracic outlet syndrome: a review of etiology and management. *Annals of Vascular Surgery*, **4**, 397–410.

Scher, L.A. *et al*. (1984). Staging of arterial complications of cervical rib: guidelines for surgical management. *Surgery*, **95**, 644–9.

Simmons, Z., Biller, J., Beck, D.W., and Keyes, W.(1986). Painless compressive cervical myelopathy with false localizing sensory findings. *Spine*, **11**, 869–72.

Simms, R.W. (1994). Controlled trials of therapy in fibromyalgia syndrome. *Baillière's Clinical Rheumatology*, **8**, 917–34.

Smythe, H.A. (1986). Referred pain and tender points. *American Journal of Medicine*, **81**,(Suppl. 3A), 90–2.

Smythe, H.A. (1994). The C6–7 syndrome: clinical features and treatment response. *Journal of Rheumatology*, **21**, 1520–6.

Sobol, S.M. and Rigual, N.R.(1984). Anterolateral extrapharyngeal approach for cervical osteophyte-induced dysphagia. *Annals of Otology, Rhinology and Laryngology*, **93**, 498–504.

Stav, A., Ovadia, L., Sternberg, A., Kaadan, M., and Weksler, N. (1993). Cervical epidural steroid injection for cervicobrachialgia. *Acta Anaesthesiologica Scandinavica*, **37**, 562–6.

Sturzenegger, M., DiStefano, G., Radanov, B.P., and Schnidrig, A. (1994). Presenting symptoms and signs after whiplash injury: the influence of accident mechanisms. *Neurology*, **44**, 688–93.

Sudarsky, L. and Ronthal, M. (1983). Gait disorders among elderly patients. A survey study of 50 patients. *Archives of Neurology*, **40**, 740–3.

Suzuki, K., Ishida, Y., and Ohmori, K. (1991). Long term follow-up of diffuse idiopathic skeletal hyperostosis in the cervical spine. Analysis of progression of ossification. *Neuroradiology*, **33**, 427–31.

Swift, T.R. and Nichols, F.T. (1984). The droopy shoulder syndrome. *Neurology*, **34**, 212–5.

Urschel, H.C., Jr. and Razzuk, M.A. (1986). The failed operation for thoracic-outlet syndrome: the difficulty of diagnosis and management. *Annals of Thoracic Surgery*, **42**, 523–8.

Van der Donk, J., Schouten, J.S., Passchier, J., van Romunde, L.K., and Valkenburg, H.A. (1991). The associations of neck pain with radiological abnormalities of the cervical spine and personality traits in a general population. *Journal of Rheumatology*, **18**, 1884–9.

Vanderburgh, D.F. and Kelly, W.M. (1993). Radiographic assessment of discogenic disease of the spine. *Neurosurgical Clinics of North America*, **4**, 13–33.

Vargo, M.M. and Flood, K.M. (1990). Pancoast tumour presenting as cervical radiculopathy. *Archives of Physical Medicine and Rehabilitation*, **71**, 606–9.

Walsh, M.T. (1994). Therapist management of thoracic-outlet syndrome. *Journal of Hand Therapy*, **7**, 131–44.

Walsh, N.E., Dumitru, D., Kalantri, A., and Roman, A.M., Jr. (1987). Brachial neuritis involving the bilateral phrenic nerves. *Archives of Physical Medicine and Rehabilitation*, **68**, 46–8.

Watkinson, A., Gargan, M.F., and Bannister, G.C. (1991). Prognostic factors in soft tissue injuries of the cervical spine. *Injury*, **22**, 307–9.

Wiesel, S.W., Tsourmas, N., Feffer, H.L., Citrin, C.M., and Patronas, N. (1984). A study of computer-assisted tomography 1. The incidence of positive CAT scans in an asymptomatic group of patients. *Spine*, **9**, 549–51.

Wilbourn, A.J. (1991). Thoracic-outlet syndromes: a plea for conservatism. *Neurosurgical Clinics of North America*, **2**, 235–45.

Wilbourn, A.J. (1993). Brachial plexus disorders. In *Peripheral neuropathy*, Vol. 2, (3rd edn)(ed. P.J. Dyck and P.K. Thomas), pp. 911–51. W.B. Saunders, Philadelphia.

Wilke, W.S. (1995). Treatment of 'resistant' fibromyalgia. *Rheumatic Diseases Clinics of North America*, **21**, 247–60.

Wober-Bingol, C. *et al*. (1992). Tension headache and the cervical spine—plain X-ray findings. *Cephalalgia*, **12**, 152–4.

Wolfe, F. *et al*. (1990). The American College of Rheumatology 1990 criteria for the classification of fibromyalgia. Report of the Multicenter Criteria Committee. *Arthritis and Rheumatism*, **33**, 160–72.

Wolfe, F., Ross, K., Anderson, J., and Russell, I.J. (1995). Aspects of fibromyalgia in the general population: sex, pain, threshold, and fibromyalgia symptoms. *Journal of Rheumatology*, **22**, 151–6.

Yiannikas, C., Shahani, B.T., and Young, R.R. (1986). Short latency somatosensory-evoked potentials from radial, median, ulnar, and peroneal nerve stimulation in the assessment of cervical spondylosis. Comparison with conventional electromyography. *Archives of Neurology*, **43**, 1264–71.

Yonenobu, K., Okada, K., Fuji, T., Fujiwara, K., Yamashita, K., and Ono, K. (1986). Causes of neurologic deterioration following surgical treatment of cervical myelopathy. *Spine*, **11**, 818–23.

Yoss, R.E., Corbin, K.B., MacCarty, C.S., and Love, J.G. (1957). Significance of symptoms and signs in the localization of the involved root in cervical disc protrusion. *Neurology*, **7**, 673–83.

5.19.1 Rheumatic diseases and neoplasia

Thomas G. Benedek

Some systemic rheumatic diseases may predispose to the development of malignant neoplasms. Conversely, some malignant neoplasms may present with rheumatic symptoms. Although the pathogenesis of these complex relationships is incompletely understood, recognition of the manifestations is necessary for the efficient diagnostic evaluation of many patients. The relations and conditions considered here may be classified thus:

1. Relation of rheumatic diseases to malignant neoplasms:
   a. rheumatoid arthritis;
   b. spondylarthropathies;
   c. Sjögren's syndrome;
   d. systemic lupus erythematosus;
   e. dermatomyositis/polymyositis;
   f. systemic sclerosis (scleroderma);
   g. polymyalgia rheumatica.
2. Rheumatic manifestations of malignant neoplasms:
   a. skeletal neoplasms;
      i. osseous metastases
      ii. synovial sarcoma
      iii. leukaemia
      iv. paraproteineamic arthropathies
   b. paraneoplastic syndromes;
      i. hypertrophic osteoarthropathy
      ii. oncogenic osteomalacia
      iii. carcinomatous polyarthritis
      iv. carcinomatous neuromyopathy
      v. Eaton–Lambert myasthenic syndrome
      vi. fasciitis syndromes
      vii. paraneoplastic Raynaud's phenomenon.

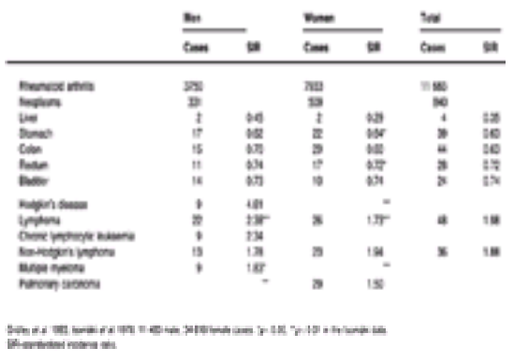
Relation of rheumatic diseases and neoplasia

Rheumatoid arthritis

The pathogenetic significance of even a high degree of association is difficult to determine and does not prove a causal relationship. This problem is exemplified by questions about the relationship between rheumatoid arthritis and cancer. The mean life expectancy is shortened by rheumatoid arthritis, according to a Finnish study (Myllykangas-Luosujärvi *et al*. 1995), by at least 3 years. The greatest proportional excess of deaths is due to infections but since infections have diminished as a primary cause of death, gastrointestinal diseases have become predominant numerically (Mitchell *et al*. 1986; Wolfe *et al*. 1994; Myllykangas-Luosujärvi *et al*. 1995). Nevertheless, most patients live into the seventh decade, when the incidence of malignancies is highest but the onset of rheumatoid arthritis is relatively infrequent. The length of the latent period before a neoplasm becomes diagnosable is unknown. Thus, a patient who has had rheumatoid arthritis for a few years may have been incubating the neoplasm before rheumatoid arthritis supervened. The best, albeit weak, evidence that neoplasms do not predispose to the occurrence of rheumatoid arthritis, rather than the reverse, is that they do not precede rheumatoid arthritis with unexpected frequency.

Epidemiological problems may include inconsistent diagnostic criteria for rheumatoid arthritis and for specific neoplasms, regional variations in the occurrence both of rheumatoid arthritis and of various neoplasms, and whether the cases of rheumatoid arthritis are entered from hospital rolls, outpatient services, or community surveys. A high degree of association with one subset of neoplasms may be hidden in a broader category. However, the narrower the subset of cancers that is being correlated with the incidence of rheumatoid arthritis, the larger the cohort of rheumatoid cases must be for statistically significant correlations to be obtained. A further confounding consideration is whether a correlation is attributable to a characteristic of the disease or of its treatment.

The best available data regarding the relationships between rheumatoid arthritis and cancer emanate from complementary epidemiological studies in Finland from 1967 to 1973 (Isomäki *et al*. 1978) and in Sweden from 1965 to 1984 (Gridley *et al*. 1993). Comparisons with national incidence data were made in both investigations. The smaller Swedish cohorts included 32.6 per cent as many men and 22.9 per cent as many women with rheumatoid arthritis, but cancer among these was diagnosed in 81.3 per cent as many men and 64.0 per cent as many women. The proportionately larger number of neoplasms in the Swedish cohorts can be attributed to the much longer period of observation (20 years rather than 7). While the total incidence of neoplasms in both investigations approximated the expected, the frequency of some neoplasms exceeded or was less than expected. Unfortunately, the Finnish data do not differentiate among lymphomas or leukaemias. Table 1 shows the neoplasms that deviated from the expected incidence in the Swedish data and asterisks indicate which deviated in the Finnish data. For Hodgkin's disease there was an excess only in men in Sweden and only in women in Finland. This demonstrates uncertainty due to small numbers. Other lymphomas occurred in excess in both investigations. Non-Hodgkin's lymphoma appears to be the most highly associated with rheumatoid arthritis among the lymphoreticular neoplasms (Symmons *et al*. 1984; Porter *et al*. 1991). Thirty-seven of 39 such cases were over the age of 50 and the interval between the onset of rheumatoid arthritis and the diagnosis of the lymphoreticular neoplasm was at least 10 years in two-thirds (Banks *et al*. 1979; Symmons *et al*. 1984). A prospective 10-year Finnish study of 500 men and 500 women with rheumatoid arthritis and age-matched, random control cases substantiated the excess of haemopoietic neoplasia (Laakso *et al*. 1986). Both apparently benign paraproteinaemia and multiple myeloma occur with increased frequency in rheumatoid arthritis (Zawadzki and Benedek 1969; Eriksson 1993). The observation that rheumatoid arthritis patients do not have an increased likelihood of a second primary neoplasm supports the data that rheumatoid arthritis, with the exception of lymphoreticular processes, does not confer an increased susceptibility to neoplasia (Hakulinen 1986).



	Men		Women		Total	
	Cases	SIR	Cases	SIR	Cases	SIR
Rheumatoid arthritis	3751		7553		11 665	
Neoplasms	331		559		840	
Liver	2	0·45	2	0·29	4	0·35
Stomach	17	0·82	22	0·54*	39	0·63
Colon	15	0·70	29	0·63	44	0·63
Rectum	11	0·74	17	0·72*	28	0·72
Bladder	14	0·73	10	0·74	24	0·74
Hodgkin's disease	9	4·01		**		
Lymphoma	20	2·39**	26	1·73**	48	1·98
Chronic lymphocytic leukaemia	9	2·34				
Non-Hodgkin's lymphoma	13	1·78	23	1·94	36	1·88
Multiple myeloma	9	1·83*		**		
Pulmonary carcinoma		**	29	1·52		

Smiley et al. 1983, Isomäki et al 1978. 11 465 males, 24 818 female cases. *p< 0.05, **p< 0.01 in the Isomäki data.
SIR=standardized incidence ratio.

Table 1 Neoplasms showing an unusual association with rheumatoid arthritis

The data of Isomake *et al*. showed a deficit of gastric and rectal carcinomas in women (Isomäki *et al*. 1978). The Swedish data demonstrated this in both sexes and for colonic carcinoma as well. Laakso *et al*. found more gastrointestinal carcinomas in their rheumatoid arthritis than in control cohorts, but better survival in the former (Laakso *et al*. 1986). Recent studies provide persuasive, albeit inferential, evidence that the deficit of at least colorectal carcinoma in rheumatoid arthritis patients is due to an effect of the chronic consumption of aspirin or other non-steroidal anti-inflammatory prostaglandin-inhibiting drugs (Thun *et al*. 1991). These drugs inhibit the chemical induction of colonic tumours in rodents. Several clinical studies have now shown the same effect, although with a wide difference in the duration of use before the effect becomes definite—possibly a decade or longer (Rosenberg *et al*. 1991; Giovannucci *et al*. 1995). In view of the substantial chronic use of these drugs in rheumatoid arthritis, the deficit of colorectal cancer may here find its explanation. Whether gastric carcinogenesis is similarly influenced and whether the larger rheumatological doses have a more rapidly demonstrable effect than the equivalent of 0.3 g of aspirin every other day have not yet been established.

The possibly carcinogenic effect of drugs used to treat rheumatoid arthritis has been examined more directly. Gold compounds and corticosteroids have not been implicated. However, the greatly increased incidence of lymphoreticular neoplasms in immunosuppressed organ transplant recipients led to concern about a similar risk in rheumatoid arthritis patients since these drugs have come into use in their treatment (Kinlen 1985). At present the implication of methotrexate and azathioprine is doubtful (Silman *et al*. 1988; Stern *et al*. 1982; Kingsmore *et al*. 1992). However, cyclophosphamide has been clearly shown to be carcinogenic in cases of rheumatoid arthritis. In a case–control study of 119 rheumatoid arthritis patients followed for 20 years, a difference in the incidence of malignancies appeared 6 years after cyclophosphamide was begun and persisted irrespective of when it was discontinued. Fifty cancers were found in 37 cyclophosphamide cases versus 26 in 25 matched, non- cyclophosphamide treated cases ($p<0.05$). All nine bladder carcinomas occurred following cyclophosphamide treatment and this was associated with the largest cumulative doses. All lymphoreticular neoplasms occurred during the first decade of observation, while bladder carcinomas were still being found two decades after cyclophosphamide treatment (Baker *et al*. 1987; Radis *et al*. 1995).

Spondylarthropathies

The spondylarthropathies as such do not predispose to the occurrence of leukaemia or solid neoplasms (Smith *et al*. 1977). However, patients whose vertebral column has received therapeutic X-irradiation, as was popular in the 1940s, are at increased risk, particularly for leukaemia. The studies begun by Court-Brown and Doll in Great Britain led to the recognition of this hazard and the eventual discontinuance of this therapy (Court-Brown and Doll 1965). Haemopoiesis may remain depressed in heavily irradiated areas more than a decade after treatment. Leukaemias, particularly the acute myeloid type, have a peak incidence 2.5 years after a course of treatment. Except for chronic lymphocytic leukaemia, these dyscrasias occur more than three times as often as expected following X-ray therapy. The excess risk then diminishes but does not disappear (Darby *et al*. 1987). Apical pulmonary fibrosis, an uncommon complication of ankylosing spondylitis (Rosenow *et al*. 1977), has been reported as the site of primary adenocarcinoma, but only following X-ray therapy (Ahern *et al*. 1982). The risk of solid neoplasms over-all following X-ray therapy is increased slightly and this has been associated with older age at the time of treatment. The occurrence of fibrosarcomas has been associated with X-ray therapy (Edgar and Robinson 1973) but not lymphoreticular neoplasms (Boice 1992).

Sjögren's syndrome

In considering the relationship of Sjögren's syndrome to neoplasia only lymphoreticular proliferation warrants consideration. Three circumstances must be defined:

1. Myoepithelial sialadenitis is the pathological term for the diagnostic alteration in salivary and lacrimal glands that is essential, but not sufficient, to substantiate the clinical diagnosis (Schmid *et al*. 1982).
2. This, in addition, requires findings of xerostomia and/or keratoconjunctivitis sicca (subnormal salivation and/or lacrimation).
3. When these findings occur in the absence of an autoimmune rheumatic disease the designation is Sjögren's disease (or primary Sjögren's syndrome). 'Sjögren's syndrome' (or secondary Sjögren's syndrome) indicates the concurrent presence of a connective tissue disease, most often rheumatoid arthritis or systemic lupus erythematosus.

Sjögren's syndrome is predominantly an ailment of middle-aged women. The mean age at the time of diagnosis of the primary and secondary form is similar (Moutsopoulos *et al*. 1980). Depending on the sites and magnitude of the characteristic proliferation of lymphocytes and epithelial cells, the disease may have various manifestations, but diminished salivary and lacrimal function predominate. The proliferation may remain within lymph nodes or extend extranodally. Because of the range of alterations within lymph nodes and extranodally, the lesions vary from clearly benign, through the potentially premalignant 'pseudolymphoma' phase, to one of the malignant lymphomas (McCurley *et al*. 1990), which usually is of a non-Hodgkin's B-cell type. Hodgkin's disease (Nagai et al., 1993), reticulum cell sarcoma (Hornbaker *et al*. 1966), T-cell lymphoma (Chevalier *et al*. 1991), monoclonal gammopathies (Sugai *et al*. 1985) including cryoglobulinaemia (Tzioufas *et al*. 1986), and angioimmunoblastic lymphadenopathy (Bignon *et al*. 1986) are rare occurrences. Latency between the onset of clinically evident salivary gland enlargement and the diagnosis of malignant lymphoma range from 1.5 to 12 years in the experience of Schmid *et al*. (Schmid *et al*. 1982), and intervals of 20 years are known.

While about 20 per cent of patients with rheumatoid arthritis also have Sjögren's syndrome, the reported prevalence of rheumatoid arthritis among cases of Sjögren's syndrome has varied, depending on the diagnostic criteria and techniques, from 13.4 per cent of a Swedish cohort (Holm 1949) and 14.3 per cent in a Scottish study (Thompson and Eadie 1956) to 58.4 per cent in a Czech investigation (Lenoch *et al*. 1967). Sjögren's syndrome has been described as a complication of all major systemic autoimmune rheumatic diseases but, aside from rheumatoid arthritis, it has been observed frequently enough only in association with systemic lupus erythematosus to conclude whether the association influences neoplasia—it does not.

Lymphocytic lesions are more likely to extend beyond the salivary and lacrimal structures in primary than in secondary Sjögren's syndrome. However, this does not alter the risk of neoplastic change (Moutsopoulos *et al*. 1980). The early diagnosis of the lymphoma is difficult since it results from an evolutionary process in the course of Sjögren's syndrome. Therefore, prevalence figures depend in large measure on the mean duration of disease in the cohort of patients studied. In the three largest published studies lymphomas were detected in a mean of 4.3 per cent of 445 cases of Sjögren's syndrome (Whaley *et al*. 1973; Kassan *et al*. 1978; McCurley *et al*. 1990). This exceeds the frequency in any other population of patients, but whether by as much as 44-fold (Kassan *et al*. 1978) has not been confirmed.

Of 20 patients in whom the routine histopathological diagnosis was myoepithelial sialadenitis, immunohistological techniques showed various evolutionary stages of neoplasia in 12, and three were frankly lymphomatous (Hyjek *et al*. 1988). Refined procedures have shown that many of these neoplasms belong to the monocytoid B-cell type and may be detectable in cases that are considered pseudolymphoma by the usual techniques (Sung *et al*. 1991). Neoplastic transformation may, in a minority of patients, result in either of two contradictory immunochemical processes. In some cases the development of a lymphoma is indicated by a decrease in serum IgM and, if rheumatoid factor has been present, by its decreasing titre. In other cases a monoclonal gammopathy develops which may be indistinguishable from Waldenström's macroglobulinaemia. In most of such cases the monoclonal protein has been IgM k, except in Japan, where an IgA paraproteinaemia appears to be as likely as the IgM type (Sugai *et al*. 1985).

Of 28 patients with Sjögren's syndrome and lymphoma recorded in three American studies (Kassan *et al*. 1978; McCurley *et al*. 1990; Sung *et al*. 1991) seven had rheumatoid arthritis and one lupus erythematosus. Contrary to the female predominance of Sjögren's syndrome, neoplastic change in Sjögren's syndrome does not appear to be sex influenced.

Lymphomas originate in salivary glands infrequently and most of these are non-Hodgkin's B-cell neoplasms—the same type that is most often associated with Sjögren's syndrome. In reviews of 33 salivary gland lymphomas in New York and 40 cases in London, each included four instances of Sjögren's syndrome, an incidence of 11 per cent (Hyman and Wolff 1978; Gleeson *et al*. 1986). A French investigation of 113 cases of non-Hodgkin's lymphoma (86 per cent B-cell, 14 per cent T-cell) found 10 (8.8 per cent) with clinical Sjögren's syndrome. It had preceded the diagnosis of lymphoma in two. Remission of the lymphoma was associated with loss of the salivary gland infiltrate in cases that met only the histopathological criteria of Sjögren's syndrome, but not in those who had clinical Sjögren's syndrome (Janin *et al*. 1992).

In conclusion, the magnitude of the association between either primary or secondary Sjögren's syndrome and the development of any of several malignant lymphomas is uncertain. However, rapid enlargement of salivary glands unaccompanied by signs of inflammation, and/or a changing immunoglobulin pattern, warrants a biopsy which should be examined immunohistochemically.

Systemic lupus erythematosus

Systemic lupus erythematosus (**SLE**) most frequently occurs between the ages of 15 and 35 and in this age group women predominate at least 8:1. Life expectancy for the patient with SLE has increased markedly since the 1950s, in part due to enhanced diagnostic awareness and improved techniques whereby milder cases, which may be expected to have a more chronic course, are being identified, and in part due to improved therapy. While the 4-year survival was about 50 per cent in 1955, the 10-year survival in the 1990s may exceed 80 per cent (Abu-Shakra *et al*. 1995). Since cancer is relatively uncommon in young women, its occurrence in a typical case of SLE may raise the suspicion of a causal relationship. It follows that if there is an association between SLE and neoplasia, more such cases would occur among the recent greater number of long-term survivors, especially since some of the immunosuppressive agents used in treating SLE are potentially carcinogenic.

The possibility of such a relationship is supported in the animal model of SLE. The offspring of matings of NZB and NZW mice manifest characteristics of an autoimmune disease resembling SLE and infrequently develop a lymphoma. Treatment of such mice with the equivalent of therapeutic doses of cyclophosphamide (Walker and Boles 1973) resulted in a high incidence of various neoplasms, and treatment with azathioprine resulted in lymphomas (Casey 1968). Results were strongly dose related.

In a review of 15 cases of solid tumours in cases of SLE, seven had been treated with one of these drugs; the neoplasm was not detected until 12 to 19 years after the diagnosis of SLE in eight, six of whom had received azathioprine or cyclophosphamide (Sulkes and Naparstek 1991). In a Finnish series of 17 cases of neoplasia in SLE, the neoplasm was detected 12 to 30 years after SLE in eight; cytostatic drugs were not considered to have been a contributory factor (Pettersson *et al*. 1992; Sweeney *et al*. 1995). The observation by Sweeney *et al*. that patients in whom a neoplasm develops have a later onset of SLE differs from the reports of others who found the diagnosis of SLE to have been made in 20 of 33 such cases at age 35 or less (Sulkes and Naparstek 1991; Pettersson *et al*. 1992) The prolongation of life has been a factor in the increased concurrence of SLE and neoplasia. However, at present neoplasia in SLE does not exceed the expected incidence (Sweeney *et al*. 1995).

Case reports give the impression that SLE conveys a susceptibility to the development of lymphoreticular neoplasms. The diagnosis of Hodgkin's disease may be obscured by the similarity of symptoms if it develops during the course of SLE (Bhalla *et al*. 1993). However, except for the aforementioned Finnish study, large series of cases of SLE do not support such an association. Four cohorts encompassing 1510 cases of SLE included only six lymphomas among 36 cases of neoplasia (Ropes 1976; Lewis *et al*. 1976; Abu-Shakra *et al*. 1995; Sweeney *et al*. 1995). Furthermore, Razis *et al*. found no instances of SLE among 1269 patients with lymphosarcoma or among 1102 with Hodgkin's disease (Razis *et al*. 1959). Miller found only one instance of SLE among 1893 patients who had a solid tissue neoplasm and one among 264 patients with a lymphoproliferative neoplasm (Miller 1967). Of 29 patients with an autoimmune rheumatic disease and lymphoma only one had SLE (Banks *et al*. 1979). Unless the SLE patient is being treated chronically with a potentially carcinogenic drug, particularly cyclophosphamide, no special surveillance for neoplasia is warranted. A recent report has emphasized the association of SLE and non-Hodgkin's lymphoma (Mellemkjaer *et al*. 1997).

Polymyositis/dermatomyositis

The first two reports of the association of dermatomyositis with cancer were published in 1916. However, whether this association might be more than coincidental initially attracted attention in 1951 when a German dermatologist concluded from a literature review that cancer occurs at least five times as frequently among cases of dermatomyositis as in the general population. The pathogenetic validity of the dermatomyositis–neoplasia association was first questioned in 1975 (Bohan and Peter 1977) and the statistical debate has continued. The only undisputed conclusion is that when the myopathy occurs during childhood it is not associated with neoplasia. The two principal epidemiological questions are:

1. May the apparent association be accounted for entirely by referral and examination biases?
2. Even assuming that the association over all is artefactual, are there hidden subsets of myopathy–neoplasia relationships that have a biological basis?

Most reports have been of individual cases of dermatomyositis and/or polymyositis, or of series of patients who were ill enough to be hospitalized, both of which make extrapolation to the entire spectrum of dermatomyositis/polymyositis cases unreliable. Some recent investigators have sought to avoid 'Berkson's (hospitalization) bias' (Masi and Hochberg 1988). Lyon *et al*. analysed 322 cases that were contributed by American rheumatologists and compared a subset of 104 with a sex-matched sibling (Lyon *et al*. 1989). Only 1.5 per cent of the total cases had had a neoplasm and there was no definite difference in neoplasia between the cases and their siblings. However, only cases that were diagnosed within a 2-year period and assessed in the third year were included. The potential longer-term risk was not addressed.

Although a Swedish investigation was based on patients who had been hospitalized with a diagnosis of dermatomyositis or polymyositis, this was not done as selectively as in most countries and therefore minimized the Berkson bias. A potential weakness was that national cancer incidence data rather than individual matching was used for controls. All cases diagnosed dermatomyositis/polymyositis between 1963 and 1983 were followed through 1987 with the help of a national cancer registry. As in most such cohorts, cancer was more prevalent among cases of dermatomyositis than of polymyositis, but during an average follow-up of 10.4 years all subdivisions demonstrated an increased risk of neoplasia: dermatomyositis male 2.4, female 3.4; polymyositis male 1.8, female 1.7. The risk of a neoplasm within 5 years after the diagnosis of dermatomyositis was: male 4.4, female 4.8 (Sigurgeirsson *et al*. 1992).

Zantos *et al*. analyzed the foregoing (Zantos *et al*. 1994) and two other differently but well controlled smaller studies (Manchul *et al*. 1985; Lakhanpal *et al*. 1986) and concluded that, indeed, both dermatomyositis and polymyositis are associated with an increased risk of cancer. Several studies have found clustering of the diagnosis of dermatomyositis or polymyositis and a neoplasm within a short time frame. Sigurgeirsson *et al*. found both diagnoses to be established within 1 year of each other in one-half of the cases of dermatomyositis and one-third of the cases of polymyositis (Sigurgeirsson *et al*. 1992). Manchul *et al*. found neoplasms in 25 per cent of 71 dermatomyositis/polymyositis cases and in 10 per cent of 142 control cases, half with inflammatory, half with non-inflammatory musculoskeletal diseases (Manchul *et al*. 1985). In eight of the 18 dermatomyositis/polymyositis cases the neoplasm was diagnosed within ±6 months of the myopathy, compared to one of 14 related to the diagnosis of the control cases. Presumably the proximity of some of the diagnoses results from particularly careful examinations stimulated by the belief that the myopathy is a paraneoplastic phenomenon. However, most of these neoplasms are discovered as a result of standard examinations (Plotz *et al*. 1989). The neoplasm which occurs in the dermatomyositis/polymyositis patient reflects the geographic prevalence of the neoplasm rather than a particular affinity to the myopathy. For example 42 per cent of the associated neoplasms in Japan have been gastric carcinomas, versus 8 per cent in the United States; conversely, 8 per cent of dermatomyositis/polymyositis associated neoplasms in Japan have been carcinomas of the breast, versus 17 per cent in the United States (Hidano *et al*. 1986; Callen 1982).

No laboratory tests reliably identify which patient with dermatomyositis/polymyositis is harbouring a neoplasm. The inflammatory and degenerative findings in muscle biopsies are not altered by the presence of a neoplasm, nor are the titres of any of the numerous autoantibodies that may be detectable in this disease helpful (Basset-Sequin *et al*. 1990). Anti-Jo-l and anti-nRNP are the most characteristic (Targoff and Reichlin 1988). Their absence helps to exclude the diagnosis of dermatomyositis/polymyositis. The most consistent biochemical abnormality is elevation of serum creatine phosphokinase, which usually reflects the acuteness of the myopathy, but is not affected by the presence of a neoplasm. It has been suggested that a normal creatine phosphokinase concentration in clinically evident dermatomyositis/polymyositis augurs a poor prognosis (Fudman and Schnitzer 1986). Interpretation must take into consideration that the normal creatine phosphokinase in comparable individuals is about 80 units higher in blacks than in whites (Black *et al*. 1986).

In regard to the association with neoplasia, polymyositis must be differentiated particularly from carcinomatous neuromyopathy (Brain and Adams 1965) and

paraneoplastic endocrine myopathies (Patel *et al*. 1993). Myasthenia gravis, which may coexist with polymyositis (Behan *et al*. 1982), may come into consideration, as well as the rarer Eaton–Lambert myasthenic syndrome (see below), and drug-induced myopathies (Zuckner 1990). Clinically, rapid response to corticosteroid therapy favours the diagnosis of dermatomyositis/polymyositis over the others.

As improved therapy prolongs survival with dermatomyositis/polymyositis, an increase in the number of neoplasm associated cases may be anticipated. Whether this is due to a causal relationship with the myopathy or with the increasingly cancer-prone older age group is uncertain. The conservative conclusion to be drawn from the available epidemiological and clinical data is that when a patient over 45 years of age is seen during the first year of dermatomyositis/polymyositis close surveillance for cancer is justified. Later in the course of the disease more than routine evaluations for cancer are not cost effective.

Diffuse scleroderma and paraneoplastic Raynaud's phenomenon

Diffuse scleroderma (systemic sclerosis) is less common than systemic lupus erythematosus, but probably occurs more frequently than polymyositis. Its peak incidence is in the fifth decade. Small cohorts have not shown an abnormal risk of neoplasia. However, an epidemiological study conducted in metropolitan Pittsburgh (United States) during 1971 to 1982 obtained an excess incidence of 1.8 over regional incidence data. During a mean observation period of 4.3 years, nine women and five men (5.3 per cent) had a malignant neoplasm. A total of 680 patients with systemic sclerosis (78 per cent women) were examined during these 12 years. Twenty-five had a malignancy, 3.6 per cent of women and 4.1 per cent of men (Roumm and Mesdger 1985). Such an excess was confirmed in a study of 248 cases of systemic sclerosis first seen between 1978 and 1992 in Toronto (66 per cent women). Compared to the population of the province of Ontario, systemic sclerosis had an increased cancer risk of 2.1 (Abu-Shakra *et al*. 1993).

It is likely that most neoplasms occur coincidentally with systemic sclerosis, including the few that antedate its development. However, the relationship of carcinomas of the breast and lung with systemic sclerosis requires special comment. The fact that breast cancer is the most common neoplasm among cases of systemic sclerosis can superficially be attributed to the female preponderance of this disease. However, 34 per cent of 90 women with systemic sclerosis and a neoplasm in four studies had breast cancer, about twice the proportion of breast cancer among North American female cancer patients (Duncan and Winkelmann 1979; Lee *et al*. 1983; Roumm and Medsger 1985; Abu-Shakra et al.). Furthermore, while the mean interval between the diagnosis of systemic sclerosis and other neoplasms was about 6 years, the interval for breast cancer was only 2 years (Duncan and Winkelmann 1979; Roumm and Medsger 1985). Of 35 cases where a neoplasm was found within 1 year of the diagnosis of systemic sclerosis, 37 per cent were carcinomas of the breast. This is not entirely attributable to the availability of more sensitive diagnostic methods for cancer of the breast than for cancer of other organs. While the explanation of the relationship remains obscure, the practical conclusion to be drawn from these observations is that women should undergo a careful breast examination at least semi-annually and annual mammography during the first 3 to 5 years after a diagnosis of systemic sclerosis has been made.

The occurrence of carcinoma of the lung in association with systemic sclerosis exhibits three peculiarities:

1. It is consistent with the female susceptibility to systemic sclerosis rather than the male susceptibility for pulmonary carcinoma.
2. Smoking is not implicated in the pathogenesis of these carcinomas.
3. The cancer occurs in cases of long-standing systemic sclerosis in which pulmonary fibrosis has become prominent (Talbott and Barrocas 1980).

Some types of pulmonary fibrosis predispose to carcinogenesis, but the precipitating factors are unidentified. Thus, cryptogenic fibrosing alveolitis is associated with a substantially increased risk of lung cancer (Turner-Warwick 1980), while rheumatoid lung disease is not. The mean duration of systemic sclerosis in 64 patients when lung cancer was diagnosed was 15 years (Talbott and Barrocas 1980). As the most optimistic 5-year survival estimate for patients with systemic sclerosis is only about 70 per cent (Steen 1990), those in whom lung cancer develops tend, contrary to those with breast cancer, to have a relatively benign course of systemic sclerosis. Among 58 patients with systemic sclerosis and lung cancer 60 per cent were less than 60 years of age and 69 per cent were women. The suggestion that alveolar cell carcinoma is peculiarly associated with systemic sclerosis may be attributable to a reporting bias. The histological types of pulmonary carcinomas in systemic sclerosis probably occur in similar proportions to those in other patients (Roumm and Medsger).

While Raynaud's phenomenon is an early manifestation of virtually all cases of systemic sclerosis, it occasionally occurs as a paraneoplastic symptom without other evidence of systemic sclerosis. The preponderance of female cases approaches the 9:1 sex ratio of Raynaud's phenomenon in general. Raynaud's phenomenon appears not to be associated with any particular neoplasm. Onset has been described as virtually concurrent with the diagnosis of the cancer, and no longer than 2 years preceding such a diagnosis. The ischaemia may be unilateral (DeCross and Sahasrabudhe 1992). In some cases it has resolved following resection of the neoplasm. The pathogenetic relationship remains obscure, although in one report cells of a cervical carcinoma were producing interleukin-6, which was suspected to be the vasospastic stimulus (Murashima *et al*. 1992). Raynaud's phenomenon may also be induced by certain antineoplastic drugs: bleomycin, vinblastine, cisplatin (Hansen and Olsen 1989).

Polymyalgia rheumatica

Polymyalgia rheumatica is the rheumatic syndrome that occurs particularly in the most cancer-prone age group. The average age of onset is about 65 years, and the disease begins below the age of 50 in fewer than 5 per cent (Huston *et al*. 1978). There is a female predominance of about 3:2. A survey of 96 patients with polymyalgia rheumatica diagnosed during 10 years in a community in the northern United States included six who had a cancer before and ten in whom the diagnosis of a neoplasm occurred after the onset of polymyalgia rheumatica. The 16 patients encompassed seven different neoplasms (Chuang *et al*. 1982). The largest investigation has been conducted in Norway with 185 cases of polymyalgia rheumatica diagnosed between 1978 and 1983 and followed to 1987. Each was matched with five control subjects. Of the polymyalgia rheumatica cases 14.6 per cent had cancer, as had 14.2 per cent of the controls; the polymyalgia rheumatica cases included 14 different neoplasms. The female preponderance of the neoplasm cases was consistent with the sex ratio of the entire cohort. Thirteen had a neoplasm, found an average of 8.3 years before polymyalgia rheumatica and in 14 a neoplasm was found an average of 4.6 years after the onset of polymyalgia rheumatica (Haga *et al*. 1993). The temporal relationships support the view that these are merely coincidental events.

Neoplasia was more strongly associated with isolated temporal arteritis (24 per cent) than with either isolated polymyalgia rheumatica (11 per cent) or polymyalgia rheumatica with temporal arteritis (10 per cent) (Haga *et al*. 1993). This observation is supported by the finding that there may be a difference between temporal arteritis with and without polymyalgia rheumatica. HLA DR4 occurs significantly less frequently in cases of isolated temporal arteritis than in the presence of polymyalgia rheumatica (Richardson *et al*. 1987). While DR4 appears to be associated with polymyalgia rheumatica as well as with rheumatoid arthritis, whether this is pertinent to neoplasia susceptibility with temporal arteritis warrants further evaluation.

Lymphomas occasionally occur in patients with polymyalgia rheumatica (Kalra and Delamere 1987; Montanaro and Bizzarri 1992). However, the occurrence in series of cases is low enough to indicate coincidence (1 of 96; 2 of 185). Chuang *et al*. found that 18 per cent of their polymyalgia rheumatica cases also met diagnostic criteria for rheumatoid arthritis (Chuang *et al*. 1982). Polymyalgia rheumatica clearly lacks the association with lymphoreticular neoplasia observed in rheumatoid arthritis. Whether the occurrence of monoclonal gammopathy resembles the prevalence in rheumatoid arthritis requires further investigation (Kalra and Delamere 1987).

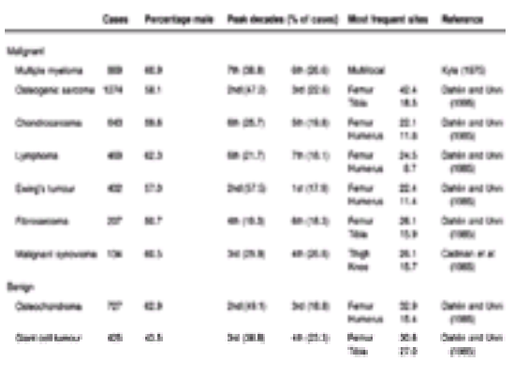
When a person in the seventh decade or beyond presents with myalgia, proximal muscle weakness, and perhaps also weight loss, both polymyalgia rheumatica and a paraneoplastic neuromyopathy should be considered. Polymyalgia rheumatica symptoms are more likely to predominate in the shoulder girdle and paraneoplastic syndromes in the pelvic girdle, but neither this or any serological findings differentiate reliably. Since carcinomatous neuromyopathy most often results from carcinoma of the lung, it is prudent to include chest radiography in the evaluation of possible polymyalgia rheumatica, but an intensive search for a neoplasm is not ordinarily warranted.

Rheumatic manifestations of malignancy

Skeletal neoplasia

Apart from hypertrophic osteoarthropathy, cancer infrequently produces skeletal symptoms that the careful clinician might mistake for a systemic rheumatic disease. Data about the principal primary neoplasms of bone are summarized in Table 2. At least 95 per cent of malignant neoplasms in bone are metastatic carcinomas. These reflect the age and sex distribution of the primary disease; on average, patients with metastatic carcinoma are about two decades older than patients with musculoskeletal sarcomas. Metastases usually result from haematogenous spread with implantation in the marrow cavity (Galasko 1982). The cortex is affected by neoplastic factors that predominantly stimulate osteoblastic activity, resulting in sclerotic lesions, even in the presence of the osteoclastic activity that causes the more

common lytic lesions (Springfield 1982). Only in myeloma and lymphoma do the lytic lesions not tend to be associated with reactive bone formation. Certain tumour cell lines are predisposed to survive in specific organs. Carcinoma of the prostate, breast, kidney, and thyroid are particularly likely to form metastases in bone. Although bronchogenic carcinoma is less likely to do so, the frequency of this neoplasm places it second only to breast as a source of skeletal metastases. The lumbosacral vertebrae, pelvis, femora, and ribs are the most frequent sites of osseous metastases. A radiographically solitary lesion occurs in fewer than 10 per cent of patients with skeletal metastases. Most solitary lesions once were considered benign. However, scintigraphy has caused this opinion to be revised. Because of its greater sensitivity, a bone scan is mandatory to confirm that a lesion probably is solitary (Merrick and Merrick 1986). Vertebrae are the most common site of metastases which are radiographically undetectable initially and manifested only by local pain (Clain 1965). In cases with normal radiographs in which a scintigram indicates metastases, the lesions generally become evident radiographically within 6 to 18 months. A scintigraphically solitary lesion proves to be metastatic in about one-half of cases (McNeill 1984). Nevertheless, a positive radiograph is occasionally contradicted by a false-negative scintigram.



	Cases	Percentage male	Peak decades (% of cases)	Most frequent sites		Reference
Malignant						
Multiple myeloma	869	60.9	7th (36.8)	6th (26.4)	Multifocal	Kyle (1975)
Osteogenic sarcoma	1274	58.1	2nd(47.2)	3rd (22.6)	Femur Tibia	42.4 18.5
Chondrosarcoma	643	59.6	6th (25.7)	5th (19.8)	Femur Humerus	22.1 11.8
Lymphoma	469	62.3	6th (21.7)	7th (16.1)	Femur Humerus	24.5 8.7
Ewing's tumour	402	57.0	2nd(57.5)	1st (17.9)	Femur Humerus	22.4 11.4
Fibrosarcoma	207	50.7	4th (18.3)	6th (16.3)	Femur Tibia	26.1 15.9
Malignant synovioma	106	60.5	3rd (29.9)	4th (26.6)	Thigh Knee	29.1 18.7
Benign						
Osteochondroma	727	62.9	2nd(49.1)	3rd (16.8)	Femur Humerus	32.9 15.4
Giant cell tumour	425	45.5	3rd (38.8)	4th (23.3)	Femur	30.8

Table 2 Principal age, sex, and site of the commoner neoplasms of bone

A primary or a metastatic neoplasm in a juxta-articular site or in synovium may imitate mono- or oligoarticular arthritis. There may be signs of synovitis, including small non-haemorrhagic effusions without actual invasion of the synovium. Conversely, villonodular synovitis, a benign neoplasm which most often affects the knee, may cause a juxta-articular lytic lesion mimicking a primary bone neoplasm (Jergensen *et al*. 1978). Neoplastic cells in joint effusions do not necessarily indicate invasion of the synovium. Despite its vascularity the synovium rarely is a metastatic site (Newton *et al*. 1984). Metastases, particularly to acral sites, may cause symptoms that mimic rheumatoid arthritis or, if there is a local inflammatory reaction and coincidental hyperuricaemia, gout. Nevertheless, acral metastases are uncommon, presumably because of the small number of tumour cells that reach the extremities of the vasculature. According to a review in 1976, 88 cases of metastases to hand bones had been published, most from pulmonary carcinomas (Uriburu *et al*. 1976). Metastases to the feet may be rarer still, 72 cases having been reported up to 1982, without as great a predominance of a particular organ of origin (Zindrick *et al*. 1982).

Synovial sarcomas are uncommon neoplasms that occur predominantly before the age of 40. Fewer than 8 per cent are diagnosed beyond age 60 (Tillotson *et al*. 1951; Cadman *et al*. 1965). There is a 3:2 male preponderance. The tumour is usually located in periarticular tissue rather than within the joint cavity. The most frequent sites in 550 cases collated from four publications are: knee 22.5 per cent; foot and ankle 21.4 per cent; and thigh 15.8 per cent (Tillotson *et al*. 1951; Cadman *et al*. 1965; Mackenzie 1966; Hajdu *et al*. 1977). A lower extremity is affected in approximately 70 per cent of the patients. However, this neoplasm occasionally occurs far from joints, such as in the hypopharynx or oesophagus (Amble *et al*. 1992).

Histologically this neoplasm may be 'biphasic,' meaning that it contains both epithelial and spindle cells, or 'monophasic,' when virtually only one cell type is present. This does not affect the prognosis. Synovioma acquired its name because of its usual proximity to the synovium. However, ultrastructual and immunochemical studies have failed to find synovial characteristics, so that it probably arises from a primitive mesenchymal precursor cell (Miettinen and Virtanen 1984).

About 60 per cent of the patients present with local pain or tenderness, but signs of inflammation are rare. Radiographically about 30 per cent of synoviomas contain calcifications and 10 per cent show signs of bone invasion. Survival is better correlated with a low prevalence of mitoses in the neoplasm than with small size (less than 5 cm diameter) (Cagle *et al*. 1987; Rööser *et al*. 1989). However, both criteria are more predictive than whether the tumour was resected locally or by amputation.

According to Mayo Clinic data, osteoarticular symptoms occur in about 14 per cent of childhood cases of acute leukaemia and in 4 per cent of adults with a leukaemia (Silverstein and Kelly 1963). Musculoskeletal symptoms in paediatric cases usually are a manifestation of acute lymphocytic leukaemia, while in adults the association is strongest with the chronic myelogenous type. Leukaemia is by far the most frequent malignancy in white children, being somewhat less common among black children. It also is the most frequent cause of neoplastic skeletal symptoms in childhood and adolescence. For example of 13 children, 2 to 14 years of age, who presented for rheumatological evaluation and proved to have a malignancy, 10 had leukaemia and the others soft tissue sarcomas (Schaller 1972). Of 28 consecutive leukaemic children (24 acute lymphocytic leukaemia), 14 had musculoskeletal symptoms, most frequently in the knees (Costello *et al*. 1983).

Musculoskeletal symptoms may be the earliest indication of leukaemia and may precede its discovery by several months. There may be bone pain and/or arthralgia, which tends to be asymmetrical. Synovial effusion is uncommon, with its leucocyte count usually below 20 000/mm ^{3}, and often without morphologically evident leukaemic cells (Holdrinet *et al*. 1989). Such identification has been made immunologically (Harden *et al*. 1984). Articular symptoms more often are due to periarticular infiltration than to synovial involvement. Scinitgraphy is more likely than radiography to detect early lesions, but no abnormality may initially be demonstrable. Radiographic findings include metaphyseal radioluscent bands (mainly in children), diffuse osteopenia (mainly in adults), osteolytic lesions, and periostitis. Analysis of synovial fluid is important because the effusion may have resulted from infection or, less frequently, contain urate or calcium pyrophosphate crystals, each of which requiring other than antileukaemic therapy. Osteoarticular leukaemic symptoms usually respond well to chemotherapy and their recurrence augurs a relapse.

Primary skeletal neoplasms are uncommon in childhood. According to a survey conducted during 1969 to 1971 in the United States, the annual incidence below the age of 15 years was about 5 per million (Young and Miller 1975). Neuroblastomas are the most frequent solid tumours to cause skeletal metastases in children. Metastasis of carcinomas to bone is rare—two of 39 in one series of children. The distribution of metastases is as in adults (Leeson *et al*. 1985).

Among skeletal neoplasms in adults multiple myeloma is most likely to present problems in rheumatological differential diagnosis because:

1. This disease occurs mainly after the age of 55, so that complaints of skeletal pain may initially be minimized as symptoms of degenerative disease of joints or intervertebral discs.
2. Myeloma may present as a polyarthritis, more or less resembling rheumatoid arthritis, in which the shoulders, wrists, and knees are affected most often.

There may be subcutaneous nodules, although usually not in the typical sites of rheumatoid nodules. Symptoms rarely are due to the infiltration of synovium by myeloma cells; the usual cause is deposition of amyloid in the synovium (Gordon *et al*. 1973). The same process and symptoms may result from primary amyloidosis (Cohen and Canoso 1975). Synovial fluid sometimes contains the paraprotein, but synovial biopsy is necessary to confirm the diagnosis of articular amyloidosis. The diagnosis of myeloma must be confirmed by the demonstration of a great excess of immature plasma cells in a bone marrow aspirate or biopsy. In rheumatoid arthritis the marrow usually shows only a slight increase in plasma cells, but a large proportion of these may resemble the immature cells seen in myeloma.

Patients with chronic rheumatoid arthritis are at an increased risk of developing multiple myeloma. However, a misdiagnosis may result from the detection of a monoclonal paraprotein, which is found two to three times as frequently among cases of rheumatoid arthritis than in the general population (Zawadzki and Benedek 1969). The significance of a paraproteinaemia may be difficult to interpret at any one time.

In 1971, 241 patients with a diagnosis of monoclonal paraproteinaemia without evidence of any of the usually associated diseases were placed under surveillance by the Mayo Clinic. After a median period of 22 years, 53 per cent had died without having developed a malignant paraproteinaemia. Multiple myeloma had developed in 39 (16 per cent) after a median of 10 years; amyloidosis had developed in 8 (3 per cent) after a median of 9 years; macroglobulinaemia in 7 (3 per cent) after a median of 8 years; and 5 (2 per cent) developed a malignant lymphoproliferative disease after a median of 10.5 years (Kyle 1993). These data suggest that about a

quarter of cases of 'benign' paraproteinaemia are at risk to develop one of these fatal diseases.

Oncogenic osteomalacia

Osteomalacia is a less frequent cause of tumour-associated skeletal pain than hypertrophic osteoarthropathy. However, since most of the tumours are benign, have no preferential location, and many are small, they may be difficult to find. Hence, symptoms may last for years unexplained. The patient presents with skeletal and/or articular pain, especially affecting the lower extremities, progressive pelvic girdle weakness leading to deterioration of the gait, there may be fractures and, with childhood onset, growth retardation. Symptoms most often begin in the third or fourth decade without a family history of rickets (Ryan and Reiss 1984). Various types of fibromas and haemangiomas are the most common; sarcomas have rarely been associated with this paraneoplastic activity, and carcinomas not at all (Weidner and Cruz 1987). Radiographic findings include pseudofractures (Looser's lines) and thickening of long bones despite osteopenia. Bone scans show scattered areas of increased tracer uptake (Schapira *et al*. 1995). The characteristic biochemical findings are hypophosphataemia, hyperphosphaturia, elevated alkaline phosphatase, normal to slightly subnormal serum calcium, normal serum creatinine and parathormone, and subnormal serum 1,25-dihydroxyvitamin D.

It has been inferred from the biochemical data that an unidentified tumour product interferes with the conversion of 25-hydroxyvitamin D to the active 1,25-dihydroxy compound (Harvey *et al*. 1992). Why this defect does not result in hyperparathyroidism is unknown. The disease is cured symptomatically and biochemically by resection of the tumour. If this is not feasible or until it is located, treatment with phosphate and calcitriol (1–3 µg/day) is beneficial.

Hypertrophic osteoarthropathy

Hypertrophic osteoarthropathy often is referred to as 'pulmonary' because most cases are associated with an intrathoracic disease. Lung cancer now accounts for at least 90 per cent of the cases in the industrialized world. However, extrathoracic circumstances as disparate as non-neoplastic liver disease (Epstein *et al*. 1979) and pregnancy (Borden and Holling 1969) may be associated with the same syndrome. The occurrence of hypertrophic osteoarthropathy in large series of cases of lung cancer has varied widely, perhaps due to inconsistent diagnostic criteria or referral bias. It is more likely to occur in males regardless of age. In the Edinburgh analysis of 4000 cases of primary lung cancer, of whom 90 per cent were men, only one of 49 cases of hypertrophic osteoarthropathy occurred in a woman (LeRoux 1968). At the Mayo Clinic, 10 per cent of 1657 cases of primary lung cancer manifested hypertrophic osteoarthropathy. Women accounted for 19 per cent of those with hypertrophic osteoarthropathy, but 26 per cent of those without this complication (Stenseth *et al*. 1967). Of 1920 cases of lung cancer seen in Cleveland (United States), hypertrophic osteoarthropathy was found in only 0.8 per cent (Segal and MacKenzie 1982), and among 200 cases in London it was found in 4 per cent (Yacoub 1965). Hypertrophic osteoarthropathy is less than one-third as likely to occur with pulmonary metastases as with primary lung cancer (Stenseth *et al*. 1967), but sarcomatous metastases are more likely than carcinomatous metastases to induce hypertrophic osteoarthropathy. Of 38 children (mean age 13) with neoplastic hypertrophic osteoarthropathy 79 per cent were boys, 37 per cent had an osteogenic sarcoma, 34 per cent a nasopharyngeal carcinoma, and 16 per cent had Hodgkin's disease (Staalman and Umans 1993).

Calculated from a sex, age, and race-specific cohort analysis of the prevalence of lung cancer in the United States during 1983 to 1986, and assuming a 6 to 9 per cent incidence of hypertrophic osteoarthropathy, the following frequency per 100 000 would be anticipated in each 55 to 64-year age group: white male 15 to 22; black male 25 to 28; white female 8 to 11; black female 9 to 13. Because of the lower female incidence of hypertrophic osteoarthropathy these figures should be reduced by at least one-third for women. For the 45 to 54-year age group the figures would be about 40 per cent of the foregoing (Devesa *et al*. 1989).

The most apparent but least specific sign of hypertrophic osteoarthropathy is digital clubbing. This begins with subungual oedema and hyperaemia, followed by proliferation of connective tissue. About 80 per cent of such individuals have some form of pleuropulmonary disease, but only about 20 per cent of cases of clubbing have lung cancer. Chronic clubbing does not predispose to the development of hypertrophic osteoarthropathy (Coury 1960).

Patients usually complain of arthralgia, especially of the knees, ankles, or wrists. However, on examination the pain typically is periarticular. Some patients have true joint pain and tenderness, and perhaps small effusions and morning stiffness. Hence, such a patient may be misdiagnosed as having rheumatoid arthritis. The erythrocyte sedimentation rate in this situation is elevated, but rheumatoid factor, with rare exceptions, is absent. Symptoms usually may be controlled with non-steroidal anti-inflammatory drugs (Schumacher 1976).

Skeletal metastases from lung cancer occur more frequently than hypertrophic osteoarthropathy and tend to produce less symmetrical symptoms. The simultaneous occurrence of hypertrophic osteoarthropathy and bony metastases is rare. The greater diagnostic problem occurs because symptoms of hypertrophic osteoarthropathy often precede the discovery of an intrathoracic neoplasm by a few months, and possibly by two or more years. Despite the frequency of rheumatoid arthritis in middle-aged women, the possibility that this disease is being mimicked by hypertrophic osteoarthropathy should be considered when a woman who has risk factors for lung cancer is being evaluated.

Radiographically, signs of periostitis with subperiosteal new bone formation along the distal and/or proximal fourths of long bones, predominantly on the extensor surfaces, is diagnostic. The most commonly and prominently affected sites are the distal and proximal portions of the tibia and fibula, and distal femur (Ali *et al*. 1980). Because the periosteal reaction is so hyperaemic, bone scanning with $^{99}Tc^{m}$-diphosphonate has proven to be diagnostically more sensitive than radiography. This technique, contrary to radiography, has demonstrated rather frequent involvement of the scapula and skull.

The occurrence of hypertrophic osteoarthropathy is influenced by the size, location, and cell type of the primary pulmonary neoplasm. The mean tumour mass in cases of hypertrophic osteoarthropathy is about twice as large as in those without hypertrophic osteoarthropathy (Stenseth *et al*. 1967). The influence of location is especially perplexing. Location in the periphery of a lung increases the chance that hypertrophic osteoarthropathy will develop, while invasion of the pleura appears not to exert this influence. The association of hypertrophic osteoarthropathy with malignant mesothelioma is also unusual, while the neoplasm with which hypertrophic osteoarthropathy is most highly associated is benign mesothelioma—35 per cent (Briselli *et al*. 1981). In contradistinction to other paraneoplastic manifestations of lung cancer, which are particularly associated with the small-cell type, hypertrophic osteoarthropathy occurs least often in this relationship.

The immediate pathophysiology of hypertrophic osteoarthropathy consists of increased blood flow in the bones and adjacent connective tissues of the legs and forearms, facilitated by the development of arteriovenous shunts (Rutherford *et al*. 1969). The same syndrome occurs in dogs, especially secondary to thoracic sarcomas (Nolling *et al*. 1963). No comprehensive explanation of the syndrome has been proved. Most of the pathogenetic hypotheses were thoroughly reviewed by Shneerson (Shneerson 1981). Impetus for a neurogenic explanation evolved from the discovery that peripheral vagotomy on the side of the neoplasm may, even without its removal, abruptly stop the pain bilaterally, followed by gradual resolution of the periosteal reaction and clubbing. Analgesia begins within a week of surgery in about three-quarters of cases, probably due to collapse of the arteriovenous anastomoses (Nolling *et al*. 1963; Stenseth *et al*. 1967; Rutherford *et al*. 1969). However, the gradual response of hypertrophic osteoarthropathy to radiation therapy of the neoplasm and the bilaterality of the syndrome suggest a humoral rather than a neurogenic mechanism. The heterogeneity of its causes proves that hypertrophic osteoarthropathy is not mediated by a uniquely neoplastic product. Gynaecomastia, which is an undoubtedly humoral manifestation that may occur in hypertrophic osteoarthropathy, is not correlated with increased serum oestrogen concentration or another identified hormone. A substance resembling growth hormone has recently been considered as a potential humoral incitor of hypertrophic osteoarthropathy (Gosney *et al*. 1990). A persuasive explanation for many, but not all, characteristics of hypertrophic osteoarthropathy has been made for platelet embolization with release of platelet-derived growth factor. This substance, among other functions, stimulates mesenchymal cell growth, increases vascular permeability, and attracts smooth muscle cells and fibroblasts (Dickinson 1993).

A combined neurohumoral pathogenesis of hypertrophic osteoarthropathy seems most plausible. Only a neural mechanism could explain the abrupt diminution of peripheral blood flow and associated analgesia that usually follows vagotomy, while the symmetrical osteogenesis is best explained by the action of one or perhaps several as yet unidentified humours.

Carcinomatous polyarthritis

Carcinomatous non-metastatic polyarthritis must be distinguished from hypertrophic osteoarthropathy, rheumatoid arthritis, and polymyalgia rheumatica. The syndrome is not associated predominantly with intrathoracic neoplasms and both sexes are equally at risk. Most cases occur above the age of 60, when the onset of polymyalgia rheumatica but not of rheumatoid arthritis is expected. Onset of the arthritis may be insidious or abrupt, but the symptoms are less likely to be symmetrical than in definite rheumatoid arthritis, and rheumatoid nodules do not develop (Pines *et al*. 1984). Pain is felt in large and small joints, rather than being periarticular as in hypertrophic osteoarthropathy. Deformities rarely develop. Symptoms are more likely than those of hypertrophic osteoarthropathy to precede the diagnosis of a neoplasm which, nevertheless, is usually made within a year of the beginning of articular complaints. On the contrary, the diagnosis of polymyalgia rheumatica and a neoplasm usually occurs several years apart. Immune complexes have rarely been detected in carcinomatous polyarthritis (Awerbuch and Brooks 1981; Bradley and

Pinals 1983).

Arthritic symptoms as well as serological abnormalities rapidly resolve following resection or successful radiation therapy of the neoplasm (Gottlieb *et al*. 1979; Simon and Ford 1980). Recurrence of the neoplasm may be indicated by recurrence of the arthritis (Egelmeijer and MacFarlane 1992). If the neoplasm persists, the clinical course of the arthropathy does not necessarily mirror the progress of the neoplasm. Symptomatic response to treatment with non-steroidal anti-inflammatory drugs or corticosteroids is similar to that in rheumatoid arthritis.

Carcinomatous neuropathy

In individual cancer patients it may be difficult to determine whether the development of weakness is due to inanition or is a specific paraneoplastic manifestation. When to suspect cancer in a patient with unexplained weakness and perhaps some muscle wasting is equally problematical. In a prospective study of 100 persons over 65 years of age who had no identified cancer but whose gait was impaired by largely unexplained weakness of the legs, eight cases of carcinoma were detected during 18 months of observation (Newman and Gugino 1964).

According to Shy and Silverstein, carcinomatous neuromyopathy is a 'clinical syndrome of symmetrical muscular weakness and wasting, associated with decrease of the appropriate myotactic reflexes. Either a myopathic or a combination of myopathic and neuropathic lesion. ... is found pathologically' (Shy and Silverstein 1965). At least initially, weakness is out of proportion to alteration of muscle mass. Pelvic girdle muscles are most affected, so that gait disturbances, particularly on ascending stairs, may be the first symptom.

The syndrome occurs most frequently in association with small-cell carcinoma of the lung, followed by other pulmonary, gastric, and ovarian carcinomas. Because of the male preponderance of lung cancer, non-myasthenic carcinomatous neuromyopathy occurs more commonly in men (Croft and Wilkinson 1965); on examination of 1465 carcinoma patients they made this diagnosis in 15.0 per cent of men and 11.6 per cent of women with lung cancer, 7.4 per cent of men and 12.5 per cent of women with gastric carcinoma, and 16.4 per cent of cases of ovarian carcinoma. No consistent relationship between the occurrence of the myopathy and either the age of the patient or known duration of the neoplasm has been determined. Compared to 'neuromyopathy' Eaton–Lambert syndrome is rare (probably less than 1:40). Autonomic dysfunctions, such as dry mouth and postural hypotension, are common findings with small-cell carcinoma and not a reliable diagnostic clue to Eaton–Lambert syndrome (Elrington *et al*. 1991). Differentiation between 'cachectic' and 'neuromyopathic' clinical findings may be uncertain. Hawley *et al*., in a prospective study of 71 patients with small-cell lung cancer, were impressed by a greater preservation of strength relative to the severity of weight loss (Hawley *et al*. 1980). The myopathy also was less severe than that found in a group of alcoholic men matched by age and weight loss.

Both histopathological and electrophysiological abnormalities are demonstrable in the absence of clinical neuromuscular findings. In a study of 100 consecutive patients with lung cancer who were assessed with muscle biopsies, 15 (mainly small-cell) had signs of proximal myopathy, while 18 (mainly non-small-cell) had cachectic (diffuse) myopathy. However, 99 patients had abnormal histological findings. Most consistent (74 cases) was atrophy of type-2 (large mean diameter, phosphorylase-rich) nerve fibres (Gomm *et al*. 1990). Contrary to some investigators, no microvascular abnormalities, such as atrophy of terminal arteriovenous anastomoses (Scelsi and Pinelli 1977) were detected. In some cases there is only scattered, non-specific atrophy or necrosis of muscle fibres with lymphocytic infiltrates.

Paul *et al*. compared the electromyographic findings of 195 non-diabetic cancer patients below the age of 65 years with 50 age-matched controls (Paul *et al*. 1978). Muscle wasting was better correlated with myographic abnormalities than with slowed nerve conduction, and both types of abnormality were more prevalent among patients with diffuse muscle wasting, most of whom presumably had a cachectic myopathy, than in those whose wasting predominantly affected the proximal musculature. Retarded conduction velocity was equally prevalent in patients with carcinomas of the lungs and of other organs, but other abnormalities were twice as frequent among cases of lung cancer. The cause of the various abnormalities and of their variability remain obscure. Treatment is non-specific until the neoplasm is found and then is focused on the neoplasm.

Rarely a more aggressive myopathy occurs. In this, weakness progresses within a few weeks and affects virtually all muscles, despite preservation of tendon reflexes. It is not associated with a particular type of carcinoma. There is extensive necrosis of terminal intramuscular nerve fibres and of muscle fibres, with consequent elevation of the serum creatine phosphokinase (Vosskämper *et al*. 1989).

Eaton–Lambert myasthenic syndrome

The Eaton–Lambert myasthenic syndrome was differentiated from myasthenia gravis in 1957. In early descriptions it was associated almost exclusively with small-cell carcinoma of the lung. The association with this neoplasm has been confirmed but it has become recognized that nearly one-half of cases occur in the absence of a neoplasm. The latter cases have brought the sex ratio from strongly male to 1:1 (Pascuzzi and Kim 1990; McEvoy 1994). Eaton–Lambert syndrome occurs in about 3 per cent of cases of small-cell carcinoma and rarely in other pulmonary or extrapulmonary neoplasms (O'Neill *et al*. 1988).

The most common presenting finding is weakness of the pelvic girdle musculature, initially revealed by an altered gait. Autonomical functions may be affected, manifested most often by dryness of the mouth and/or orthostatic hypotension (Khurana *et al*. 1988). The main differential diagnosis is with myasthenia gravis. While one can be fairly confident that no carcinoma is impending if none has been found after 3 years of Eaton–Lambert syndrome, one must consider that myasthenia gravis is also associated with an increased incidence of neoplasia, albeit less so. In contrast to Eaton–Lambert syndrome, at least half of these cancers are found after myasthenia gravis has been present for longer than 3 years. Most patients with either syndrome who are found to have cancer are between 55 and 65 years of age (Papatestas *et al*. 1971). One simple differentiating symptom from myasthenia gravis is that extraocular muscles typically are not affected in Eaton–Lambert syndrome.

If Eaton–Lambert syndrome is associated with a carcinoma, there rarely is more than a 2-year interval between the onset of muscular symptoms and discovery of the neoplasm. In the 21 patients with small-cell lung cancer of O'Neill *et al*. the tumour was diagnosed within 1 year of the syndrome's onset in 16 (O'Neill *et al*. 1988). The maximum interval between the onset of Eaton–Lambert syndrome and detection of the neoplasm has been less than 4 years and in such cases the association may be coincidental. The age distribution is wider in the absence of neoplasia but the neuromuscular manifestations are the same.

The simplest screening test for Eaton–Lambert syndrome consists of evoking a single muscle action potential with an electromyograph before and after voluntary effort. The potential should at least double, and may increase 17-fold. Repetitive stimulation of a nerve at a tetanic rate also results in increase of the action potential, while in myasthenia gravis the potential gradually diminishes. There are also pharmacological differences. Neostigmine and other anticholine esterases improve strength in myasthenia gravis, but have little effect in Eaton–Lambert syndrome. Consequently, in Eaton–Lambert syndrome a test dose of edrophonium fails to elicit a definite response as it does for myasthenia gravis (Brown and Johns 1974).

Both conditions result from a defect in the neuromuscular transmission of nerve impulses. In myasthenia gravis acetyl choline reaches the motor endplates in normal increments, but the sensitivity of its receptors on the endplates and contiguous muscle cell membrane is diminished. Each receptor site can accept two molecules of acetyl choline. Release of the quanta of acetyl choline in the depolarization phase of the electric potential is mediated by the influx of calcium ions into the presynaptic membrane. In Eaton–Lambert syndrome an inadequate amount of acetyl choline is released due to a complex mechanism which, at least in part, is an antibody blockade of the calcium phase (Leys *et al*. 1991; Maselli 1994). Injection of the IgG serum fraction from patients with Eaton–Lambert syndrome into mice may temporarily reduce their muscle action potential. This has been interpreted to indicate that there is an autoantibody which, when transferred in sufficient concentration, can block neuromuscular transmission (Lambert and Lennon 1988). Such an antibody appears to be more prevalent in non-neoplastic cases of Eaton–Lambert syndrome than in the presence of small-cell carcinoma (Leys *et al*. 1991). How this antibody (or antibodies) is generated remains obscure. It appears unlikely that this mechanism is related to the generation of other pathogenic autoantibodies. The occurrence of Eaton–Lambert syndrome has been reported in association with rheumatoid arthritis (Peris *et al*. 1990), systemic lupus erythematosus (Hughes and Katirji 1986), and Sjögren's syndrome (Tsuchiya *et al*. 1993), but is so rare as to be presumed coincidental.

Treatments endeavour to regain the effectiveness of acetyl choline. Plasmapheresis produces transient clinical improvement, although more gradually in Eaton–Lambert syndrome than in myasthenia gravis. Therefore, Newsom-Davis and Murray speculated that improved strength in Eaton–Lambert syndrome results from newly synthesized acetyl choline rather than from removal of blocking antibody (Newsom-Davis and Murray 1984). Resection of the carcinoma may result in neuromuscular improvement. Corticosteroid therapy may be beneficial, but the effect develops more gradually than in myasthenia gravis, possibly over months. The first relatively specific medication for Eaton–Lambert syndrome was guanidine, 20 to 30 mg/kg per day orally; it facilitates the release of acetyl choline. However, there is risk of bone marrow depression and other toxicities (Henriksson *et al*. 1977). 3,4-diaminopyridine enhances acetyl choline release by prolonging the action potential

and calcium influx. In dosages from 15 to 100 mg/day it is at present the most efficient medication. Its main, infrequent, side-effect is seizures (McEvoy 1994). In the absence of a fatal neoplasm Eaton–Lambert syndrome is a chronic disease.

Fasciitis syndromes

Several varieties of non-infectious fasciitis have been described in recent years: eosinophilic fasciitis (1974); palmar fasciitis (1982); fasciitis–panniculitis (1992). The main pathological difference between eosinophilic fasciitis and fasciitis–panniculitis seems to be the lack of an eosinophilic infiltrate in the latter. The strongest clue that the neoplasms described in a minority of cases of eosinophilic fasciitis and fasciitis–panniculitis are not merely coincidental events is that they have almost exclusively been haematological (Lakhanpal *et al*. 1988; Naschitz *et al*. 1994).

Medsger *et al*. described a 'palmar fasciitis-arthritis syndrome' associated with ovarian carcinoma (Medsger *et al*. 1982). This may be the neoplasia induced variant of reflex sympathetic dystrophy, which it resembles. Cases published as 'shoulder-hand syndrome associated with cancer' probably have been instances of the same process (Michaels and Sorber 1984). However, there are differences from typical reflex sympathetic dystrophy: palmar fasciitis is always bilateral, may involve the lower extremities, and tends to advance to severe disability more rapidly. The skin of the hands becomes so tense that a misdiagnosis of scleroderma may be made. Involvement of the plantar fascia has also been reported. The shoulders are affected consistently, but the arthritis may also involve extremity joints. Most patients have been women above the age of 55, and about half of the neoplasms have been ovarian carcinomas (Pfinsgraff *et al*. 1986). The syndrome develops a few weeks to a year before discovery of the neoplasm and may first occur with a recurrence (Mesdger *et al*. 1982). Treatment of the neoplasm usually does not alleviate the syndrome.

Chapter References

Abu-Shakra, M., Guillemin, F., and Lee, P. (1993). Cancer in systemic sclerosis. *Arthritis and Rheumatism*, **36**, 460–4.

Abu-Shakra, M., Urowitz, M.B., Gladman, D.D., *et al*. (1995). Mortality studies in systemic lupus erythematosus. Results from a single center. 1. Causes of death. *Journal of Rheumatology*, **22**, 1259–64.

Ahern, M.J., Maddison, P., Mann, S., and Scott, C.A. (1982). Ankylosing spondylitis and adenocarcinoma of the lung. *Annals of the Rheumatic Diseases*, **41**, 292–4.

Ali, A., Tetalman, M.R., Fordham, E.W., *et al*. (1980). Distribution of hypertrophic pulmonary osteoarthropathy. *American Journal of Radiology*, **134**, 771–80.

Amble, A.R., Olsen, K.D., Nascimento, A.G., *et al*. (1992). Head and neck synovial cell sarcoma. *Otolaryngology, Head and Neck Surgery*, **107**, 631–7.

Awerbuch, M.S. and Brooks, P.M. (1981). Role of immune complexes in hypertrophic osteoarthropathy and non-metastatic polyarthritis. *Annals of the Rheumatic Diseases*, **40**, 470–2.

Baker, G.L., Kahl, L.E., Zee, B. *et al*. (1987). Malignancy following treatment of rheumatoid arthritis with cyclophosphamide: a long-term case-control follow up study. *American Journal of Medicine*, **83**, 1–9.

Banks, P.M., Witrak, G.A., and Conn, D.L. (1979). Lymphoid neoplasia following connective tissue disease. *Mayo Clinic Proceedings*, **54**, 104–8.

Basset-Sequin, N., Roujeau, J., Gherardi, R., *et al*. (1990). Prognostic factors and predictive signs of malignancy in adult dermatomyositis. *Archives of Dermatology*, **126**, 633–7.

Behan, W.M., Behan, P.O., and Doyle, D. (1982). Association of myasthenia gravis and polymyositis with neoplasia, infection and autoimune disorders. *Acta Neuropathologica*, **57**, 221–9.

Bhalla, R., Ajmani, H.S., Kim, W.W., *et al*. (1993). Systemic lupus erythematosus and Hodgkin's lymphoma. *Journal of Rheumatology*, **20**, 1316–20.

Bignon, Y.J., Janin-Mercier, A., Dubost, J.J., *et al*. (1986). Angioimmunoblastic lymphadenopathy with dysproteinaemia (AILD) and sicca syndrome. *Annals of the Rheumatic Diseases*, **45**, 519–22.

Black, H.R., Quallich, H., and Gereleck, C.B. (1986). Racial differences in serum creatine kinase levels. *American Journal of Medicine*, **81**, 479–86.

Bohan, A. and Peter, J.B. (1977). A computer-assisted analysis of 153 patients with polymyositis and dermatomyositis. *Medicine (Baltimore)*, **56**, 255–86.

Boice, J.D. (1992). Radiation and non-Hodgkin's lymphoma. *Cancer Research*, **52**(suppl.), 5489s-91s.

Borden, E.C. and Holling, H.E. (1969). Hypertrophic osteoarthropathy and pregnancy. *Annals of Internal Medicine*, **71**, 577–80.

Bradley, J.D. and Pinals, R.S. (1983). Carcinoma polyarthritis: role of immune complexes in pathogenesis. *Journal of Rheumatology*, **10**, 826–8.

Brain, R. and Adams, R.D. (1965). A guide to the classification and investigation of neurological disorders associated with neoplasms. In *The remote effects of cancer on the nervous system* (ed. R. Brain and F.H. Norris), pp. 216–21. Grune and Stratton, New York.

Briselli, M., Mark, E.J., and Dickensin, R. (1981). Solitary fibrous tumors of the pleura. *Cancer*, **47**, 2678–88.

Brown, J.D. and Johns, R.J. (1974). Diagnostic difficulties encountered in the myasthenic syndrome sometimes associated with carcinoma. *Journal of Neurology, Neurosurgery and Psychiatry*, **37**, 1212–24.

Cadman, N.L., Soule, E.H., and Kelly, P.J. (1965). Synovial sarcoma: an analysis of 154 tumors. *Cancer*, **18**, 613–27.

Cagle, L.A., Mirra, J.M., Storm, K., *et al*. (1987). Histologic features relating to prognosis in synovial sarcoma. *Cancer*, **59**, 1810–14.

Callen, J.P. (1982). The value of malignancy evaluation in patients with dermatomyositis. *Journal of the American Society of Dermatology*, **6**, 253–9.

Casey, T.P. (1968). The development of lymphomas in mice with autoimune disorders treated with azathioprine. *Blood*, **31**, 396–9.

Chevalier, X., Gaular, P., Voisin, M.-C., *et al*. (1991). Peripheral T-cell lymphoma with Sjögren's syndrome; a report with immunologic and genotypic studies. *Journal of Rheumatology*, **18**, 1744–6.

Chuang, T.-Y., Hunder, G.G., Ilstrup, D.M., *et al*. (1982). Polymyalgia rheumatica: a 10-year epidemiologic and clinical study. *Annals of Internal Medicine*, **97**, 672–80.

Clain, A. (1965). Secondary malignant disease of bone. *British Journal of Cancer*, **19**, 15–29.

Cohen, A.S. and Canoso, J.J. (1975). Rheumatologic aspects of amyloid disease. *Clinics in Rheumatic Disease*, **1**, 149–61.

Costello, P.B., Beecher, M.L., Starr, J.J., *et al*. (1983). A prospective analysis of the frequency, cause, and possible prognostic significance of the joint manifestations of childhood leukaemia. *Journal of Rheumatology*, **10**, 753–7.

Court-Brown, W.M. and Doll, R. (1965). Mortality from cancer and other causes after radiotherapy for ankylosing spondylitis. *British Medical Journal*, **2**, 1327–32.

Coury, C. (1960). Hippocratic fingers and hypertrophic osteoarthropathy. A study of 350 cases. *British Journal of Diseases of the Chest*, **54**, 202–9.

Croft, P.B. and Wilkinson, M. (1965). The incidence of carcinomatous neuromyopathy in patients with various types of carcinoma. *Brain*, **88**, 427–34.

Dahlin, D.C. and Unni, K.K. (1985). *Bone tumors: general aspects and data on 8542 cases*, (4th edn), pp. 12–15. Thomas, Springfield, IL.

Darby, S.C., Doll, R., Gill, S.K., and Sith, P.G. (1987). Long term mortality after a single treatment course with spondylitis. *British Journal of Cancer*, **55**, 179–90.

DeCross, A.J. and Sahasrabudhe, D.M. (1992). Paraneoplastic Raynaud's phenomenon. *American Journal of Medicine*, **92**, 571–2.

Devesa, S.S., Blot, W.J., and. Fraumeni, J.F. (1989). Declining lung cancer rates among young men and women in the United States: a cohort analysis. *Journal of the National Cancer Institute*, **81**, 1568–71.

Dickinson, C.J. (1993). The aetiology of clubbing and hypertrophic osteoarthropathy. *European Journal of Clinical Investigation*, **23**, 330–8.

Duncan, S.C. and Winkelmann, R.K. (1979). Cancer in scleroderma. *Archives of Dermatology*, **115**, 950–5.

Edgar, M.A. and Robinson, M.P. (1973). Post-radiation sarcoma in ankylosing spondylitis. *Journal of Bone and Joint Surgery*, **55B**, 183–8.

Egelmeijer, F. and MacFarlane, J.D. (1992). Polyarthritis as the presenting symptoms of the occurrence and recurrence of a laryngeal carcinoma. *Annals of the Rheumatic Diseases*, **51**, 556–7.

Elrington, G.M., Murray, N.M., Spiro, S.G., *et al*. (1991). Neurological paraneoplastic syndromes in patients with small cell lung cancer. A prospective survery of 150 patients. *Journal of Neurology, Neurosurgery and Psychiatry*, **54**, 764–7.

Epstein, O., Adjukiewicz, A,B,M., Dick, R., *et al*. (1979). Hypertrophic hepatic osteoarthropathy. *American Journal of Medicine*, **67**, 88–97.

Eriksson, M. (1993). Rheumatoid arthritis as a risk factor for multiple myeloma: a case-control study. *European Journal of Cancer*, **29A**, 259–63.

Fudman, E.J. and Schnitzer, T.J. (1986). Dermatomyositis without creatine kinase evaluation. *American Journal of Medicine*, **80**, 329–32.

Galasko, C.S. (1982). Mechanisms of lytic and blastic metastatic disease of bone. *Clinical Orthopedics and Related Research*, **169**, 20–7.

Giovannucci, E., Egan, K.M., Hunter, D.J., *et al*. (1995). Aspirin and the risk of colorectal cancer in women. *New England Journal of Medicine*, **333**, 609–14.

Gleeson, M.J., Bennett, M.H., and Clawson, R.A. (1986). Lymphomas of salivary glands. *Cancer*, **58**, 699–704.

Gomm, S.A., Thatcher, N., Barber, P.V., *et al*. (1990). A clinicopathological study of the paraneoplastic neuromuscular syndromes associated with lung cancer. *Quarterly Journal of Medicine*, **75**, 577–95.

Gordon, D.A., Pruzanski, W., Ogryzlo, M.A., *et al*. (1973). Amyloid arthritis simulating rheumatoid disease in five patients with multiple myeloma. *American Journal of Medicine*, **55**, 142–54.

Gosney, M.A., Gosney, J.R., and Lye, M. (1990). Plasma growth hormone and digital clubbing in carcinoma of the bronchus. *Thorax*, **45**, 545–7.

Gottlieb, M., Hoppe, R.T., Calin, A., *et al*. (1979). Arthritis in a patient with mycosis fungoides: complete remission after radiotherapy. *Arthritis and Rheumatism*, **22**, 424–5.

Gridley, G., McLaughlin, J.K., Ekbom, A., *et al*. (1993). Incidence of cancer among patients with rheumatoid arthritis. *Journal of the National Cancer Institute*, **85**, 307–11.

Haga, H.-J., Eide, G.E., Brun, J., *et al*. (1993). Cancer association with polymyalgia rheumatica and temporal arteritis. *Journal of Rheumatology*, **20**, 1335–9.

Hajdu, S.I., Shiu, M.H., and Former, J.G. (1977). Tendosynovial sarcoma. A clinico-pathological study of 136 cases. *Cancer*, **39**, 1201–17.

Hakulinen, T. (1986). Similar survival rates for rheumatoid and non-rheumatoid cancer patients. *Scandinavian Journal of Rheumatology*, **15**, 285–9.

Hansen, S. and Olsen, N. (1989). Raynaud's phenomenon in patients treated with cisplatin, vinblastine, and bleomycin for germ cell cancer: measurement of vasoconstrictor response to cold. *Journal of Clinical Oncology*, **7**, 940–2.

Harden, E.A., Moore, J.O., and Haynes, B.F. (1984). Leukemia-associated arthritis: identification of leukemic cells in synovial fluid using monoclonal and polyclonal antibodies. *Arthritis and Rheumatism*, **27**, 1306–8.

Harvey, J.M., Gray, C., and Belchetz, P.E. (1992). Oncogenous osteomalacia and malignancy. *Clinical Endocrinology*, **37**, 379–84.

Hawley, R.J., Cohen, M.H., Saini, N., *et al*. (1980). The carcinomatous neuromyopathy of oat cell lung cancer. *Annals of Neurology*, **7**, 65–72.

Henriksson, K.G., Nilsson, O., Rosen, I., *et al*. (1977). Clinical neurophysiological and morphological findings in Eaton Lambert syndrome. *Acta Neurologica Scandinavica*, **56**, 117–40.

Hidano, A., Kanako, K., Arai, Y., *et al*. (1986). Survey of the prognosis for dermatomyositis, with special reference to its association with malignancy and pulmonary fibrosis. *Journal of Dermatology*, **13**, 233–41.

Holdrinet, R.S., Corstens, F., van Horn, J.R., *et al*. (1989). Leukemic synovitis. *American Journal of Medicine*, **86**, 123–5.

Holm, S. (1949). Keratoconjunctivitis sicca and the sicca syndrome. *Acta Ophthalmologica*, **33**(suppl.), 44–59.

Hornbaker, J.H., Foster, E.A., Williams, G.S., *et al*. (1966). Sjögren's syndrome and nodular reticulum cell sarcoma. *Archives of Internal Medicine*, **118**, 449–52.

Hughes, R.L. and Katirji, B. (1986). The Eaton-Lambert (myasthenic) syndrome in association with systemic lupus erythematosus. *Archives of Neurology*, **43**, 1186–7.

Huston, K.A., Hunder, G.E., Lie, J.T., *et al*. (1978). Temporal arteritis, A 25-year epidemiologic, clinical, and pathologic study. *Annals of Internal Medicine*, **88**, 162–7.

Hyjek, E., Smith, W.J., and Isaacson, P.G. (1988). Primary B-cell lymphoma of salivary glands and its relationship to myoepithelial sialadenitis. *Human Pathology*, **19**, 766–76.

Hyman, G.A. and Wolff, M. (1976). Malignant lymphoma of the salivary glands. *American Journal of Clinical Pathology*, **65**, 421–38.

Isomäki, H.A., Hakulinen, T., and Joutsenlahti, U. (1978). Excess risk of lymphomas, leukemia and myeloma in patients with rheumatoid arthritis. *Journal of Chronic Diseases*, **31**, 691–6.

Janin, A., Morel, P., Quiguandon, I., *et al*. (1992). Non-Hodgkin's lymphoma and Sjögren's syndrome. An immunopathological study of 113 patients. *Clinical and Experimental Rheumatology*, **10**, 565–70.

Jergensen, H.E., Mankin, H.J., and Schiller, A.L. (1978). Diffuse pigmented villonodular synovitis of the knee mimicking primary bone neoplasm. *Journal of Bone and Joint Surgery*, **60A**, 825–9.

Kalra, L. and Delamere, J.P. (1987). Lymphoreticular malignancy and monoclonal gammopathy presenting as polymyalgia rheumatica. *British Journal of Rheumatology*, **26**, 458–9.

Kassan, S.S., Thomas, T.L., and Moutsopoulos, H.M. (1978). Increased risk of lymphoma in sicca syndrome. *Annals of Internal Medicine*, **89**, 888–92.

Khurana, R.K., Koski, C.L., and Mayer, R.F. (1988). Autonomic dysfunction in Lambert-Eaton myasthenic syndrome. *Journal of Neurological Sciences*, **85**, 77–86.

Kingsmore, S.F., Hall, B.D., Allen, N.B., *et al*. (1992). Association of methotrexate, rheumatoid arthritis and lymphoma: report of 2 cases and literature review. *Journal of Rheumatology*, **19**, 1462–5.

Kinlen, L.J. (1985). Incidence of cancer in rheumatoid arthritis and other disorders after immunosuppressive treatment. *American Journal of Medicine*, **78** (Suppl. 1A), 44–9.

Kyle, R.A. (1975). Multiple myeloma: review of 869 cases. *Mayo Clinic Proceedings*, **50**, 29–40.

Kyle, R.A. (1993). 'Benign' monoclonal gamopathy - after 20 to 35 years of follow-up. *Mayo Clinic Proceedings*, **68**, 26–36.

Laakso, M., Mutru, O., Isomäki, H., *et al*. (1986). Cancer mortality in patients with rheumatoid arthritis. *Journal of Rheumatology*, **13**, 522–6.

Lakhanpal, S., Bunch, T.W., Melton, L.J., *et al*. (1986). Polymyositis-dermatomyositis and malignant lesions: does an association exist? *Mayo Clinic Proceedings*, **61**, 645–53.

Lakhanpal, S., Ginsburg, W.W., Michet, C.J., *et al*. (1988). Eosinophilic fasciitis: clinical spectrum and therapeutic response in 52 cases. *Seminars in Arthritis and Rheumatism*, **17**, 221–31.

Lambert, E.H. and Lennon, V.A. (1988). Selected IgG rapidly induces Lambert-Eaton myasthenic syndrome in mice: complement independence and EMG abnormalities. *Muscle and Nerve*, **11**, 1133–45.

Lee, P., Alderdice, C., Wilkinson, S., *et al*. (1983). Malignancy in progressive systemic sclerosis - association with breast carcinoma. *Journal of Rheumatology*, **10**, 665–6.

Leeson, M.C., Makley, J.T., and Carter, J.R. (1985). Metastatic skeletal disease in the pediatric population. *Journal of Pediatric Orthopedics*, **5**, 261–7.

Lenoch, F., Bremova, A., Kankova, D., *et al*. (1967). The relation of Sjögren's syndrome to rheumatoid arthritis. *Acta Rheumatologica Scandinavica*, **10**, 297–304.

LeRoux, B.T. (1968). Bronchial carcinoma with hypertrophic pulmonary osteoarthropathy. *South African Medical Journal*, **42**, 1074–5.

Lewis, R.B., Castor, C.W., Knisely, R.E., *et al*. (1976). Frequency of neoplasia in systemic lupus erythematosus and rheumatoid arthritis. *Arthritis and Rheumatism*, **19**, 1256–60.

Leys, K., Lang, B., Johnston, I., *et al*. (1991). Calcium channel autoantibodies in the Lambert-Eaton myasthenic syndrome. *Annals of Neurology*, **29**, 307–14.

Lyon, M.G., Bloch, D.A., Hollak, B., *et al*. (1989). Predisposing factors in polymyositis-dermatomyositis: results of a nationwide survey. *Journal of Rheumatology*, **16**, 1218–24.

Mackenzie, D.H. (1966). Synovial sarcoma. A review of 58 cases. *Cancer*, **19**, 169–80.

Manchul, L.A., Jin, A., Pritchard, K.I., *et al*. (1985). The frequency of malignant neoplasms in patients with polymyositis-dermatomyositis. *Archives of Internal Medicine*, **145**, 1835–9.

Maselli, R.A. (1994). Pathophysiology of myasthenia gravis and Lambert-Eaton syndrome. *Neurologic Clinics of North America*, **12**, 285–303.

Masi, A.T. and Hochberg, M.C. (1988). Temporal association of polymyositis-dermatomyositis with malignancy: methodologic and clinical considerations. *Mount Sinai Journal of Medicine*, **55**, 471–8.

McCurley, T.L., Collins, R.D., Ball, E., *et al*. (1990). Nodal and extranodal lymphoproliferative disorders in Sjögren's syndrome; a clinical and imunopathological study. *Human Pathology*, **21**, 482–92.

McEvoy, K.M. (1994). Diagnosis and treatment of Lambert-Eaton myasthenic syndrome. *Neurologic Clinics of North America*, **12**, 387–99.

McNeill, B.J. (1984). Value of bone scanning in neoplastic diseases. *Seminars in Nuclear Medicine*, **14**, 277–86.

Medsger, T.A., Dixon, J.A., and Garwood, V.F. (1982). Palmar fasciitis and polyarthritis associated with ovarian carcinoma. *Annals of Internal Medicine*, **96**, 424–31.

Mellemkjaer, L., Andersen, V., Linet, M.J., Gridley, G., Hoover, R., and Olen, J.H. (1997). Non-Hodgkin's lymphoma and other cancers among a cohort of patients with systemic lupus erythematosus. *Arthritis and Rheumatism*, **40**, 761–9.

Merrick, M.V. and Merrick, J.M. (1986). Bone scintigraphy in lung cancer; a reappraisal. *British Journal of Radiology*, **59**, 1185–94.

Michaels, R.M. and Sorber, J.A. (1984). Reflex sympathetic dystrophy, as a probable paraneoplastic syndrome: case report and literature review. *Arthritis and Rheumatism*, **27**, 1183–5.

Miettinen, M. and Virtanen, I. (1984). Synovial sarcoma - a misnomer. *American Journal of Pathology*, **117**, 18–25.

Miller, D.G. (1967). The association of immune disease and malignant lymphoma. *Annals of Internal Medicine*, **66**, 507–21.

Mitchell, D.M., Spitz, P.W., Young, D.Y., *et al*. (1986). Survival, prognosis, and causes of death in rheumatoid arthritis. *Athritis and Rheumatism*, **29**, 706–14.

Montanaro, M. and Bizzarri, F. (1992). Non-Hodgkin's lymphoma and subsequent acute lymphoblastic leukaemia in a patient with polymyalgia rheumatics. *British Journal of Rheumatology*, **31**, 277–8.

Moutsopoulos, H.M., Chused, T.M., Mann, D.L., *et al*. (1980). Sjögren's syndrome (sicca syndrome): current issues. *Annals of Internal Medicine*, **92**, 212–26.

Murashima, A., Takasaki, Y., Hashimoto, H., *et al*. (1992). A case of Raynaud's disease with uterine cancer producing interleukin-6. *Clinical Rheumatology*, **11**, 410–12.

Myllykangas-Luosujärvi, K., Aho, K., Kautiainen, H., *et al*. (1995). Shortening of life span and acuses of excess mortality in a population-based series of subjects with rheumatoid arthritis. *Clinical and Experimental Rheumatology*, **13**, 149–53.

Nagai, M., Sasaki, K., Tokuda, M., *et al*. (1993). Hodgkin's disease and Sjögren's syndrome. *European Journal of Haematology*, **50**, 180–2.

Naschitz, J.E., Yeshurun, D., Zuckerman, E., *et al*. (1994). Cancer-associated fasciitis-panniculitis. *Cancer*, **73**, 231–5.

Newman, M.K. and Gugino, R.J. (1964). Neuropathies and myopathies associated with occult malignancies. *Journal of the American Medical Association*, **190**, 575–7.

Newsom-Davis, J. and Murray, N.M. (1984). Plasma exchange and immunosuppressive drug treatment in the Lambert-Eaton myasthenic syndrome. *Neurology*, **34**, 480–5.

Newton, P., Freemont, A.T., Noble, J., *et al*. (1984). Secondary malignant synovitis: report of three cases and review of the literature. *Quarterly Journal of Medicine*, **53**, 135–43.

Nolling, H.E., Danielson, G.K., and Hamilton, R.W. (1963). Hypertrophic pulmonary osteoarthropathy. *Journal of Thoracic and Cardiovascular Surgery*, **46**, 310–21.

O'Neill, J.H., Murray, N.M., and Newsom-Davis, J. (1988). The Lambert-Eaton myasthenic syndrome: a review of 50 cases. *Brain*, **111**, 577–96.

Papatestas, A.E., Osserman, K.F., and Kark, A.E. (1971). The relationship between thymus and oncogenesis: a study of the incidence of non-thymic malignancy in myasthenia gravis. *British Journal of Cancer*, **25**, 635–45.

Pascuzzi, R.M. and Kim, Y.I. (1990). Lambert-Eaton syndrome. *Seminars in Neurology*, **10**, 35–41.

Patel, A., Davila, D.G., and Peters, S.G. (1993). Paraneoplastic syndromes associated with lung cancer. *Mayo Clinic Proceedings*, **68**, 278–87.

Paul, T., Katiyar, B.C., Misra, S., *et al*. (1978). Carcinomatous neuromuscular syndromes. *Brain*, **101**, 53–63.

Peris, P., Del Orme, J., Gratacos, J., *et al*. (1990). The Lambert-Eaton myasthenic syndrome in association with rheumatoid arthritis. *British Journal of Rheumatology*, **29**, 75–7.

Pettersson, T., Pukkala, E., Teppo, L., *et al*. (1992). Increased risk of cancer in patients with systemic lupus erythematosus. *Annals of the Rheumatic Diseases*, **51**, 437–9.

Pfinsgraff, J., Buckingham, R.B., Keister, S.R., *et al*. (1986). Palmar fasciitis and arthritis with malignant neoplasms: a paraneoplastic syndrome. *Seminars in Arthritis and Rheumatism*, **16**, 118–25.

Pines, A., Kaplinsky, N., Olchovsky, D., *et al*. (1984). Rheumatoid arthritis-like syndrome: a presenting symptom of malignancy. Report of 3 cases and review of the literature. *European Journal of Rheumatic Inflammation*, **7**, 51–5.

Plotz, P.H., Dalakas, M., Leff, R.L., *et al*. (1989). Current concepts in the idiopathic inflammatory myopathies: polymyositis, dermatomyositis, and related disorders. *Annals of Internal Medicine*, **111**, 143–57.

Porter, D., Madhok, O., and Vapell, H. (1991). Non-Hodgkin's lymphoma and rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **50**, 275–6.

Radis, C.D., Kahl, L.E., Baker, G.L., *et al*. (1995). Effects of cyclophosphamide on the development of malignancy and on long-term survival of patients with rheumatoid arthritis. *Arthritis and Rheumatism*, **38**, 1120–7.

Razis, D.V., Diamond, H.D., and Craver, L.F. (1959). Hodgkin's disease associated with other malignant tumors and certain non-neoplastic diseases. *American Journal of Medical Sciences*, **238**, 327–35.

Richardson, J.E., Gladman, D.D., Fam, A, *et al*. (1987). HLA-DR4 in giant cell ateritis: association with polymyalgia rheumatica syndrome. *Arthritis and Rheumatism*, **30**, 1293–7.

Rööser, B., Willen, H., Hugoson, A., *et al*. (1989). Prognostic factors in synovial sarcoma. *Cancer*, **63**, 2182–5.

Ropes, M.W. (1976). *Systemic lupus erythematosus*, p. 74. Harvard University Press, Cambridge, MA.

Rosenberg, L., Palmer, J.R., Zauber, A.G., *et al*. (1991). A hypothesis: non-steroidal anti-inflammatory drugs reduce the incidence of large-bowel cancer. *Journal of the National Cancer Institute*, **83**, 355–8.

Rosenow, E.C., Strimlan, C.V., and Muhm, R. (1977). Pleuropulmonary manifestations of ankylosing spondylitis. *Mayo Clinic Proceedings*, **52**, 641–9.

Roumm, A.D. and Medsger, T.A. (1985). Cancer and systemic sclerosis: an eoidemiologic study. *Arthritis and Rheumatism*, **28**, 1336–40.

Rutherford, R.B., Rhodes, B.A., and Wagner, H.N. (1969). The distribution of extremity blood flow before and after vagotomy in a patient with hypertrophic pulmonary osteoarthropathy. *Diseases of*

the Chest, **56**, 19–23.

Ryan, E.A. and Reiss, E. (1984). Oncogenous osteomalacia. *American Journal of Medicine*, **77**, 501–12.

Scelsi, R. and Pinelli, P. (1977). Subclinical myopathic findings in patients affected by malignant tumours. *Acta Neuropathologica*, **38**, 103–8.

Schaller, J. (1972). Arthritis as a presenting manifestation of malignancy in children. *Journal of Pediatrics*, **81**, 793–7.

Schapira, D., Ishak, O.B., Nachtigal, A., *et al*. (1995). Tumour-induced osteomalacia. *Seminars in Arthritis and Rheumatism*, **25**, 35–46.

Schmid, U., Helbron, D., and Lemert, K. (1982). Development of malignant lymphoma in myoepithelial sialadenitis (Sjögren's syndrome). *Virchow's Archives of Pathological Anatomy*, **395**, 11–43.

Schumacher, H.R. (1976). Articular manifestations of hypertrophic pulmonary osteoarthropathy. *Arthritis and Rheumatism*, **19**, 629–36.

Segal, A.M. and Mackenzie, A.H. (1982). Hypertrophic osteoarthropathy: a 10-year retrospective analysis. *Arthritis and Rheumatism*, **12**, 220–32.

Shneerson, J.M. (1981). Digital clubbing and hypertrophic osteoarthropathy: the underlying mechanisms. *British Journal of Diseases of the Chest*, **75**, 113–31.

Shy, G.M. and Silverstein, I. (1965). A study of the effects upon the motor unit by remote malignancy. *Brain*, **88**, 515–28.

Sigurgeirsson, B., Lindelof, B., Edhag, O., *et al*. (1992). Risk of cancer in patients with dermatomyositis or polymyositis. *New England Journal of Medicine*, **326**, 363–7.

Silman, A.J., Petrie, J., Hazelman, B., *et al*. (1988). Lymphoproliferative cancer and other malignancy in patients with rheumatoid arthritis treated with azathioprine: a 20 year follow up study. *Annals of the Rheumatic Diseases*, **47**, 988–92.

Silverstein, M.N. and Kelly, P.J. (1963). Leukemia with osteoarticular symptoms and signs. *Annals of Internal Medicine*, **59**, 637–45.

Simon, R.D. and Ford, L.E. (1980). Rheumatoid-like arthritis associated with a colonic carcinoma. *Archives of Internal Medicine*, **140**, 698–700.

Smith, P.G., Doll, R., and Radford, E.P. (1977). Cancer mortality among patients with ankylosing spondylitis not given X-ray therapy. *British Journal of Radiology*, **50**, 728–34.

Springfield, D.S. (1982). Mechanisms of metastasis. *Clinical Orthopedics and Related Research*, **169**, 15–19.

Staalman, C.R. and Umans, U. (1993). Hypertrophic osteoarthropathy in childhood malignancy. *Medical and Pediatric Oncology*, **21**, 676–9.

Steen, V.D. and Medsger, T.A. (1990). Epidemiology and natural history of systemic sclerosis. *Rheumatic Disease Clinics of North America*, **16**, 1–10.

Stenseth, J.H., Clagett, O.T., and Woolner, C.B. (1967). Hypertrophic pulmonary osteoarthropathy. *Diseases of the Chest*, **52**, 62–8.

Stern, R.S., Zierler, S., and Parish, J.A. (1982). Methotrexate used for psoriasis and the risk of non-cutaneous malignancy. *Cancer*, **50**, 869–72.

Sugai, S., Shimizu, S., Hirose, Y., *et al*. (1985). Monoclonal gammopathies in Japanese patients with Sjögren's syndrome. *Journal of Clinical Immunology*, **5**, 90–101.

Sulkes, A. and Naparstek, Y. (1991). The infrequent association of systemic lupus erythematosus and solid tumors. *Cancer*, **68**, 1389–93.

Sung, S.S., Sheibani, K., Fishleder, A., *et al*. (1991). Monocytoid B-cell lymphoma in patients with Sjögren's syndrome; a clinicopathologic study of 13 patients. *Human Pathology*, **22**, 422–30.

Sweeney, D.M., Manzi, S., Janosky, J., *et al*. (1995). Risk of malignancy in women with systemic lupus erythematosus. *Journal of Rheumatology*, **22**, 1478–82.

Symmons, D.P., Ahern, M., Bacon, P.A., *et al*. (1984). Lymphoproliferative malignancy in rheumatoid arthritis: a study of 20 cases. *Annals of the Rheumatic Diseases*, **43**, 132–5.

Talbott, J.H. and Barrocas, M. (1980). Carcinoma of the lung in progressive systemic sclerosis: a tabular review of the literature and a detailed report of the roentgenographic changes in two cases. *Seminars in Arthritis and Rheumatism*, **9**, 191–217.

Targoff, I.N. and Reichlin, M. (1988). Humoral immunity in polymyositis and dermatomyositis. *Mount Sinai Journal of Medicine*, **55**, 487–93.

Thompson, M. and Eadie, S. (1956). Keratoconjunctivitis sicca and rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **15**, 21–5.

Thun, M.J., Namboodiri, M.M., and Heath, C.W. (1991). Aspirin use and reduced risk of fatal colon cancer. *New England Journal of Medicine*, **325**, 1593–6.

Tillotson, J.F., McDonald, J.R., and Janes, J.M. (1951). Synovial sarcomata. *Journal of Bone and Joint Surgery*, **33A**, 459–73.

Tsuchiya, N., Sato, M., Uesaka, Y., *et al*. (1993). Lambert-Eaton myasthenic syndrome associated with Sjögren's syndrome and discoid lupus erythematosus. *Scandinavian Journal of Rheumatology*, **22**, 302–4.

Turner-Warwick, M., Lebowitz, M., Burrows, B., *et al*. (1980). Cryptogenic fibrosing alveolitis and lung cancer. *Thorax*, **35**, 496–9.

Tzioufas, A.G., Manoussakis, M.N., Costello, R., *et al*. (1986). Cryoglobulinemia in autoimmune rheumatic diseases. *Arthritis and Rheumatism*, **29**, 1098–104.

Uriburu, I.J., Moschio, F.J., and Marin, J.C. (1976). Metastases of carcinoma of the larynx and thyroid gland to the phalanges of the hand. *Journal of Bone and Joint Surgery*, **58A**, 134–5.

Vosskämper, M., Korf, B., Franke, F., *et al*. (1989). Paraneoplastic necrotizing myopathy: a rare disorder to be differentiated from polymyositis. *Journal of Neurology*, **236**, 489–92.

Walker, S.E. and Boles, G.G. (1973). Augmented incidence of neoplasia in NZB/NZW mice treated with cyclophosphamide. *Journal of Laboratory and Clinical Medicine*, **82**, 619–33.

Weidner, N. and Cruz, D. (1987). Phosphaturic mesenchymal tumours. *Cancer*, **59**, 1442–54.

Whaley, K., Williamson, J., Chisholm, D.M., *et al*. (1973). Sjögren's syndrome. *Quarterly Journal of Medicine*, **42**, 279–304.

Wolfe, F., Mitchell, D.M., Sibley, J.T., *et al*. (1994). The mortality of rheumatoid arthritis. *Arthritis and Rheumatism*, **37**, 481–94.

Yacoub, M.H. (1965). Relation between the histology of bronchial carcinoma and hypertrophic pulmonary osteoarthropathy. *Thorax*, **20**, 537–9.

Young, J.L. and Miller, R.W. (1975). Incidence of malignant tumors in U.S. children. *Journal of Pediatrics*, **86**, 254–8.

Zantos, D., Zhang, Y., and Felson, D. (1994). The overall and temporal association of cancer with polymyositis and dermatomyositis. *Journal of Rheumatology*, **21**, 1855–9.

Zawadzki, Z.A. and Benedek, T.G. (1969). Rheumatoid arthritis, dysproteinemic arthropathy, and paraproeinemia. *Arthritis and Rheumatism*, **12**, 555–68.

Zindrick, M.R., Young, M.P., Daley, R.J., *et al*. (1982). Metastatic tumors of the foot. *Clinical Orthopedics and Related Research*, **170**, 219–26.

Zuckner, J. (1990). Drug-induced myopathies. *Seminars in Arthritis and Rheumatism*, **19**, 259–68.

5.19.2 Algodystrophy (reflex sympathetic dystrophy syndrome)

Geoffrey O. Littlejohn

Introduction

Algodystrophy is an important pain syndrome characterized by variable dysfunction of musculoskeletal, skin, and vascular systems. The most characteristic feature is persistent pain, and this plus the associated disability, may have profound psychosocial effects. Moreover, the clinical features may be recalcitrant to medical intervention. Algodystrophy presents in a variety of clinical states and may be mild and short-lived or severe and prolonged. It is important to recognize the syndrome in the early and often incomplete phase if currently available therapeutic approaches are to be most effective.

Nomenclature

Algodystrophy is a chronic pain syndrome. Pain is the dominant symptom and, although the clinical features are characteristic and reproducible, the exact cause of the problem is unclear. Algodystrophy may occur in a variety of clinical conditions, affect different regions of the body to a greater or lesser extent, and the clinical manifestations vary around the central core features. As such there are a large number of descriptive names that are synonyms for algodystrophy. Reflex sympathetic dystrophy syndrome is the most appropriate alternative. Even so, this term is only used in a descriptive sense and does not imply specific underlying mechanisms (Janig *et al*. 1991). Table 1 summarizes some of the synonyms for algodystrophy. Some prefer to use the original nomenclature in situations where the precipitating cause is obvious: for instance, the syndrome is often called causalgia when it follows injury to a peripheral nerve, or Sudeck's atrophy where it follows wrist fracture. The term algodystrophy, derived from the Greek *algos* meaning pain and 'dystrophy' meaning a disorder related to poor nourishment, best encapsulates the principal clinical features and serves to remind us of the essential nature of the syndrome.

Reflex sympathetic dystrophy syndrome
Causalgia
Shoulder–hand syndrome
Sudeck's atrophy
Transient osteoporosis
Regional migratory osteoporosis
Post-traumatic painful osteoporosis
Complex regional pain syndrome

Table 1 Selected alternative terms for algodystrophy

The distinctive but varied clinical features of algodystrophy have provided classic clinical descriptions for over 100 years. Indeed, Hunter was one of the earliest to allude to this condition (Hunter 1843). By the middle of the nineteenth century, formal clinical descriptions were available in the context of traumatic lesions of peripheral nerves. Mitchell *et al*. (1864) described the deep burning pain of causalgia but also emphasized other features such as 'the skin affected in these cases was deep red or mottled, or red and pale in patches, ... the surface of all the affected part was glossy and shining as though it had been skilfully varnished ..., in some form, pain has been invariably attendant upon the disease state of skin which we have tried to describe. ... In the great mass of cases, it has been of that peculiar burning character of which we have spoken, ... in other instances, there was associated with this, acute or aching pain which extended beyond the disease tissues.' With the advent of radiology the clinical manifestations were greatly expanded as the breadth of the syndrome was recognized (Doury *et al*. 1981).

More recently the putative role of the sympathetic nervous system in this syndrome has been incorporated through use of the terms sympathetic-maintained (or mediated) pain and sympathetically independent pain (Roberts 1986). These subgroups are differentiated on the response of the pain to sympathetic blockade (Campbell *et al*. 1992). In a further development the term complex regional pain syndrome has been proposed, based entirely on clinical features (Merskey and Bogduk 1994)

Epidemiology

Algodystrophy is a common rheumatological disorder. It has been seen in all races and geographical regions. It affects both sexes and may occur at any age. Historically the published records have concentrated on the elderly patient who has sustained trauma, for instance a Colles' fracture, but over the last decade the occurrence of algodystrophy in children, particularly adolescent girls, has been well characterized (Silber and Majd 1988; Cicuttini and Littlejohn 1989; Sherry *et al*. 1991). In the adolescent, girls are affected far more than boys and in the adult, men seem to be affected more than women. The most frequent age group affected is between 40 and 60 years of age.

As with many rheumatic diseases the 'classical case' is easily characterized and recognized but occurs far less commonly than *formes frustes*. Minor variants tend to be less troublesome or may be regarded, in some instances, as part of the initiating cause. For example, minor algodystrophy following trauma may be regarded as a part of the normal response to injury. Minor forms may also occur in other painful, peripheral rheumatic disorders such as inflammatory arthritis and thus be hidden with the clinical features of the inflammation itself. A true estimate of the prevalence of the condition is thus extremely hard to provide.

In its classical form, 1 in 200 people presenting to a trauma unit will develop algodystrophy (Plewes 1956), but in other traumatic states up to 1 in 20 will develop it (Kozin 1986). Before the intensive mobilization of patients with myocardial ischaemia or hemiplegia became acceptable, between 5 and 20 per cent of such patients developed algodystrophy (Davis 1977). In a study of 109 unselected patients with Colles' fracture, 25 per cent had two or more features of algodystrophy at 9 weeks, while at 6 months 62 per cent showed some residual abnormalities (Atkins *et al*. 1989). It is likely that the prevalence of algodystrophy will always be underestimated

because of its minor variants. In order to identify mild or early forms of algodystrophy, it should be anticipated where there are triggering events known to be associated with the syndrome.

Clinical features

Like other chronic pain syndromes the chief complaint and that which draws the syndrome to the attention of the medical practitioner is pain. This not only manifests as spontaneous, often burning, pain but also as a pain amplification state. Here otherwise innocuous stimuli produce pain (allodynia) and there is increased pain perception to a given painful stimulus (hyperalgesia). Hyperpathia, where there is delayed overreaction particularly to a normal repetitive cutaneous stimulus, is also found. The pain will usually occur some days to weeks after a triggering factor, if present, and will generally be constant. Although burning pain is a characteristic of the syndrome, more commonly the pain is described as a deep, dull aching sensation. Paroxysms of pain may occur, lasting for seconds. The pain may disturb sleep, and even any slight movement that mechanically stimulates the sensitized region will produce worsening of the symptom. The patient may report the use of cold, wet compresses to relieve the discomfort.

Typically the syndrome principally involves the distal part of a limb; for instance, a hand and lower forearm region (Fig. 1) or a foot and lower leg. With time, in the majority, much less apparent but definitely abnormal clinical features occur in the opposite limb. The syndrome may involve one digit or rarely all four limbs. Shortly after the onset of pain, swelling, the second most common feature, occurs (Blumberg *et al*. 1994). This may be intermittent in the early stages and is usually associated with a change in the texture of the overlying skin, producing the reticular or lividoid appearance over the involved part (Fig. 2). Palmar erythema may be noted. Early in the course the involved region is warmer than the surrounding region. Varying degrees of sweating may occur but are not essential to the diagnosis. There may be piloerection.

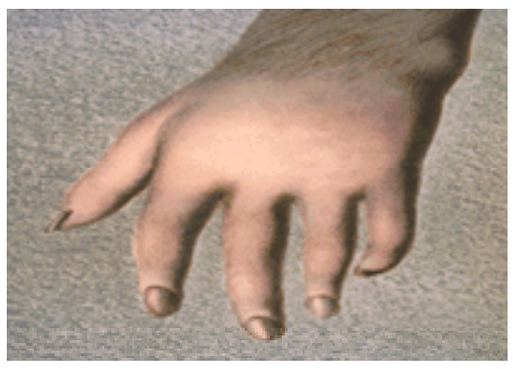


Fig. 1 Reflex sympathetic dystrophy syndrome of the left hand in a 21-year-old man after trauma, showing diffuse swelling, cutaneous blood-flow change, and dystrophic skin changes. One of the earliest reported pictures of this condition (from Otis (1877), by courtesy of Dr R.L. Travers).

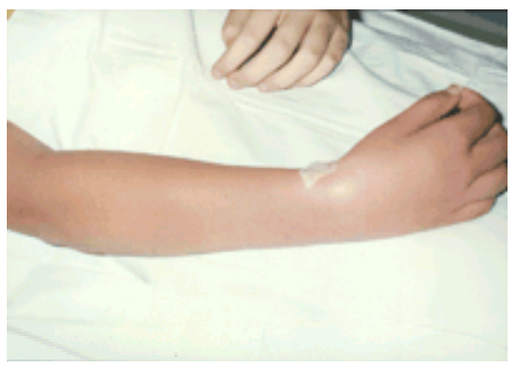


Fig. 2 Diffuse swelling of lower right forearm and hand in a 13-year-old girl with reflex sympathetic dystrophy syndrome. There is extreme hyperalgesia. Note flexed wrist and hand posture and skin discoloration. (By courtesy of Dr R. Allan, Royal Children's Hospital, Melbourne, Australia.)

Examination confirms the presence of a decreased pain threshold, particularly to mechanical stimuli, leading to abnormal tenderness as the principal clinical finding. Early in the course the hyperalgesia tends to be periarticular, but later regions well away from joints and particularly those over bone also become exquisitely tender. This distribution is non-(neuro-) anatomical and found over a wide region of the involved parts. For instance, if the principal involvement is in the mid- tarsal region of the foot, one would usually find tenderness on palpation of a metatarsophalangeal joint and regions of the lower leg, as well as in intervening areas. Mild oedema may be noted, either pitting or not. Early on, joint movement in the region will be restricted by pain but with care it be demonstrated that the range of motion is often near normal, despite the initial impression (Fig. 3). There may be a prominent motor component with tremor, weakness, or muscle tightness.

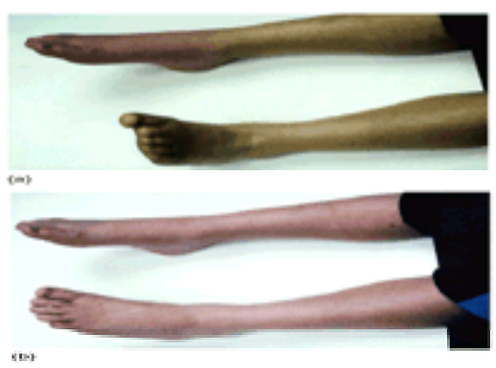


Fig. 3 Right lower leg and foot in a 15-year-old female patient with reflex dystrophy syndrome showing skin discoloration, regional swelling, and inability to extend the right ankle (a) and the adoption of a plantar flexion (b) assumed in order to reduce pain. The lower third of the leg, including the foot was extremely hyperalgesic. (By courtesy of Dr R. Allan, Royal Children's Hospital, Melbourne, Australia.)

Triggering events

A number of well-characterized events may precede algodystrophy. It is best to regard these events as triggers rather than as causative because of the wide variety of different associations. Table 2 summarizes the important disorders associated with algodystrophy.

	Approximate frequency of association (%)
Trauma	50
Idiopathic	25
Disorders of the nervous system	
Central conditions	
Painful peripheral conditions	25
Medication	
Others	

Table 2 Selected disorders associated with algodystrophy

In most series, depending on the criteria for classifying algodystrophy, trauma will be identified as a trigger in around half of subjects (Doury *et al*. 1981; Kozin 1986). In many instances it is likely that a prior traumatic episode will have been recalled by the patient after the development of algodystrophy rather than the event being a true trigger. However, major and even minor, apparently trivial, traumas, particularly to a distal part of a limb, have been clearly linked to the development of the syndrome. In about one-half of cases the trauma will have led to bone fracture and indeed many accounts of the syndrome come from observations made on peripheral fractures, such as Colles' fracture at the wrist. Other important traumatic triggers include simple sprains, strains, contusions, or jarring injuries, particularly to the distal part of the limbs. Surgical procedures, particularly of limb joints, such as arthroscopy, can induce algodystrophy (Small 1993). Other triggering traumas include peripheral burns and frost-bite. Trauma to articular more than diaphyseal regions predisposes to this syndrome (Fontaine *et al*. 1957). Although trauma is a common trigger, in the context of the large number of traumatic events found in any community, the subsequent development of algodystrophy affects less than 1 per cent of all such events. Perhaps a common thread among injuries that result in algodystrophy is the individual's reaction to pain and the frightening nature of the injury, be it major or apparently trivial.

A variety of disorders of the nervous system may act as triggers. These include conditions resulting in hemiplegia and other central diseases, including cerebral tumour, meningitis, or syringomyelia, among others. Peripheral nerve injury is the classical cause of causalgia but other peripheral nerve disturbances such as those caused by *herpes zoster*, nerve-root impingement, or peripheral neuropathy may result in this syndrome. Pain arising from a visceral or deep somatic structure, such as myocardial ischaemia in the former instance and mechanical pain from the cervical spine in the latter, may be associated. Pain arising from a peripheral origin, such as that seen in inflammatory arthritis or after deep venous thrombosis, may provoke the syndrome. Certain medications, particularly barbiturates and isoniazid, have been reported as triggers.

There are a variety of other conditions that may be linked with algodystrophy. These include pregnancy (with particular involvement of the hip), metastatic tumours, acrodermatitis continua, and prolonged immobilization of a peripheral limb (Schwartzman and McLellan 1987). The suggestion that there is an increased prevalence of diabetes and hypertriglyceridaemia in this group of patients (Amor *et al*. 1980) has been refuted by others (Eurly 1992).

At least 25 per cent of cases have no easily identifiable trigger. Here the search for central factors or psychological susceptibility is often made. In children it is common to find an unresolved stress or an unsatisfactory psychosocial state in the background (Sherry and Weisman 1988). In adults, this is more difficult to identify clearly. A small percentage of patients developing algodystrophy will have a defined psychiatric condition, such as major depression, or a chronic anxiety state, but many more are likely to have abnormal reactions to psychosocial stresses. These factors may be difficult to identify accurately but there is often a strong clinical impression that they are playing an important part in the triggering or expression of the syndrome. However, in many patients, no such factors can be identified. There is no definite evidence that particular personality traits predispose a person to develop algodystrophy (Bruehl and Carlson 1992; Lynch 1992).

Course

In most instances the clinical features persist, fluctuate, or gradually resolve according to the natural history or treatment intervention, without further sequelae. In some patients there is a dystrophic phase, usually some months after the start of the problem. Here the limb becomes cool, and cutaneous pallor or cyanosis replace the previous erythema. There is decreased growth of dermal appendages, such as hair, and the nails become brittle. Skin and subcutaneous tissues may atrophy. Increased sweating becomes more prominent and the pain persists and often worsens. There may be dysaesthesia (painful, abnormal skin sensation) in a non-anatomical distribution. Radiological evaluation will show change in the mineral content of bone in the area. Joints become tighter, this time due to contracture of surrounding structures. After several months there may be an atrophic phase. Pain usually decreases at this stage but can remain intractable in some. There may be marked atrophy of subcutaneous tissue, such as to cause tapering of digits, and flexion contractures of peripheral joints become prominent. The limb is characterized by vasoconstriction and coolness and the skin is cyanotic, smooth, and glossy. Hair growth can be increased or decreased.

While the clinical features of algodystrophy are easy to recognize, there are, as mentioned above, more patients who have variations or incomplete forms than have the typical condition. Variants are particularly common in children where the painful limb is usually cool and slightly swollen from the outset and the syndrome does not tend to go through the classical phases, often resolving much more quickly than in an adult.

Traditionally the disorder has been staged according to observations by Steinbrocker and Argyros (1958), where stage 1 comprises the 'acute' clinical features with pain, tenderness, swelling, and vasomotor and pseudomotor changes predominating. With time, stage 2 will supervene and this is dominated by dystrophic features and will last for several months. Stage 3 comprises atrophic changes and may be long lasting. In extreme cases this classification is of use but in the average situation most patients do not progress beyond stage 1 or early stage 2. Perhaps this is because of early intervention or perhaps we are now diagnosing the condition in its milder and incomplete forms more often. Many patients also have features derived from different components, and investigations and clinical features often do not correlate, irrespective of the chosen staging. The success of treatment does not depend on the patient's clinical stage except that profound stage 3 changes are difficult to reverse.

Classification

There is no validated classification for diagnosis in the clinical setting or for research purposes because of the polymorphic nature of algodystrophy. The classification of Kozin *et al*. (1981) is often very useful clinically. Here, definite algodystrophy is present when there is pain and tenderness in an extremity, swelling of that extremity, and vasomotor or pseudomotor changes. Dystrophic change may also be present. Probable algodystrophy consists of pain and tenderness, together with vasomotor and pseudomotor changes or swelling; possible algodystrophy comprises vasomotor or pseudomotor changes; and doubtful algodystrophy is considered where there is pain and tenderness out of keeping with any preceding injury or other organic disease process.

Regions involved

Several common areas of involvement have been defined (Doury *et al*. 1981). The arm may be involved in a unipolar fashion, with either shoulder (frozen shoulder) or hand being affected. Involvement of the shoulder is associated with marked retraction of the joint capsule, restricted range of motion, and pain persisting for several months. On resolution of the syndrome, loss of range of motion of the shoulder is common, although usually not clinically significant. Peripheral involvement may include a single digit (Laukaitis *et al*. 1989), a few metacarpal rays (Lequesne *et al*. 1977), or more typically the whole hand and lower forearm may be involved. Bipolar forms include the shoulder–hand syndrome. Up to 25 per cent of upper-limb involvement is bilateral, the changes in the opposite side often being more evident after investigation (Kozin *et al*. 1976).

Leg involvement, although more common, is less dramatic than that of the arm.. Here the features are usually confined to a part of the limb such as the foot, knee, or hip. Bipolar and bilateral involvement are less common than in the arm. Involvement of the knee, either of the whole or of part, including the patella, has only more recently been characterized (Tietjen 1986; Coughlan *et al*. 1987).

Algodystrophy of the hip is less commonly recognized because the hip joint is deeply positioned and there is a lack of the cutaneous features that are more common in peripheral and superficial sites. The duration of symptoms in this region is shorter, possibly because early mobilization is easier. Some (for example, Lakhanpal *et*

al. 1987) suggest that this variant is distinct enough to be termed transient regional osteoporosis. Absence of precipitating trauma, good outcome, and a propensity for recurrent episodes and involvement of multiple regions is characteristic of this variant. Signs of cutaneous or vascular change are uncommon. Algodystrophy may also affect the spine and occasionally other areas of the skeleton.

Investigations

Laboratory investigations

There is no abnormality of acute-phase reactants nor of standard biochemical markers in algodystrophy. In some patients with early disease, 24-h urinary hydroxyproline excretion may be increased, reflecting bone demineralization (Doury 1988).

Imaging

The principal contribution to diagnosis remains clinical. However, a number of radiological and nuclear medicine techniques may show characteristic abnormalities. As there is no specific 'gold standard' for diagnostic criteria it is difficult to establish the clinical usefulness of these investigations. The main problem arises from the many variations of the syndrome, and most studies have considered only severe algodystrophy in evaluating investigations, thus producing a distorted view of their clinical utility. Table 3 summarizes selected investigations.

Investigation	Changes	Usefulness
Plain radiography	Patchy osteopenia early; diffuse late	Moderate predictive value
Fine-detail radiography	Diffuse and periarticular osteopenia common	80% abnormal
	Cortical bone resorption	
	Juxta-articular/subchondral erosion	
Dual-photon absorptiometry	Loss of up to 30% of bone mass	Unknown
Dual-X-ray absorptiometry	Bone loss shown	Unknown
Magnetic resonance imaging	Regional/diffuse bone loss	Unknown, probably low
Scintigraphy	Abnormal in early, pool, or late phases	High specificity
Thermography	Change in cutaneous temperature	Unknown

Table 3 Investigations in algodystrophy

Sudeck (1900) first showed the value of radiology in this condition. Plain radiographic changes reflect bone demineralization; these may take several weeks to months to appear after the onset of the syndrome. These changes persist longer than other clinical features on resolution. Some patients never have radiological changes and others have only subtle changes. It is essential that good technique and meticulous examination of radiographs, with comparison of the affected side to the other side, be performed. Characteristically children have far less radiological changes than adults, with only the minority showing demineralization. Patchy osteopenia may be seen in the early stages and diffuse changes characterize the later stages. Typically in the hip or knee, changes are most obvious in the subchondral region where loss of subchondral bone and preservation of subchondral plate will produce a significant finding. Severe juxta-articular changes might include erosions. However, the joint space is never affected. The sensitivity and specificity of radiological appearances are around 70 per cent for well-defined algodystrophy (Kozin *et al*. 1981). Plain radiography is an essential investigation but has only moderate predictive value for diagnosis. Fine-detailed radiography better shows the demineralization, which is accentuated around epiphyseal regions, reflecting hyperaemia at this site. Subperiosteal, endosteal, and intracortical bone resorption together with juxta-articular and subchondral erosions, are observed frequently with this technique (Genant *et al*. 1975; Griffiths and Virtama 1988).

Magnetic resonance imaging (MRI) may also show changes characteristic of bone loss (Fig. 4). For instance, in the hip, low intensity signal on T_1-weighted images and high intensity on T_2-weighted images may be seen. The clinical usefulness of this technique and absorptiometry, which also may reflect bone loss, is yet to be established.

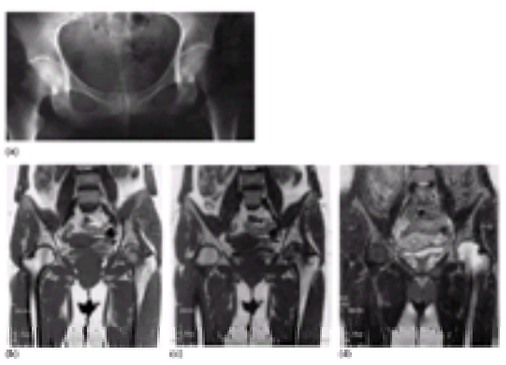


Fig. 4 (a)–(d) A 31-year-old female patient with transient regional osteoporosis of the left hip. Plain films (a) indicate regional osteoporosis of the left femoral head. MRI scan (b,c,d) shows changes of diffuse osteopenia in the left femoral head, compared with normal changes on the right side. (By courtesy of Dr John Stuckey, Victorian Imaging Group, Melbourne, Australia.)

Scintigraphic studies using technetium have a high specificity but a similar sensitivity to that of plain radiographs (Fig. 5). Typically a three-phase study (Kozin *et al*. 1976) is made, with the early phase showing regional blood flow over the first 2 to 3 min, the second phase showing the blood-pool image, and the third phase (some 2 to 4 h later) showing the standard bone-uptake findings. While any of these three phases may be abnormal it is more common to find increased flow (early phase) and uptake (late phase) in around 80 per cent of abnormal studies. Diminished flow and uptake are more commonly seen in children and adolescents than in adults (Silber and Majd 1988; Goldsmith *et al*. 1989). Scintigraphy appears more accurate if symptoms have been present for less than 6 months or if the patient is older than 50 years (Werner *et al*. 1989). Multiple investigations are often needed (Fig. 6).

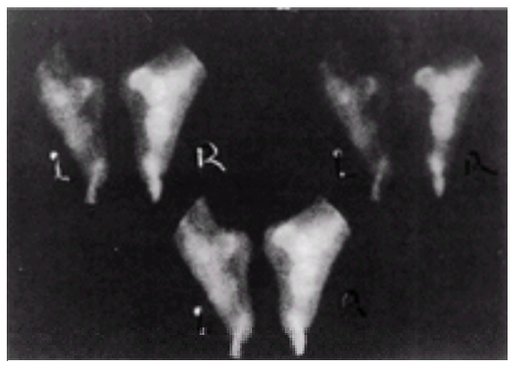


Fig. 5 A 28-year-old female patient with diffuse increase in isotopic uptake of right foot and lower end of right tibia, in a foot which was diffusely swollen and extremely hyperalgesic. The reflex sympathetic dystrophy in this case was associated with Reiter's disease initially affecting the right ankle. Inflammatory arthritis, trauma, or

other pain stimuli may associate with the onset of reflex sympathetic dystrophy.

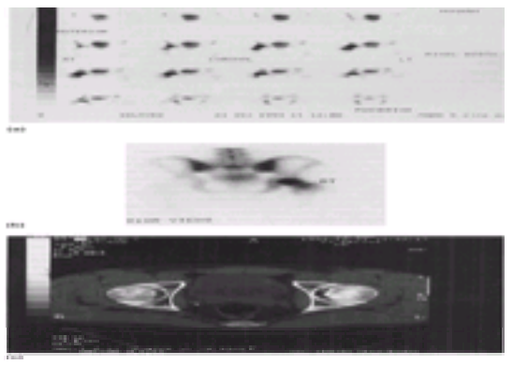


Fig. 6 A 39-year-old male with 4 months of pain from right hip. The nuclear scan (a and b) with spectroscopy reveals marked increase in isotopic uptake in the right femoral head and neck. The CT scan (c) shows osteopenia. Femoral head biopsy showed osteopenia of the neck and head of femur. The diagnosis was transient regional osteoporosis of the femoral head. (By courtesy of Dr John Stuckey, Victorian Imaging Group, Melbourne, Australia.)

Thermography may show changes in regional cutaneous temperatures compared with the unaffected side (Perelman *et al*. 1987). Typically there is an early increased or, particularly in children, decreased cutaneous heat emission. Critical evaluation of this technique is required before it can be usefully applied clinically (Fig. 7).

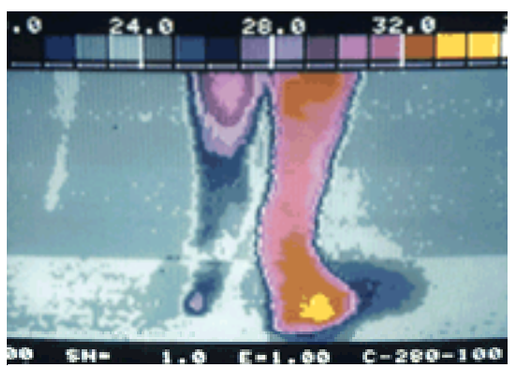


Fig. 7 Thermogram indicating lowered skin temperature (blue colours, compared with normal temperatures pink-red) of right leg and foot in a 15-year-old female patient with reflex dystrophy syndrome. (By courtesy of Dr R. Allan, Royal Children's Hospital, Melbourne, Australia.)

Histopathological changes

Synovial biopsies of joints in the region of algodystrophy will often show low-grade synovitis characterized by proliferation of synovial cells and small blood vessels, together with a mild chronic perivascular infiltrate. There have only been limited studies on the bone demineralization, periarticular fibrosis, and dermal atrophy (Arlet *et al*. 1978; Doury *et al*. 1981).

Diagnosis

The diagnosis of algodystrophy is based on the clinical features. A high degree of clinical suspicion is required, particularly in early diagnosis. Focal disease may mimic inflammatory or infectious arthritis or bone disease, trauma, or osteonecrosis. Characteristic clinical features, elimination of other causes for such symptoms, and appropriate imaging will usually allow for a clinically robust diagnosis. The most important thing in diagnosis is to think of the possibility of algodystrophy.

Pathophysiology

Algodystrophy is caused by a regional change in function of the pain system. This involves both peripheral and central mechanisms in a complex manner. The cause for this pain syndrome is poorly understood. However, many of the clinical features reflect change in function of various components of the peripheral nervous system. Many fibre types with different functions appear to contribute in varying degrees to the syndrome.

Table 4 summarizes the principal clinical characteristics of the syndrome and possible causative mechanisms. Input from two afferent fibre types, the small diameter, non-myelinated C-fibres and the large, myelinated A-b-fibres, together with change in function of sympathetic efferents, appear the likely mediators of many of the peripheral features. Sympathetic activity, through release of noradrenaline and the inflammatory prostaglandins, may sensitize peripheral nociceptors, thus decreasing threshold to peripheral mechanical and chemical stimuli. This may result in the hyperalgesia. It has been suggested that a $_1$-adrenergic receptors may become expressed on such nociceptors, particularly in a post-injury situation, and that these receptors will respond to sympathetic fibre release of noradrenaline (Campbell *et al*. 1988). Subsequent release of proinflammatory neuropeptides, such as substance P, by activated C-fibres would contribute to regional neurogenic inflammation with increase in blood flow and oedema. The associated synovitis may be due to this mechanism. Levine *et al*. (1984) have demonstrated that substance P contributes to experimental synovitis and this system is further modulated by alteration of sympathetic input. Periarticular osteoporosis may relate to the hyperaemia in the epiphyseal vessels in particular, or to neuropeptide effects on bone mineral metabolism. Allodynia is likely to reflect abnormality within the large myelinated afferent fibre system, as blocking these fibres by local anaesthetics will abolish this finding in other situations (Meyer *et al*. 1972; Roberts 1986).

TABLE cellSpacing=0 cellPadding=0 align=left border=0 hspace="10" vspace="5">

Characteristic	Mechanism
Pain	
Spontaneous	Sensitized peripheral/central nociceptors
Allodynia	Large myelinated fibre input to sensitized dorsal horn transmission neurones
Hyperalgesia	C-fibre sensitization in periphery+sensitized 2nd order neurones
Movement pain	Large myelinated fibre input to sensitized dorsal horn transmission neurones
Swelling	Neuropeptides from C-fibres; sympathetic effects on post-capillary venules
Bone change	Hyperaemia secondary to neuropeptide release; sympathetic effects
Synovitis	Neuropeptides
Dystrophy	Unknown neurological mechanism; mechanical effects

Table 4 Pathophysiological mechanisms in algodystrophy

In sophisticated tests, sympathetic efferents can be shown to have abnormal activity in the majority of patients with this syndrome. Many patients, though not all, do respond to sympathetic blockade, highlighting the role of sympathetic activity. The interaction between the products of the sympathetic nervous system and the peripheral nociceptors is also highlighted by the knowledge that injection of noradrenaline will exacerbate pain and this is blocked by intravenous administration of the a_1-adrenergic antagonist, phentolamine. The phentolamine test can be used diagnostically to help predict patients who have a large sympathetically maintained component to their syndrome and those who are likely to benefit from sympathetic blockade (Arner 1991; Campbell *et al*. 1992).

The cause for these functional changes in the peripheral pain system is still unclear. Livingstone (1943) suggested that there was a persisting peripheral afferent stimulus which led to abnormal activation of internuncial neurones situated in the dorsal horn at the relevant level. Efferent sympathetic activity was felt to follow that sequence. The realization that the majority of patients with algodystrophy do not have an identifiable persisting afferent stimulus in the region of pain, has led to an appreciation that a functional change within the dorsal horn itself is the most important cause for the syndrome. This functional dorsal horn abnormality is usually triggered by minor trauma, often in an emotional context. This was accounted for by Lankford and Thompson's theory (Lankford and Thompson 1977).

Roberts (1986) has suggested that initial activation of the unmyelinated C-fibre nociceptors, through unknown mechanisms, leads to sensitization of wide dynamic range neurones in the dorsal horn. Such neurones receive not only nociceptive input but input from other sources, including low-threshold, large myelinated afferents that otherwise serve functions such as proprioception. Input from these fibres resulting, say, from change in joint position or movement will impinge on the sensitized neurone with wide dynamic range in the dorsal horn and this input will be perceived as pain (allodynia). Central sensitization mechanisms mediated by excitatory amino acids and their interaction with the *N*-methyl-D-aspartate receptor, as well as the neuropeptides, are relevant to this process (Schwartzman 1993).

With activation of the system, automatic reflex changes at a segmental level lead to both muscle change and stimulation of the sympathetic system. The former may result in various neuromuscular features that can accompany algodystrophy, such as dystonia or other movement disorders (Schwartzman and Kerrigan 1990). The latter will affect functions in peripheral tissues and may result in a number of the dystrophic features of the syndrome. In addition, the cycle is completed whereby sympathetic efferent activity maintains peripheral fibre sensitization, which in turn feeds back into the dorsal horn system initiating further sympathetic activity. Such neurogenic reflexes are usually bilaterally represented, explaining the bilateral findings in many patients with algodystrophy.

The reason for change in function of these interactions in the dorsal horn is unclear. It seems naïve to explain such intense pain syndromes on purely segmental and peripheral pathophysiological changes. The complex and hierarchical nature of the pain system includes essential connections to the higher centres including the cortex. Important descending influences from higher centres impinge on the dorsal horn and modulate many components of the pain system. Thus central events which might include a variety of cortical and psychological factors may further influence this system. These might include the stress of pain itself and its effects on the central nervous system and the hypothalamic–pituitary axis, the patient's beliefs, mood, emotions, and even their quality of sleep. A change in any of these factors may affect the homeostatic mechanisms involved in pain control and sensitization phenomena in the dorsal horn. It is through such mechanisms that the emotional and affective components of the painful stimulus modify the peripheral reception and processing of further information relating to pain (Harvey 1987).

Algodystrophy shares many features with fibromyalgia syndrome, another common chronic pain syndrome (Bengtsson and Bengtsson 1988; Vaeroy *et al*. 1989). When this syndrome is localized a significant component of patients' symptoms relate to sympathetic-maintained pain (Fig. 8). Other features shared between the two syndromes are allodynia, hyperalgesia, and dermatographia. Table 5 compares selected features of these syndromes.

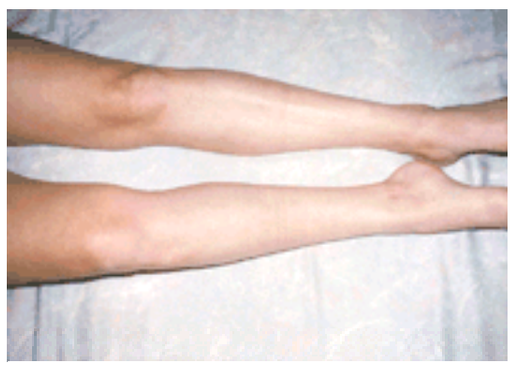


Fig. 8 A 28-year-old female with reflex sympathetic dystrophy syndrome of right lower leg from knee to foot. This area was extremely hyperalgesic compared with the left side. Signs of overt swelling, colour change, or other change 'sympathetic' features were lacking, as can be the case. This patient later developed a regional pain syndrome involving low back and all of the right leg. This illustrates the spectrum of reflex sympathetic dystrophy syndrome in that not all patients have the classic features.

Syndrome	Nociceptor	Mechanoreceptor	Sympathetic nervous system	Motor
Algodystrophy	++	++	+++	+
Fibromyalgia	++	+	+	+
Dystonia				+++

+ to +++ indicates range of qualitative changes.

Table 5 Abnormalities in different pain syndromes and dystonia

Other factors that may be important in the pathophysiology have been variously reviewed (Doury *et al*. 1981; Carlson and Jacobs 1986; Escobar 1986; Fields 1987; Kozin 1994; Janig 1996).

Management

Algodystrophy must been seen as a pain syndrome if an effective treatment programme is to be provided. This involves treating the whole patient and not just the obviously abnormal area of complaint in the periphery.

Preventive strategies include the recognition of situations that are likely to provoke the syndrome. Thus early mobilization after myocardial infarction, cerebrovascular accident, hand surgery, or mild peripheral injury is essential. Appropriate reassurance and direction in the handling of patients in any post-traumatic setting, particularly in emotionally charged, work-related events is essential.

Many authorities indicate that prognosis for full recovery relates inversely to the duration of symptoms before the onset of treatment; however, others feel that the syndrome can reverse, to a large extent, at any time (Kozin *et al*. 1987).

In milder forms of algodystrophy the principles of management (Table 6) include adequate pain relief, reassurance, and explanation. A positive approach in regard to outcome is appropriate and necessary. Careful explanation is required to ensure that the holistic concept of a chronic pain state is understood by the patient, according to their level of understanding. This is essential; it leads to a co-operative management plan that has the patient as the key person in the team. This

syndrome is not well served by the adoption of the 'injury' model where a powerful external 'treater' is required to give the 'curative' treatment. Such well-intentioned approaches often lead to a more dense ingraining of the syndrome, probably through blocking of positive benefits that would come from modulation of the central pathways on to the dorsal horn. Similarly, certain systems of compensation and other medicolegal events appear to inhibit resolution of the syndrome.

Accurate, early diagnosis
Explanation, reassurance
Adequate analgesia (medication, transcutaneous electrical stimulation)
Prevention of dystrophic change—exercise, movement
Attention to sleep disturbance
Attend to pain behaviour, psychosocial stresses, including compensation, litigation issues
Transient interruption of sympathetic activity
Skilled pain management/psychological counselling
Other

Table 6 Principles of management of algodystrophy

Counselling the patient, attending to the associated anxiety, and correcting any sleep disturbance, often through the use of low-dose tricyclic medication, are beneficial and prevent further amplification of the syndrome.

Useful treatments in milder algodystrophy, particularly in children, are exercise programmes which might include hydrotherapy. Activation of muscles and mechanoreceptors probably inhibits the activated dorsal-horn pain system on a local, segmental basis.

Numerous other treatment programmes have been suggested for the more severe or persistent types of algodystrophy. Some find systemic corticosteroids to be beneficial, even when the syndrome has been long-lasting (Kozin *et al*. 1981; Christensen *et al*. 1982). A typical course might commence at around 50 mg of prednisolone a day in divided doses and decrease to zero over 3 to 4 weeks. Subsequent courses may be needed. Some have advocated calcitonin, for example, 100 to 160 IU/day over 10 to 14 days, which has been shown to have benefits over placebo (Gobelet *et al*. 1986; Gobelet *et al*. 1992).

Interruption of the sympathetic efferent system to the region through various techniques has proved beneficial to many patients. Typically, a regional sympathetic ganglion block is made, using up to five short-acting sympathetic blocks, on a daily or alternate-day basis. If pain relief results, the patient is started on a more vigorous physiotherapy or exercise programme. Such approaches may result in 40 per cent of patients having a better result after 3-years follow-up; however, the majority have a less useful outcome (Wang *et al*. 1985).

Regional sympathetic blocks compare favourably with ganglion blockade (Bonnelli *et al*. 1983). These may be achieved using the Bier technique, whereby occlusion of venous outflow to the region, usually the lower arm or leg, is followed by intravenous installation of guanethidine or another similar agent. The exact mechanism of action of sympathetic modulation is unclear. Peripheral nerve blocks in the axilla will also modify afferent nerve transmission and may be useful. Again these procedures often need to be repeated to achieve sustained results (Schwartzman and McLellan 1987).

Surgical sympathectomy is only done if there is definite, but short-lived, improvement with the previous procedures.

Many other approaches have been used; these range from epidural opioids to cutaneous clonidine patches. Controlled studies are lacking for most of these approaches (Dotson 1993). The large variety of therapies emphasizes the difficulty in treating established algodystrophy. My approach to management is as follows:

1. *Young patient, any severity, any duration*:

   Always

   a. accurate diagnosis, explanation, reassurance as to expected good outcome;
   b. identification and management of psychosocial stressors especially family, school, or other;
   c. establishment of a regular activity programme—walking, bicycle riding, swimming, etc.
   d. resumption of all previous activities, especially those related to school;
   e. careful supervision of programme and close liaison with parents, especially the mother if a female patient.

   Often

   a. transcutaneous nerve stimulation to help pain control and allow entry into activity programme;
   b. hydrotherapy if limb sensitivity to movement is extreme. An empathetic physiotherapist is often the most powerful tool in this programme.

   Sometimes

   a. tricyclic medication to correct sleep disturbance and increase the pain threshold;
   b. a clinical psychologist to help manage pain and consider psychosocial background dynamics;
   c. regional temporary sympathetic/ganglion blocks;
   d. other approaches.
2. Adult patient, mild severity, and duration:

   Always

   a. accurate diagnosis, explanation, reassurance as to probable good outcome;
   b. identification and management of any obvious psychosocial stressors including work-related or legal issues;
   c. establishment of regular activity programme using involved limb;
   d. planned return to all previous activities currently abandoned due to pain;
   e. careful supervision of the programme with attention to the 'blocks' to progress that often occur.

   Often

   a. supervision of the physical programme by a physiotherapist with use of transcutaneous nerve stimulation, hydrotherapy, and similar tactics to control pain and encourage activity;
   b. short-term use of tricyclic medication.

   Sometimes

   a. pain-management counselling by clinical psychologist or similar skilled person;
   b. regional temporary sympathetic/ganglion blocks;
   c. other approaches, such as short-course corticosteroids.
3. Adult patient, moderate/severe, prolonged duration:

Always

a. use all the above approaches, sometimes sequentially and sometimes concomitantly;
b. put emphasis on pain-management counselling early in programme, with frequent use of interventions to block efferent sympathetic outflow, and/or corticosteroids;
c. stay with the patient—suboptimal therapeutic responses require longer-term support.

Prognosis

The outcome for patients with algodystrophy is quite varied. Minor forms seem to have an excellent outcome and children with the condition are expected to regain normal function and lose their pain. Some adults with severe forms of the condition may have protracted chronic pain states, significant dystrophic tissue change and longer-term disability. Pain usually eases over time even in those patients.

Summary

Algodystrophy is a common and significant musculoskeletal condition that results from a complex neurophysiological change in the pain transmission pathways. The syndrome is characteristic and reproducible but many variants occur clinically. The best treatment is through early recognition and management along principles used in other chronic pain syndromes, but with particular attention to the characteristic increase in sympathetic-maintained pain.

Chapter References

Amor, B. G. A. de, Saporta, L., Abergel, S., and Delbarre, F. (1980). Algodystrophies et hyperlipidemies. *Revue de Rhumatisme et des Maladies Osteoarticulaires*, **47**, 353–8.

Arlet, J., Ficat, P., Durroux, R., Theallie, J.P., Mazieres, B., and Bouteiller, G. (1978). Histopathologie des lesions osseuses dans 9 cas d'algodystrophie de la hanche. *Revue de Rhumatisme et des Maladies Osteoarticulaires*, **45**, 691–8.

Arner, S. (1991). Intravenous phentolamine test: diagnostic and prognostic use in reflex sympathetic dystrophy. *Pain*, **46**, 17–22.

Atkins, R.M., Duckworth, T., and Kanis, J.A. (1989). Algodystrophy following Colles' fracture. *Journal of Hand Surgery*, **14**, 161–4.

Bengtsson, A. and Bengtsson, M. (1988). Regional sympathetic blockade in primary fibromyalgia. *Pain*, **33**, 161–7.

Blumberg, H., Hoffman, U., Mohadjer, M., and Scheremet, R. (1994). Clinical phenomenology and mechanisms of reflex sympathetic dystrophy: emphasis on edema. In *Proceedings of the 7th World Congress on Pain. Progress in Pain Research and Management*, Vol. 2, (ed. G.F. Gebhart, D.L. Hammond, and T.S. Jensen), pp. 455–81. IASP Press, Seattle.

Bonnelli, S., Conoscente, F., and Movilia, P.G. (1983). Regional intravenous guanethidine vs stellate ganglion block in reflex sympathetic dystrophies; a randomized trial. *Pain*, **16**, 297–307.

Bruehl, S. and Carlson, C.R. (1992). Predisposing psychological factors in the development of reflex sympathetic dystrophy. *Clinical Journal of Pain*, **8**, 287–99.

Campbell, J.N., Raja, S.N., and Meyer, R.A. (1988). Painful sequelae of nerve injury. In *Proceedings of the 5th World Congress on Pain*(ed. R. Dubner, G.F. Gebhart, and M.R. Bond), pp. 135–43. Elsevier, Amsterdam.

Campbell, J.N., Meyer, R.A., and Raja, S.N. (1992). Is nociceptor activation by alpha-1 adrenoreceptors the culprit in sympathetically maintained pain. *American Pain Society Journal*, **1**, 3–11.

Carlson, T. and Jacobs, A.M. (1986). Reflex sympathetic dystrophy syndrome. *Journal of Foot Surgery*, **25**,149–53.

Christensen, K., Jensen, E.M., and Noer, I. (1982). The reflex dystrophy syndrome: response to treatment with systemic corticosteroids. *Acta Chirurgica Scandinavica*, **148**, 653–5.

Cicuttini, F. and Littlejohn, G.O. (1989). Female adolescent rheumatological presentations: the importance of chronic pain syndromes. *Australian Paediatric Journal*, **25**, 21–4.

Coughlan, F.J., Hazleman, B.L., and Page-Thomas, D.P. (1987). Algodystrophy: a common unrecognized cause of chronic knee pain. *British Journal of Rheumatology*, **26**, 270–4.

Davis, S.W. (1977). Shoulder-hand syndrome in a hemiplegic population: 5 year retrospective study. *Archives of Physical Medicine and Rehabilitation*, **58**, 353–9.

Dotson, R.M. (1993). Causalgia—reflex sympathetic dystrophy—sympathetically maintained pain: myth and reality. *Muscle and Nerve*, **16**, 1049–55.

Doury, P. (1988). Algodystrophy: reflex sympathetic dystrophy syndrome. *Clinical Rheumatology*, **7**, 173–80.

Doury, P., Dirheimer, Y., and Pattin, S. (1981). *Algodystrophy*. Springer, Berlin.

Escobar, P.L. (1986). Reflex sympathetic dystrophy. *Orthopaedic Review*, **15**, 646–51.

Eurly, F., Chevalier, X., Crozes, P., Prudat, M., Lechevalier, D., Larget-Pret, B. (1992). Is hyperlipidaemia a contributing factor to algodystrophy (reflex sympathetic dystrophy)? *Clinical Rheumatology*, **11**, 526–8.

Fields, H.L. (1987). Efferent activity and pain: the reflex sympathetic dystrophy syndrome. In *Pain*, pp. 145–55. McGraw-Hill, New York.

Fontaine, R. *et al*. (1957). Contribution a la physiopathologie de l'osteoporose post-traumatique. *Acta Chirurgica Belgica*, **1**(Suppl.), 173–201.

Genant, H.K., Kozin, F., Bekerman, C., McCarty, D.J, and Sims, J. (1975). The reflex sympathetic dystrophy syndrome: a comprehensive analysis using fine-detail radiography, photon absorptiometry, and bone and joint scintigraphy. *Radiography*, **117**, 21–32.

Gobelet, C., Meier, J-L., Schaffner, W., Bischof-Delaloye, A., Gerster, J-C., and Burckhardt, P. (1986). Calcitonin and reflex sympathetic dystrophy syndrome. *Clinical Rheumatology*, **5**, 382–8.

Gobelet, C., Waldburger, M., and Meier, J.L. (1992). The effect of adding calcitonin to physical treatment on reflex sympathetic dystrophy. *Pain*, **48**, 171–5.

Goldsmith, D.P., Vivino, F.B., Eichenfield, A.H., Athreya, B.H., and Heyman, S (1989). Nuclear imaging and clinical features of childhood reflex neurovascular dystrophy: comparison with adults. *Arthritis and Rheumatism*, **32**, 480–5.

Griffiths, H.J and Virtama, P. (1988). Juxta-articular erosions in reflex sympathetic dystrophy. *Acta Radiologica*, **29**, 183–7.

Harvey, A.R. (1987). Neurophysiology of rheumatic pain. *Clinical Rheumatology*, **1**,1–26.

Hunter, J. (1843). Lessons on the principles of surgery (1766). In *Oeuvres complete*, Vol. 1, (ed. J. Hunter). Tortin-Masson, Paris.

Janig, W. (1996). The puzzle of 'reflex sympathetic dystrophy': mechanisms, hypotheses, open questions. In: *Reflex sympathetic dystrophy: a reappraisal* (ed. W. Janig and M. Stanton-Wicks), pp. 1–24. IASP Press, Seattle.

Janig, W., Blumberg, H., Boas, R.A., and Campbell, J.N. (1991). The reflex sympathetic dystrophy syndrome: consensus statement and general recommendations for diagnosis and clinical research. In *Proceedings of the 6th World Congress on Pain*(ed. M.R. Bond, J.E. Charlton, and E.J. Woolf), pp. 373–6. Elsevier, Amsterdam.

Kozin, F. (1986). Reflex sympathetic dystrophy syndrome. *Bulletin of the Rheumatic Diseases*, **36**,1–8.

Kozin, F. (1994). Reflex sympathetic dystrophy syndrome. *Current Opinions in Rheumatology*, **6**, 210–6.

Kozin, F., McCarty, D.J., Sims, J., and Genant, J. (1976). The reflex sympathetic dystrophy syndrome. I. Clinical and histologic studies: evidence for bilaterality, response to corticosteroids, and articular involvement. *American Journal of Medicine*, **60**, 321–31.

Kozin, F., Fyan, L.M., Carrera, G.F., Soin, J.S., and Wortmann, F.L. (1981). The reflex sympathetic dystrophy syndrome. III. Scintigraphic studies, further evidence for the therapeutic efficacy of

systemic corticosteroids and proposed diagnostic criteria. *American Journal of Medicine*, **70**, 23–30.

Lakhanpal, S., Ginsburg, W.W., Luthra, H.S., and Hunder, G.G. (1987). Transient regional osteoporosis. *Annals of Internal Medicine*, **106**, 444–50.

Lankford, L. and Thompson, J. (1977). Reflex sympathetic dystrophy, upper and lower extremity: diagnosis and management. *American Academy of Orthopaedic Surgeons Instructional Course Lectures*, **26**, 163–78. Mosby, St Louis.

Laukaitis, J.P., Varma, V.M., and Borenstein, D.G. (1989). Reflex sympathetic dystrophy localized to a single digit. *Journal of Rheumatology*, **16**, 402–5.

Lesquesne, M., Kerboull, M., Bensasson, M., Perez, C., Deirser, R., and Forest, A. (1977). Partial transient osteoporosis. *Skeletal Radiology*, **2**, 1–9.

Levine, J.D., Clark, R., Devor, M., Helms, C., Moskowitz, M.R., and Basbaum, A.I. (1984). Intravenous substance P contributes to the severity of experimental arthritis. *Science*, **226**, 547–9.

Livingston, W.K. (1943). *Pain mechanisms: a physiologic interpretation of causalgia and its related states*. Macmillan, New York.

Lynch, M.E. (1992). Psychological aspects of reflex sympathetic dystrophy: a review of the adult and paediatric literature. *Pain*, **49**, 337–47.

Merskey, H. and Bogduk, N. (1994). *Classification of chronic pain: descriptions of chronic pain syndromes and definition of pain terms*, 2nd edn. IASP Press, Seattle.

Meyer, R.A., Campbell, J.N., and Raja, S. (1972). Peripheral neural mechanisms of cutaneous hyperalgesia. *Advances in Pain Research and Therapeutics*, **9**, 53–71.

Mitchell, S.W., Morehouse, G.R., and Keen, W.W. (1864). *Gunshot wounds and other injuries of nerves*. Lippincott, Philadelphia.

Otis, G.S. (1877). *The medical and surgical history of the war of rebellion*, Part 2, Vol. 2, p. 1020, plate 52. US Government Printing Office, Washington.

Perelman, R.B., Adler, D., and Humphreys, M. (1987). Reflex sympathetic dystrophy: electronic thermography as an aid in diagnosis. *Orthopaedic Review*, **16**, 561–6.

Plewes, L.W. (1956). Sudeck's atrophy in the hands. *Journal of Bone and Joint Surgery*, **38B**, 195–201.

Roberts, W.J. (1986). An hypothesis on the physiological basis for causalgia and related pains. *Pain*, **24**, 297–311.

Schwartzman, R.J. (1993). Reflex sympathetic dystrophy. *Current Opinions in Neurology and Neurosurgery*, **6**, 531–6.

Schwartzman, R.J. and Kerrigan, J. (1990). The movement disorder of reflex sympathetic dystrophy. *Neurology*, **40**, 57–61.

Schwartzman, R.J. and McLellan, T.L. (1987). Reflex sympathetic dystrophy: a review. *Archives of Neurology*, **44**, 555–61.

Sherry, D.D. and Weisman, M.A. (1988). Psychologic aspects of childhood reflex neurovascular dystrophy. *Paediatrics*, **81**, 572–8.

Sherry, D.D., McGuire, T., Mellins, E., Salmonson, K., Wallace, C.A., and Nepom, B. (1991). Psychosomatic musculoskeletal pain in childhood: clinical and psychological analyses of one hundred children. *Pediatrics*, **88**, 1093–9.

Silber, J. and Majd, M. (1988). Reflex sympathetic dystrophy syndrome in children and adolescents. *American Journal of Diseases of Children*, **142**, 1325–30.

Small, N.C. (1993). Complications in arthroscopic surgery of the knee and shoulder. *Orthopedics*, **16**, 985–8.

Steinbrocker, O. and Argyros, T.G. (1958). The shoulder-hand syndrome: present status as a diagnostic and therapeutic entity. *Medical Clinics of North America*, **42**, 1538–53.

Sudeck, P. (1900). Über die akute entzundliche knochenatrophie. *Archiv für Klinische Chirurgie*, **62**, 147–56.

Tietjen, R. (1986). Reflex sympathetic dystrophy of the knee. *Clinical Orthopaedics*, **209**, 234–43.

Vaeroy, J., Qiao, Z-G., Morkrid, L., and Forre, O. (1989). Altered sympathetic nervous system response in patients with fibromyalgia (fibrositis syndrome). *Journal of Rheumatology*, **16**, 1469–5.

Wang, J.K., Johnson, K.A., and Ilstrup, D.M. (1985). Sympathetic blocks for reflex sympathetic dystrophy. *Pain*, **23**, 13–7.

Werner, R., Davidoff, G., Jackson, D., Cremer, S., Ventocilla, C., and Wolf, L. (1989). Factors affecting the senstivity and specificity of the three-phase technetium bone scan in the diagnosis of reflex sympathetic dystrophy syndrome in the upper extremity. *Journal of Hand Surgery*, **14A**, 520–3.

5.19.3 Rheumatic complications of drugs and toxins

Robert M. Bernstein

Congratulations if you have reached this chapter or happened upon it by chance. The author felt obliged to meet you here, because Francis Bacon held 'every man a debtor to his profession.' Bacon was also of the opinion that 'books must follow sciences, and not sciences books', but often books follow books: they contain not truth but copies of bad maps, so it is up to the student to read the scientific literature directly as well as learn by exploration and experience.

This chapter deals briefly with rheumatological presentations of adverse drug reactions and poisoning; more extensive reviews are available (Shoenfeld and Isenberg 1989; Kahn 1991). With so much public whim and paranoia it is easy to close our minds in defence of what we do, particularly when it comes to prescribing strange chemicals or eating them. Assessment of a potential reaction should involve clinical observation during therapy, after withdrawal of the drug, and, if possible, upon rechallenge; where the reaction is subjective rechallenge should be blinded.

Immunological reactions

A long list of drugs can trigger autoimmune diseases in many ways similar to their idiopathic counterparts (Table 1). These serve as important models for the study of idiopathic disease. Apart from anaphylaxis, which is immediate, and rashes, which often occur early in treatment, most of these reactions develop gradually after an appreciable exposure to the drug (Perry 1973). Some reactions are rare, but others may affect over 10 per cent of patients with sufficient drug exposure. Factors influencing susceptibility include cumulative dose, renal and hepatic function, genetic polymorphisms influencing drug metabolism (particularly acetylation and sulphoxidation), sex, and immunological response (HLA-DR type and complement null alleles). Grounds for imputing an immunological mechanism include infiltration with lymphocytes (hepatitis, myositis) and the presence of autoantibodies (haemolytic anaemia, myasthenia gravis, lupus, Goodpasture's syndrome, pemphigus). The autoantibodies are usually similar to those arising in the relevant idiopathic autoimmune condition (Table 1), but generally occur much more frequently than any clinical expression of disease. Thus, hydralazine induces antinuclear antibodies in up to 60 per cent of patients but the lupus syndrome in only 2 to 10 per cent, and Coombs' antibody need not cause haemolysis.

Syndrome	Autoantibody specificity
Lupus	Histones, Single-stranded DNA, Poly (ADP-ribose), Mitochondria (venocuran-induced) Myeloperoxidase
Myositis	Nuclear
Myasthenia gravis	Acetylcholine receptor
Scleroderma	None recognized*
Haemolytic anaemia	I antigen of red cell
Thrombocytopenia	Platelet membrane
Pemphigus	Skin basement membrane
Goodpasture's syndrome	Kidney basement membrane

*Topoisomerase I in silica mimics

Table 1 Drug-induced autoimmune syndromes and autoantibodies

Drug-induced lupus

Starting with observations by Perry in the 1950s on a late toxic reaction to hydralazine therapy (Perry 1973) and then similar observations with procainamide, many drugs have been reported to induce systemic lupus erythematosus or a lupus-like syndrome (Table 2). To implicate a drug with certainty the syndrome should remit after drug withdrawal and recur on rechallenge, but often there are just a few case reports and sometimes the connection is doubtful since idiopathic lupus may have been developing at the time. In particular, oestrogens and sulphonamides may exacerbate or bring on idiopathic systemic lupus erythematosus, and hair dyes containing aromatic amines have been implicated in a cluster of cases of systemic lupus erythematosus and scleroderma in a small town in Georgia (Freni-Titulaer *et al*. 1988). Minocycline, a semisynthetic tetracycline used for a variety of infections, and for treatment of acne vulgaris, has been reported to induce a lupus-like syndrome (Gough *et al*. 1996).

Definite	Hydralazine Isoniazid Procainamide
Probable	Penicillamine Sulphasalazine Acebutalol Labetalol Methyldopa Captopril Phenytoin Carbamazepine Chlorpromazine Lithium Propylthiouracyl Quinidine Psoralen/ultraviolet A (PUVA) Venocuran Minocycline

For a more comprehensive list see Slinger (1988).

Table 2 Drugs reported to induce a lupus syndrome

Clinical features

The clinical features of drug-induced lupus are shown in Table 3, with data on idiopathic systemic lupus erythematosus for comparison. Common, early manifestations (after several months of drug exposure) are arthralgia, aching, malaise, and elevated erythrocyte sedimentation rate. Arthritis, rash, lymphadenopathy, pleurisy and pleural effusion, pericardial effusion and hepatosplenomegaly occur less often, and Raynaud's phenomenon, central nervous system involvement, and renal disease are rare.

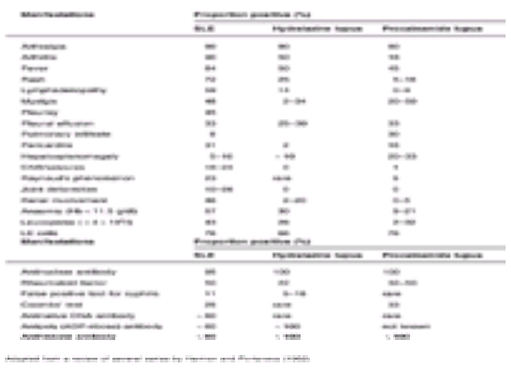


Table 3 Clinical and laboratory manifestations of idiopathic systemic lupus erythematosus and drug-induced lupus

Drug-induced lupus is clearly distinguished from idiopathic systemic lupus erythematosus by the dearth of renal, central nervous system, and Raynaud's involvement, the lower frequency of rash, a narrower autoantibody profile, the rarity of hypocomplementaemia, and a different HLA background (Table 3 and Table 4). However, these are not all necessarily fundamental differences, since drug-induced lupus is usually curtailed within the first couple of years of clinical expression by withdrawal of the drug or (if misdiagnosed) by treatment with corticosteroid. A few cases do go on to cutaneous vasculitis (Bernstein *et al*. 1980) and may progress to glomerulonephritis. After stopping the drug, clinical features generally improve within days or weeks and resolve within months of stopping the offending drug; the antinuclear antibody titre wanes over a year or two (Mansilla-Tinoco *et al*. 1982).

	Hydralazine-treated controls	Hydralazine lupus	Odds ratio
Mean hydralazine intake	65 g	150 g	
Female	31%	81%	9.9
White	63%	96%	8.2
Slow acetylator	55	96	8.2
HLA-DR4	25%	72%	8.1
C4 null	43%*	76%	4.3

Based on Batchelor et al. (1980), Mansilla-Tinoco et al. (1982), and Speirs et al. (1989).
*Healthy controls not hydralazine treated.

Table 4 Risk factors for hydralazine lupus

Antibody profile

The antinuclear antibody profile in drug-induced lupus is much narrower than in idiopathic systemic lupus erythematosus. Antibodies to extractable nuclear antigens are rare, and to cardiolipin and native DNA uncommon. The characteristic antinuclear antibody is of homogenous pattern and high titre—the sort producing the LE cell phenomenon. The specificity of these antibodies is for histones and poly(ADP-ribose), sometimes for single-stranded DNA, but only rarely for native double-stranded DNA (Hobbs *et al*. 1987). Serial measurement of DNA binding may show a slight rise and then a fall on drug withdrawal but this usually remains within the normal range of DNA binding (which laboratories set quite high to exclude the modest levels of anti-DNA antibody seen in quite a wide range of autoimmune conditions). Antihistone antibody titres rise much higher in drug-induced lupus than in idiopathic systemic lupus erythematosus, whereas levels of antibody to poly(ADP-ribose) are similar. Antibodies to myeloperoxidase (anti-MPO), giving a pANCA pattern of immunofluorescence on alcohol-fixed neutrophils, have also been reported in drug-induced lupus (Nässberger *et al*. 1990). The presence of autoantibodies is not *per se* a reason to discontinue treatment.

Risk factors

Dose and metabolism

The risk factors for developing hydralazine induced lupus are fairly clear (Table 4). A modest dose must be taken for a year or two or a large dose for several months, and almost always the patient is a 'slow acetylator'. Drug acetylation occurs immediately after absorption on first pass through the liver, and we know that administration of acetylprocainamide (itself an effective antiarrhythmic drug) never leads to the lupus syndrome even though a little of the dose is deacetylated (Woosley *et al*. 1978). The antinuclear antibody frequency rises faster in slow than fast acetylators. Indeed, it was thought for a time that acetylation of dietary toxins might be relevant to the genesis of idiopathic systemic lupus erythematosus but this is not the case.

Sex and race

As in idiopathic systemic lupus erythematosus, females are more at risk than men (though drugs like hydralazine and procainamide are used more often in men), but, in contrast to idiopathic systemic lupus erythematosus, Perry observed that black patients are highly resistant to the development of hydralazine lupus.

Major histocompatibility complex

Hydralazine lupus is associated with HLA-DR4 (Batchelor *et al*. 1980) and with null alleles of the complement component C4 (Speirs *et al*. 1989). HLA-DR4 and C4B null are known to be in linkage disequilibrium, just as in idiopathic systemic lupus erythematosus the HLA-DR2 and -DR3 associations may reflect the frequency of haplotypes containing C4A null alleles together with these DR antigens.

Pathogenesis

Complement dysfunction induced by drugs

Serum complement levels are usually normal in drug-induced lupus, but the drug may interfere with its function. In immune complex disease, complement is important in the solubilization of immune complexes and in their removal on the CR1 receptor of red cells (Schifferli *et al*. 1986). Indeed, many patients with idiopathic systemic lupus erythematosus have a null allele of C4 or occasionally C2, and homozygous deficiency of any of the early components of complement often leads to systemic lupus erythematosus. Reidenberg made the point that pharmacological rather than small immunizing doses are required for a drug to induce and maintain the lupus syndrome (Reidenberg 1981). One pharmacological mechanism demonstrated *in vitro* is that several lupus-inducing drugs block the transient activated form of C4. Inhibition of C4 activity is seen with hydralazine, isoniazid, penicillamine, and the hydroxylamine form of procainamide but not acetylprocainamide (Sim *et al*. 1984). Despite these findings, hydralazine does not exacerbate idiopathic systemic lupus erythematosus *in vivo*, and it has to be determined whether inhibition of C4A or C4B is more important.

Modulation of lymphocyte function

Drugs inducing autoimmune disease may also have a pharmacological action on lymphocytes. T lymphocytes from patients treated with methyldopa show reduced suppression of lymphocyte responsiveness to phytohaemagglutinin and immunoglobulin synthesis. Conversely, procainamide may have an inhibitory effect on T-cell responses to mitogens. Lymphocytotoxic antibodies are frequently induced by procainamide and hydralazine therapy, but their functional significance is unclear since they kill best below body temperature and can be found whether or not the lupus syndrome has developed (Bluestein *et al*. 1979).

Induction of autoantibodies

Autoantibodies are necessary for the genesis of drug-induced lupus (though not sufficient, whatever the titre, class, or complement-fixing ability), and much thought has been given to how drugs might induce autoantibodies. The specificity of the response rules out a general polyclonal B-cell activation, nor can tissue damage be blamed since there is no inflammation when the antibodies first appear.

The favoured explanation (Table 5) is that the drug interacts directly with the autoantigen. Hydralazine and procainamide bind to chromatin, and this may alter the autoantigen so as to overcome T-cell tolerance and thereby release autoreactive B cells from suppression. The fine specificity of autoantibodies in drug-induced lupus is in keeping with this model; histones and poly(ADP-ribose) are components of the nucleosome (unit of chromatin structure) and the histone epitopes recognized are those exposed on the surface of the nucleosome (Craft *et al*. 1987). This pattern of immune response can be imitated by immunization of rabbits with whole nucleosomes, whereas naked histones induce antibodies to epitopes that are normally buried.

1. Binding of drug to autoantigen, so overcoming T-cell tolerance
2. Antidrug antibody shows fortuitous cross-reaction with auto-antigen.
3. Anti-idiotype to antidrug antibody mimics the drug and binds to its receptor (the autoantigen)

The autoantigen might be on the cell membrane or released from dead cells.

Table 5 Models for the induction of autoantibodies by drugs

Alternative models involve antidrug antibodies that are cross-reactive either directly (rheumatic fever model) or through an anti-idiotype response (Table 5). These options demand fortuitous cross-reaction with one epitope on the autoantigen, followed by a spreading induction of autoantibodies to other epitopes.

Drug-induced myositis

Toxic myopathies occur occasionally with various drugs including clofibrate, emetine, chloroquine, e-aminocaproic acid, vincristine, lithium, amphotericin B, salbutamol, colchicine, and nitroxoline (Le Quintrec and Le Quintrec 1991). A true drug-induced myositis is seen occasionally with D-penicillamine and also reported with penicillin, sulphonamides, procainamide, hydralazine, phenytoin, propylthiouracil, phenylbutazone, cimetidine, and tamoxifen. In penicillamine therapy for rheumatoid arthritis, the frequency of myositis is 0.2 to 1.2 per cent rather than the expected coincidence of less than 0.001 per cent. Most cases of drug-induced myositis are antinuclear-antibody positive, and there seems to be a high frequency of dysphagia and muscle weakness but little myalgia. Myositis is mediated mainly by T cells, and in a case associated with cimetidine therapy there was a marked increase in the proportion of cytotoxic/suppressor T cells in the peripheral blood similar to the inflammatory infiltrate in muscle.

Drug-induced myasthenia gravis

Rheumatoid arthritis and idiopathic myasthenia gravis are weakly associated, but in 1975 Bucknall *et al*. described four cases of myasthenia gravis developing during penicillamine therapy and remitting after treatment was discontinued (Bucknall *et al*. 1975). This reaction may occur in up to 1 per cent of rheumatoid arthritis patients after several months of penicillamine therapy at daily doses of 500 mg or more. Antibodies to the acetylcholine receptor are found in all cases but their specificity is restricted to the human receptor. HLA typing shows more BW35 and DR1 and less B8 and DR3 than in idiopathic myasthenia gravis, and less DR4 than in rheumatoid arthritis controls (Dawkins *et al*. 1981).

Drug-induced scleroderma

Scleroderma-like disease with the typical microvascular and fibrotic changes can follow exposure to a variety of drugs and chemicals. The drugs include bleomycin, 5-hydroxytryptophan, carbidopa, pentazocine, penicillamine, phytonadione, diethyopropion, and mazindol (appetite suppressants), and intravenous cocaine abuse. In bleomycin-induced scleroderma, there is increased collagen synthesis by dermal fibroblasts *in vitro* and administration of bleomycin to rats produces skin thickening with increased collagen synthesis (Rush *et al*. 1984; Bourgeois and Aeschlimann 1991).

Chemicals inducing scleroderma-like disease

Just as there are chemical causes of cirrhosis and pulmonary fibrosis, there are chemicals other than drugs that induce a scleroderma-like disease. These include silicone implants, vinyl chloride, and the outbreak of 'Spanish oil disease'.

Silicone implants

Silicones are polymers mainly of dimethylsiloxane (chains of alternating silicon and oxygen atoms with two methyl groups attached to each Si atom). The silicone forms a liquid, gel, or rubber-like material (elastomer) depending on the number of cross-links between the polymer chains. Being fairly inert within the body, silicone is used to lubricate syringes and has been incorporated in intraoccular lenses, ventriculoperitoneal shunts, and heart valves, as well as forming breast implants and artificial testicles.

Breast implants consists of silicone gel, and sometimes a space for saline, enclosed within a silicone rubber envelope. Implants are used chiefly for breast augmentation and reconstruction after surgery for breast cancer. In animals, injection of gel can induce local chronic inflammatory reactions and unsightly local reactions can occur in humans after direct injection of gel and occasionally after rupture of an implant. Augmentation of the breasts with silicone became popular from

the 1960s but is now being squeezed out by the lawyers.

Miners exposed to inhalation of crystalline silica have an increased risk of scleroderma (Erasmus 1957), but the question of scleroderma and other chronic illness as a response to silicone polymers is now a contentious issue.

Anecdotal reports of rheumatic symptoms arising after silicone injection or implantation have been published since the 1960s (Kamugai *et al*. 1984) and in a welter since medicolegal interest arose (Sanchez-Guerrero *et al*. 1994). Most are vague complaints, to which the controversial term human adjuvant disease has been applied (because silicone is not itself immunogenic but appeared to have adjuvant activity). Only a minority of cases have involved a clear cut autoimmune rheumatic disease and these have been indistinguishable on clinical and serological grounds from idiopathic cases.

There is no good epidemiological evidence for an increased frequency of autoimmune rheumatic diseases in relation to silicone although the overall community prevalence of scleroderma may have increased in the past 50 years (Sanchez-Guerrero *et al*. 1994).

Despite the lack of scientific evidence, some patients may wish to have their implants removed for peace of mind. If the manufacturers survive bankruptcy, their settlement offer may reward those with common complaints like fatigue, widespread aching, paraesthesia, and the like. We must guard against this windfall making our patients feel worse.

Vinyl chloride disease

Workers exposed to vinyl chloride for prolonged periods in the manufacture of polyvinylchloride may develop an illness characterized by breathlessness, Raynaud's phenomenon, and contracture of the hands with thickening of the skin. Deposits of complement and fibrinogen are present in blood vessel walls (emphasizing that, like systemic sclerosis, this is a disease of the microvasculature as well as fibrosis), but anticentromere and anti-Scl-70 antibodies are not found. It is suggested that HLA-DR5 influences susceptibility and DR3 influences the severity of vinyl chloride disease (Black *et al*. 1983).

Spanish oil disease (see also Chapter 5.8)

The sale of contaminated rape-seed oil as cooking oil to about 20 000 people in the Madrid area led to a 'toxic oil syndrome' with an initial acute phase of fever, rash, gastrointestinal upset, neurological disturbance, acute interstitial pneumonia, and sometimes death. Many recovered, but several hundred went on to develop hardened, thickened skin, Raynaud's phenomenon, dysphagia, pulmonary hypertension, alopecia, dry eyes, dry mouth, arthritis, and flexion contractures. Autoantibodies were not a feature, but there was microvascular damage with endothelial proliferation and infiltration of vessel walls by lymphocytes and macrophages (Spurzem and Lockey 1984).

Eosinophilia–myalgia syndrome (see also Chapter 5.9.1)

First reported to the Centers for Disease Control in November 1989, almost 1500 cases had been notified 4 months later and it was clear that the ingestion of tryptophan as a health-food supplement was responsible. The amino acid all came from one manufacturer and a contaminant in the manufacturing process has been implicated. Long-term follow-up is underway but new cases have ceased occurring. The syndrome is characterized by the abrupt onset of malaise, myalgia, weakness, contractures, induration of fascia, morphoea-like lesions, and blood eosinophilia ($1–30\times10^{9}$/l). Neuromyopathy is frequent and cardiac abnormalities have been seen, with some deaths. The histological findings resemble eosinophilic fasciitis with an interstitial and perivascular inflammatory infiltrate (Le Quintrec and Le Quintrec 1991).

Serum sickness and hypersensitivity vasculitis

Serum sickness is an immune-complex disease. First seen following the injection of antiserum for the treatment of bacterial infections such as diphtheria and tetanus, it now occurs rarely with some drugs, particularly penicillin, sulphonamides, penicillamine, and thiouracil. The first indication is usually fever, appearing 7 to 12 days after the beginning of treatment, followed by urticaria, joint pains, and occasionally glomerulonephritis or myocarditis (Rich 1942); this may progress to vasculitis.

Hypersensitivity or leucocytoclastic vasculitis is characterized by an infiltrate of dead or dying polymorphonuclear white cells in the walls of small blood vessels. There is often no obvious cause but in some cases drugs have been implicated, particularly sulphonamides, penicillin, thiouracil, iodides, organic arsenicals, oestrogens, and hydantoins (Dubost *et al*. 1991). For more on vasculitis see Chapter 5.11.1, Chapter 5.11.2, Chapter 5.11.3 and Chapter 5.11.4.

Food and arthritis

Despite much folklore and several books, there is little convincing evidence that immune reactions to what we eat have any bearing on the pathogenesis of arthritis or the autoimmune rheumatic diseases (Walport *et al*. 1982). Anecdotal reports of food allergy have involved foods such as wheat, eggs, beef, and pork, and one diet recommends avoiding 'acidic' foods. Features often associated with arthritis in 'allergic' cases are migraine headaches, rhinitis, and gastrointestinal symptoms. After 50 years of open study, double-blind trials of diet are now in progress. It is possible that a diet of fish oil alters prostaglandin synthesis in a way that reduces the intensity of inflammation.

Specific syndromes occasionally associated with food allergy include palindromic rheumatism (reports of provocation by nitrates and menthol), vasculitis (various foodstuffs and in one case a particular brand of beer), hydrarthrosis of the knees (a case induced by English walnuts), and seronegative rheumatoid arthritis (a case exacerbated by milk and cheese). Most of these studies involved withdrawal of the foodstuff and rechallenge in open fashion.

Could rheumatoid arthritis be caused or exacerbated by absorption of antigens from the bowel? Pigs develop arthritis and nodules if fed a diet high in fish protein, and the onset of arthritis is associated with increased isolation of *Clostridium perfringens* from the gut flora. Arthritis is commoner in people with selective IgA deficiency, again highlighting mucosal defence. After intestinal bypass (once in favour as a treatment for morbid obesity) arthritis, often accompanied by features of Behçet's syndrome, may develop, and radiolabelled fragments of *E. coli* administered by mouth have been demonstrated in joints.

Metabolic reactions to drugs and toxins

Drugs and toxins causing gout

The relationship of alcohol to gout is complex in that many who drink well also eat well (with a high intake of purines), but it does seem that a high alcohol intake stimulates endogenous urate synthesis while high doses of alcohol temporarily reduce urate excretion (Scott 1991).

Lead rather than alcohol may have been responsible for the frequency of 'saturnine' gout in Georgian times when pewter was in fashion (lead-induced renal tubular damage leading to reduced excretion of uric acid). Nowadays, thiazide and loop diuretics are a common cause of hyperuricaemia, and gout is well recognized not just in bucolic men but in elderly women; the average increase in serum urate is about 70 µmol/l.

Salicylates have a uricosuric effect at higher doses but at low dose (with a low urine salicylate concentration) the excretion of uric acid is actually inhibited. Even the uricosuric agent probenecid has this paradoxical effect if given at a tiny dose. Among non-steroidal anti-inflammatory drugs, azapropazone has the most clinically useful uricosuric effect.

Hyperuricaemia, through an effect on the kidney, can also occur with pyrazinamide, ethambutol, and cyclosporin. Increased production of urate is caused by nicotinic acid and various cytotoxic agents, including vincristine, busulphan, thiotepa, cytarabine, 6-mercaptopurine, chlorambucil, cyclophosphamide, and the like. Increased tissue destruction leads to the release of purines and their metabolism to uric acid; this pathway can be blocked by the xanthine oxidase inhibitor allopurinol but with the warning that allopurinol increases the bioavailability of the azathioprine metabolite 6-mercoptopurine.

Fluoride and bone pain

In certain areas of India very high levels of fluoride in water lead to increased bone density and hyperostosis that goes on to cause widespread nerve root entrapment. Ingestion of moderate amounts of fluoride, as sometimes used in the treatment of osteoporosis, can cause severe pain in the legs, felt mainly around the joints. This lower limb pain occurs in up to 25 per cent of patients and remits when fluoride therapy is stopped or the dose reduced (Reeve 1990); microfractures may be the cause (Laroche and Mazieres 1991; Rooney *et al*. 1991).

Rheumatological effects of retinoids, quinolones, and proton pump inhibitors

Retinoids are derivatives of Vitamin A used in the treatment of severe acne. Chronic administration whether as a food fad or as dermatological treatment can cause arthralgias, arthritis, bone pain, hypercalcaemia, and periosteal new bone formation. Hyperostosis of the appendicular skeleton can occur, especially in children, and premature closure of the epiphyses has been reported (Kaplan and Haettich 1991).

Quinoline antibiotics include nalidixic acid and ciprofloxacin; very rarely these cause arthralgias and tenosynovitis in children, while puppies and the young of other susceptible species show surface blistering of the articular cartilage (Ribard and Kahn 1991).

Omeprazole and lansoprazole may cause arthralgias in occasional patients.

Transplant arthropathy

A painful, self-limiting, pseudoinflammatory arthropathy of lower limb joints can occur a few months after organ transplantation. There are clinical, radiological, and scintigraphic similarities to reflex sympathetic dystrophy. Knees or ankles are affected most often, but hip and wrist involvement has been seen. The clinical features are joint pain and tenderness and there may be effusion, periarticular oedema, and erythema. Blood tests for inflammation and autoantibodies are unhelpful; joint fluid cytology is non-inflammatory and does not reveal crystals. Radiographs show no specific abnormality, but isotope bone scans show greatly increased uptake on both sides of the joint (as well as in clinically unaffected joints sometimes). Avascular necrosis does not develop and spontaneous resolution after several months is the rule. This arthropathy has been recognized since the widespread adoption of cyclosporin for immunosuppression in organ transplantation. Cyclosporin levels have often been high when the arthropathy began and resolution tends to follow a reduction in cyclosporin levels. For instance, in our series of eight cases, mean cyclosporin levels were 458 mg/l at onset and 175 mg/l at resolution (Jones and Bernstein 1994).

Toxicity of antirheumatic treatment

Simple analgesics

Simple analgesics such as phenacetin, though probably not paracetamol, can cause renal papillary necrosis; this is commoner in hot climates where the urine is likely to be concentrated. Analgesics have a narrow safety range: eight tablets of paracetamol daily are safe, yet 30 tablets can ruin the liver.

Non-steroidal anti-inflammatory drugs

Gastrointestinal damage

Non-steroidal anti-inflammatory drugs commonly irritate the gastrointestinal tract (Henry *et al*. 1996), possibly through inhibition of prostaglandins which are important in protecting the mucosa and in the cellular mechanism for mending mucosal breaches. Endoscopy studies have shown up to 20 per cent of patients have ulcers and a further 30 per cent minor gastroduodenal lesions at any one time, and the risk of an acute bleed, perforation, or death rises from under 0.1 per cent in middle life to about 2.5 per cent in the elderly (Beardon *et al*. 1989). The synthetic prostaglandin E_1 misoprostol reduces non-steroidal anti-inflammatory-induced damage in the stomach, duodenum, and small intestine by about 75 per cent, whereas H_2-antagonists are protective only in the duodenum. Selective inhibitors of the inducible cyclooxygenase (COX-2) are likely to prove safer anti-inflammatory drugs.

Renal, haematological, and skin reactions

Non-steroidal anti-inflammatory drugs can reduce renal blood flow causing a tendency to hypertension and renal insufficiency (especially in the presence of hypovolaemia or pre-existing renal damage). Only sulindac is said to spare the kidneys. Other complications of non-steroidal anti-inflammatory drugs therapy include thrombocytopenia caused by increased platelet destruction, aplastic anaemia, and thrombocytopenia with phenylbutazone (now restricted in the United Kingdom to the hospital treatment of ankylosing spondylitis). Rashes are uncommon and occur most often with fenbufen.

Effects on cartilage

Shortly after the introduction of indomethacin there were reports of accelerated osteoarthritis of the hip, and in a clinical study of osteoarthritis of the hip, deterioration to the end-point of joint replacement was rather faster in patients treated with a strong inhibitor of prostaglandin synthesis (indomethacin) than with a weak inhibitor (azapropazone) (Rashad *et al*. 1989). Experimental data suggest various ways in which non-steroidal anti-inflammatory drugs might damage or even protect articular cartilage, and new chondroprotective agents are being sought by the pharmaceutical industry.

Disease-modifying, second-line therapy

Ocular toxicity from the deposition of antimalarial drugs such as chloroquine is well known. Rashes are common with most of the second-line drugs but with gold and penicillamine there is a particular risk of exfoliative dermatitis. Gold and penicillamine also cause membranous glomerulonephritis with nephrotic syndrome and, rarely, renal failure, while bone marrow toxicity with leucopenia, thrombocytopenia, or occasionally aplastic anaemia can occur with gold, penicillamine, sulphasalazine, and cytotoxic agents such as methotrexate, azathioprine, and cyclophosphamide; though onset can be abrupt, it is often gradual, so regular monitoring of the blood count and urine is recommended. Gold can also cause, rarely, pneumonitis and, not uncommonly, a 'post-injection flare' of arthritis (Rooney *et al*. 1991).

Penicillamine

Penicillamine used in the treatment of rheumatoid arthritis and Wilson's disease can trigger the whole range of drug-induced autoimmune reactions (membranous glomerulonephritis presenting as nephrotic syndrome, Goodpasture's syndrome, myasthenia gravis, polymyositis, pemphigus, lupus, and scleroderma), accompanied by the appropriate autoantibody (to kidney or skin basement membrane, acetylcholine receptor, double-stranded DNA, and so on). Susceptibility to at least some of these reactions is increased by the HLA antigen DR3 and by slow sulphoxidation of penicillamine (Emery *et al*. 1984).

Corticosteroids

Corticosteroid therapy (Geusens and Dequeker 1991) reduces resistance to infection whether administration is systemic or intra-articular, and this must be borne in mind in all sick patients on steroids. A hot, red joint following a steroid injection may be infected or (more often, one hopes) a gout-like reaction to particles in the steroid preparation responsive to non-steroidal anti-inflammatory drug therapy. Repeated steroid injections can lead to avascular necrosis. In the shoulder there is the added risk of rotator cuff degeneration, so injections there should be limited to a few.

Osteonecrosis (avascular necrosis) is also a serious complication of systemic steroid therapy, affecting usually one or both hips but sometimes other joints such as knees and shoulders. It may occur in 5 per cent or more of patients on long-term steroid therapy, and there is a particular risk with high-dose intravenous steroids used in the treatment of transplant rejection and sometimes for connective tissue diseases.

Osteoporosis is a concern with systemic corticosteroid therapy. In younger women treated with prednisolone at doses of 5 to 7.5 mg/day there was no bone loss detectable by dual photon absorptiometry over a 1 year period, but exacerbation of bone loss after the menopause is more of a problem. Calcium supplements may

help and female hormone replacement therapy or three monthly cycles of calcium for 11 weeks followed by a diphosphonate for 2 weeks encourages deposition of calcium on remaining bone trabeculas (Reeve 1990). More effective diphosphonates will be available shortly.

Monoclonal antibody and recombinant protein therapies

The next edition of this textbook will be full of optimism for the treatment of autoimmune diseases and arthritis by new methods. This paragraph is reserved for a host of adverse reactions—such as serum sickness (from mouse Fab), infection (as with steroids), enhanced autoimmunity (reported with interferons), and novel phenomena as yet undreamt of—that may arise from these new therapies, and also for debate with taxpayers who must pay for the therapies, as science brings home the bacon.

Chapter References

Batchelor, J.R., Welsh, K.I., Tinoco, R.M., Bernstein, R.M. *et al*. (1980). Hydralazine-induced systemic lupus erythematosus: the influence of HLA-DR and sex upon susceptibility. *Lancet*, **i**, 1107–9.

Beardon, P.H.G., Brown, S.V., and McDevitt, D.G. (1989). Gastrointestinal events in patients prescribed non-steroidal anti-inflammatory drugs: A controlled study using record linkage in Tayside. *Quarterly Journal of Medicine*, **71**, 497–505.

Bernstein, R.M., Egerton-Vernon, J., and Webster, J. (1980). Hydralazine induced cutaneous vasculitis. *British Medical Journal*, **1**, 156–7.

Black, C.M., Walker, A.E., Welsh, K.I., Bernstein, R.M., Catoggio, L.J. McGregor, A.R., and Lloyd-Jones, J.K. (1983). Genetic susceptibility to scleroderma-like disease induced by vinyl chloride. *Lancet*, **i**, 53–5.

Bluestein, H.G., Zvaifler, N.J., Weisman, M.H., and Shapiro, R.F. (1979). Lymphocyte alteration by procainamide: relation to drug-induced lupus erythematosus syndrome. *Lancet*, **ii**, 816–19.

Bourgeois, P. and Aeschlimann, A. (1991). Drug-induced scleroderma. *Clinical Rheumatology*, **5**, 13–20.

Bucknall, R.C., Dixon, A.S.J., Glick, E.N., *et al*. (1975). Myasthenia gravis associated with penicillamine treatment for rheumatoid arthritis. *British Medical Journal*, **1**, 600–2.

Craft, J.E., Radding, J.A., Harding, M.W., Bernstein, R.M., and Hardin, J.A. (1987). Autoantigenic histone epitopes: a comparison between procainamide- and hydralazine-induced lupus. *Arthritis and Rheumatism*, **30**, 689–94.

Dawkins, R.L., Zilko, P.J., Carrano, J., *et al*. (1981). Immunobiology of D-penicillamine. *Journal of Rheumatology*, **7**, 56–61.

Dubost J.-J., Souteyrand, P., and Sauvezie, B. (1991). Drug-induced vasculitides. *Clinical Rheumatology*, **5**, 119–38.

Emery, P., Panayi, G.S., Huston, G., Welsh, K.I., Mitchell, S.C., Idle, J.K., Smith, R.L., and Waring, R.H. (1984). D-penicillamine toxicity in rheumatoid arthritis. The role of sulphoxidation status and HLA-DR3. *Journal of Rheumatology*, **11**, 626–32.

Erasmus, L.D. (1957). Scleroderma in gold-miners on the Witwatersrand with particular reference to pulmonary manifestations. *South African Journal of Laboratory and Clinical Medicine*, **3**, 209–31.

Freni-Titulaer, L.W.J., Kelley, D.B., Grow, A.C., Hochberg, M.C., and Arnett, F.C. (1988). Clustering of connective tissue diseases in a small Georgia community. III. A search for environmental factors. *Arthritis and Rheumatism*, **31**, s75.

Geusens, P. and Dequeker, J. (1991). Locomotor side-effects of corticosteroids. *Clinical Rheumatology*, **5**, 99–118.

Gough, A., Chapman, S., Wagstaff, K., Emery, P., and Elias, E. (1996). Minocycline induced autoimmune hepatitis and systemic lupus erythematosus-like syndrome. *British Medical Journal*, **312**, 169–72.

Harmon, C.E. and Portanova, J.P. (1982). Drug-induced lupus. *Clinics in Rheumatic Disease*, **8**, 121–35.

Henry, D., Lim, L.L-Y., Garcia Rodriguez, L.A., *et al*. (1996). Variability in risk of gastrointestinal complications with individual non-steroidal anti-inflammatory drugs: results of a collaborative meta-analysis. *British Medical Journal*, **312**, 1563–6.

Hobbs, R.N., Clayton, A-L., and Bernstein, R.M. (1987). Antibodies to the five histones and poly(adenosine diphosphate-ribose) in drug induced lupus: implications for pathogenesis. *Annals of the Rheumatic Diseases*, **46**, 408–16.

Jones, P.B.B. and Bernstein, R.M. (1994). Joint pain after organ transplantation. *Arthritis and Rheumatism*, **37**, S273.

Kahn, M-F. (ed.) (1991). Drug-induced rheumatic diseases. *Clinical Rheumatology*, **5**, 1–196.

Kamugai, Y., Shiokawa, Y., Medsger, T.A. Jnr., *et al*. (1984). Clinical spectrum of connective tissue disease after cosmetic surgery. Observations on eighteen cases and a review of the Japanese literature. *Seminars in Arthritis and Rheumatism*, **27**, 1–12.

Kaplan, G. and Haettich, B. (1991). Rheumatological symptoms due to retinoids. *Clinical Rheumatology*, **5**, 77–97.

Laroche, M. and Mazieres, B. (1991). Side effects of fluoride therapy. *Clinical Rheumatology*, **5**, 61–76

Le Quintrec, J-S. and Le Quintrec, J-L. (1991). Drug-induced myopathies. *Clinical Rheumatology*, **5**, 21–38.

Mansilla-Tinoco, R., Harland, S.J., Ryan, P.J., Bernstein, R.M., *et al*. (1982). Hydralazine, antinuclear antibodies and the lupus syndrome. *British Medical Journal*, **284**, 936–9.

Nässberger, L., Sjöholm, A.G., Jonsson, H., Sturfelt, G., and Åkesson, A. (1990). Autoantibodies against neutrophil cytoplasm components in systemic lupus erythematosus and in hydralazine induced lupus. *Clinical and Experimental Immunology*, **81**, 380–3.

Perry, H.M. (1973). Late toxicity to hydralazine resembling systemic lupus erythematosus or rheumatoid arthritis. *American Journal of Medicine*, **54**, 58–72.

Rashad, S., Hemmingway, A., Rainsford, K., Revell, P., Low, F., and Walker, F. (1989). Effect of non-steroidal anti-inflammatory drugs on the course of osteoarthritis. *Lancet*, **ii**, 519–22.

Reeve, J. (1990). Restoring trabecular bone mass in established osteoporosis. In *Osteoporosis 1990*, (ed. R. Smith), pp. 143–55. RCP Publications, London.

Reidenberg, M.M. (1981). Clinical induction of SLE and lupus like illness. *Arthritis and Rheumatism*, **24**, 1004–9.

Ribard, P. and Kahn, M-F. (1991). Rheumatological side-effects of quinolones. *Clinical Rheumatology*, **5**, 175–92.

Rich, A.R. (1942). The role of hypersensitivity in periarteritis nodosa: as indicated by seven cases developing during serum sickness and sulphonamide therapy. *Bulletin of the John Hopkins Hospital*, **71**, 123–40.

Rooney, P.J., Balint, G.P., Szebenyi, B., and Petrou, P. (1991). Rheumatic syndromes caused by antirheumatic drugs. *Clinical Rheumatology*, **5**, 139–73.

Rush, P.J., Bell, M.J., and Fam, A.G. (1984). Toxic oil syndrome (Spanish oil disease) and chemically induced scleroderma-like conditions. *Journal of Rheumatology*, **11**, 262–4.

Sanchez-Guerrero, J., Schur, P.H., Sergent, J.S., and Liang, M.H. (1994). Silicone breast implants and rheumatic disease: clinical, immunologic and epidemiologic studies. *Arthritis and Rheumatism*, **37**, 158–68.

Schifferli, J.A., Ng., Y.C., and Peters, D.K. (1986). The role of complement and its receptor in the elimination of immune complexes. *New England Journal of Medicine*, **315**, 488–95.

Scott, J.T. (1991). Drug-induced gout. *Clinical Rheumatology*, **5**, 39–60.

Shoenfeld and Isenberg, D. (1989). *The mosaic of autoimmunity*, pp. 321–43 and 397–400. Elsevier, Amsterdam.

Sim, E., Gill, E.W., and Sim, R.B. (1984). Drugs that induce systemic lupus erythematosus inhibit complement component C4. *Lancet*, **ii**, 422–4.

Solinger, A.M. (1988). Drug-related lupus. *Clinics in Rheumatic Disease*, **14**, 187–202.

Speirs, C., Fielder, A.H.L., Chapel, H., Davey, N.J., and Batchelor, J.R. (1989). Complement system protein C4 and susceptibility to hydrazaline-induced systemic lupus erythematosus. *Lancet*, **i**, 922–4.

Spurzem, J.R. and Lockey, J.E. (1984). Toxic oil syndrome. *Archives of Internal Medicine*, **144**, 249–50.

Walport, M.J., Parke, A.L., and Hughes, G.R.V. (1982). Food and the connective tissue diseases. *Clinics in Immunology and Allergy*, **2**, 113–20.

Woosley, R.L., Drayer, D.E., Reidenberg, M.M., Nies, A.S., Carr, K., and Oates, J.A. (1978). Effects of acetylator phenotype on the rate at which procainamide induces antinuclear antibodies and the lupus syndrome. *New England Journal of Medicine*, **298**, 1157–60.

6.1 Surgery in adults

Justin Cobb and C. B. D. Lavy

Introduction

The surgeon is an indispensable member of the team involved in the treatment of a patient with arthritis. He or she should be available not only as a technician to perform operations but ideally should also be involved in the decision making processes in the management of the patient, such as the timing of surgery, the order in which affected joints should be treated, and of course the procedures undertaken. Most large units have combined clinics where the experience of rheumatologist, surgeon, physiotherapist, occupational therapist, and patient can be shared. In patients with polyarthritis, where several joints are likely to need surgery, a valuable relationship between patient and surgeon can be developed—a relationship which gives the patient trust and confidence and gives the surgeon a greater understanding of the patient's specific needs than can be gained by a single meeting.

This chapter will firstly list the options available to the surgeon dealing with rheumatic disease. Secondly, it will cover specific pre- and postoperative considerations that must be addressed by the surgeon. The third and largest part of the chapter will focus on specific anatomical regions in detail, outlining the procedures most commonly used. Operating theatre technique will not be covered. For this the reader is referred to specific operative texts, some of which are included in the references.

Surgical technique

The armamentarium of surgical treatment of arthritis includes the following procedures:

1. arthroscopy and arthroscopic debridement;
2. synovectomy;
3. soft tissue release, realignment, or repair;
4. tendon transfer;
5. osteotomy;
6. excision arthroplasty;
7. prosthetic arthroplasty;
8. arthrodesis and stabilization.

Not all the above are relevant in every anatomical region but they are listed here for completeness. The particular techniques that are commonly used will be outlined under individual joints and regions below.

Special considerations in planning surgery in rheumatic diseases

When to operate

It is difficult to give specific guidelines as to when surgery should be considered for a particular joint as no joint can be considered in isolation from the other joints in the same limb or the other limbs, the patient his or herself, and the patient's position in his or her family and community. However, with all the above in mind, the two factors that most affect the decision to proceed with surgery are pain and loss of function. If the patient's pain cannot be adequately controlled by conservative means and their life, or more particularly their sleep, is being affected then surgery should be considered (Seyfer 1993). Similarly if the patient's function is being significantly impaired by joint destruction or deformity then surgery to reconstruct or replace the joint should be considered.

Prophylactic surgery has a limited place in the treatment of rheumatic diseases. Synovectomy is certainly of use in incipient tendon rupture. It has also been performed in an effort to prevent deterioration of inflammatory arthritis; however, no long-term study has shown significant slowing of joint destruction. Prophylactic surgery also has a role in stabilizing the potentially unstable spine.

Which joint to operate on first

The order of surgery in patients who have several joints that will benefit from surgery is a common problem in polyarthritic conditions such as rheumatoid arthritis. In deciding which joints to start with there are no hard and fast rules but there are several considerations that must be taken into account. Firstly, when surgery to upper and lower limbs is being considered the lower limb is usually done first as this restores important independent mobility. Rehabilitation following such surgery, however, often requires the use of crutches and if the arm function is so poor that even gutter crutches cannot be used, then surgery to the upper limb may have to precede that to the lower limb. Secondly, when two joints in a limb are equally affected it is usual to operate on the proximal joint before the distal joint. This helps to reduce pain and improve the proximal stability in the limb before surgery to the distal joint. Where two joints in a limb are both equally affected, for example the hip and the knee, it may be found that surgery to the hip so improves lower limb function that surgery to the knee is no longer needed. Similarly, surgery to the shoulder may dramatically reduce pain and functional demand on the elbow so that pain is improved and surgery to the elbow is not needed. In other patients the opposite may be found—surgery is so successful in relieving pain in the most painful joint that the patient's attention now turns to a less affected joint, one that previously had not troubled them much which now becomes the focus of their attention.

The final consideration in the order of surgery is the concept of surgical success. When a number of procedures are likely to be necessary it is often helpful to the patient's morale to start with a procedure that carries a high statistical likelihood of success. An example of this is excision of the distal ulna and synovectomy of the wrist.

Patient's consent and preparation

Patients must be adequately consented for surgery. This includes not only understanding the operation and its important complications but also the nature of the rehabilitation process involved which may be long. It is helpful if the physiotherapist and occupational therapist both meet the patient beforehand. It can also be of help for the patient to meet others who have had similar surgery.

Anaesthetic considerations

Patients with rheumatic diseases often have associated medical conditions that must be considered during surgery (Skues and Welchew 1993). For example in rheumatoid arthritis cervical spine instability means that great care must be taken with endotrachial incubation and, if in doubt, fibreoptic intubation used. With some rheumatic diseases temporomandibular joint function is also limited and cricoarytenoid cartilages can be involved—both also causing airway difficulties. Patients with chronic inflammatory disorders are also often anaemic and some require preoperative transfusion. Pulmonary, cardiovascular, and renal complications in rheumatic diseases can also affect anaesthesia. These are discussed elsewhere.

Drugs affecting surgery

Patients who have been on steroids or antimetabolic drugs may have thin, delicate skin that tears or bruises easily. Great care must therefore be taken during surgery and in positioning and handling the anaesthetized patient. Skin may also be sensitive to the adhesive compound used in some dressings. Steroids and antimetabolic drugs may also reduce the patient's resistance to infection and also delay wound healing, although they are not in themselves contradictions to surgery. Specific medical complication to steroids and other disease-modifying agents will be considered elsewhere.

Positioning the patient

Often joint stiffness and restriction of movement will not allow normal positioning to be used. For example shoulder and elbow stiffness may not allow the hand to be flat on a table beside the patient. A degree of flexibility on the part of the surgeon is therefore needed. Great care must always be taken of the anaesthetized patient to ensure that joints are not over stressed. This is especially the case with rheumatoid arthritis where the soft tissues and bones are fragile. Care must also be taken at pressure points to avoid skin necrosis, bruising, and damage to peripheral nerves.

Surgical treatment in specific anatomical regions

Shoulder

The shoulder is commonly involved in inflammatory arthritis with clinical features of pain in 47 per cent (Laine *et al*. 1954). Petersson reports radiographic changes in the shoulders in 83 per cent of patients with rheumatoid arthritis (Petersson 1986). There is, however, little correlation between clinical signs and symptoms and radiographic appearance and it is not uncommon for a grossly destroyed shoulder joint to present with few symptoms.

The shoulder comprises four main anatomical areas; the sternoclavicular joint, the acromioclavicular joint, the subacromial region, and the glenohumeral joint. Pain arising from the sternoclavicular joint is usually easy to diagnose as it is localized to the sternoclavicular region. However, it is more difficult to be specific about the anatomical origin of pain and tenderness in the area around the shoulder itself. Physical signs can point to one of the three sites as the origin of the pain but most surgeons rely on the additional diagnostic information given by local anaesthetic tests. One per cent lignocaine is injected sequentially into the subacromial space, the acromioclavicular joint, the region of the biceps tendon, and the glenohumeral joint until the pain is abolished. Many patients will have some involvement at more than one, or indeed all, of these sites (Johnston and Kelly 1990).

Sternoclavicular joint

If conservative measures such as injection of steroids cannot control the pain arising from this joint then synovectomy can be considered or a limited excision of the medial end of the clavicle can be performed, keeping the excision medial to the costoclaviclar ligaments. Care must be taken to avoid instability by preserving as much of the capsule as possible. The surgeon should also be cautious as the immediate posterior relation of the left joint is the innominate vein.

Acromioclavicular joint

This joint is involved in 70 per cent of rheumatoid shoulders. If pain becomes severe and cannot be controlled by conservative treatment or injection, the distal end can be excised lateral to the coracoclavicular ligaments which provide stability.

Subacromial region

Pain commonly arises from this region due to subacromial impingement which may be associated with inflammation in the subacromial bursa or a partial or complete rotator cuff tear. If pain from this area is abolished by local anaesthetic injection yet not controlled by steroid injection, then a subacromial decompression can be performed. This can be done as an open operation or, where the expertise exists, as an arthroscopic procedure using a powered burr. Rotator cuff repair is usually unrewarding in late rheumatoid arthritis because of the atrophied nature of the supraspinatus tendon.

Glenohumeral joint

This is involved in more than two-thirds of patients with rheumatoid arthritis (Ennevaara 1967). In early cases synovectomy can be performed either as an open procedure or arthroscopically (Bennett and Gerber 1994). Ogilvie-Harris, and Wiley had good results in nine out of eleven patients having this procedure arthroscopically (Ogilvie-Harris and Wiley 1986). In moderate to severe bony destruction a double osteotomy, as described by Benjamin, can be performed (Benjamin *et al*. 1979). This comprises an osteotomy of the humerus at the surgical neck and a similar osteotomy of the scapula at the glenoid neck. The mechanism of action of this procedure is poorly understood but it has been shown to have good result in terms of pain relief and increased movement in 25 out of 29 patients in Benjamin's series and in 29 out of 32 patients in Jaffe and Learmonth's series (Benjamin 1987; Jaffe and Learmonth 1989).

Joint replacement

Arthroplasty of the shoulder has not been as successful as in the knee or the hip in terms of pain relief or restoration of range of movement. This has been due to poor bone stock in which to anchor prostheses and poor quality soft tissues, especially the rotator cuff. Despite these difficulties, shoulder replacement is a useful therapeutic tool in rheumatoid arthritis giving pain relief in over 90 per cent of patients (Friedman *et al*. 1989).

Early shoulder replacements were constrained, for example the Stanmore shoulder, which had a plastic glenoid cup holding a humeral head fully captive (Lettin *et al*. 1982), or the Kessel shoulder which had the ball joint the other way around (Kessel and Bayley 1988). More recent total shoulder replacements such as the Neer (Neer *et al*. 1982), the Global, or the Copeland have an anatomically shaped humeral head which can articulate with the patient's own glenoid if this is not severely damaged, or with a prosthetic glenoid which is usually cemented in place. The range of movement following total shoulder replacement depends largely on the state of the rotator cuff. Where the rotator cuff is intact the reconstruction of the anatomical contours of the shoulder joint following arthroplasty can give a very good range of movement with abduction to over 90°. In severe bony destruction, however, there is often virtually no rotator cuff remaining and range of movement is minimally increased by surgery. Indeed the main benefit of the procedure is pain relief.

Arthrodesis

The success of arthroplasty means that arthrodesis is hardly ever indicated in rheumatic diseases. Even in rare cases where because of infection total shoulder replacements need to be removed a stable fibrous pseudoarthrosis usually gives a better functional result than an arthrodesis.

Elbow

Extra-articular procedures commonly performed around the elbow in rheumatoid arthritis include removal of nodules and transposition of the ulnar nerve. Rheumatoid nodules are common around the elbow and they may become painful, unsightly, or ulcerated. Ulnar nerve dysfunction may be motor or sensory and can result both from mechanical and pressure factors due to the presence of swollen and inflamed synovium and a deforming joint. Ischaemic factors also have a role. Ulnar nerve function may be improved by a simple release or, if it is clearly tight, then transposition anterior to the medial epicondyle is recommended.

Synovectomy

Synovectomy of the elbow joint itself is difficult as surgical access to all parts of the intact joint requires several incisions; however, it gives some benefit at least in the short term when there is gross swelling. Porter reported relief of pain and return of function in 12 out of 16 patients over a 5-year period (Porter *et al*. 1974). Arthroscopic synovectomy is a new tool that has to be performed with care as the radial and median nerves are closely related to the anterior capsule.

Excision of radial head

When the radial side of the joint is giving rise to the majority of symptoms, excision of the radial head can be performed at the same time as synovectomy (Copeland and Taylor 1979). This is an excellent procedure in terms of pain relief. Stability of the joint should be considered but is seldom significantly compromised in low demand patients. Silastic replacement of the radial head has not proved universally successful.

Arthroplasty

When there is severe pain in the presence of destructive changes in the elbow joint, both in rheumatoid arthritis and in osteoarthritis, arthroplasty can be performed. The simplest and oldest comprises excision arthroplasty which relieves pain but has unreliable stability. Inter position arthroplasty using silicon or metal had limited popularity (Coates *et al*. 1991) but most surgeons now use a form of total joint replacement. The available options are a simple hinge such as the Stanmore replacement, a semiconstrained hinge such as the GSB (Fig. 1), or a minimally constrained surface replacement such as the Souter (Souter 1990; Pritchard 1991). The increased stability of a hinge has been sited as increasing the risk of loosening by increasing forces at the prosthesis–bone interface. However, in up to 12-year follow-up GSB elbows continue to be considered excellent by 90 per cent of patients.

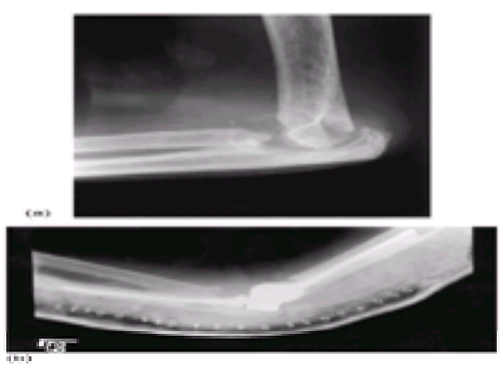


Fig. 1 Rheumatoid elbow (a) pre- and (b) postarthroplasty using a GSB semiconstrained hinge prosthesis.

Wrist

Osteoarthritis

Osteoarthritis of the of the wrist is not common and is usually post-traumatic, following either fracture of the distal radius or carpal bones, especially the scaphoid. When pain cannot be controlled, conservatively surgery should be considered. Joint replacement, as described below for rheumatoid arthritis, is an option in very low demand patients. Excision of the proximal carpal row can be performed where the disease is confined to this region. The space previously occupied by the excised proximal row becomes a pseudoarthrosis, relieving pain at the expense of some stability. Limited carpal fusions can also be performed particularly when the osteoarthritis is confined to the radial side of the wrist. This procedure does not have a high rate of success but if pain persists or recurs it can be revised to a complete wrist fusion.

Rheumatoid arthritis

Rheumatoid arthritis can affect the wrist joint as well as the flexor and extensor tendons and surgery to the tendons, such as synovectomy or transfer of the extensor tendons, is often combined with surgery to the wrist; however, in this text we will consider them separately. Rheumatoid arthritis can affect the radioulnar as well as the radiocarpal joints and often it is this joint that gives much of the pain, both from instability and from impingement on the carpus. Treatment by excision of distal 10 to 15 mm of the ulnar is very effective in terms of relief of pain and also removes an unsightly lump (Posner and Ambrose 1991; Clawson *et al*. 1991). When the pain is also arising from the radiocarpal joint, the surgical options are fusion or arthroplasty. Limited wrist fusion such as the Chamay technique (Chamay *et al*. 1983), which uses graft from the distal ulnar to fuse the radius to the lunate, are of use when there is good bone stock remaining but in most late cases there is such distortion of the normal anatomy that complete arthrodesis is performed.

The commonest method used involves insertion of a Rush pin (Fig. 2) or a Steinmann pin down the third metacarpal and across the wrist. This can be supplemented with a staple to improve stability. The position of stabilization is slight radial deviation and dorsiflexion. However, if both wrists are to be fused then perineal toilet is often improved if one is fused in slight flexion.

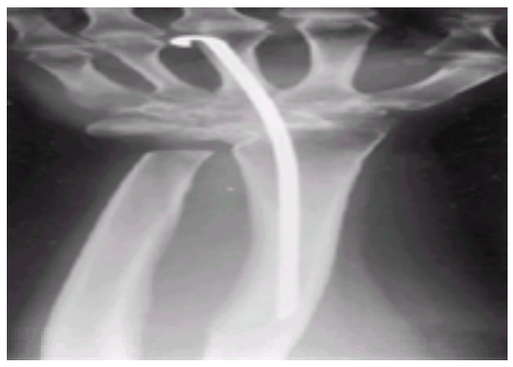


Fig. 2 Rheumatoid wrist following fusion using an intramedullary Rush in.

Joint replacement

The Swanson design silastic prosthesis is the most popular form of arthroplasty in Britain. The proximal carpal row is excised and the prosthesis fills the gap with one stem that fits into the distal radius and one that fits into a hole made through the capitate into the third metacarpal. Metal grommets can be used to protect the prosthesis from the sharp edges of the bone. This procedure gives good results in terms of relief of pain in low demand patients (Jolly 1993) but range of movement is usually limited to an arc of 40 to 50°. There are other designs of wrist replacement, such as the Meuli ball and socket joint which is popular in continental Europe. Where both wrists are involved in rheumatoid arthritis, many surgeons fuse the dominant wrist which is likely to take more force and confine arthroplasty to the non-dominant side.

Surgery to extensor tendons

In early disease, synovectomy of the extensor tendons is often combined with surgery to the wrist. The tendons are exposed by reflecting a flap of the dorsal extensor retinaculum, then after synovectomy the flap is replaced deep to the extensor tendons thus protecting them from the wrist joint. Synovectomy relieves pain and swelling and also reduces the risk of rupture.

The most common tendons to be ruptured are the extensor pollicis longus and the extensor digiti minima. The former is commonly treated by transfer of the extensor indicis proprius tendon from the index finger, which conveniently has two extensor tendons, and the remaining extensor indicis communes is adequate to extend the index. Treatment of rupture of extensor digiti minimi is usually by joining the distal end of the tendon to the extensor tendon of the ring finger. Realignment of the extensor tendons can be performed in rheumatoid arthritis. This commonly occurs at the level of the metacarpophalangeal joints where all the tendons tend to shift in an ulnar direction but also at the wrist where the extensor carpi ulnaris can migrate in a volar direction around the ulnar head.

Surgery to the flexor tendons

Synovitis also occurs around the flexor tendon in the synovial sheaths at the wrist and the fingers. When this interferes with function because of pain or by restricting flexion, synovectomy can be performed to improve movement and reduce the risk of rupture. In late cases rupture can occur. Surgical treatment of such rupture is complex and involves tendon grafting, often with a staged procedure. In mild cases of synovitis, triggering can occur around the neck of the flexor sheath. This is treated by simple division of the tight neck of the sheath. A limited synovectomy can also be performed.

Carpal tunnel syndrome

Median nerve compression is common in rheumatoid arthritis and can be relieved by carpal tunnel release. This is easily performed as a day case procedure under local anaesthetic. Endoscopic carpal tunnel release is popular in Europe and in the United States but because of early reports of nerve injury it has had a cautious reception so far in the United Kingdom.

Hand

Surgery to the fingers in rheumatoid arthritis

Boutonniere or Swan neck deformity can be corrected surgically by soft tissue procedures provided the joints are passively mobile. Once fixed deformity of the proximal interphalangeal (PIP) joints occurs, correction is much harder and surgical improvement may only be achieved by fusion in a position of function or arthroplasty using a silastic replacement joint. The position of fusion is usually 45 to 90° of flexion but depends on the range of movement at the metacarpophalangeal joints.

Surgery to the thumb in rheumatoid arthritis

Rheumatoid arthritis commonly causes a Z-deformity of the thumb and can cause instability of either the metacarpophalangeal or interphalangeal joint. Such instabilities are seldom helped bysoft tissue procedures and a fusion is usually necessary (Toledano *et al*. 1992).

Surgery to the metacarpophalangeal joints

When rheumatoid arthritis involves the metacarpophalangeal joint there is often a tendency to subluxation of these joints in an ulnar and volar direction. In early cases this can be treated with synovectomy and soft tissue realignment (Wynn Parry and Stanley 1993). This does slightly decrease range of movement but improves function because of decreased pain and increased power. In late cases where there is bony destruction of the metacarpal head, metacarpophalangeal joint replacement using Swanson silastic prostheses may be necessary (Fig. 3). This is always performed in conjunction with soft tissue release and realignment. Such reconstructive surgery, however, should only be done in order to improve function in the hand and should not be performed purely to improve the cosmetic appearance of a deformed rheumatoid hand. The postoperative care of metacarpophalangeal joint replacement and soft tissue realignment is extensive and involves the close co-operation of a hand therapist with experience in splinting as well as physical exercise techniques (Stanley 1992).

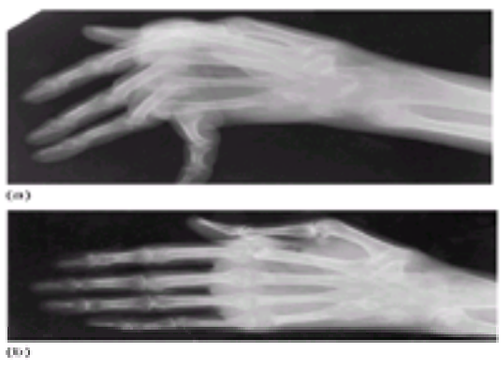


Fig. 3 Rheumatoid hand (a) before and (b) after silastic metacarpophalangeal joint replacements.

Cervical spine

Rheumatoid arthritis can affect all the synovial joints in the cervical spine including the bursa between the odontoid process and the transverse ligament. This can result not only in destruction of joints but also in attenuation of ligaments and consequent instability of the cervical spine (Heywood *et al*. 1988). The commonest site for involvement in the cervical spine are:

1. atlantoaxial subluxation or anterior subluxation of C1 on C2;
2. atlantaoccipital impaction, also known as cranial settling;
3. subluxation of the cervical vertebras below C2.

Surgery is not indicated in all cases of cervical instability but it is very important that instability is fully investigated with flexion and extension radiographs, a full neurological examination, and, where there are any neurological signs, a magnetic resonance imaging scan. Neurological impairment can be caused not only by the mechanical effect of the instability but also by pressure from proliferation of soft tissues.

The nature of surgery indicated depends on the pathology demonstrated. If there is progressive instability then prophylactic stabilization can be performed to prevent later deterioration. There are several methods of stabilization of the cervical spine including bone grafting and fixation with plates or preformed wire rectangles or

loops (Ranawat *et al*. 1979; Ransford *et al*. 1986). Stabilization is also indicated when there is intermittent or persistent neurological deficit. In these cases it may also be necessary to decompress the spinal cord or nerve roots (Johnston and Kelly 1990).

Foot and ankle

In the foot, as in other areas, meticulous attention to the history and clinical signs is essential. Pain may arise from impingement or entrapment of tendons or nerves, or from joints that are unstable not just destroyed. The judicious use of soft tissue procedures, selective arthrodeses, and joint arthroplasty may maintain fore- and hindfoot function for some years. When destruction and deformity have progressed, it is still usually possible to salvage a painless, albeit flat, foot by forefoot arthroplasty and triple arthrodesis of the hindfoot.

Synovectomy and soft tissue procedures

While little is published regarding synovectomy alone in the foot, there is some evidence that early synovectomy may delay deterioration in the ankle as in other joints (Hecker *et al*. 1982; Kvien *et al*. 1987; Mohing *et al*. 1982). Pain around the lateral malleolus in the valgus hindfoot may be due to impingement of the lateral malleolus, or to peroneal tendon entrapment. Injection of the tendon sheath is a simple diagnostic and therapeutic procedure that may give relief for some years. A conservative approach to the foot and ankle may also include early decompression of the tibialis posterior tendon and other medial structures to prevent rupture and subsequent deformity (Cracchiolo 1984).

Excision arthroplasty

Damage to the foot, and especially the forefoot, occurs very early in rheumatoid arthritis. A painful and deformed forefoot with hallux valgus, dorsal dislocation of the lesser metatarsophalangeal joints together with significant synovitis is the rule. The buttonholing of the metatarsal heads through the plantar fascia makes walking exquisitely painful. The surgical goal is a foot that is comfortable for slow pedestrian life, as part of the management of a systemic condition. Forefoot arthroplasty is the procedure of choice: the metatarsal heads are resected, allowing the hammering toes to drop down into the plane of the foot. The hallux valgus may be corrected by excising the metatarsal head or the base of the proximal phalanx (Fowler 1959; Gainor *et al*. 1988; Kates *et al*. 1967).

Arthrodesis

Isolated painful destruction of the ankle joint in rheumatoid arthritis is best treated by arthrodesis (Fig. 4). Several methods of internal fixation are available (Dent *et al*. 1993; Holt *et al*. 1991; Iwata *et al*. 1980; Moran *et al*. 1991). These have a lower complication rate and are better tolerated than the traditional method using Charnley external fixation clamps which have a significant infection rate (Moeckel *et al*. 1991).

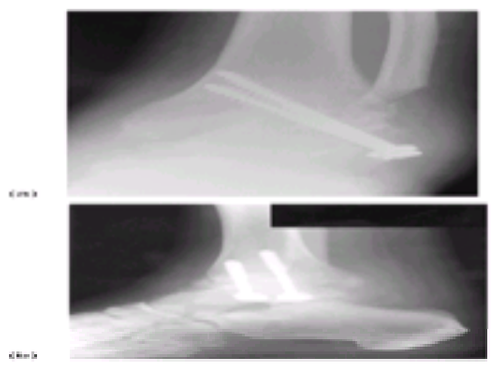


Fig. 4 Anterioposterior (a) and lateral (b) radiographs showing an ankle joint arthrodesed using parallel screws. The subtalar joint is normal.

More commonly the significant valgus deformity of the hindfoot is accompanied by pantalar arthritis. In these circumstances triple or pantalar arthrodesis is the treatment of choice (Vahvanen 1967). In the badly damaged hindfoot, a nail may be inserted from the sole up into the tibia, with or without the use of bone graft. A painless ankylosis can usually be obtained (Stone and Helal 1991). Talonavicular joint destruction with valgus deformity should also be corrected early and before the forefoot, if symptomatic: it may reduce symptoms in the foot by reducing the valgus shearing forces (Cracchiolo 1984).

Joint replacement

While ankle joint replacement is technically possible, the clinical results are disappointing compared to fusion (Jensen and Kroner 1992; Stauffer 1977; Wagner 1982). It may offer some early advantage in elderly patients with low physical demands, but these benefits are outweighed by the complication rate and difficulty in salvage (McGuire *et al*. 1988).

Prosthetic replacement of the first metatarsophalangeal joint has its advocates but the long-term results are not good enough to justify the additional complications (Hasselo *et al*. 1987).

Knee

Synovectomy

Arthroscopic synovectomy can effectively abolish the pain and stiffness in knees with a florid synovitis that is refractory to conservative measures (Ogilvie-Harris and Basinski 1991). It has considerable advantages over open synovectomy, in terms of hospital stay and morbidity, while the results of the two methods appear comparable with both groups regaining an average of 75° of movement (Matsui *et al*. 1989). This improvement is maintained for about 5 years although in that time the radiographic appearances continue to deteriorate (Paus and Dale 1993), but even at 14 years 67 per cent have remained in remission from the inflammatory element of their condition (Ishikawa *et al*. 1986). Simpler arthroscopic lavage and debridement may provide temporary relief but nothing more than this.

Joint replacement

The stiff and painful knees of the inflammatory arthropathies are best treated by total knee replacement. Double osteotomies were used in the past with varying effects, but excision of the entire articular surface and a thorough synovectomy is possible during total knee replacement reducing the rate of reactivation following the operation (Low *et al*. 1994). There is a considerable biological advantage to this removal of all antigenic stimulation as well as the great advantage of excellent mechanical function.

Prosthetic design

The current designs allow only the joint to be resurfaced with a minimum of bone resected. The great advantage this has over the hinge replacements, such as the Stanmore, is that of bone stock preservation. By resecting as little bone as possible, the prosthesis loads the juxta-articular cancellous bone, preventing stress shielding and subsequent bone loss. Ligaments may be spared if present. In rheumatoid arthritis where they are invariably absent, their absence may be accommodated by constraining the knee with a moulded polyethylene insert (Fig. 5). This will give the stability needed for a normal gait without the problems of loosening, fracture, and massive loss of bone stock encountered by the hinge knees of the 1970s and 1980s.

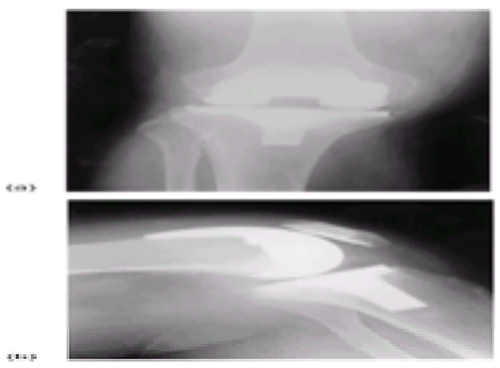


Fig. 5 Anterioposterior (a) and lateral (b) radiographs showing a cemented total condylar knee replacement. The patella has been resurfaced.

The present condylar type knee replacements have had significant problems with excessive wear of the tibial and patellar insert followed by secondary loosening. Many prostheses have been withdrawn from use following relatively unsuccessful trials (Kim and Oh 1995) and their failure rates vary considerably (Knutson *et al*. 1994).

Cemented or uncemented

There is little to choose between cemented and uncemented in total knee replacement at present although loosening seems slightly less in the tibial component if cemented (Knutson *et al*. 1994).

Outcome studies

Total knee replacement has evolved over the last three decades into a procedure that rivals total hip replacement for reliability and safety. In the Swedish knee arthroplasty register, which holds the records of over 30 000 knee replacements since 1976, the revision rate at 5 years has fallen from 10 per cent to 3 per cent (Knutson *et al*. 1994). Despite the softened bone and often deformed joints, the success rate in rheumatoid arthritis is just as gratifying as in osteoarthritis, with no significant difference in outcome (Briggs and Augenstein 1995; Hsu *et al*. 1995).

Complications

Aseptic loosening

This remains the principle problem in the long term. Poor surgical technique resulting in malalignment of the prosthesis is a significant cause of early loosening (Harvey *et al*. 1995).

Other knees loosen after excessive wear of the tibial insert. This leads to eccentric movement and abnormal loading. Early designs of replacement prevented normal joint motion and the constraints themselves caused the abnormal loads and early failure (Rickhuss *et al*. 1994). All knee replacements now allow some rotation as well as unlimited flexion, and this freedom from constraints has been a major factor in preventing early failure. The results of revision knee replacements were very disappointing when the initial prosthesis was a hinge type joint. The huge loss of bone stock led to rapid loosening once again, and many patients in the end faced excision arthroplasty or amputation (Ahlberg and Lunden 1981). Today, the more conservative joint resurfacing procedures have a lower revision rate and last longer so while the annual revision rate is still rising, the percentage of operations that are revisions has actually fallen over the last decade.

Arthrodesis

Joint replacement has a good track record now and the operation of knee fusion is only indicated as a salvage procedure in osteoarthritis affecting few joints. If properly performed, the mechanical function of the individual can be excellent (Behr *et al*. 1985; Figgie *et al*. 1987). If many joints are affected, the impact of knee fusion is so mechanically serious that this should not be considered.

Excision arthroplasty

Failed revision arthroplasty may occasionally result in a flail leg without a functioning joint but with so little bone stock that an arthrodesis would be difficult or impossible. This is a procedure for a very low level of activity and only really appropriate if many other joints are involved and repeated infection or poor soft tissue cover cause revision surgery to fail (Adam *et al*. 1994).

Hip

When the hip has been damaged by an inflammatory arthropathy, the principal surgical intervention is joint replacement. Arthroscopy and arthroscopically assisted synovectomy are technically difficult and of limited benefit. Corrective osteotomies around the hip are rarely appropriate, as the primary pathology is not mechanical. The possible beneficial effects of osteotomy in inflammatory arthropathy, which are poorly understood and inconsistent, been superseded by the more reliable effects of joint replacement.

Synovectomy

Synovitis in the hip although present in up to 40 per cent of patients as shown on ultrasound, does not correlate well with clinical findings. Many of the joint with florid synovitis on ultrasound will have few hip symptoms while other symptomatic hips have little synovial thickening (Eberhardt *et al*. 1995). The hip itself is not simply accessible for open surgery and carries with it a high risk of avascular necrosis of the femoral head. Open synovectomy requiring hip dislocation has been reported in younger patients with some success (Albright *et al*. 1975) but is not common practice. Arthroscopically assisted synovectomy has also been reported (Gondolph-Zink *et al*. 1988) but this is a difficult procedure with limited application.

Joint replacement

Total hip replacement has specific problems related to each of the inflammatory arthropathies, and their biological manifestations. Ankylosing spondylitis causes progressive ankylosis that may continue after the operation, while rheumatoid arthritis sufferers will usually have very poor bone stock and may have eroded the acetabulum. While these specific problems may make the technical aspects of the replacement demanding, the procedure is as successful for patients with rheumatoid arthritis as those with osteoarthritis. Despite being a major operation, involving pain and 10 days in hospital, joint replacement remains the most important intervention in a rheumatoid patient's disease process and, by their perception, well ahead of methotrexate and early aggressive management (Fries 1988).

Prosthetic design

Numerous designs of prosthesis exist, with little to recommend one over another. None has performed as well as John Charnley's original design (Wroblewski 1986) but several others have a proven record such as the Exeter (Fowler *et al*. 1988).

Cemented or uncemented?

Aseptic loosening remains the major cause of failure in total hip replacement. Improvements in cementing technique have reduced the rate at which early signs of loosening now appear, but the erosion of bone by the loosening process remains a concern. Various surface treatments have been used to attempt to stabilize the

prosthesis-bone interface without the use of polymethyl-methacrylate. Extensive use of porous surfaces coating the prostheses have failed to demonstrate any advantage over cement in terms of overall survival and symptom control. There may, however, be some improvement in bone stock, with less loss of bone mass owing to stress shielding. Hydroxyapaptite coating has been available for 8 years and seems very promising (Fig. 6). There may be a significant improvement in bone mass in the uncemented hydroxyapaptite coated group, making the revision a more successful operation, but this is not yet proven. The uses of a cemented stem and an uncemented cup, a 'hybrid' hip replacement, is another acceptable compromise.

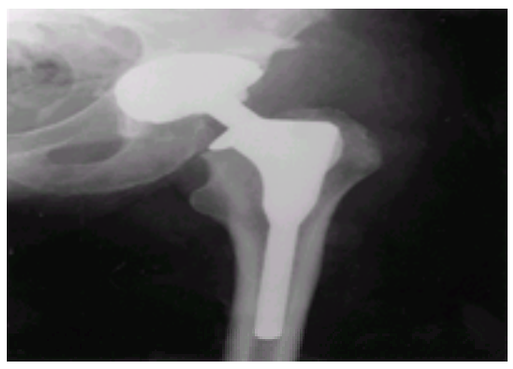


Fig. 6 A hip 1 year after hydroxyapatite-coated, uncemented total hip replacement. Radiolucent lines are visible indicating solid fixation.

Biomaterials

Polyethylene wear particles from the artificial joint have been implicated in the loosening process. Small particles excite an inflammatory reaction causing erosion of bone, visible on radiographs as radiolucent lines. Improvements in the density of the polyethylene and the smoothness of the cobalt-chrome alloy femoral head have reduced this. Ceramic bearings have theoretically superior wear characteristics and should further reduce the volume of particulate debris. Long-term clinical results are not available.

Outcome

Rheumatoid arthritis sufferers consider joint replacement to be the most successful and significant intervention in their disease, being rated above any of the pharmacological therapies. Total hip replacements is certainly reliable and durable; 96 per cent of Charnley cemented total hip replacements were reported as being good to excellent at 15 to 21 years (Wroblewski 1986) in Charnley's own unit, while at the Mayo clinic 79 per cent were good or excellent at 15 years (Kavanagh *et al*. 1989). The Exeter hip has also a good record with 5.5 per cent needing revision at 13.5 years (Fowler *et al*. 1988).

In rheumatoid arthritis the reduction in bone density seems to be offset by the low level of exercise and most authors find the outlook worse for those patients with rheumatoid arthritis. One single-arm study showed a 91 per cent chance of the hip still functioning well at 11 years in rheumatoid arthritis (Severt *et al*. 1991).

Complications of total hip replacement

Infection

The systemic administration of prophylactic antibiotics, together with ultra clean air theatres have made the infection rate in primary hip replacement very low with figures of less than 1 per cent. However, in rheumatoid arthritis the rate is higher, at around 3 per cent (Severt *et al*. 1991), owing to immunocompromise and poor skin healing. The infected total hip replacement can often be salvaged by extensive debridement, and one or two stage revision followed by long-term antibiotics. In the long term however, repeated infection will lead to an excision arthroplasty.

Aseptic loosening

Hips may loosen for a number of reasons:

1. excessive wear leading to eccentric motion of the head and thus high peakloads;
2. stress shielding of the bone leading to resorption of the proximal bone and loss of bone stiffness;
3. faulty technique;
4. aggressive granulomatous reaction.

If the interface between the implant and bone is not solid, progressive motion will cause slow resorption of bone and an increasing zone of soft tissue between the implant and the bone. The painful total hip replacement should be revised early rather than left for as long as possible, as the enlarging radiolucent line around a prosthesis on radiographs represents resorption of bone and thus progressive weakening of the remaining bone stock. The results of revision surgery are fair: most people are helped considerably but the life expectancy of the revised hip is less than for the primary total hip replacement. A 10 per cent failure rate at 5 years would be average (Lord *et al*. 1988).

For this reason, cementless revisions with porous ingrowth or hydroxyapaptite coating is being tried in many centres. Bone graft may be needed, either in the shape of cancellous chips from the pelvis or allograft from femoral heads harvested during primary total hip replacements. These are used to augment the failing bone stock of either proximal femur or acetabulum. Inevitably, if the process of joint replacement is started young, and the patient has an oligoarthropathy, with a relatively active lifestyle, then repeated revisions will culminate in excision arthroplasty. Without a hip joint an otherwise fit person will walk with two sticks, but a typical polyarthritic rheumatoid arthritis sufferer will become comfortable but wheelchair bound.

Arthrodesis

Arthrodesis of the hip is the operation of choice in active young people with a single very painful and stiff joint. It will give good service for 20 years and may be revised safely. It does not have a role in polyarticular disease, where the change in biomechanics increases the stress to the knee joint unacceptably.

Excision arthroplasty

Excision arthroplasty of the hip is not necessary in uncomplicated IA. While it will give considerable relief of pain and will not deteriorate, it is only indicated where infection or very poor bone stock make further reconstructive attempts unwise. While the range of motion is excellent, the power is minimal and walking without two sticks is difficult, even for an otherwise fit young person with no other mechanical disability.

Chapter References

Adam, R.F., Watson, S.B., Jarratt, J.W., Noble, J., and Watson, J.S. (1994). Outcome after flap cover for exposed total knee arthroplasties. A report of 25 cases. *Journal of Bone and Joint Surgery*, **76**, 750–3.

Ahlberg, A. and Lunden, A. (1981). Secondary operations after knee joint replacement. *Clinical Orthopaedics and Related Research*, **156**, 170–4.

Albright, J.A., Albright, J.P., and Ogden, J.A. (1975). Synovectomy of the hip in juvenile rheumatoid arthritis. *Clinical Orthopaedics and Related Research*, **77**, 48–55.

Behr, J.T., Chmell, S.J., and Schwartz, C.M. (1985). Knee arthrodesis for failed total knee arthroplasty. *Archives of Surgery*, **120**, 350–4.

Benjamin, A. (1987). Double osteotomy of the shoulder. *Scandinavian Journal of Rheumatology*, **3**, 65.

Benjamin, A., Hirschowitz, D., and Arden, G.P. (1979). The treatment of arthritis of the shoulder. *International Orthopaedics*, **3**, 211–16.

Bennett, W.F. and Gerber, C. (1994). Operative treatment of the rheumatoid shoulder. *Current Opinion in Rheumatology*, **6**, 177–82.

Briggs, J.R. and Augenstein, J.S. (1995). Tricon hybrid total knee arthroplasty: a review of 81 knees followed for 2 to 4 years. *Orthopedics*, **18**, 341–6.

Chamay, A., Delia Santa, D., and Vilaseca, A. (1983). Radiolunate arthrodesis; a factor of stability for the rheumatoid wrist. *Annales de Chirugie de la Main*, **2**, 5–17.

Clawson, M.C. and Stern, P.J. (1991). The distal radioulnar joint complex in rheumatoid arthritis: an interview. *Hand Clinics*, **7**, 373–81.

Coates, C.J., Bolton-Maggs, B.G., and Helal, B.H. (1991). Interpositional arthroplasty in the management of rheumatoid arthritis of the elbow. In *Rheumatoid arthritis surgery of the elbow* (ed. M. Hamalainen and F.W. Hagewa), pp. 52–9. Karger, Basel.

Copeland, S.A. and Taylor, J.G. (1979). Synovectomy of the elbow in rheumatoid arthritis. The place of excision of the head of the radius. *Journal of Bone and Joint Surgery*, **61B**, 69–73.

Cracchiolo, A.D. (1984). Foot abnormalities in rheumatoid arthritis. *Academy of Orthopaedic Surgeons Instructional Course Lecture*, **33**, 386–404.

Dent, C.M., Patil, M., and Fairclough, J.A. (1993). Arthroscopic ankle arthrodesis. *Journal of Bone and Joint Surgery*, **75**, 830–2.

Eberhardt, K., Fex, E., Johnsson, K., and Geborek, P. (1995). Hip involvement in early rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **54**, 45–8.

Ennevaara, K. (1967). Painful shoulder joint in rheumatoid. *Acta Rheumatologica Scandinavica*, **11**, 1–116.

Figgie, H.E.D., Brody, G.A., Inglis, A.E., Sculco, T.P., Goldberg, V.M., and Figgie, M.P. (1987). Knee arthrodesis following total knee arthroplasty in rheumatoid arthritis. *Clinical Orthopaedics and Related Research*, **224**, 237–43.

Fowler, A.W. (1959). A method of forefoot reconstruction. *Journal of Bone and Joint Surgery*, **41B**, 507–13.

Fowler, J.L., Gie, G.A., Lee, A.J., *et al*. (1988). Experience with the Exeter total hip replacement since 1970. *Orthopedic Clinics of North America*, **19**, 477–89.

Friedman, R.J., Thornhill, T.S., Thomas, W.H., and Sledge, C.B. (1989). Non-constrained total shoulder replacement in patients who have rheumatoid arthritis and class 4 function. *Journal of Bone and Joint Surgery*, **71A**, 494–8.

Fries, J.F. (1988). Milestones in rhematologic care (1965–1985). In *Milestones in management: rheumatoid arthritis*, (ed. J.F. Fries). Syntex, Puerto Rico.

Gainor, B.J., Epstein, R.G., Henstorf, J.E., *et al*. (1988). Metatarsal head resection for rheumatoid deformities of the forefoot. *Clinical Orthopaedics and Related Research*, **230**, 207–13.

Gondolph-Zink, B., Puhl, W., and Noack, W. (1988). Semiarthroscopic synovectomy of the hip. *International Orthopedics*, **12**, 31–5.

Grossman, J.F. and Valance, R. (1980). Clinical and radiological features of the shoulder joint in rheumatoid arthritis. *Journal of Bone and Joint Surgery*, **62B**, 116.

Gswend, N. (1991). The case for a linked elbow prosthesis. In *Rheumatoid arthritis surgery of the elbow* (eds M. Hamalainer and F.W. Hagana), Vol 15, pp. 98–112. Basel.

Harvey, I.A., Manning, M.P., Sampath, S.A., Johnson, R., and Elloy, M.A. (1995). Alignment of total knee arthroplasty: the relationship to radiolucency around the tibial component. *Medical Engineering and Physics*, **17**, 182–7.

Hasselo, L.G., Willkens, R.F., Toomey, H.E., *et al*. (1987). Forefoot surgery in rheumatoid arthritis: Subjective assessment of outcome. *Foot and Ankle*, **8**, 148–51.

Hecker, R.L., Furness, I.C., and Gostich, C.M. (1982). Ankle synovectomy: an approach to the rheumatoid ankle. *Journal of Foot Surgery*, **21**, 4–6.

Heywood, A.W., Learmonth, I.D., and Thomas, M. (1988). Cervical spine instability in rheumatoid arthritis. *Journal of Bone and Joint Surgery*, **70A**, 702–7.

Holt, E.S., Hansen, S.T., Mayo, K.A., and Sangeorzan, B.J. (1991). Ankle arthrodesis using internal screw fixation. *Clinical Orthopaedics and Related Research*, **268**, 21–8.

Hsu, R.W., Fan, G.F., and Ho, W.P. (1995). A follow-up study of porous-coated anatomic knee arthroplasty. *Journal of Arthroplasty*, **10**, 29–36.

Ishikawa, H., Ohno, O., and Hirohata, K. (1986). Long-term results of synovectomy in rheumatoid patients. *Journal of Bone and Joint Surgery*, **68A**, 198–205.

Iwata, H., Yasuhara, N., Kawashima, K., Kaneko, M., Sugiura, Y., and Nakagawa, M. (1980). Arthrodesis of the ankle joint with rheumatoid arthritis: experience with the transfibular approach. *Clinical Orthopaedics and Related Research*, **153**, 189–93.

Jaffe, R. and Learmonth, I.D. (1989). Benjamin double osteotomy for arthritis of the glenohumeral joint. *Rheumatology*, **12**, 52–7.

Jensen, N.C. and Kroner, K. (1992). Total ankle joint replacement: a clinical follow up. *Orthopedics*, **15**, 236–9.

Johnston, R.A. and Kelly, I.G. (1990). Surgery of the rheumatoid cervical spine. *Annals of the Rheumatic Diseases*, **49** (suppl. 2), 845–50.

Jolly, S.L., Ferlic, D.C., Clayton, M.L., Dennis, D.A., and Stringer, E.A. (1992). Swanson silicone arthroplasty of the wrist in rheumatoid arthritis: a long-term follow-up. *Journal of Hand Surgery—American Volume*, **17**, 142–9.

Kates, A., Kessel, L., and Kay, A. (1967). Arthroplasty of the forefoot. *Journal of Bone and Joint Surgery*, **49B**, 552–7.

Kavanagh, B.F., Dewitz, M.A., Ilstrup, D.M., Stauffer, R.N., and Coventry, M.B. (1989). Charnley total hip arthroplasty with cement. Fifteen-year results. *Journal of Bone and Joint Surgery*, **71A**, 1496–503.

Kessel, L. and Bayley, I. (1982). Prosthetic replacement of the shoulder joint. *Journal of the Royal Society of Medicine*, **72**, 748–52.

Kim, Y.H. and Oh, J.H. (1995). Evaluation of the anatomic patellar prosthesis in uncemented porous-coated total knee arthroplasty: seven-year results. *American Journal of Orthopedics*, **24**, 412–9.

Knutson, K., Lewold, S., Robertsson, O., and Lidgren, L. (1994). The Swedish knee arthroplasty register. A nation-wide study of 30,003 knees 1976–1992. *Acta Orthopaedica Scandinavica*, **65**, 375–86.

Kvien, T.K., Pahle, J.A., Hoyeraal, H.M., and Sandstad, B. (1987). Comparison of synovectomy and no synovectomy in patients with juvenile rheumatoid arthritis. A 24-month controlled study. *Scandinavian Journal of Rheumatology*, **16**, 81–91.

Laine, V.A.I., Vainio, K.J., and Pekanmaki, K. (1954). Shoulder affections in rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **13**, 157–60.

Lettin, A.W.F., Copeland, S.A., and Scales, J.T. (1982). The Stanmore total shoulder replacement. *Journal of Bone and Joint Surgery*, **64B**, 47–51.

Lord, G., Marotife, J.-H., Guillamon, J.-L., and Blanchard, J.-P. (1988). Cementless revisions of failed aseptic cemented and cementless total hip arthroplasties. *Clinical Orthopaedics and Related Research*, **235**, 67–74.

Low, C.K., Tan, S.K., Satku, K., and Kumar, V.P. (1994). Reactivation of rheumatoid arthritis in knees following total knee replacement. *Annals of Academic Medicine, Singapore*, **23**, 887–90.

Matsui, N., Taneda, Y., Ohta, H., Itch, T., and Tsuboguchi, S. (1989). Arthroscopic versus open synovectomy in the rheumatoid knee. *International Orthopaedics*, **13**, 17–20.

McGuire, M.R., Kyle, R.F., Gustilo, R.B., and Premer, R.F. (1988). Comparative analysis of ankle arthroplasty versus ankle arthrodesis. *Clinical Orthopaedics and Related Research*, 174–81.

Moeckel, B.H., Patterson, B.M., Inglis, A.E., and Sculco, T.P. (1991). Ankle arthrodesis. A comparison of internal and external fixation. *Clinical Orthopaedics and Related Research*, **268**, 78–83.

Mohing, W., Kohler, G., and Coldewey, J. (1982). Synovectomy of the ankle joint. *International Orthopaedics*, **6**, 117–21.

Moran, C.G., Pinder, I.M., and Smith, S.R. (1991). Ankle arthrodesis in rheumatoid arthritis. 30 cases followed for 5 years. *Acta Orthopaedica Scandinavica*, **62**, 538–43.

Neer, C.S. II, Watson, K.C., and Stanton, F.J. (1982). Recent experience in total shoulder replacement. *Journal of Bone and Joint Surgery*, **64A**, 319–37.

Ogilvie-Harris, D.J. and Basinski, A. (1991). Arthroscopic synovectomy of the knee for rheumatoid arthritis. *Arthroscopy*, **7**, 91–7.

Ogilvie-Harris, D.J. and Wiley, A.M. (1986). Arthroscopic surgery of the shoulder. *Journal of Bone and Joint Surgery*, **68B**, 201–7.

Paus, A.C. and Dale, K. (1993). Arthroscopic and radiographic examination of patients with juvenile rheumatoid arthritis before and after open synovectomy of the knee joint. A prospective study with a 5-year follow-up. *Annales Chirurgiae et Gynaecologie Fenniae*, **82**, 55–61.

Petersson, C.J. (1986). Painful shoulders in patients with rheumatoid arthritis. *Scandinavian Journal of Rheumatology*, **15**, 275–9.

Porter, B.B., Richardson, C., and Vianio, K. (1974). Rheumatoid arthritis of the elbow: the results of synovectomy. *Journal of Bone and Joint Surgery*, **56B**, 427–37.

Posner, M.A. and Ambrose, L. (1991). Excision of the distal ulnar in rheumatoid arthritis. *Hand Clinics*, **7**, 383–90.

Pritchard, R.W. (1991). Total elbow joint arthroplasty in patients with rheumatoid arthritis. *Seminars in Arthritis and Rheumatism*, **21**, 24–9.

Ranawat, C.S. *et al*. (1979). Cervical spine fusion in rheumatoid arthritis. *Journal of Bone and Joint Surgery*, **61A**, 1003.

Ransford, A.O., Crockard, H.A., Pozo, J.L. *et al*. (1986). Craniocervical instability treated by contoured loop fixation. *Journal of Bone and Joint Surgery*, **68B**, 173–7.

Rickhuss, P.K., Gray, A.J., and Rowley, D.I. (1994). A 5–10 year follow-up of the Sheehan total knee endoprosthesis in Tayside. *Journal of the Royal College of Surgeons, Edinburgh*, **39**, 326–8.

Severt, R., Wood, R., Cracchiolo, A., and Amstutz, H.C. (1991). Long-term follow-up of cemented total hip arthroplasty in rheumatoid arthritis. *Clinics in Orthopedics*, **265**, 129–36.

Seyfer, A.E. (1993). Indications for upper extremity surgery in rheumatoid arthritis patients. *Seminars in Arthritis and Rheumatism*, **23**, 125–34.

Skues, M.A. and Welchew, E.A. (1993). Anaesthesia and rheumatoid arthritis. *Anaesthesia*, **48**, 989–97.

Souter, W.A. (1990). Surgery of the rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **49** (suppl 2), 871–82.

Stanley, J.K. (1992). Conservative surgery in the management of rheumatoid disease of the hand and wrist. *Journal of Hand Surgery—British Volume*, **17**, 339–42.

Stauffer, R.N. (1977). Total ankle joint replacement. *Archives of Surgery*, **112**, 1105–9.

Stone, K.H. and Helal, B. (1991). A method of ankle stabilization. *Clinical Orthopaedics and Related Research*, **268**, 102–6.

Toledano, B., Terrono, A.L., and Millender, L.H. (1992). Reconstruction of the rheumatoid thumb. *Hand Clinics*, **8**, 121–9.

Vahvanen, V.A. (1967). Rheumatoid arthritis in the pantalar joints. A follow-up study of triple arthrodesis on 292 adult feet. *Acta Orthopaedica Scandinavica*, (Suppl 107), 3–80.

Wagner, F.W., Jr (1982). Ankle fusion for degenerative arthritis secondary to the collagen diseases. *Foot and Ankle*, **3**, 24–31.

Wroblewski, B.M. (1986). Fifteen to twenty-one year results of the Charnley low friction arthroplasty. *Clinics in Orthopedics*, **211**, 30–5.

Wynn Parry, C.B. and Stanley, J.K. (1993). Synovectomy of the hand (review). *British Journal of Rheumatology*, **32**, 1089–95.

6.2 Surgery in children

Malcolm Swann

Introduction

Surgery now has a firmly established place in the management of juvenile chronic arthritis and in the rheumatic diseases of childhood (Arden and Ansell 1978; Swann 1983a; Rydholm 1990). This chapter is principally concerned with surgery of juvenile chronic arthritis and includes the scope and indications, problems and outcome. The number of children being offered surgical relief varies considerably not only between different countries but between different units within a country (Granberry and Brewer 1978). Much will depend on the attitude of the rheumatologist about seeking such help. It will then depend upon the immediate availability of an orthopaedic surgeon with an expert team conversant with the particularly challenging problems that this condition presents. Experience drawn from the Medical Research Council unit at the Canadian Red Cross Hospital at Taplow and at Wexham Park Hospital, Slough (England) forms the basis for the conclusion drawn here (Ansell and Swann 1983). There are some 5000 patients with juvenile chronic arthritis on the records and of these, 10 per cent have had a surgical procedure and many of them multiple procedures. In this surgically treated group, half have seropositive arthritis.

Surgery should be seen as complementing the other forms of treatment. Thus all the patients must have full exposure to conservative treatment including medication, physiotherapy, and splintage. At some stage it may become evident that these methods are failing and that further help must be sought. The clinician has a difficult task in managing this condition, whose cause is unknown and whose natural course is unpredictable both in the patient in general and in the individual joints. There are no markers to indicate when the disease will undergo remission or finally burn out. The problem is compounded by the unpredictable response to medication, which in itself may have side-effects; by physiotherapy, which can be painful if inappropriate; and by structural changes, which cannot be overcome by conservative means. For instance, splints need to be moulded, comfortable, and changed frequently if there is any alteration in a position. Thus, in our present state of knowledge, some problems cannot be surmounted by conservative means alone and surgical intervention will be required.

Many centres lack the specific expertise and team required for the disciplined and perhaps intensive approach that is necessary to overcome the many problems in treating these children. Failure of this vigilant guidance and its application, and in some cases failure of compliance by the patient (or their parents) may also result in unsuccessful therapy, with the establishment of loss of movement and deformity. Surgical help may be required in such cases in order to recover lost ground. There can be no doubt that the management of juvenile chronic arthritis is best supervised in special centres (Woo *et al*. 1990).

The main reasons for undertaking orthopaedic surgery are relief of pain, correction of deformity, and to overcome loss of motion. The last of these reasons applies not only to the individual joints but to the patient as a whole. These principles apply in particular to the surgical management of juvenile chronic arthritis. There are, however, two principles on which this surgical practice is based. First, the patients must have the opportunity to undergo corrective procedures early in the course of the disease in order to prevent the development of deformity and loss of movement. No surgeon would wish to be presented with a patient in a wheelchair who he or she is seeing for the first time with no options but to advise the total replacement of a destroyed joint. Second, the selection of a particular operative procedure is based on a thorough understanding of the cause of deformity, lack of movement and pain. A more detailed analysis of this will be made later but an example is a misshapen overgrown patella, stuck by fibrous adhesions to the underlying femoral condyle. Here there is deformity and loss of movement where medication, physiotherapy, and splintage may be fruitlessly and painfully applied until the adhesions are ultimately divided by operation.

The first of these principles can only be applied if the surgeon is invited to see the patient early in the course of the disease. Ideally, joint consultation should take place in the outpatients department, and on ward rounds so that cases can be presented and discussed, and a plan of management drawn up. The team must include an experienced physiotherapist, an occupational therapist, and nursing staff and others to reflect the special needs of children with arthritis. The second of these principles can only be applied after full investigation and appreciation of the precise anatomical, physiological, and pathological changes that have occurred (Swann 1987).

Contraction, contracture, and spasm

A number of loose descriptive phrases used in the clinical assessment of patients are often misleading. Muscle spasm is associated with any painful condition in the musculoskeletal condition and in essence serves to prevent movement that is painful, holding the joint in a position of comfort and maximum capacity. There is a contraction of the muscles but there is not a contracture. If the pain can be relieved by anaesthetic or other means, the affected joint will be seen to have a full range of movement. A contracture of muscles does occur after its denervation, when there is true replacement of muscle by fibrosis, or in conditions such as arthrogryphosis. The replacement of muscle with fibrous tissues does not occur in arthritis. There is, however, shortening of the muscle fibres and in established cases, it may take some while for these to be stretched out completely by appropriate treatment, including physiotherapy. However, contracture does occur in the periarticular tissues, such as in the capsule and ligaments, and is particularly noticeable in seronegative arthritis. This reaction may be so severe that it appears to be part of the primary pathology of the disease. The two states are interrelated, for a continuation of muscle spasm, either voluntary or involuntary, will allow a periarticular contracture to become established, with the shortening of the muscles and tissues on the concave side compounded by the wasting and stretching of the muscles and tissues on the convex side. The latter is particularly noticeable in relation to the quadriceps muscles at the knee.

If one considers a knee the clinical record may state, 'flexion range 30 to 90 degrees' or alternatively, 'flexion contracture 30 degrees'. This reading may be totally misleading as it takes no account of the pathology causing the loss of movement; nor is there usually standardization of the conditions under which such a record is made. The issue of the cause of a flexion contracture has been discussed and it should also be appreciated that a knee with a true range of movement under anaesthetic of, say, 30 to 90° has both an extension contracture as well as a flexion contracture, that is, it will neither bend fully nor straighten. If, however, the loss of movement is because of other causes such as pain, for example, the range will depend on many factors including the disease activity, and the level of analgesia and medication at the time (joints being noticeably stiffer on the morning that alternate-day steroids are due), and not least the confidence and co-operation of the patient.

Surgical intervention fits into this picture as being absolutely indicated in the presence of an established, true contracture.

Selection of surgical methods

This can be decided after a thorough analysis of the specific cause of the problem that is being addressed. Thus we need to examine the issues as to why a particular joint may be painful, why it is deformed and why it loses movement (Fig. 1). In order to appreciate this an understanding of the pathology is required in terms of both the intra- and periarticular aspects. There is also a clear distinction between loss of movement and stiffness. It is thus not only the arthritic process within the joint but its effect on lack of mobility, muscle wasting, osteoporosis, general debility of the patient, and many other factors that determine the outcome.

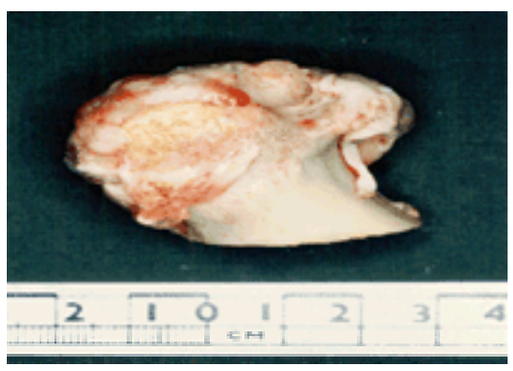


Fig. 1 The femoral head of a 14-year-old patient with juvenile chronic arthritis removed at the time of arthroplasty.

In general the cause of deformity, loss of movement, and pain follows a particular pattern. Pain in a joint is often associated with effusion so the patient will derive comfort in holding the position of maximum capacity, while attempts at movement are countered by muscle spasm. In time, a contracture of the periarticular tissues with shortening of the muscles on the concave side and lengthening of the muscles on the convex side of the joint occurs. Intra- and periarticular ligaments are affected and may be destroyed or fibrosed and together with intra-articular destruction of a joint surface, subluxation is often observed. Meanwhile this direct damage to the joint leads to loss of cartilage, and bone destruction and collapse, and this is sometimes accompanied by avascular necrosis and finally fibrous ankylosis and bone ankylosis. The greatest tendency to contracture is seen in the patients with seronegative arthritis, particularly those with the so-called dry-type arthritis. Loss of movement in a joint means that some areas of articular cartilage are never normally opposed and are thus denied the intermittent compression necessary for their normal nutrition. At the same time, opposed surfaces held under continuous pressure will also lack the normal stimulation to nutrition and in both cases degeneration of the hyaline cartilage follows (Salter and Field 1960).

These patients not only suffer a generalized retardation of growth but also suffer local defects of growth. Hypertrophy, irregular growth, or premature fusion of an epiphysis occur and are probably the result of the local increased blood supply. Frequently these children are unable to walk because of the pain caused by polyarthritis or because of their constitutional illness. This lack of activity contributes to the failure of joints, particularly of the hip, to develop properly. Bone porosis and muscle weakness add to the problem.

A range of surgical treatment is available and designed to tackle the problem at various stages of this pathological process (Fig. 2). Thus conservative measures including medication, splintage, and physiotherapy are adequate in the mildly affected, soft-tissue operations are indicated at the stage of early contractures, and operations on bones and joints (where even an arthroplasty may be recommended) in those more severely affected.

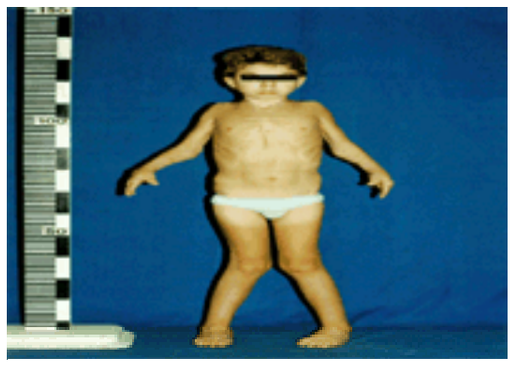


Fig. 2 Photograph of a 14-year-old boy who developed systemic-onset juvenile arthritis aged 6 years. Note the flexion of hips, flexion and valgus deformity of the knees, femoral internal rotation and tibial external rotation, and varus deformity of the feet. The cause of these deformities is analysed in the text. Some upper-limb problems are also apparent. He has since had bilateral soft-tissue releases of the hips followed later by total hip replacements, bilateral supracondylar osteotomies of the femurs, and a wedge tarsectomy.

Any synovial joint can be affected by inflammatory change but it is the knee and hip that are most commonly affected. Emphasis on the surgical management will therefore be directed to these two main joints but the principles outlined apply to other sites.

The hip

Cause of deformity

All the components that cause pain, loss of movement, and deformity are seen when the hip is affected. Pain and swelling of the joint lead to early flexion deformity and some adduction, with restricted movement. The psoas and adductor muscles are particularly dominant in this respect. In time erosion, destruction, sometimes subluxation, and often necrosis of the femoral head are identified and sometimes growth disturbance (Hastings *et al*. 1994). The persistent spasm and increased vascularity lead to a disturbance of normal growth. The capital epiphysis fuses early and there is some broadening and shortening of the femoral neck, while unequal forces may alternatively induce a marked valgus and anteversion. The greater trochanter grows at a normal or accelerated pace and the lesser trochanter is frequently overgrown because of traction from the psoas muscle. The femoral shaft becomes excessively bowed and the lumen of the medullary cavity has its greatest diameter in the sagittal plane.

The attitude assumed by the hip in early-onset arthritis is therefore dominated not only by the joint destruction and lack of proper use but also by the growth derangement brought about by the epiphyses. If there is anteversion of the femoral neck the femoral head tends to sublux anteriorly out of the socket and this leads to the patient assuming a flexed hip position, particularly when standing, in order to ensure stability and comfort. The valgus anteverted neck also means that the abductor lever is reduced and a Trendelenburg or waddling gait will occur.

The clinician must be alert to these causative factors when dealing with a flexion deformity of the hip because if the joint cannot be straightened it must be determined whether this is because of the spasm of the muscles, a soft-tissue contracture, or an abnormality of the bone development. Likewise the factors causing a limp must be sought.

A hip joint unaffected itself by disease can develop deformity secondary to a fixed flexed knee below or a scoliotic spine above, and with the chronicity of this illness

such a hip may develop a fixed deformity itself.

Specimens of femoral heads removed at operation have shown definite evidence of repair by fibrocartilage in patients whose disease is controlled or quiescent. It is therefore important that the hip is maintained in a congruous position with uninhibited motion so that there is adequate moulding if attempts at repair are to be successful.

Surgical management

This is best considered in three stages dependent on the degree of the problem:

Stage 1: hips in which there is almost a full range of movement under anaesthesia and with minimum radiological change.

Stage 2: hips in which there is still some fixed deformity and restriction of full movement under anaesthesia and in which radiographs, arthrograms, computerized tomographic (CT) scans, magnetic resonance imaging (MRI) or ultrasonography confirm the presence of a joint space.

Stage 3: hips lacking movement under anaesthesia with persistent deformity and advanced radiological change.

Stage 1

The establishment of fixed deformities and loss of movement can be aggressively countered at this stage by a well-supervised conservative regimen. These children are admitted to the hospital for full evaluation of their problem. Medication includes analgesics and possibly an intra-articular steroid injection. Skin traction is applied, with the pull toward abduction, and maintained for 24 h a day interspersed with physiotherapy. When the condition permits, prone lying for part of each day may help reduce flexion deformity of the hip. However, it should be appreciated that in a hip which has an effusion, attempts at extension may actually increase the intra-articular pressure and thus the pain. There is some evidence to suggest that a rise of intra-articular pressure will also lead to avascular necrosis (Soto-Hall 1964). Early surgical decompression of these hips is therefore indicated before attempts at extension using traction or prone lying. This will be discussed later. When the hips are affected the knee in particular needs to be watched and may require splints to prevent concomitant flexion. As the patients begin to improve, hydrotherapy should be instituted daily in addition to the traction. Sitting, such as in a wheelchair, is forbidden. When the pain abates, walking with aids can start progressively, with traction maintained at night.

Stage 2

When there is an established flexion contracture, movement is limited even under anaesthesia, and the joint space can be demonstrated, then the situation can be improved by a soft-tissue release operation (Mogensen *et al*. 1983). While complete release operations may sometimes be necessary (Alvarez *et al*. 1992; Witt and McCullough 1994), the tightness of the adductors and psoas is the main obstacle to correction (Swann and Ansell 1986). An open adductor tenotomy can be done and through the same incision the psoas tendon can easily be reached and divided at its attachment to the lesser trochanter.

After the operation the hips are again put on traction in abduction and extension, and the regimen, continued as described for stage 1 except that hydrotherapy, of necessity, is delayed until the wounds are healed. Experience of this procedure in over 100 cases has produced gratifying results that include lessening of the flexion contracture (Fig. 3), an increase in the total range of movement, and a very dramatic relief of pain, probably due to decompression of the joint. Radiographs have shown an improvement in terms of a decreased osteoporosis, a clearer definition of the joint line, and some widening of the joint space in more than half the patients (Fig. 4). The longer-term results at 3 years and more have been shown to mirror the activity of the subsequent disease. Nevertheless, by improving the movement and function of the joint a congruous, contained femoral head may be possible with some repair, so if the disease activity abates a serviceable hip remains. In those cases where the continuing disease activity results in a less favourable outcome a period of relief from pain is usually enjoyed and the structural anatomy is in a better position for hip arthroplasty at a later state.

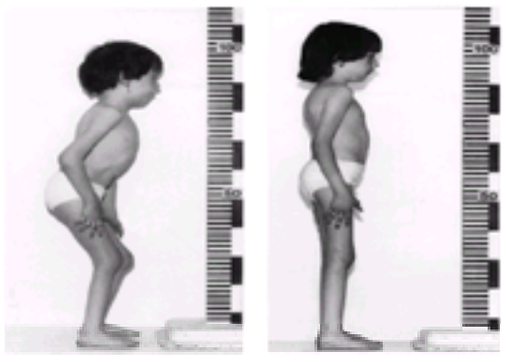


Fig. 3 This patient developed pauciarthritis aged 15 months, which subsequently extended. (a) At age 3 years: appearance after a full regimen of medication, physiotherapy, and splintage; she had active disease in both hips and knees with persistent flexion deformities at these joints. (b) After soft-tissue release of both hips and both knees a few months later—the posterolateral knee scar can just be seen.

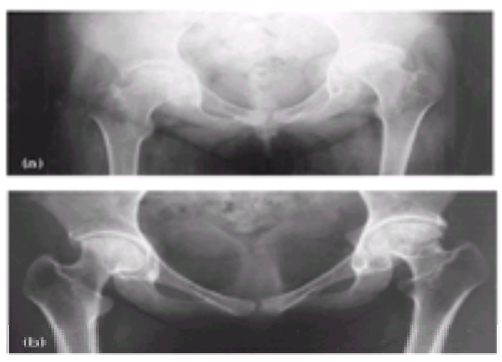


Fig. 4 This patient developed a severe systemic disease followed by polyarthritis aged 4 years. (a) Radiographs of the hips when the patient was aged 13 during an active phase of the disease and immediately before soft-tissue release; there is a marked flexion contracture. (b) The same hips 3 years later following the soft-tissue release: the disease had also become quiescent but the operation permitted mobility to be regained and some healing has occurred; these hips have remained painless and the patient is now 29 years old and walks without a limp.

The role of osteotomy to correct deformity is limited, but occasionally muscle imbalance, joint effusion, and an increased valgus of the femoral neck lead to anterolateral subluxation of the hip and theoretically a varus osteotomy is indicated under these circumstances. However, the porosity of bone will often preclude internal fixation and external immobilization will lead to intractable stiffness. An osteotomy can occasionally be used in a patient whose disease is no longer active and in whom there is good bone stock. Care must be exercised before proceeding with a containment osteotomy because sometimes a coxa magna precludes the femoral head being fully restored to the acetabulum. This containment may also be prevented if there is a mass of thickened synovial tissue in the acetabular floor—a situation which may be identified by MRI. In addition, care in alignment of the shaft is essential to allow the possibility of total arthroplasty at a later date.

Stage 3

Total hip arthroplasty may be indicated in a patient with a severely destroyed joint who may be condemned to a painful existence in a wheelchair. Some of the patients have had to have custom-made prostheses to fit their mini skeletons (Fig. 5). These are time-consuming procedures because of the bleeding and the frailness of the bone, and our long-term studies have shown they carry, as expected, a higher rate of infection and prosthetic loosening than in other conditions (Learmonth *et al*. 1989; Williams and McCullough 1993). A current review of 96 hip replacements with a follow-up of 6 to 18 years (average 10 years) revealed that 25 per cent had failed, and the appearance of translucent lines and hip migration even in the asymptomatic patient suggests that this rate will rise with time (Witt *et al*. 1991). Of the 42 children reviewed in 1986 following total hip replacement (Ruddlesdin *et al*. 1986), four have since died of amyloid disease. Nevertheless, those who have reached late teenage or early adulthood have clearly been transformed by being physically able to integrate better with the community in terms of work, pleasure, and social contact. All the patients and their parents, save one, felt the operation had been worthwhile, despite those that had failed and the need for a revision arthroplasty in a large number of cases. However, modern techniques including the use of bone grafting, have enhanced the success of these more difficult procedures. I feel that despite the overall high failure rate the operation remains well worthwhile in carefully selected cases as the only alternative to a painful wheelchair existence.

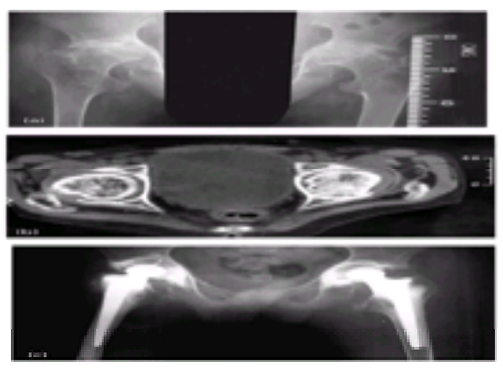


Fig. 5 This patient developed systemic-onset juvenile chronic arthritis aged 7 years. He also suffers from diabetes and myasthenia gravis. (a) Hips at age 16 immediately before surgery: the patient was unable to walk because of the painful disability; note the severe acetabular erosion and the overgrowth of the lesser trochanters; the sunken hips prevent proper abductor function. (b) A CT scan of the same patient used to define the pelvic bone stock available for anchoring the acetabular component of a total hip replacement; note the thin medial wall. (c) The postoperative appearance: note this patient has subsequently had both his knees replaced (see Fig. 6).

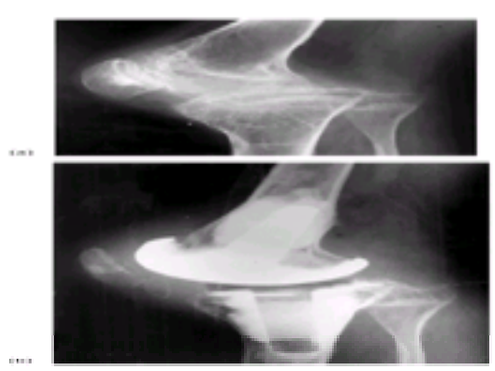


Fig. 6 The same patient as depicted in Fig. 5; after hip arthroplasty further deterioration occurred in both knees. (a) Preoperative radiograph at age 16 years: note the severe degree of osteoporosis, overgrowth of the patella, posterior subluxation of the tibia, and the marked depression of the tibial plateau. An anteroposterior view showed the knee to be in a marked valgus position as a result of the collapse of the lateral tibial plateau but also an overgrowth of the medial femoral condyle. (b) Total knee replacement using an Attenborough prosthesis. The small size, destroyed surface, and osteoporosis are best suited to an unrestrained, stemmed prosthesis with cement. Some newer designs of total condylar replacement are proving satisfactory after adolescence when the disease has become inactive. This patient has now regained mobility having lost the flexion contractures in both hips and both knees, and in each case has a flexion range of 90° or more.

The knee

Cause of deformity

The knee is frequently affected early in several of the subgroups of juvenile arthritis, possibly because of its large synovial surface. When this becomes inflamed the joint is flexed into a position of comfort and maximum capacity, and hamstring spasm will prevent it straightening. The stretched quadriceps muscle rapidly wastes and this may also be a reflex phenomenon associated with the effusion within the joint. This weakness of the quadriceps adds to the inability of the patient actively to straighten the knee. The periarticular structures undergo fibrosis and it is important to note that this includes the broad expanse of the quadriceps expansion on each side because this in turn will limit flexion. The joint thus develops a flexion contracture and an inability to flex so that it loses movement early, i.e. an extension contracture.

These effects combine to lead to a characteristic posterior subluxation of the tibia. The articular cartilage becomes eroded with continuing activity but because this is particularly thick in the younger child, the erosion must be well advanced before significant joint narrowing can be seen. Marked effects are imposed on the growth of the bone because of the hyperaemia and venous hypertension. There is an overgrowth produced by the epiphyseal plate and in patients with only one knee affected, characteristically in pauciarthritis, the limb grows longer than its fellow. The medial side of the femoral growth plate appears to have a propensity to be stimulated in the greatest manner and this is one of the causes of a valgus deformity (see Fig. 2). Valgus may also be related to collapse of the lateral side of the joint, and this may be in response to forces of the adducted hip above and possibly a varus foot below. A complex state develops in these patients, leading to a rotation deformity. Anteversion of the femoral neck produces internal rotation deformity of the femur, which is clinically evident by the squinting patellae. Secondary external rotation of the tibia at the knee plus some torsion of the tibia itself may then begin to develop to keep the axis of the ankle in line with the direction of progress (see Fig. 2). In addition, when a foot affected by the disease has painful articulations and is held in eversion and valgus when the patient walks, this will also induce a torsional force on the tibia. Finally, an iliotibial band contracture causes flexion and external rotation of the knee and contributes to this complex mechanism of deformity. Lateral subluxation of the patella may follow, which in itself may require surgical correction because the line of pull of the quadriceps has thus become mechanically inefficient.

The patella itself often develops overgrowth into a misshapen and mechanically unacceptable form, whilst closure of the patellofemoral joint by fibrous ankylosis may be a cause of a fixed deformity that will certainly require surgical release.

Management

In terms of surgical treatment it is helpful to consider the knee in three stages, depending on the extent of the pathology present.

Stage 1: A knee in which there is a full range of movement under anaesthetic and minimal changes on radiographs.

Stage 2: Knees with some restriction of movement and slight deformity that cannot be corrected under anaesthetic but where a joint space is still present.

Stage 3: Gross loss of movement with a severe deformity and marked radiographic changes.

An overall appraisal of the other joints affected, particularly in the same limb, must be made before beginning treatment. Correcting the flexed deformity of the knee if the ipsilateral hip exhibits a similar problem is useless and probably the hip should be dealt with first. When planning surgery it has to be decided whether the deformity can be corrected by a soft-tissue operation and thus increase the total range of movement, or by osteotomy, leaving the range of movement unchanged. The former is ideal but can only be achieved when the disease activity is controlled and the joint destruction is minimal. an osteotomy realigns the limb by producing a secondary deformity that masks the first. However limited the movement in a knee, it must be possible to extend it. The patient can then stand in comfort, wear a calliper if necessary, avoid quadriceps fatigue, and prevent the induction of a secondary deformity of the hip.

When fixed flexion is considerable, several complications of treatment must be recognized. First, as the leg is extended, traction will occur on the popliteal artery and nerves, and may lead to a vascular embarrassment or neurological damage. Second, passive conservative treatment using traction or plaster may lead to problems such as subluxation of the knee, supracondylar fracture of porotic bone, or anterior compression across the joint when the posterior capsule is unyielding and acts as a hinge at the back. This compression itself may cause necrosis of the cartilage and subarticular microfracture. If the method employed requires prolonged immobilization, further disuse porosis will occur and the articular surface will suffer, as hyaline cartilage requires intermittent compression to maintain effective nutrition.

Treatment

Stage 1

The patient is examined under a general anaesthetic and if this reveals a normal range of movement and no fixed deformity, physiotherapy and rest splints are indicated. If pain is severe, slow correction by reversed dynamic slings is used if the deformity does not exceed 25° but these must not be applied under anaesthetic. Medication can be helpfully reinforced by intra-articular steroid injections (Earley *et al*. 1988). Synovectomy is indicated where there is permanent thickening and effusion in the knee, often leading to overgrowth of the limb. Synovectomy, however, has a limited role in polyarthritis because the patients who might benefit most are often too sick for surgery because of systemic disease. It should be therapeutic rather than prophylactic in a disease whose future progress in an individual patient and an individual joint is unpredictable. Arthroscopic synovectomy is the method of choice if the joint can be distended (Vilkki *et al*. 1991).

Stage 2

Soft-tissue release of the posterior structures may result in an immediate correction of a flexion deformity or be used as a prelude to further serial plasters or reversed dynamic traction (see Fig. 3). It is a satisfactory procedure when there is no concomitant valgus deformity. The operation involves lengthening of the hamstring muscles and division of the posterior capsule of the knee joint right across its axis and around on either side (Clarke *et al*. 1988). It is a difficult procedure and if the tibia is found to be subluxed the contracted cruciate ligament may need to be divided. After the surgery, plaster is applied for 3 days, following which it is bivalved and movement encouraged to overcome the tendency to postoperative stiffness. The plaster shells are kept as rest splints and used for early walking until strength is regained.

If the deformity is fixed and bone destruction is more advanced, then a supracondylar osteotomy is an appropriate procedure, and can be used to correct valgus at the same time. The operation is done through a short medial incision and the porosity of the bone often precludes internal fixation so that a plaster cylinder or similar external splint is applied and weight bearing encouraged at the outset. The plaster is removed as early as possible, usually at about 4 weeks to obviate stiffness. A calliper or alternative support is advised for immediate use as a night splint to maintain correction. The bone can be expected to unite satisfactorily and the position of the leg is maintained in correct alignment by direct inspection. Intensive physiotherapy is required to prevent postoperative stiffness. It is occasionally necessary to repeat the osteotomy if the forces that necessitated it in the first place continue to act, particularly if there is an exceptional increase in growth of the patient.

Epiphyseal stapling has occasionally been used to correct valgus of the knee and some surgeons use this method to arrest excessive longitudinal growth, particularly in cases of pauciarthritis (Rydholm *et al*. 1987).

Stage 3

These knees are often totally adherent, with fibrous adhesions not only between the main articular surfaces but in particular between the patella and the femoral condyle. They are usually associated also with considerable periarticular contracture. For this reason, access to the knee is difficult through an arthroscope as the joint cannot be distended with fluid in the normal fashion in order to get a clear view. Some of these patients can benefit from an extensive surgical approach that divides the adhesions, followed by an immediate postoperative regimen using the continuous passive-motion machine under adequate analgesic cover. However, many joints are clearly so badly destroyed that nothing short of total knee replacement is indicated (Fig. 6). Technological design and improvements in techniques have now produced a generation of knee prostheses that can be manufactured in a smaller size and are suitable for these juvenile patients. The experience being gained with these is encouraging and for any child who has a severely painful joint with deformity and destruction that precludes any other form of treatment, this method must be considered. The indications are similar to those suggested for a hip prosthesis.

The foot and ankle

Studies have shown (Ansell and Wood 1976) that within 1 year of onset of the disease, the ankle joint was affected in 40 per cent of cases, the subtalar and midtarsal in 27 per cent, and the metatarsophalangeal in 9 per cent. At the 15-year follow-up only 0.9 per cent of children felt that the foot was the most limiting factor. The early onset of the disease in the foot leads to growth abnormalities that present clinical features often requiring surgical remedy. The multiple and complex synovial articulations fall prey to the severity and duration of the disease. Problems present as elsewhere, with pain, deformity, and loss of function. Pain is associated with active disease but it must be appreciated that a deformity causing pressure or stiffness may throw abnormal strains on other, possibly unaffected joints. Foot problems may also arise secondarily to those of the knee. Finally, inflamed synovial sheaths or bursae may require attention.

Management

This is best considered under the headings of the forefoot, the hindfoot, and the ankle (Ansell and Swann 1996).

The forefoot

The main indication for surgical intervention in the foot is pain. Deformity *per se* is not an indication to operate unless it is causing pain or loss of function, and many foot problems can be dealt with by simple conservative means, including special shoes, orthoses, splints, or even a calliper. However, when the problem cannot be controlled by these simpler means or if the older patient is reluctant to wear a surgical shoe or an orthosis, surgery is indicated. Specific examples where surgical intervention will, however, be required as a primary source of relief to the patient include significant toe deformities produced by epiphyseal abnormalities, and severe metatarsalgia with clawing of the toes associated with disruption and dislocation of the metatarsophalangeal joints, as seen in seropositive disease. In these patients, severe hallux valgus deformity may occur, which will require a metatarsal osteotomy for realignment of the great toe. The reader should be aware that overall excellent results can be obtained by careful selection of patients for surgery, particularly for instance where a local pressure point in the forefoot or toes is the source of severe pain. The operations include surgical trimming of a bony prominence or irregularity, straightening of the toes, and excision arthroplasty of the metatarsophalangeal joints (Fowler 1959).

The hindfoot

This includes the subtalar, talonavicular, and calcaneocuboid joints. No set pattern of deformity is seen here but the deviation may be into either varus or valgus. Whichever way they go, both the heel and the forefoot usually diverge in the same direction, setting up abnormal foot imprints and consequently secondary strain-pressure areas and callosities. Valgus deformity can be produced by a spasm of the peroneal muscles and this can be a presenting feature of juvenile chronic arthritis, where it is labelled a peroneal spastic flat foot. Valgus is also caused by some collapse of the foot to spare the painful articulations affected by disease, or secondarily to compensate deformity at the knee. Loss of joint congruity and fixation by joint adhesion or contracture will later prevent passive correction. A varus

deformity is particularly disabling because the patient walks on the outer border of the foot (see Fig. 2). Not infrequently this is associated with a valgus deformity of the knee and as such may be partly compensatory, accentuated by taking a full load of body weight. This combined deformity may seriously affect the patient's gait.

The indications for treatment are pain and deformity: however, a deformity that is painless but interfering with the child's gait should be corrected. In the majority of patients, conservative treatment suffices in as much as one is able to obtain a good correction of the deformity, but it is necessary to do this under a general anaesthetic. At the same time an intra-articular injection of triamcinolone is given into the dominantly affected joint and, if the exact location is not clear, it is put into the sinus tarsi. After this, the foot is put into a corrective, plaster-of-Paris walking cast for 4 weeks; the plaster is then removed and the position maintained with a calliper (Mavidrou *et al*. 1991).

Surgical management

In some patients where passive correction even under anaesthetic is not possible, some form of surgical correction will be indicated. In principle this consists of the removal of bone wedges to correct alignment of the foot and fusion of painful joints, in particular the talonavicular joint, which is often the predominant cause of trouble. Realignment of the heel in the presence of a good triple joint can be achieved by an osteotomy of the os calcis.

The ankle joint

The ankle joint itself may undergo erosive disease or be secondarily affected by avascular necrosis of the talus. It may also become mechanically stressed when normal movement is lost at the subtalar or midtarsal joints.

There is sometimes difficulty in identifying the source of pain in the hindfoot region, especially in differentiating between the ankle joint and the subtalar joint. Clinical examination and radiographs may help in identifying the joint at fault. However, the solution is complicated by the fact that a radiologically destroyed and stiff subtalar joint may itself be relatively painless but throw strain on an ankle joint that is not the site of primary disease. It is therefore recommended that a conservative approach is taken for as long as possible, including the use of gentle examination and repositioning of the foot under anaesthetic, injection of triamcinolone, and immobilization in plaster for 4 weeks. If these measures fail and it is necessary to proceed further the surgical options consist of arthroplasty or arthrodesis of the ankle joint.

Soft-tissue lesions around the foot

Tenosynovitis may induce pain, swelling, and spasm; in particular the sheath of the tibialis posterior is most commonly affected. The paratenon and bursae associated with the Achilles tendon are involved in patients with ankylosing spondylitis. Local injections of steroid into the inflamed tendon sheath are helpful but if this method fails, surgical synovectomy may have to be considered.

Intensive conservative treatment of feet affected by disease should preclude the need for surgery in the majority of cases. Nevertheless, conservative treatment may be instituted too late or fail to control the condition, particularly in persistently active disease. Surgery may then have some part to play in correcting the deformity, which is painful and retards walking. With rare exceptions it should only be employed when growth has ceased.

The upper limb

The most common problems that will need to be considered by the surgeon in juvenile chronic arthritis are those in the lower limb, and these have therefore been addressed at some length. However, the principles of surgery in the hip, knee, and foot are also invoked in the management of the upper limb and a short summary of the position with regard to the shoulder, elbow, hand and wrist follows.

The shoulder joint

This tends to be involved in those patients who have severe, generalized disease with prolonged activity. Its onset is often insidious and the patient is usually more concerned with the arthritis that affects the weight-bearing joints. It is when other joints in the upper limb, such as the elbow or wrist, are involved that the patient will become aware of the shoulder.

Shoulder problems are often exacerbated by the need to use crutches, and earlier and more satisfactory relief of problems in the leg joints have been helpful in this respect. There is very little conservative treatment that is effective for the shoulder, although it is worth trying an intra-articular steroid injection. If both shoulders are seriously affected, then everyday tasks such as dressing and feeding become well-nigh impossible and under these circumstances surgical relief may be sought. On the whole, synovectomy is ineffective, although if there is a very big subacromial bursa this can be excised. Osteotomy of the shoulder has an amazingly good outcome in terms of pain relief. This results in better use of the shoulder and scapula, although the actual movement in the glenohumeral joint itself may not be demonstrably increased. Arthrodesis is rarely indicated in a disease where stiffness itself is a problem. However, the present generation of shoulder arthroplasties is showing promise and some surgeons with an interest in this joint are reporting good results.

The elbow joint

This joint is more frequently involved than the shoulder, 45 per cent of the patients in a 15-year follow-up having problems. Assuming a failure of conservative treatment there are a number of surgical procedures that have a significant place in the management of this joint. Recurrent attacks of pain and swelling may derive benefit from a synovectomy; if the radial head is damaged and growth has reached maturity, it too can be removed. Arthrodesis is impractical but arthroplasty has a significant place. After excision of the bone ends, either an interposition of skin or fascia or a prosthetic replacement if enough bone stock is available are both very successful in relieving pain and improving movement and function.

The wrist and hand

These have a multitude of synovial joints so it is to be expected that most patients suffering from juvenile chronic arthritis in one of its forms should at some time have at least one of these joints involved (Harrison 1978). If the synovial sheaths of the flexor and extensor tendons are also included, few patients will escape some manifestation, and if the many growth plates and epiphyseal centres are involved, then the stage can be set for severe disability. Probably nowhere else is better rewarded by early splinting and physiotherapy to prevent deformity and stiffness. There are, however, indications where surgery can be helpful in this condition, but for the most part it is advised that this should be undertaken by a hand surgeon who is conversant with this minefield of problems. Indeed, the systemic illness of a child may militate against surgical intervention and when tendon sheaths are affected their involvement may also regress as the disease itself comes under control. We have found that the use of corticosteroids in persistent disease of the synovial sheath has been particularly helpful; intrathecal injections of hydrocortisone are given under general anaesthesia because they are painful. Occasionally, patients with a persistent thickening from proliferative synovitis in a limited field may benefit from a simple synovectomy followed by rapid mobilization. Extensive synovectomy of the flexor sheaths is never undertaken because it can lead to intractable stiffness. However, limited excision of the synovium of the dorsal sheaths, particularly in seropositive arthritis, does have some place, particularly where there is a danger of tendon rupture. The problems presented by a trigger finger or carpal tunnel compression are dealt with in the conventional fashion with the use of local steroid or, failing response, surgical decompression.

Wrist and carpal joints

There is often early destruction and deformity from epiphyseal involvement at these sites and the outcome is dictated by the degree of disease, age of onset, and the particular extraneous forces about the joint. A flexion deformity may occur early, and forward subluxation and ulnar deviation of the wrist must be countered by vigorous splinting and exercises. If stiffness does occur, one must ensure that it does so in a good position. Subluxation of the distal ulna often causes a painful swelling with some risk of tendon rupture; provided that the distal growth plates are closed, ulna styloidectomy may be indicated. For the severely damaged wrist and carpus, the choice will lie between arthrodesis and arthroplasty. Arthrodesis is well tried and, if indications are right, it can be most successful, mimicking that which nature often attempts. However, careful appraisal of the whole position of the arm must be made, as in some patients who already have stiff shoulders and elbows the retention of a small amount of movement, particularly in radial deviation, may make a disproportionate difference to their quest for independence.

Finger joints

The metacarpophalangeal joints when affected particularly in seropositive arthritis may benefit from synovectomy. If totally destroyed, and there is ulnar drift, a realignment procedure is worth undertaking with the use of some form of arthroplasty. When the problem is as severe in the interphalangeal joints, stabilization is again worthwhile, using a silastic implant.

The spine

The cervical spine is affected in both seropositive and seronegative juvenile arthritis (Swann 1983b), the dorsal spine exhibits crush fractures when steroids have been exhibited, and the lumbar spine takes the brunt of ankylosing spondylitis.

Patients with seronegative juvenile chronic arthritis may present with a painful torticollis and if this is neglected the patient will be left with a permanent disability with the head tilted rigidly on one side. These patients should be given a general anaesthetic, if necessary, in order to relieve the spasm, and whilst the head and neck are held straight a firm collar is applied. This is taken off for daily exercise and then reapplied, for although it will not prevent fusion occurring it does permit it to take place in a more satisfactory position. Seropositive arthritis, on the other hand, leads to instability and it is particularly at the atlantoaxial level where the subluxation or dislocation may occur. There is an absolute indication for surgical fusion at this level if there is a neurological involvement or severe pain or both.

It is of particular importance that the anaesthetist should be aware of these problems in the cervical spine in both types of disease. The rigidity of the neck, associated with a small airway, may make intubation difficult or impossible by conventional means. Subluxation or dislocation at the atlantoaxial level will equally be a hazard because attempts at intubation may lead to cord damage. It is essential that an up-to-date radiograph should be available immediately before the patient is given a general anaesthetic. The use of the fibreoptic laryngoscope and other airway techniques have obviated the worries that existed previously; many patients who are having lower limb surgery can be managed well with spinal or epidural anaesthesia (Smith 1990).

Scoliosis

Structural scoliosis occurs more commonly in patients with juvenile chronic arthritis than in the normal population (Ross *et al*. 1987). This may arise from the postural curves associated with asymmetrical involvement of the lower limb joints causing pelvic tilting. Timely surgical relief of the primary cause in the lower limbs has led to a lessening of the spinal curve. Asymmetrical involvement of the apophyseal joints may also contribute to this problem. In patients with scoliosis, careful appraisal of the underlying pathology should be made, but if there is a need, conventional methods of correction and stabilization can be used.

Fractures

The limited activity of these patients and their inability to participate in sport protects them largely from this problem. Nevertheless they are prone to pathological fractures because of their osteoporosis and the stiffness of their joints. This is particularly so with supracondylar fractures of the knee and underlines the inadvisability of manipulating the joints under anaesthetic. Examination under anaesthetic is permissible but must be attended by the utmost care.

After fractures, immobilization of the cast must be maintained for the minimal time consistent with union to prevent additional stiffness of the adjacent joint. The degree of osteoporosis usually precludes internal fixation of the fragments.

A number of patients have suffered crush fractures of the spine, particularly if they are on corticosteroids. No special treatment is indicated and recovery with remarkable reformation of the vertebrae can be expected.

Summary

All synovial joints are vulnerable to the deformity and functional loss that can occur in juvenile chronic arthritis. The principles outlined in this chapter should be applied intensively in this polyarthritic condition. The prognosis in terms of complete remission for many sufferers makes the treatment all the more demanding and worthwhile so that they may grow into physically functional adults.

Chapter References

Alvarez, M.J.M., Espada, G., Maldonado-Cocco, J.A., and Gagliardi, S.A. (1992). Long term follow up of hip and knee soft tissue release in juvenile chronic arthritis. *Journal of Rheumatology*, **19**, 1608–10.

Ansell, B.M. and Swann, M. (1983). The management of chronic arthritis of children. *Journal of Bone and Joint Surgery*, **65B**, 536–43.

Ansell, B.M. and Swann, M. (1996). Juvenile arthritis. In *Surgery of disorders of the foot and ankle*, vol. 1, (ed. B. Helal, D.I. Rowley, A. Cracchiolo III, and M.S. Myerson), pp. 263–70. Martin Dunitz, London.

Ansell, B.M. and Wood, P.H. (1976). Prognosis of juvenile chronic polyarthritis. *Clinics in Rheumatic Disease*, **2**, 397–412.

Arden, G. and Ansell, B.M. (ed.) (1978). *Surgical management of juvenile chronic arthritis*. Academic Press, London.

Clarke, D.W., Ansell, B.M., and Swann, M. (1988). Soft tissue release of the knee in children with juvenile chronic arthritis. *Journal of Bone and Joint Surgery*, **70B**, 224–7.

Earley, A., Cuttica, G., McCullogh, C., and Ansell, B.M. (1988). Triamcinalone into the knee joint in juvenile chronic arthritis. *Clinical and Experimental Rheumatology*, **6**, 153–9.

Fowler, A.W. (1959). A method of forefoot reconstruction. *Journal of Bone and Joint Surgery*, **41B**, 507–13.

Granberry, G.M. and Brewer, E.J. (1978). The combined pediatric–orthopedic approach to the management of juvenile rheumatoid arthritis. *Orthopedic Clinics of North America*, **9**, 481–507.

Harrison, S.H. (1978). Wrist and hand problems and their management. In *The surgical management of juvenile chronic arthritis*, (ed. G.P. Arden and B.M. Ansell), pp. 161–83, Academic Press, London.

Hastings, D.E., Orsini, E., Myers, P., and Sullivan, J. (1994). An unusual pattern of growth disturbance of the hip in juvenile rheumatoid arthritis. *Journal of Rheumatology*, **21**, 744–47.

Learmonth, I.D., Heywood, A.W.B., Kay, J., and Dall, D. (1989). Radiological loosening after cemented hip replacement for juvenile chronic arthritis. *Journal of Bone and Joint Surgery*, **71B**, 209–12.

Mavidrou, A., Klenerman, L., Swann, M., Hall, M.A., and Ansell, B.M. (1991) Conservative management of the hindfoot in juvenile chronic arthritis. *The Foot*, **1**, 139–43.

Mogensen, B., Brattstrom, H., Svantesson H., and Lidgren, L. (1983). Soft tissue release of the hip in juvenile chronic arthritis. *Scandinavian Journal of Rheumatology*, **12**, 17–20.

Ross, A.C., Edgar, M.A., Swann, M., and Ansell, B.M. (1987). Scoliosis in juvenile chronic arthritis. *Journal of Bone and Joint Surgery*, **69B**, 175–82.

Ruddlesdin, C., Ansell, B.M., Arden, G.P., and Swann, M. (1986). Total hip replacement in children with juvenile chronic arthritis. *Journal of Bone and Joint Surgery*, **68B**, 218–22.

Rydholm, U. (ed.) (1990). *Surgery for juvenile chronic arthritis*. Ortolani, Lund.

Rydholm, U., Brattstrom, H., Bylander, B., and Lidgren, L. (1987). Stapling of the knee in juvenile chronic arthritis. *Journal of Paediatric Orthopaedics*, **7**, 63–8.

Salter, R.B. and Field, B. (1960). The effects of combining compression on living articular cartilage: an experimental investigation. *Journal of Bone and Joint Surgery*, **42A**, 31–49.

Smith, B.L. (1990). Anaesthesia in paediatric rheumatology. In *Paediatric rheumatology update*, (ed. P. Woo, P.H. White, and B.M. Ansell), pp. 124–30. Oxford University Press, Oxford.

Soto-Hall, R. (1964). Variations in the intraarticular pressure of the hip joints in injury and disease. *Journal of Bone and Joint Surgery*, **46A**, 509–16.

Swann, M. (1983*a*). Juvenile chronic arthritis: surgical aspects. In *Clinical orthopedics*, (ed. N.H. Harris), pp. 249–67. Wright PSG, Bristol.

Swann, M. (1983*b*). Surgical treatment of the cervical spine in rheumatoid arthritis. *Annals of the Academy of Medicine of Singapore*, **12**, 233–42.

Swann, M. (1987). Juvenile chronic arthritis. *Clinical Orthopaedics and Related Research*, **219**, 38–49.

Swann, M. and Ansell, B.M. (1986). Soft tissue release of the hips in children with juvenile chronic arthritis. *Journal of Bone and Joint Surgery*, **68B**, 404–8.

Vilkki, P., Virtanen, R., and Makela, A.K. (1991). Arthroscopic synovectomy in the treatment of patients with juvenile chronic arthritis. *Acta Universitatis Carolinae Medica*, **37**, 84–6.

Williams, W.W. and McCullough, C.J. (1993). *Journal of Bone and Joint Surgery*, **75B**, 872–4.

Witt, J.D. and McCullough, C.J. (1994). Anterior soft tissue release of the hip in juvenile chronic arthritis. *Journal of Bone and Joint Surgery*, **76B**, 267–70.

Witt, J.D., Swann, M., and Ansell, B.M. (1991). Total hip replacement for juvenile chronic arthritis. *Journal of Bone and Joint Surgery*, **73B**, 770–73.

Woo, P., White, P.H., and Ansell, B.M. (ed.) (1990). *Paediatric rheumatology update*. Oxford University Press, Oxford.

6.3 Rehabilitation of adults

Lynne Turner-Stokes

Introduction

Arthritis produces a wide spectrum of disability. Some aspects, such as deformity or florid synovitis of the hands, are clearly visible. Others, for example variable fatigue, are less so, and can lead to real social handicap as they are likely to be unsympathetically received by friends and family.

Until recently, 'rehabilitation' to the average rheumatologist meant wax baths and wheelchairs. In the last decade, the development of a multiprofessional approach to rehabilitation has transformed the way we think about disability. The purpose of this chapter is to describe that multiprofessional approach, as applied to the management of inflammatory joint disease.

Epidemiology

Epidemiological surveys in the United Kingdom (Harris 1971; Badley *et al*. 1978) have shown that the largest single group of impaired people have locomotor disease, amounting to one-third of the total impaired and to 40 per cent in the elderly population.

In a population of 100 000 people over 1000 will have rheumatoid arthritis, of whom 120 will be severely handicapped and 25 will be chair- or bed-bound. The same population will have 300 or more patients with severe osteoarthritis, of whom 50 will be totally immobile (Clarke *et al*. 1987)

Impairment, disability, and handicap

In the past, terms such as disability and handicap have been loosely applied. For the purpose of this chapter I shall use the *International classification of impairments, disabilities and handicaps (ICIDH)* (World Health Organization 1980). In essence, 'impairment' refers to the symptoms and signs of arthritis, 'disability' is the resulting loss of function, and 'handicap' is the disadvantage experienced as a result of that disability.

Handicap is determined largely by social and environmental factors, and by the individual's expectations and aspirations. Thus the same level of impairment may lead to very different degrees of handicap in different individuals. For example, relatively mild arthritis in the hands may be devastating to the concert pianist but of little consequence to the cross-country runner, while similarly mild arthritis of the knees could have the opposite impact.

As health professionals, we tend to measure what we can alter. Doctors traditionally measure impairment, while therapists measure functional loss or disability. Patients and their families tend to be more concerned with handicap, but until recently, very little effort has gone into measurement of handicap or quality of life.

Handicap is only indirectly affected by reducing disability. The main limiting factors are social and environmental—housing, education, employment for example—and are largely beyond the control of healthcare workers.

In the United States the Americans with Disabilities Act now ensures that all public buildings must be fully accessible to disabled people, and an individual cannot be excluded from employment because of disability alone. In the United Kingdom and Europe we still fall a long way behind, but if we are to be the best advocates for our patients, we must measure these aspects and continue to apply pressure to improve the lot of our disabled patients.

Measurement of disability

If we are to manage disability effectively, it is essential that we measure it in order to be able to assess the outcome of our endeavours. There are two main types of measure for disability, specific and generic. Generic measures take a global view of disability and record it in a standardized manner that allows comparison. It is essential to collect generic information to allow comparison of different programmes, populations, and practices, but we must also realize that these scores often show little relevance to how individual patients function in their own environment. To assess this, we require specific measurements.

For example, incontinence is a major cause of disability and handicap, something we all dread. Stiffness, immobility, and poor dexterity combine to make it a common problem for patients with arthritis, many of whom are middle-aged to elderly women with a degree of urinary urgency and frequency. If we want to know whether a patient can get to the toilet in time, it matters not how long it takes them to walk a standard 10 m, or undo three standard buttons on a standard garment. What matters is how long it takes them to get to their toilet from their usual chair in the daytime, or from their bed at night, and once in the toilet, how long it takes them to undress and position themselves appropriately to pass urine without spillage (Turner-Stokes and Frank 1992). The presence or absence of a downstairs toilet, or a commode in the bedroom, may make the critical difference between continence and incontinence.

Measurement of disability therefore requires a combination of specific and generic measures. No one instrument will suffice. A large number of disability scores have been produced, some well validated, others not. It is important to select a scale that is set at the right level. The Barthel score, for example, is used widely for patients with neurological disability. It has pronounced floor and ceiling effects and, although reliably scored between different observers, it is relatively unresponsive to

change, which makes it unsuitable for assessing functional outcome in most patients with arthritis.

Some instruments require an independent observer, others have been developed for self-completion. One of the most widely used functional scores is the Stanford Health Assessment Questionnaire, which was developed in the United States, but has subsequently been translated into English practice and validated for self-completion by Kirwan and Reeback (1986). If routinely completed by patients while they wait for a clinic appointment, Health Assessment Questionnaire scores can provide a useful serial record of disability and its change with time.

Measurement of handicap and quality of life

As mentioned earlier, disability is not the only outcome measure that should be recorded. Handicap and quality of life are more important to the patient and their family, but more often than not we have no indication of whether our treatment is having any impact at all in these areas (see also Chapter 1.1.5).

An abbreviated list of potentially useful generic scales is given in Table 1. The London Handicap Scale was developed for stroke patients, and the Community Integration Questionnaire for patients with brain injury, but both collect information in a standardized way that is potentially useful for patients with arthritis.

Modality to be measured	Instrument	Reference
Impairment	Ritchie index	Ritchie *et al.* (1968)
Disability	Health Assessment Questionnaire (HAQ)	Kirwan and Reeback (1986)
	Functional Independent Measure (FIM)	Hamerman *et al.* (1993)
Handicap	London Handicap Scale	Harwood *et al.* (1994)
	Community Integration Questionnaire	Willer *et al.* (1994)
Quality of Life	SF-36	Garratt *et al.* (1993)
	General Health Questionnaire	Goldberg and Hillier (1979)
	Nottingham Health Profile	Hunt *et al.* (1981)
Depression	Beck depression score	Beck *et al.* (1979)
Anxiety	Spielberger score	Spielberger *et al.* (1970)
Psychological response to pain and disability	Sickness Impact Profile (SIP)	Bergner *et al.* (1976)
	Multi-dimensional pain inventory	Kerns *et al.* (1985)

Table 1 Modalities and examples of available measures for assessing patients with arthritis

A number of instruments have been developed to assess quality of life in relation to health, including the General Health Questionnaire, the Nottingham Health Profile, the Sickness Impact Profile, and the SF-36. The SF-36 in particular has been designed for ease of use and is gaining popularity. It could potentially provide useful information on quality of life for patients with arthritis and their families. However, as yet it has not been validated for use with disabled patients or their carers, so any information obtained should be treated with caution.

Patients in the early stages of arthritis may not be significantly disabled other than by pain, and standard disability scores may not be suitable. Alternatives such as the Multidimensional Pain Inventory may be useful, not only for measuring the impact of their symptoms but also the extent to which they use positive coping strategies to combat them.

These are just a few of the scales that have been developed. None of them is perfect and they all have somewhat different emphasis. When selecting an instrument, careful examination of all of the items included will indicate which is the most relevant to the population under study. Although some scales apply weighting to different items, so that they perform as interval scales, it should be remembered that most measures provide ordinal data, which should not be submitted to simple mathematical manipulation such as summing or averaging. Rasch analysis is a method of applying a mathematical model to ordinal data, so that they behave as interval data. Its role is currently being explored in this area, but it is too early to say if this will prove useful. Whichever instrument is used, it should be applied only with reference to the original manual, and careful thought should be given to the mode of analysis employed.

Functional anatomy and disability in arthritis

The hand

Functional dexterity of the hand is more advanced in man than in any other species, and it is our ability to manipulate objects in our environment that has allowed us to capitalize on our large cerebral capacity. The two major functions of the hand are the power grip the precision or pincer grip. If these are limited by weakness or pain, the hand rapidly becomes non-functional. Opposition, required for the pincer grip, is often lost, but patients will learn to oppose the side of the thumb, which provides adequate function.

Rupture of extensor tendons is commonly regarded as a rheumatological emergency, but surgical repair is necessary not so much to provide active extension as to prevent the dropped fingers getting in the way of hand function. The importance of hand function cannot be overstressed. Patients and their medical attendants often crave cosmesis, but cosmetic repair undertaken at the expense of function is almost universally disastrous.

The wrist

The wrist provides a mechanism to extend the range of activity of the hand, and a fulcrum to stabilize the grip. Some 30°of wrist extension is required for optimal power grip, and if function is limited by painful wrists, arthrodesis may be considered. Careful attention should be paid, however, to the position of arthrodesis. To be able to wipe their own bottom, a patient requires at least 30° of wrist flexion, so if bilateral arthrodesis is undertaken, one wrist should be in flexion and one in extension. If one wrist has spontaneously fused, its position must be considered when fusing the other.

The elbow and shoulder

These joints are often restricted in rheumatoid arthritis but, in functional terms, it is their combined range that is important. Is it sufficient to allow the hand to reach to the mouth for feeding, to the back of the head for grooming, and to the perineum for maintaining personal hygiene?

Flexion deformities of the elbow are usually well documented but restriction of 20 to 30° is functionally unimportant on its own. Meanwhile, the radioulnar joints are often forgotten, but they allow supination and pronation, which are essential to optimal hand function.

The spine

Involvement of the cervical spine may result in subluxation and dislocation, which is not only a source of painful disability and disturbed sleep, but may result in cord compression. Patients with atlantoaxial subluxation must be taught to protect the neck from sudden movements, for example when motoring or using public transport.

The hip

Hip restriction results in immobility and stiffness that can be severely limiting. Climbing stairs, getting in and out of the bath, and sitting comfortably on the toilet all require hip flexion. For young women, abduction as well as flexion is required for sexual intercourse and childbirth.

The knee

The locking mechanism of the knee, which allows us to stand for long periods without using much muscle power, requires not only full extension but also rotation in

the last 5 to 10°. Failure to achieve this severely limits standing, and flexion deformities of 20° or more make walking extremely difficult.

The ankle and foot

The true ankle joint allows plantar and dorsiflexion, and the subtalar joint, abduction and adduction. Restriction of the subtalar joints make walking on rough ground difficult.

Metatarsalgia, which is often a hallmark of rheumatoid arthritis, results from dropped metatarsal heads. If, as a doctor, you do only one thing for a patient, education about appropriate footwear early in the course of disease is probably the greatest contribution you can make to maintaining mobility (see ' Footwear' below).

Systemic disease

In addition to specific joint problems the effects of systemic disease itself can be severely limiting. Fatigue, anaemia, muscle wasting all take their toll, quite apart from morning stiffness, which may require the patient to get up several hours earlier to prepare for work or get the family off to school, for example. General advice on sensible eating, exercise, and sleeping patterns is important and should not be overlooked.

Psychological reaction (seeChapter 1.1.6)

Traditionally last on the list, the psychological impact of arthritis should in reality come first. Few of us can readily imagine what it is like to live with constant pain. Response to pain can vary enormously between patients. We are all familiar with the 'typus robustus' who denies pain despite gross destructive disease, while others are totally incapacitated but have relatively minor visible signs.

A variety of psychological and social factors may interact to sensitize or desensitize patients to the effects of pain. Depression, altered body image with poor self-esteem, altered sexuality, role reversal in marriage, as well as intellect, comprehension, and expectations, will all affect a patient's perception of pain and disability. The standard 15-min medical outpatient appointment in the United Kingdom National Health Service does not allow for sensible appraisal of these factors even though it is they, and not drugs or blood tests, that will ultimately determine outcome.

Medical and social models of rehabilitation

Many areas of medicine are currently going through a period of substantial change and rehabilitation is no exception. The days are gone when the doctor reigned supreme over the patient's care and wrote detailed prescriptions of exactly what physiotherapy should entail, usually based on very little understanding. Instead we have interdisciplinary teams made up of highly trained professionals, each of whom brings specialist knowledge of their own field to the group. In most cases, gone too are the days when the patient was expected simply to be a passive recipient of therapy. It is essential that any programme remains as far as possible within the patient's control. Since it will ultimately be their responsibility to maintain progress by themselves it is important to engage then from the start and failure to do so is the single most important cause of failure to demonstrate long-term benefits of therapy.

There is currently much debate in the United Kingdom over whether rehabilitation services should be provided in the hospital or community setting. This is the wrong question. Services should clearly be provided in both settings. The question should be how to identify the most suitable service for a given individual and how to ensure that they have access to it. Some patients are ill and require rehabilitation in the context of their medical management. Others are not ill and reasonably demand the right to exercise choice over the services they wish to use. The medical and social models of rehabilitation are not, therefore, mutually exclusive, it is more a question of 'horses for courses' (appropriateness).

The principles of rehabilitation

Whichever model of rehabilitation is adopted, the principles are essentially the same. Rehabilitation has been defined in many different terms, but ultimately it involves restoring an individual to their maximal functional independence. It consists of two main parts:

1. restitution: restoring lost function by minimizing impairment;
2. substitution: compensation for residual functional loss by using alternative techniques, aids, or appliances.

In practice these two approaches are often applied simultaneously and a problem-orientated approach will combine restorative therapy, teaching compensatory techniques, and the use of a variety of aids and appliances. The important components of rehabilitation for patients with arthritis can be summarized as in Table 2.

Table 2 Important components of rehabilitation for patients with arthritis

Problem-orientated approach and the interdisciplinary team

The interdisciplinary team

Members of the interdisciplinary team are highly trained professionals in their own right and each brings a specialist knowledge of their own field to the group. Some of the professions that may be included in the rheumatological interdisciplinary team are listed in Table 3.

Doctor
Nurse
Physiotherapist
Occupational therapist
Psychologist
Counsellor
Orthotist
Dietitian
Chiropodist
Social worker
Rehabilitation engineer

Table 3 Professions included in a rheumatological interdisciplinary team

The key to good rehabilitation is teamwork (Fig. 1) and this requires coordination and a team leader. The doctor may not necessarily be that leader—having good medical skills is not synonymous with having team management skills—but is nevertheless an essential part of the team. Interdisciplinary team management allows individual members of the team to refer to any other, without going back through the coordinator. Opinion and action are thus obtained more quickly, and are more likely to be relevant to the original question.

Fig. 1 An effective interdisciplinary team requires a coordinated approach both with, and away from, the patient. (a) A doctor, physiotherapist, and occupational therapist join forces to splint a hand in order to meet both therapeutic and functional requirements; (b) the key to coordinated therapy is integrated timetabling.

The problem-orientated approach

Rather than members working individually within their own professional boundaries, the well-coordinated interdisciplinary team provides a unified, problem-orientated approach. The individual problems that limit a patient's independence are identified and tackled systematically by achieving a series of staged goals set by the patient and therapists together. If at all possible, goals should be set, together with a date at which they will be reviewed. This helps to focus both patient and therapists on exactly what has to be done. It is immediately clear when progress is not being made, so that the reasons for this can be explored. It is essential that objectives and treatment are functionally relevant to the patient in their normal environment, and very often this will involve the active participation of the carer as well.

Contents of the rehabilitation programme

Exercise

Although some textbooks in circulation still have 'bed rest' at the top of their list of management, it is now almost universally accepted that bed rest is bad for patients with arthritis, resulting in flexion contractures, osteopenia, and muscle wasting. Instead it is recognized that active use and weight bearing through joints maintains bone and muscle strength. Nevertheless, patients must be taught to apply these principles sensibly. An exercise programme should contain the four elements listed in Table 4. Wherever possible, exercise should be put in a social context, taking advantage of local leisure facilities that patients can have access to outside the hospital environment. Once again, advice on appropriate footwear for exercise is an essential part of the programme.

To maintain:	Type of exercise	For example:
Muscle strength	Isometric exercise to build up muscle bulk and strength	Quadriceps drill But the programme must be tailored to the age and fitness of the patient
Joint mobility	Active exercise to maintain joint range	Pendular and wall-climbing exercises for the shoulder
Limbering up	Programme of daily exercises to combat early morning stiffness	Working the hands in a basin of warm water
Cardiovascular fitness	Physical stamina, cardiac and respiratory function	Swimming, cycling, walking, dancing

Table 4 The main components of an exercise programme

Education

Education of the patient and family is vital to the continued success of the rehabilitation process. They need to know what to expect from the disease, what to expect from rehabilitation, what they can do to maximize benefit, and what they must do to avoid exacerbating the condition. Education must begin one-to-one, although it can be continued in group sessions, with the help of leaflets and audio- and videotapes to digest at home. Introduction to patient support and self-care groups can also be extremely helpful, not only for obtaining advice and information, but for the comfort of knowing they are not alone in their suffering.

Joint protection

Trauma exacerbates inflammatory arthritis, whereas careful use of muscle and joints can reduce pain and fatigue. The main objectives of a joint protection programme are therefore to reduce pain, prevent damage and deformity, and conserve energy. The principles are outlined in Table 5.

1. Use the strongest and largest joints possible to accomplish a task
2. Spread the load of carrying/lifting over several joints
3. Use each joint in its most stable functional position
4. Maintain joint mobility and function
5. Avoid maintaining a joint in one position for a long time—especially in positions that exacerbate deformity, e.g. avoid sleeping with a pillow under the knees
6. Avoid excessive activity
7. Pace activity: punctuate periods of activity with regular rest periods

Table 5 The principles of joint protection

Energy conservation

While bed rest is to be avoided, fatigue is one of the major causes of disability in arthritis. In order to avoid the inappropriate blitzing the housework and then retiring to bed for 2 days, patients need to be taught to pace their activities and to conserve their energy where possible. Some practical examples of ways to save energy are given in Table 6.

(a) *Personal care*
Dress sitting down
Avoid back-fastening or tight-fitting clothes
Shower rather than bathing if possible
Avoid low chairs

(b) *Domestic*
Reorganize the kitchen so that all frequently used equipment is readily to hand
Do all activities in sitting where possible, using a perch stool
Avoid lifting and carrying—move heavy objects using a trolley
Use both hands
Use a jug to fill the kettle, rather than lifting the kettle itself
Steam or microwave vegetables to avoid lifting heavy pans
Soak pans immediately after cooking
Use duvets on the beds instead of blankets
Plan domestic chores sensibly, for example:
Complete upstairs tasks and then the downstairs ones to avoid frequent trips up and down the stairs
Intersperse heavier tasks with lighter ones
Use modern labour-saving devices whenever possible

Table 6 Practical example of ways to conserve energy

It is not always necessary to prescribe special equipment. Some examples of standard modern labour-saving devices that may be helpful for patients with arthritis are listed in Table 7. These are relatively inexpensive, readily come by, and more often acceptable to patients than having a house full of special adaptations that advertise their disability to every visitor.

Food processor, blender
Electric tin-opener
Portable telephone
Remote control for television, video, hi-fi etc.
Dishwasher
Front-loading automatic washer/dryer
Computer/word-processor
Automatic car—power steering

Table 7 Standard modern labour-saving devices that may be helpful for patients with arthritis

Community-orientated supervision

Whatever techniques are instituted in hospital, follow-up into the home is necessary to ensure that advice given is appropriate. It is common experience that specific advice given on joint protection and energy conservation changes when the patient is seen in their own environment. Time-consuming though a home visit may be, it will save time in the long run.

Likewise, exercises will only be continued if they are functionally appropriate, and this too is much easier to gauge in the context of the home environment. Reassessment at home every 2 to 3 months allows plans to be made in the context of the normal environment and avoids painful, slow, and unreliable ambulance journeys to hospital.

Home assessment is also vital in considering the need for structural changes or adaptations. Some areas to consider are given in Table 8, but first some basic information is needed about mobility and transfers.

Access
Can they get in and out of the house through the front or back door?
Can they move from room to room inside the house?
If living on more than one level, can they get up and downstairs:
If not, could a stairlift or through-floor lift be installed?
How will they get to the toilet by day and during the night?
Can they get to washing/showering facilities?
Inside the home
Is furniture suitable and conveniently placed?
Remove potential hazards, e.g. slip rugs
Are plugs and switches at reachable height?
Security
Can they control their own front door?
How will they summon help in an emergency?
Kitchen
Are frequently used items stored conveniently?
Are work surfaces at a suitable height?
Can they use the basic equipment?
Bathroom and lavatory
Can they get into the shower or bath?
Is any special equipment needed such as:
Bathing aids
Raised toilet seat
Rails, tap handles?
Bedroom
Height of bed—can they get in and out of it?
Is the mattress suitable—are they at risk of pressure sores?
Is the mattress needed for bedbound patients?

Table 8 Some areas to consider in assessment of the home environment

Mobility, transfers, and stairs

Is the patient in a wheelchair? If so, consider (i) whether ramps are needed, (ii) are the door frames wide enough to pass through without trapping their fingers, and (iii) is there room to manoeuvre. Are they able to transfer themselves from the wheelchair independently; if so, do they need equipment such as grab rails or a sliding board? If on their feet, can they manage stairs? If so, do they need one rail or two? Note: if a walking stick is used, only one hand is free to hold a rail to go upstairs, but a rail is needed on the opposite side coming down unless they are able to use the walking stick in the opposite hand.

Environmental control

For severely disabled people who live alone, the use of a computer to control the home environment has now become commonplace, but the availability of this facility is often forgotten. Its use requires neither dexterity nor high-powered cognitive function, and, using scanning software and an appropriate on–off switch, almost any movement that is under reliable voluntary control may be used to operate an environmental control unit (Fig. 2). A sample of available switch types is given in Table 9,

and the sort of functions they may be used to control in Table 10.

Fig. 2 An environmental control unit may be used with a range of single switches and operated by almost any movement that is under reliable voluntary control. Remote control by radio waves or infrared avoids the need for fixed wiring. (a) A wheelchair-mounted Steeper FOX system; (b) the Steeper Persona.

Switch type	Movement that operates it
Plate switch	Pressure from hand or foot
Button switch	Pressure from thumb or finger
Chin or head switch	Pressure from chin or head
Eye blink	Movement of eyelid
Suck-and-puff	Sucking and blowing
Tongue plate	Touch of tongue on palate plate

Table 9 Some types of on–off switch that may be used for controlling an environmental control unit

Alarm
Intercom
Control of front-door lock
Telephone
Opening/closing curtains
Control of electric sockets for heaters/fans/lights
Control of radio/television
Communication aid
Computer

Table 10 Common functions that may be controlled by an environmental control unit

Psychosocial aspects

Despite the best endeavours of a multiprofessional team, the functional outcome from rehabilitation of patients with arthritis remains dependent on psychosocial factors, which include:

1. personality: coping style, self-esteem and motivation;
2. section of community: class, intellect, and education;
3. patient's reaction to disability: e.g. depression, denial;
4. family's reaction: sympathy, overprotection, anger;
5. supportive network: family, friends, workmates;
6. environment, housing, transport;
7. financial security, employment.

Active psychological support is essential to allow patients and families to work through the changes that have occurred in their lives and to re-establish their relationships on a positive basis.

Cognitive behavioural rehabilitation techniques have been used extensively in patients with chronic pain. They involve teaching them how to use coping strategies to increase their activity despite the pain and regain control over their lives. The approach is shown to be beneficial not only in increasing activity levels but also in reducing depression and anxiety. More recently the approach has been successfully applied to patents with arthritis (O'Leary *et al*. 1988) but it is essential that it is introduced in the early stages of disease. It is also essential to involve the family and carers because they must be prepared to stand back and allow the patient to regain control of his or her own life.

Sexual problems

Sexual problems are extremely common when one member of the partnership has arthritis and may arise from a number of factors listed in Table 11.

Altered body image in the face of physical deformity
Joint restriction—limiting usual positioning for sexual intercourse, in particular the missionary position
Fatigue and loss of libido
Poor cardiovascular fitness—inability to sustain activity for long enough to achieve orgasm
Hand deformity—impeding foreplay and fitting of contraceptive devices
Fear of pain leading to frigidity and impotence

Table 11 Factors that may contribute to sexual dysfunction in patients with arthritis

Young patients with arthritis often encounter difficulties with forming relationships because of the above problems and also due to lack of access to dances, clubs, and other traditional courting scenes.

It is important to remember that sexual problems are not confined to marriage, heterosexuality or to the under-60s. Patients often have to be given the right to talk about it, but a busy clinic is not the time or the place. A number of organizations provide trained sexual counselling specifically for people with disabilities (for example, in the United Kingdom, SPOD, is the Association to Aid the Sexual and Personal Relationships of People with a Disability). The wise physician will keep to hand the telephone number of their local service.

Aids and appliances

In general, aids help to compensate for either pain and weakness or for lack of joint range. Some aids that are particularly useful to patients with arthritis are listed in Table 12 and illustrated in Fig. 3. However, they must be carefully chosen with specific needs in mind. Articles must be lightweight, strong and reliable, cheap, and acceptable to the patient and family.

Fig. 3 Some aids that may be useful to a patient with arthritis. (a) Thick-handled utensils; (b) rubber grip-pad; (c) kettle rocker-stand; (d) angled chopping knife; (e) plug with grip-handle; (f) Easi-reach device; (g) ejector chair.

To compensate for weakness or pain

Thick-handled utensils—require less effort to grip

Key turners and other easy-to-grip devices

Lever tap handles

To compensate for lack of joint range

Long-handled utensils—extend effective reach

Raised chair, toilet seat—easier to stand up from

Velcro fasteners on clothes—require less dexterity than zips and buttons

Table 12 Aids that are likely to be useful to the patient with arthritis

Orthoses

An orthosis is a device worn outside the body that aids function, protects, or prevents pain. One easy way to waste money is to provide orthoses that patients never wear. Worse, an inappropriately prescribed orthosis can actually enhance damage and deformity.

In order to avoid this great care must be paid in prescription to ensure that:

1. the device is well fitting and does the right job;
2. the patient is able to put it on and take it off easily;
3. both patient and carers know how, when, and why to use it;
4. the device is acceptable to the patient.

Because of these factors it is vital that the orthosis be checked after fitting and reviewed at each prescription. 'Stabilized' prescriptions that replicate the device when it wears out should only be signed when a trained orthotist is available to supervise subsequent fittings and has the experience and authority to alter the specifications in accordance with need. Unless this is so, each orthosis should be prescribed and checked by a suitable experienced physician—which by and large does not include junior medical staff.

Splints and orthoses have a special role in childhood arthritis. Some forms of juvenile arthritis are likely to remit by the time the time the child reaches adulthood. Splints form a vital part of management in preventing deformity so that when synovitis does resolve the joint still has optimal range and function. Hand splints (working and resting), knee splints, and shoe raises should be regularly reviewed to ensure this outcome as the child grows.

Adults tends to be less tolerant of non-functional splints, usually not having parents at home to bully them into wearing splints for certain parts of the day or night. Therefore, in adult arthritis, the emphasis on splint wearing is to increase function by stabilizing the joint and reducing pain more than preventing deformity.

Orthoses may be custom-made or provided ready-made or 'off the shelf'.

Off-the-shelf orthoses

Commonly prescribed items in rheumatology include the following.

Futura wrist splints

These stabilize the wrist in 30° extension and in theory act as a working splint. In practice they either fail to support the wrist adequately or the extensor bar effectively prevents useful function. Many physicians do not realize that the extensor bar is malleable and should be specifically adjusted when the splint is prescribed to ensure a good fit.

Cervical collars

Soft collars

Soft collars can be used to support the neck relieving painful muscle spasm, especially at night.

Hard collars

For patients with cervical subluxation, a hard collar is recommended while motoring or using public transport, to avoid dislocation with sudden jarring movements. In practice, they are so uncomfortable that patients rarely comply with this advice; a soft collar may be less protective but at least it is more likely to be worn.

Knee braces

Lightweight SK knee braces (Fig. 4(a)) are often prescribed for unstable knees, but are rarely able to provide adequate support. A range of more sophisticated devices is now available, which can be adjusted to allow a defined range of flexion and extension as well as providing mediolateral support. These heavier braces are more supportive, but many patients who need to use them will then need help getting them on and off.

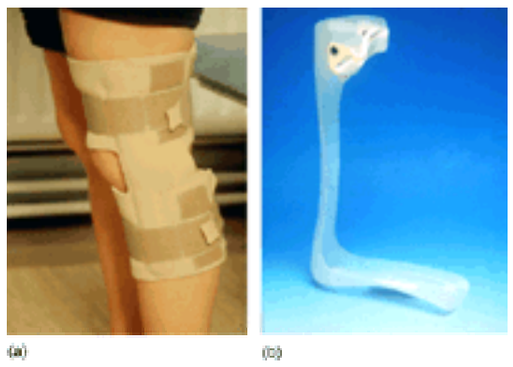


Fig. 4 'Off-the-shelf' orthoses: (a) lightweight SK knee brace; (b) ankle–foot orthosis.

Ankle–foot orthoses

Ankle–foot orthoses (Fig. 4(b)) provide dorsiflexion at the ankle, for example where foot drop results from vasculitic nerve damage. The orthosis fits inside the shoes and usually requires a shoe that is one size larger than normal. Although ankle–foot orthoses are available off the shelf, if one is to be worn for any length of time a custom-made one that is moulded to the foot is more satisfactory.

Custom-made orthoses

In general, custom-made orthoses (principally splints) are more satisfactory if someone of sufficient experience is available to make them. Splints can help to reduce deformity or increase function by:

1. stabilization of a joint;
2. reducing pain by immobilization;
3. maintaining passive joint range;
4. functional positioning during activity;
5. protection leading to greater confidence in use.

Working splints are designed to be used during functional activity, for example cock-up wrist splints that leave the fingers free and support the wrist in 30° extension. Dynamic splints have a moving part that allows either free movement or movement against resistance provided by a spring or a piece of elastic. These may be used for example to correct flexion contractures. Resting splints hold the joint or limb in an optimal position and many be used to combat pain or prevent deformity. For example, paddle splints help to maintain finger extension, but because they render the hand useless, they are usually worn at night.

Footwear

Advice on appropriate footwear may be the single most useful thing you can provide for patients. Patients with arthritis should not wear silly shoes, except on very special occasions, and all sloppy slippers should be thrown away.

Appropriate footwear has the following characteristics:

1. adequate width and height to accommodate the toes;
2. soft rubber sole and padded insole to provide shock absorption for the metatarsal heads;
3. arch-support insole and strong medial side of shoe to prevent valgus deformity of the ankle and spread weight evenly across the weight-bearing areas of the foot.

Trainers or sneakers provide all of these features and are often acceptable to younger patients. Older women may prefer a leather casual, but Ecco and Roehide are two shoe manufacturers whose lines include all of the above features. In the United Kingdom, their products are obtainable in some high-street shoe shops, e.g. Clark's and K Shoes.

Alternatively, an arch support insole may be fitted into an existing pair of strong shoes (Fig. 5(a)), but this will tend to be bulky and reduce the depth of the shoe so that the toes may rub on the leather upper. This is particularly a problem with flexion deformities of the toes and may lead to callous formation and ulceration over the interphalangeal joints. Extra-depth shoes can be prescribed to combat this problem and are available off the shelf in styles that are not readily distinguishable from ordinary footwear (Fig. 5(b)).

Fig. 5 Footwear: (a) arch-support insoles with built-in metatarsal pads; (b) extra-depth shoes may be needed to accommodate arch supports.

Metatarsal pads are often prescribed but are rarely of any help as they fail to support the weight-bearing surface of the foot. The standard arch-support insole provided by most orthotics departments is made of expanded polystyrene. Some have three little discs in the metatarsal areas, which are designed to cushion the metatarsal heads but more often than not merely exacerbate metatarsalgia. Sorbithané is a highly shock-absorbent material and arch-support insoles made of Sorbithane are available in many sports shops. Although more expensive, they are generally much more satisfactory and worth the initial outlay as they are resilient and can be moved from shoe to shoe.

Mobility

Although doctors spend much of their time and effort trying to relieve pain, when patients are asked which aspect of their arthritis affects them the most, their reply is most commonly 'immobility' (Chamberlain and Buchanan 1978). The major problem is mobility outside. Public transport is often totally inaccessible, and community transport systems for disabled people are limited. Unless they have a car, immobility can severely limit a patients ability to lead a normal life.

Walking aids

The choice of a walking aid must take into account a number of factors including the extent of involvement of lower and upper limbs, and the terrain to be negotiated. Often a patient may require a variety of walking aids to use under different circumstances.

A walking stick can provide confidence as well a mechanical advantage if it is the right length and the patient is taught to use it correctly—usually held in the opposite hand to the affected lower limb. Attention should be paid to the comfort of the handle for patients with hand involvement—a moulded Fischer handle helps to spread the weight evenly in the presence of hand deformity (Fig. 6).

Fig. 6 Walking aids: a moulded Fischer handle (a) helps to distributed the weight more evenly across the hand than with a standard walking stick (b).

Elbow crutches give a firmer base and a gutter crutch allows patients with severe involvement of the hand and elbow to bear weight on the forearm. It should be remembered, however, that bilateral use of sticks or crutches effectively impedes carrying anything other than a bag slung over the shoulder and the width of the base of a two-crutch user is wider than that of a wheelchair. Walking frames provide a more stable alternative (Fig. 7), although some patients feel that they carry a greater stigma of disability, and they cannot give support while going up and down stairs.

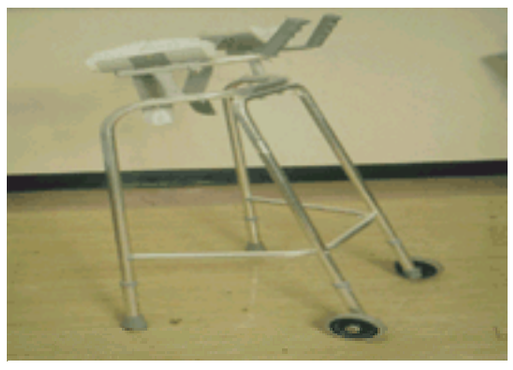


Fig. 7 A gutter rollator walking frame provides stability without putting pressure on the hands.

Any walking aid that is prescribed should be fitted and checked by an experienced physiotherapist, who will ensure that it is the most suitable aid, correctly adjusted, and in safe working order. The physiotherapist will also instruct the patient in its use and maintenance; in particular, ferrules should be checked regularly for wear.

Wheelchairs

Wheelchairs may be regarded as admission of defeat but in reality they can increase mobility, broaden horizons, and conserve energy for more productive use. Attendant-operated (8L) chairs have small wheels that can be difficult to negotiate over kerbs. The larger wheels on a 9L self-propelling chair (Fig. 8) make it easier to push, and therefore it is usually recommended even though most patients with rheumatoid arthritis who require a wheelchair will be unable to self-propel over any distance because of shoulder involvement.

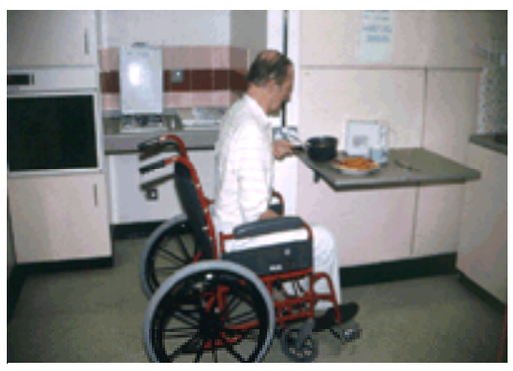


Fig. 8 The large wheels on a self-propelling 9L wheelchair make it easier for an attendant to push; most patients with rheumatoid arthritis are unable to propel much

because of shoulder problems.

Electric wheelchairs allow independent mobility for the severely disabled patient (Fig. 9) but a smooth proportional control is essential as jerky movement is not only painful but also dangerous if there is risk of cervical subluxation.

Fig. 9 An electric wheelchair allows independent mobility for severely disabled patients; a smooth proportional control is important where there is risk of cervical subluxation.

Patients who spend any length of time in a wheelchair are at risk of pressure sores. They require appropriate cushioning and may also need a special seating package, which should be prescribed by a specialist clinic, to correct their position.

Driving for disabled motorists

A number of adaptations are available for patients who drive, or are regularly driven in a family car. Swivel seats can facilitate getting in and out of the vehicle (Fig. 10). Wide-angled mirrors offer a panoramic rear view without having to turn the head. Headrests are essential to protect patients from rear-shunt whiplash injuries, whether or not they have atlantoaxial subluxation. In the majority of cases, automatic gear shift and powered steering may suffice, but a range of adapted controls is available, and is rapidly becoming more sophisticated (Fig. 11). If considered as potential users of adapted cars, patients should be professionally assessed at a specialist centre such as, in the United Kingdom, Banstead Mobility Centre. Patients unable to walk more than 100 yards (approx. 90 m) may be entitled to a disabled parking permit—known in the United Kingdom as the 'orange badge'—to facilitate convenient parking. This may be applied for through the local authority.

Fig. 10 Swivel seats facilitate getting in and out of a car.

Fig. 11 The range of devices for controlling a car is becoming more sophisticated. (a) A complex mechanical control system; (b) its modern electronic counterpart.

Social factors

While healthcare workers may not have direct influence over factors such as housing, finance, employment, and education, they may be the best advocate that a patient has to help them achieve at least a minimum standard of living. Many patients with arthritis are young, most are cognitively intact and should have good employment prospects given appropriate retraining. Failing that, patients need to be prepared for increased hours of leisure time. Contrary to the views of some doctors, rehabilitation does not end with being able to wash, dress, and make a cup of tea.

Acknowledgements

I am grateful for the help of Dr Luay Zebouni and David Hawkins (Steeper Ltd) in collecting illustrative material, and to the Luff Foundation for financial support.

Chapter References

Badley, E. M., Thompson, R. P., and Wood, P. H. N. (1978). The prevalence and severity of major disability conditions—a reappraisal of the Government Social Survey of the Handicapped and Impaired in Great Britain. *International Journal of Epidemiology*, **7**, 145–51.

Beck, A. T., Ward, C. H., Mendelssohn, M., Mock, J., and Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, **4**, 561–71.

Bergner, M., Bobbitt, R. A., and Pollard, W. (1976). The sickness impact profile. Validation of the health status measure. *Medical Care*, **14**, 56–67.

Chamberlain, M. A. and Buchanan, J. (1978). Mobility in arthritis. *Reports of the Rheumatic Diseases*, **66**, 1–2.

Clarke, A. K., Allard, L., and Braybrooks, B. A. (1987). *Rehabilitation in rheumatology: the team approach*. Martin Dunitz, London.

Garratt, A. M., Ruta, D. A., Abdalla, M. L., Buckingham, J. K., and Russell, I. T. (1993). The SF-36 Health Survey Questionnaire: an outcome measure suitable for routine use within the NHS? *British Medical Journal*, **306**, 1440–4.

Goldbeg, D. P. and Hillier, V. F. (1979). A scaled version of the General Health Questionnaire. *Psychological Medicine*, 9, 139–45.

Harris, A. I. (1971). *Handicapped and impaired in Great Britain*. HMSO, London.

Harwood, R. H., Rogers, A., Dickinson, E., and Ebrahim, S. (1994). Measuring handicap—the London Handicap Scale. A new outcome measure in chronic disease. *Quality in Healthcare*, **3**, 11–16.

Heineman, A. W., Linacre, J. M., Wright, B. D., Hamilton, B. B., and Granger, C. V. (1993). Relationships between impairment and physical disability as measurement by the Functional Independence Measure. *Archives of Physical Medicine and Rehabilitation*, **74**, 566–73.

Hunt, S. M., McKenna, S. P., and Williams, J. (1981). Reliability of a population survey tool for measuring perceived health problems: a study of patients with osteoarthritis. *Journal of Epidemiology and Community Health*, **35**, 297–300.

Kerns, R. D., Turk, D. C., and Rudy, T. E. (1985). The West-Haven Yale Multi-dimensional Pain Inventory (WHYMPI). *Pain*, **23**, 245–56.

Kirwan, J. R. and Reeback, J. S. (1986). Stanford l Health Assessment Questionnaire modified to assess disability in British patients with rheumatoid arthritis. *British Journal of Rheumatology*, **25**, 206–9.

O'Leary, A., Shoor, S., Lorig, K., and Holman, H. R. (1988). A cognitive behavioural treatment for rheumatoid arthritis. *Health Psychology*, **7**, 527–44.

Ritchie, D. M. *et al*. (1968). Clinical studies with an articular index for the assessment of joint tenderness in patients with rheumatoid arthritis. *Quarterly Journal of Medicine*, **37**, 393–406.

Spielberger, C. D., Gorsuch, A., and Lushane, R. (1970). *Manual for the State-Trait Anxiety Inventory*. Consulting Psychologists Press, Palo Alto CA.

Turner-Stokes, L. and Frank, A. O. (1992). Urinary incontinence among patients with arthritis—a neglected disability. *Journal of the Royal Society of Medicine*, **85**, 389–93.

Willer, B., Ottenbacher, K. J., and Coad, M. L. (1994). The Community Integration Questionnaire. A comparative examination. *American Journal of Physical Medicine and Rehabilitation*, **73**, 103–11.

World Health Organization (1980). *The international classification of impairments, disabilities and handicaps (ICIDH)—a manual of classification relating to the consequences of disease*. WHO, Geneva.

6.4 Rehabilitation of children

Renate Häfner and Marianne Spamer

Introduction

Inflammatory arthritis is the most commonly seen feature in rheumatic disorders of children and adolescents. It occurs most often as idiopathic arthritis with different subtypes (systemic, polyarthritis, oligoarthritis, etc.) but may also be a symptom of an underlying disease such as systemic lupus erythematosus, sarcoidosis, or Behçet's syndrome. Rehabilitation follows the same principles as for all arthritic disorders. The therapy offered must give careful consideration to the child's age and developmental status, the pattern of joint involvement, and the individual disease course.

Some autoimmune rheumatic diseases such as juvenile dermatomyositis where muscular involvement is predominant or skin fibrosis, such as scleroderma, can lead to functional impairment and joint contractures without the presence of an overt arthritis. Patients with these conditions benefit from different therapy protocols.

Aims of rehabilitation—a multidisciplinary team approach

The different aspects of rehabilitation require the co-operation of several health-care professionals who work together to improve the child's function, independence, and self-esteem. This includes caring for the whole family and overseeing integration into the community. The team is usually guided and co-ordinated by a paediatric rheumatologist. It includes a physiotherapist, occupational therapist, social worker, and psychologist. The team members co-operate with each other and interact with community and school staff (Table 1).

Achievements	Team members
Improved joint function and mobility (proper joint positioning, correction of deformities)	Physictherapist, occupational therapist
Self-care and independence	Occupational therapist
Integration into school	Social worker, occupational therapist, school staff (teacher, school nurse)
Guidance of the whole family	Social worker, psychologist
Promoting self-esteem	All team members

Table 1 Achievements of a multidisciplinary health-care team; the paediatric rheumatologist co-ordinates all team activities

The physiotherapist works together with the occupational therapist to improve function of individual joints as well as general mobility. They are also responsible for correction of joint deformity, which often requires individual splinting. Increased mobility forms the basis for greater independence. Children with marked impairment benefit from additional self-care support by an experienced occupational therapist. Most children can learn to master daily activities like hair washing, combing, dressing, feeding, and going to the toilet without extra help. However, they may need some specially adapted equipment. Instructions for joint protection should always be part of the therapy programme.

The child's rheumatic disease always has an impact on the whole family. Physiotherapy, doctor's visits, transportation to school, and other activities are time-consuming. Financial hardship in such families is common. Siblings may feel neglected and parents are usually overburdened with the extra tasks and worries, so rehabilitation must include care for the whole family, which is where the social worker has an important part to play. The social worker gives advice on financial entitlements, may help organize transportation, and assist with individual social problems. A psychologist can be helpful to handle conflict situations.

A very important aspect of rehabilitation involves integrating the child into school life. To achieve this, it is necessary for several health-care professionals to co-ordinate contact with teachers and other school staff to highlight the child's needs and help organize school life with the minimum of disruption for all concerned.

Perhaps the team's most important role is psychological—to create the highest possible level of integration so that children cease to see themselves as sick and different and acquire a sense of belonging.

Rehabilitation of chronic arthritis

Management of arthritis in children requires a wide-ranging knowledge of how functional impairment and deformities develop in individuals. The therapist must always keep in mind the progress of deformity and try to thwart this vicious circle as early as possible. It is much more beneficial to prevent deformities than to treat them.

Development of joint deformities

Inflammatory arthritis is very painful. Pain in children with arthritis, however, is often neglected as small children in particular rarely complain of painful joints. It is therefore important to note non-verbal expressions of pain (Scott *et al*. 1977; Melvin 1989; Altenbockum *et al*. 1993; Truckenbrodt 1993). Unfortunately children

rapidly adapt to painful arthritis with a reflex pain-relieving positioning of the affected joint. This is always a malposition which in turn produces a muscular imbalance. Those muscles which draw the joint into the relieving position become hypertonic, and the antagonists become weak (Altenbockum *et al*. 1993; Truckenbrodt 1993. If caught at an early stage the deformity can still be corrected passively. If inflammation and pain persist the incorrect position becomes permanent and a normal part of daily activities. The deformity usually increases during physical activity. Finally neither active nor passive correction is possible; a fixed deformity has developed. Incorrect weight distribution on the joints favours arthritic destruction which increases pain and deformity; hence, the vicious circle. At this stage a complete restoration of joint function is nearly impossible. Unaffected neighbouring joints can develop secondary deformities through compensatory over- or misloading. If neighbouring joints are also affected, primary and secondary deformities combine.

Principles of physiotherapy

It is the aim of all therapeutic approaches to keep or restore joint function and alignment as much as possible and achieve a normal pattern of mobility (Jarvis 1980; Erlandson 1989; Melvin 1989; Altenbockum *et al*. 1993). Even slight restrictions of joint function must be taken seriously and treated. In this regard it is important to appreciate that children have a greater joint mobility than adults. Joint function which is normal in adults may already be impaired in a child and must be treated accordingly.

As a precondition for effective physiotherapy the child must relax during treatment, as fear and pain increase muscle tone and intensify the reflex pain-relieving position. A relationship of trust between child and therapist is of the utmost importance in achieving the most beneficial environment and results.

Physiotherapy starts with passive-assisted movement of the joint which is non-weight-bearing. This manoeuvre is carried out with the utmost care within the pain-free range of movement and the best possible corrected axis. Thus even acutely inflamed and painful joints can be treated and protected from functional impairment. Careful movement reduces pain and improves mobility through relaxation of the hypertonic muscle groups.

Next the hypertonic and shortened muscle groups, which keep the joint in its incorrect position, must be stretched together with the other shortened joint structures. This procedure requires a relief position, and joint protection must be considered. Strong leverage to the joint should be avoided and positions such as hand prop, heel posture, or squatting are contraindicated as they increase intra-articular pressure. Stretching must be maintained over a long period of time to become effective. Individual splints may be used to maintain the corrected position. However, active stretching procedures such as hold–relax techniques should be restricted to non-inflamed joints, since they produce high intra-articular pressure.

A further step to extend mobility involves activation of the muscle groups which counteract the deformity. The child learns first with support by the therapist, and then on his or her own to rectify the incorrect position by tensing specific muscles.

Finally, the child must learn to integrate the regained mobility into daily life, overcoming the pathological patterns of movement. Muscular co-ordination is first trained by slow, consciously controlled movements. Frequent repetition of simple actions helps to integrate these into daily activities. In the beginning the therapist will need to support and correct the joint movements. Later more complex and faster movements can be achieved unaided. It is important, however, to adapt the training to the child's abilities. If the patient is overburdened by movements that are too complex, he or she will fall back into the unphysiological pattern of motion. This can lead to deterioration in the function of the involved joints.

Some therapists recommend strengthening of the weak muscles as a first step to overcome weakness and joint deformities (Jarvis 1980; Erlandson 1989; Melvin 1989). In our experience, however, this is ineffective as long as inflammation persists, joint alignment has not been regained adequately, and the child is still accustomed to a pathological pattern of motion. If, however, all these conditions have been restored, training of muscle co-ordination will simultaneously increase muscular power. Forced strengthening of individual muscle groups is rarely necessary and exercises to combat resistance should be discouraged since they increase pressure on the joints (Altenbockum *et al*. 1993).

The cervical spine

Involvement of the cervical spine mostly affects children with systemic onset arthritis and in those with seronegative polyarthritis, rarely in pauciarticular disease. The typical pain-relieving position is in mild flexion with restricted extension. All other movements are also impaired. Cervical spine involvement is easily detected when the child is asked to look up or around. Extensive eye movements compensate for the restricted extension and rotation (Fig. 1). Neck pain can be especially severe in children with systemic disease. It is induced by playing, writing, or working in a sitting position with the head slightly flexed. The cervical spine tends to early ankylosis in juvenile arthritis. Radiographic lesions usually start at the C2/C3 level and progress downwards. Fusion of the whole cervical spine can develop (Fig. 2). Atlantoaxial subluxation is rare in children (Ansell and Kent 1977).

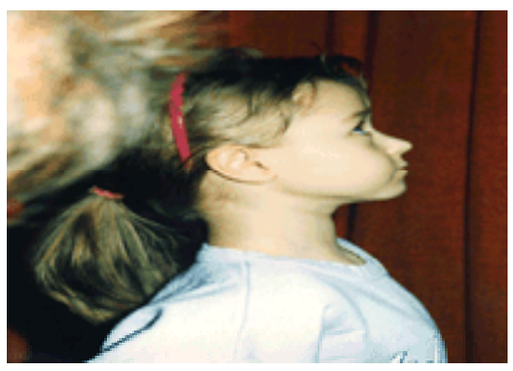


Fig. 1 Child with systemic onset disease and involvement of the cervical spine. Impaired extension is compensated by upward movement of the eyes.

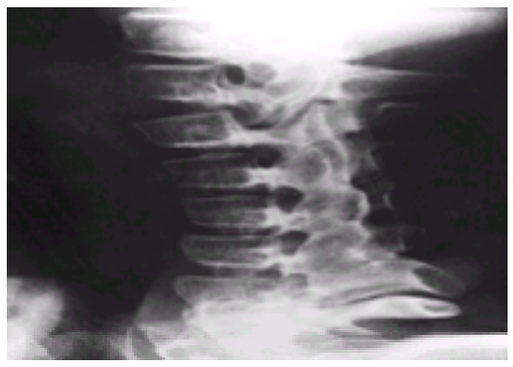


Fig. 2 Spinal ankylosis in a child with systemic juvenile chronic arthritis.

Therapy

Treatment of the cervical spine requires extreme caution and should be carried out in a relief position, preferably with the child lying on his or her back. Careful passive movement in all directions with slight traction can relieve pain and improve function. The dorsolateral neck muscles can be stretched by moving the shoulder girdle while the cervical spine remains in a fixed position. However, if spinal ankylosis has already developed no manipulation should be attempted.

A soft collar relieves pain and muscular tension of the cervical spine. It should be worn during long sedentary periods and as soon as pain starts. Children with severe neck problems may need to wear a collar all day. To aid comfortable sleep, pillows with a neck pad are helpful (Jarvis 1980; Erlandson 1989; Melvin 1989).

The temporomandibular joints

These joints are often neglected since they are mainly affected in children with polyarthritis where many other joint problems predominate. However, the sequelae of temporomandibular joint involvement can have severe impact on the child's well-being. If these joints become affected in early life significant growth disturbance of the mandible can occur (Bache 1964; Stabrun *et al*. 1988; Melvin 1989). This micrognathia creates malocclusion and disturbs the facial appearance. Restricted mouth opening together with impaired extension of the cervical spine can cause problems if intubation becomes necessary.

Movement of the temporomandibular joint comprises a caudal and ventral gliding of the mandibular condyle to permit mouth opening. Arthritis leads to pain and restriction of both movements, and opening the mouth and lateral displacement of the jaw become impaired (Bache 1964; Melvin 1989) (Fig. 3).

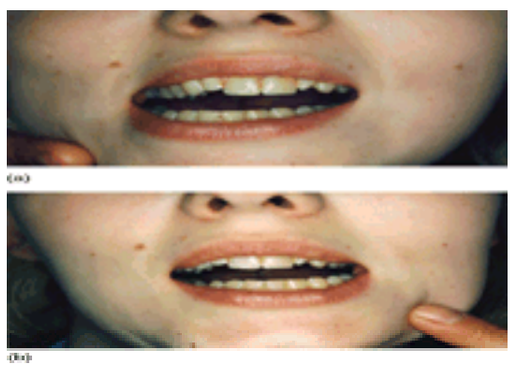


Fig. 3 Asymmetric involvement of the left temporomandibular joint. Lateral movement of the jaw is normal to the right (a) but markedly impaired to the left side (b).

Therapy

Cautious traction with the therapist's thumbs on the dorsal dental rows can mobilize the caudal–ventral motion of the jaw (Fig. 4). Mobilizing grips from outside the mouth are rarely possible since the temporomandibular joint area is very painful and effective therapy would require significant pressure. Active or actively supported exercises with ventral and lateral moving of the jaw are indicated to improve function. Temporomandibular joint involvement needs close collaboration with the dentist or orthodontist since dental or jaw regulation may be required to improve function and occlusion.

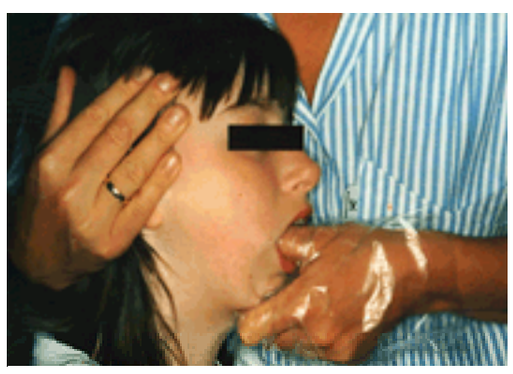


Fig. 4 Mobilization of the caudal–ventral motion of the jaw.

The shoulder

Arthritis of the shoulder occurs mainly in children with polyarticular disease. Since synovitis is difficult to detect and pain is usually greater in other joints it is easily overlooked.

Shoulder involvement impairs elevation and abduction. These movements involve a caudal gliding of the humeral head which is restricted in arthritis. Early co-operation of the shoulder blade during elevation and abduction compensates for impaired range of motion. The pain-relieving position of the glenohumeral joint tends towards adduction and internal rotation. The muscles for external rotation, abduction, and elevation become weak. Joints of the shoulder girdle, such as the acromioclavicular and the sternoclavicular joint, can also be affected. Their position of comfort comprises elevation and protraction of the shoulder girdle.

Therapy

Mobilization of the restricted external rotation, elevation, and abduction comes first (Fig. 5). Manipulation of the caudal gliding of the humeral head is helpful as well as stretching of the tense adductor and internal rotator muscles. Postisometric relaxation can also improve range of motion (Melvin 1989). When adequate shoulder mobility has been achieved muscular activation becomes possible which concentrates on the weak scapular muscles, the external rotators, and abductors. The scapular muscles must be activated towards a caudal and dorsal movement of the shoulder blade.

Fig. 5 Passive-assistive moving of the shoulder towards external rotation.

Exercises in a sling suspension are beneficial for shoulder involvement. They enable active mobilization while the arm is protected from gravity to eliminate joint

stress. The child can swing the arm into elevation and retroversion while lying on the side and into horizontal abduction and adduction while sitting upright.

The elbow

Synovitis of the elbow joint first impairs extension. A flexion contracture up to 30° may develop unnoticed since it hardly interferes with normal activities. Further contracture, however, significantly reduces arm length and can eventually lead to problems with dressing, toilet care, and other daily tasks (Erlandson 1989; Melvin 1989). Severe elbow involvement may also impair flexion which then mainly interferes with eating, combing, and facial care.

Forearm rotation favours pronation as a position of comfort while supination is usually restricted.

Therapy

Mobilization is directed towards the restricted function, usually extension and supination. Since synovial tissue tends to fill the fossa olecrani, moving the elbow may be very painful. Manual traction vertical to the forearm can relieve pain and improve extension, while stretching of the flexor and pronator muscles is important to restore function. Postisometric relaxation of the flexor muscles may also improve extension. Plaster splints which are used for about half an hour after manual stretching help to maintain the improvement.

The hand

In polyarthritis symmetrical involvement of wrist and finger joints occurs in over 90 per cent of patients. In oligoarthritis (or pauciarticular disease) asymmetric wrist involvement affects about 10 to 20 per cent, while single finger joints or digits may also be implicated (Melvin 1989; Altenbockum *et al*. 1993). The typical pain-relieving position of the wrist in children is flexion and ulnar deviation (Table 2) (Chaplin *et al*. 1969; Granberry and Mangum 1980; Findley *et al*. 1983; Erlandson 1989; Melvin 1989; Altenbockum *et al*. 1993). The extensors weaken while the flexor carpi ulnaris muscle becomes hypertonic. First active and later also passive dorsal wrist extension is limited. Secondary hyperextension of the metacarpophalangeal joints may occur (Fig. 6). To compensate for ulnar wrist deviation the metacarpophalangeal joints often drift radially even when they are not affected. The result is a juvenile hand scoliosis often described as 'Z deformity' (Fig. 7), in contrast to adult hand deformity with radial deviation of the wrist and ulnar drift of the fingers (Granberry and Mangum 1980; Altenbockum *et al*. 1993). Occasionally, and particularly in older children with disease positive for rheumatoid factor, an adult-type hand deformity can develop. Finally, the disturbed muscular balance together with loosening of capsule and ligaments results in a volar subluxation of the carpus (Table 2).

Fig. 6 Arthritis of the wrist joint with a typical malposition in flexion. Impaired extension is compensated by hyperextension in the metacarpophalangeal joints.

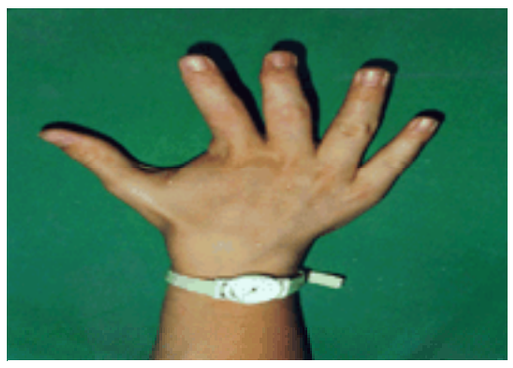


Fig. 7 Juvenile hand scoliosis with ulnar deviation of the wrist and compensatory radial drift of the fingers.

	No. of wrists affected (%)
Ulnar deviation	142 (59)
Juvenile hand scoliosis	82 (34)
Volar subluxation carpus	164 (68)

Study from the Children's Hospital for Rheumatic Diseases, Garmisch-Partenkirchen. 133 children with 239 involved wrist joints.

Table 2 Deformities of the wrist in juvenile chronic arthritis

Involvement of finger joints can lead to swan-neck or boutonnière deformities. A few children with polyarthritis develop the same malposition in all fingers. Often, however, combinations occur in the same hand. Swan-neck deformity develops mainly in index and middle fingers, boutonnière deformity in the fourth and fifth finger (Altenbockum *et al*. 1993). The thumb tends towards opposition and adduction in its saddle joint and flexion deformity in the metacarpophalangeal joint. A compensatory hyperextension often develops in the interphalangeal joint (Fig. 8).

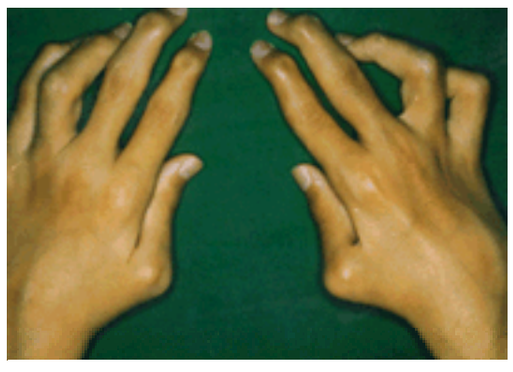


Fig. 8 Multiple finger deformities in a patient with juvenile polyarthritis: subluxation and adduction of the thumb at the metacarpophalangeal joint with hyperextension of the interphalangeal joint; swan-neck deformity of the index and middle fingers; and boutonnière deformity of the fourth and fifth fingers.

Flexortenosynovitis of the fingers is a common feature in juvenile arthritis. It can be very painful and may result in the adoption of a pain-relieving position with all three finger joints in flexion (Melvin 1989; Altenbockum *et al*. 1993). Due to impaired tendon gliding active, flexion and extension are often reduced although passive movement is still possible (Fig. 9). Finally shortening of the finger flexors develops with a significant impairment of hand function.

Fig. 9 Flexortenosynovitis of the middle finger. Active flexion is markedly impaired (a) while passive flexion is still complete (b).

Therapy

Treatment of the wrist must include restoration of passive and later active dorsal extension as well as radial abduction. The therapist's grip during passive-assistive movement should support the carpus and correct the hand axis. Slow movements which take account of the pain threshold reduce tension of the flexor carpi ulnaris muscle and enable careful stretching (Fig. 10). When joint inflammation has subsided, activation of the extensor muscles becomes part of the programme. Since abduction of the thumb counteracts ulnar wrist deviation it is important to relearn and practise spreading the thumb. It is also important that the children become aware of the physiological course of movement with the therapist's assistance. He or she must correct inappropriate movements such as hyperextension of the fingers or flexion of the elbow that compensate for impaired dorsal extension.

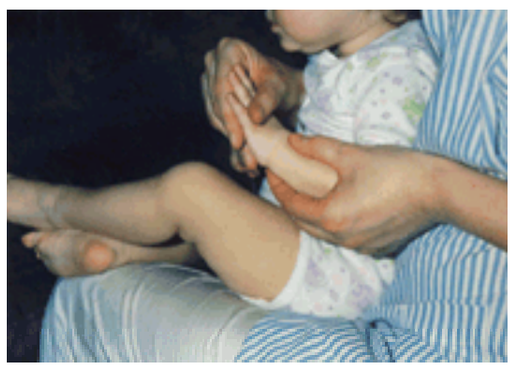


Fig. 10 Cautious passive-assistive moving of the wrist in a small child with polyarthritis.

For small children play activities can help restore active wrist extension. For instance wiping shaving cream over a mirror or painting a sun on the palm and volar side of the fingers. The child is then asked to let the sun shine by raising the hand and spreading and extending the fingers (Fig. 11). Children also like to throw small balls; here the therapist must control the movement to prevent compensation by other joints, especially the elbow (Fig. 12).

Fig. 11 A child demonstrates full range wrist extension and spreading of fingers by letting the sun shine.

Fig. 12 A boy trains active wrist extension by throwing balls. The therapist assists to prevent elbow movement.

Therapy of affected finger joints must take account of individual deviation. In swan-neck deformity it is important to mobilize impaired extension of the metacarpophalangeal joints and then continue with flexion of the proximal interphalangeal joints (Fig. 13). Muscular co-ordination of the extensor digitorum longus with the flexor digitorum superficialis muscle must be trained to achieve a physiological grip. In boutonnière deformity the often hyperextended metacarpophalangeal joints must be mobilized towards flexion. In a further step, stretching of the flexor digitorum superficialis muscle improves proximal interphalangeal flexion contracture.

Fig. 13 Treatment of swan-neck deformity. Mobilization of extension at the metacarpophalangeal joint (a) is followed by flexion at the proximal interphalangeal joint (b).

Flexotenosynovitis is treated by stretching of the flexors which, unfortunately, must exceed the pain threshold in order to be effective. The impaired flexion must also be trained. In treatment of flexortenosynovitis both passive and active movements are necessary to prevent adhesion of the tendons. To reduce pain, cold applications prior to physiotherapy are helpful (Jarvis 1980; Melvin 1989; Altenbockum *et al*. 1993).

Hand orthoses

Therapy of wrist and finger deformities benefits from individual orthoses. The wrist joint can be stabilized by a working splint to prevent or correct flexion position, ulnar (or radial) drift, and carpal subluxation (Chaplin *et al*. 1969; Granberry and Mangum 1980; Jarvis 1980; Findley *et al*. 1983; Erlandson 1989; Melvin 1989; Altenbockum *et al*. 1993) (Fig. 14). These splints should be worn most of the day and especially during manual activities since all active work with the hand increases the deviation. Stabilization of the carpus will improve power transfer in the finger–hand area and will also protect the inflamed joint structures from over- and misloading

Fig. 14 Benefit of working splints. The malposition in flexion and ulnar deviation (a) is corrected with a stabilizing splint (b).

Resting splints are indicated for finger deformities. They should be worn for several hours during the night and contribute to a careful passive stretching of shortened muscles and joint structures (Fig. 15).

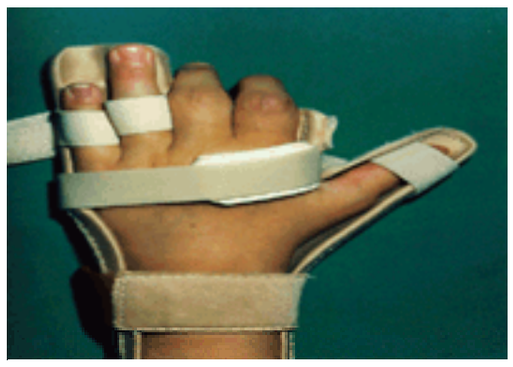


Fig. 15 Individual resting splint for a hand with combined finger deformities (see Fig. 8). The thumb rests in abduction; swan-neck deformity of the second and third fingers is corrected with extension at the metacarpophalangeal joint allowing flexion of proximal interphalangeal joints. Boutonnière deformity of the fourth and fifth finger requires proximal interphalangeal extension.

The hip

Hip disease often becomes a major problem in children with chronic arthritis. Frequency of hip involvement differs among the subgroups and varies with time (Fig. 16). It is most common in seronegative polyarthritis at onset and as the disease progresses. In systemic disease about one-fifth of these children start with hip arthritis which increases with follow-up. These patients also run a high risk of hip destruction (Fig. 17). Hip problems can also arise in patients with HLA B27-related arthritis or juvenile spondylitis.

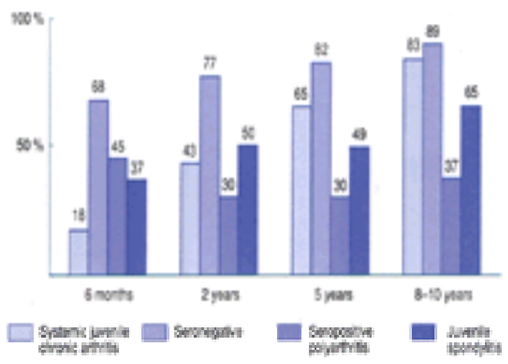


Fig. 16 Frequency of hip involvement in juvenile chronic arthritis within the first 6 months from onset and during follow-up. (Numbers derived from patient studies of the Children's Hospital for Rheumatic Diseases in Garmisch-Partenkirchen).

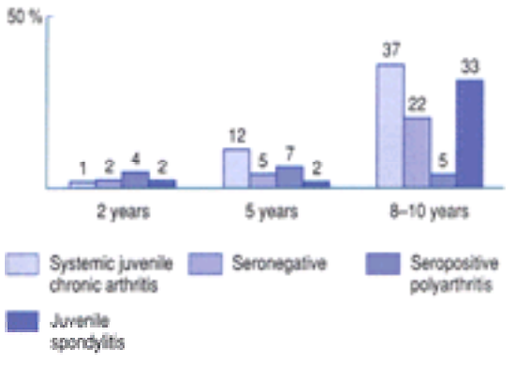


Fig. 17 Development of hip destruction in the subgroups of juvenile chronic arthritis (the same patients as in Fig. 16). The figure demonstrates the percentage of significant destructive lesions in children with hip involvement during follow-up.

Hip mobility in children exceeds that of adults. Young children with free hip function can bend their hips until the knee reaches the belly and external rotation may reach 90°.

The first sign of hip involvement is pain during full range flexion with adduction (Fig. 18). Later full extension, flexion, and inner rotation in 90° flexion become painful. When the children stand or walk they usually deviate into adduction and inner rotation. The abductor and external rotator muscles become weak while adductor and internal rotators as well as flexor muscles tend to shorten. Severe hip involvement is characterized by flexion, adduction, and inner rotation contractures (Swann 1978a; Melvin 1989). In such patients flexion position is compensated for by an increased lumbar lordosis as well as knee flexion (Swann 1978a; Erlandson 1989; Melvin 1989).

Fig. 18 Combined movement of flexion and adduction which is painful early in the course of hip involvement.

Therapy

Arthritis of the hip can be very painful especially in children with systemic disease. In acute stages only careful passive moving and mild traction is tolerated. This, however, should be done as often as possible to prevent contractures. If pain interferes with a full hip extension it is important to position the child in bed, supporting the upper legs and knees by pillows in a mild flexion to allow the muscles to relax. Regular, cautious passive extension, however, is compulsory to prevent flexion contractures.

When pain subsides stretching should be done, concentrating on mainly the flexors and probably the adductor muscles. Later on, active exercises of extensor and abductor muscles is important to improve the range of movement of the joint.

Exercises in a sling suspension are indicated to mobilize the joint. Continuous or intermittent traction for 1 or 2 h a day may relieve pain and contracture. Periods of lying in a prone position are beneficial to increase hip extension (Ansell 1978; Swann 1978a; Jarvis 1980; Melvin 1989; Lloyd and Aldrich 1993; Hayem *et al*. 1994). It is, however, important that the child does not compensate for impaired hip extension with an increased lumbar lordosis when lying prone.

Weight-bearing exercises for children with hip arthritis is controversial (Lloyd and Aldrich 1993). Some therapists encourage ambulation to restore the joint structures and support joint congruity (Bernstein *et al*. 1977; Ansell 1978; Melvin 1989; Lloyd and Aldrich 1993; Hayem *et al*. 1994). Absence of weight bearing may interfere with a normal development of the femur and acetabulum in young children. Valgus deformity and lateralization of the femoral head and acetabular underdevelopment are often seen in children with onset of hip disease at an early age (Bernstein *et al*. 1977; Ansell 1978; Blane *et al*. 1987).

Conversely, weight bearing increases joint stress enormously and can promote hip destruction. Therefore, partial weight bearing is a compromise for children with hip arthritis (Rombouts and Rombouts-Lindemans 1971; Ansell 1978; Jarvis 1980; Garcia-Morteo *et al*. 1981). Depending on the child's age and the condition of the other joints we recommend crutches, bicycles, tricycles, or scooters with a saddle (Fig. 19) for mobility. Wheelchair use is to be avoided since sitting increases flexion

contractures of hips and knees (Ansell 1978; Swann 1978a; Melvin 1989).

Fig. 19 Special vehicle for children with involvement of the lower limbs to avoid weight bearing.

Children have a good potential for restoration of joint cartilage and bone (Rombouts and Rombouts-Lindemans 1971; Bernstein *et al*. 1977; Melvin 1989) (Fig. 20). In our experience walking is not necessary for repairing damage to cartilage or bone and may even delay it due to increased stress on the joints. We therefore prefer partial weight bearing and regular, continuous moving of the hips in a relief position. Such children benefit from using a sling suspension at home for daily exercises (Fig. 21).

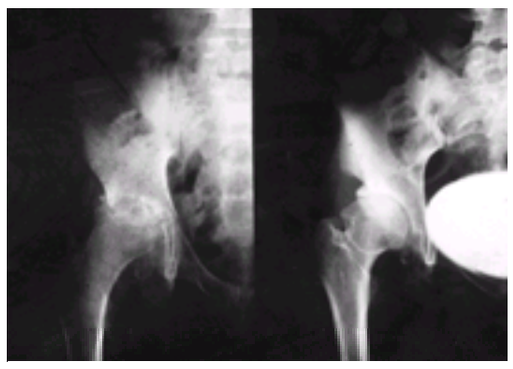


Fig. 20 Remodelling of hip joint destruction in a child with seronegative polyarthritis.

Fig. 21 Sling suspension for home exercises.

The knee

In all forms of juvenile arthritis the knee is most frequently involved. However, destructive lesions are less prominent compared with joints such as the wrist or hip. Deformities, however, can develop rapidly.

The typical pain-relieving position is flexion of the knee joint which is stabilized by an increased tension of the hamstring muscles (Swann 1978a; Jarvis 1980; Erlandson 1989; Melvin 1989; Altenbockum *et al*. 1993). The quadriceps muscle becomes hypotonic. Knee flexion contractures can develop into a very severe problem in young children. Each surgical procedure, even a small biopsy or arthroscopy, produces deterioration and should therefore be avoided in children under 5 years of age.

Flexion position of the knee enables rotatory movements to be made. Predominant activity of the biceps femoris muscle in childhood arthritis results in an outer rotation of the lower leg that is often compensated for by inner rotation of the hip joint, which then gives the impression of a valgus position of the knee but is really a pseudovalgus deformity (Altenbockum *et al*. 1993) (Fig. 22). Small children are particularly prone to develop a true valgus position which exceeds the physiological valgus for that age (Swann 1978a; Erlandson 1989; Melvin 1989; Altenbockum *et al*. 1993). It is caused by increased tension of the iliotibial band with ensuing instability of capsule and ligaments. Permanent knee flexion with dominant activity of the hamstring muscles favours dorsal subluxation of the tibia.

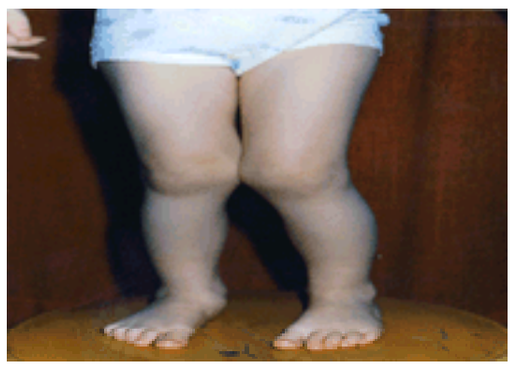


Fig. 22 Pseudovalgus deformity in a child with arthritis of the left knee. Flexion contracture with inner rotation of the leg gives the impression of a valgus position.

All these knee deformities are especially noticeable in young children. In a study of 63 patients with early-onset oligoarthritis who had 73 affected knee joints we have

seen flexion contracture and quadriceps atrophy in almost all. About 60 per cent also showed outer rotation of the lower leg and tibial subluxation (Table 3).

	No. of knees affected (%)
Flexion contracture	73 (100)
Outer rotation, lower leg	39 (53)
Subluxation tibia	40 (55)

Study of 63 children with early onset oligoarthritis. 73 involved knees.

Table 3 Deformities of the knee in juvenile chronic arthritis

Asymmetric knee involvement often results in a discrepancy in leg length due to increased growth at the affected knee joint.

Therapy

The first step in treatment of the knee should be to consider how to achieve full extension as well as the physiological hyperextension. When sitting with knee joints stretched the child should be able to lift the heel.

When treating a knee flexion contracture it is especially important that the therapist is careful to respect the pain threshold, for only if the child is totally relaxed does it become possible to stretch the hypertonic hamstrings, in particular the biceps femoris muscle.

Once there is improvement of passive extension the quadriceps muscle can be mobilized. A playful trick helps small children to find the right muscle: a face drawn on the child's knee with the mouth in the skin fold above the patella will start to smile when the quadriceps tenses.

A more difficult way to activate the quadriceps is when children throw balloons which have been put on their feet. Fast muscular co-ordination is best trained by letting the child kick about in the air or in water. However, the therapist must watch how far the full active extension is integrated in the course of movement. When muscular balance between hamstrings and quadriceps muscle has been restored the physiological movement pattern alone builds up muscle strength without extra power training (Altenbockum *et al*. 1993).

Gait training for affected knee joints must ensure a correct heel strike with fully extended knees and knee flexion during the swing phase.

Additional aids

Stretching of the hamstring muscles can be supported by splints (Ansell 1978; Swann 1978a; Jarvis 1980; Erlandson 1989; Melvin 1989; Altenbockum *et al*. 1993). They are put in place after physiotherapy in the best position of maximum extension.

As for hip treatment, arthritis of the knee also benefits from partial weight bearing (Swann 1978a; Jarvis 1980; Melvin 1989; Altenbockum *et al*. 1993). It reduces stress to the joint, enables relaxation of the hypertonic muscles, and helps to avoid the wrong pattern of motion during walking.

A leg length discrepancy must be corrected by a combined sole and heel lift, otherwise pelvic tilting and scoliosis of the spine can develop. The increased leg length also promotes fixed knee flexion of the affected knee.

The foot

The numerous foot joints and variety in the pattern of joint involvement usually lead to a combination of different axial deformities and in most patients the gait becomes disturbed (Swann 1978b; Rana 1982; Gschwend and Ivosevic-Radovanovic 1986; Lechner *et al*. 1987; Erlandson 1989; Melvin 1989; Altenbockum *et al*. 1993; Truckenbrodt *et al*. 1994). A mild dorsal extension is the typical position of comfort when the ankle or talonavicular joint is inflamed (Melvin 1989; Altenbockum *et al*. 1993) (Fig. 23). The muscular imbalance includes a hypertonic tibialis anterior and a weak triceps surae and peroneus longus muscle. A rheumatic heel foot can develop. Plantar flexion and especially cranial movement of the heel become impaired (Lechner *et al*. 1987; Melvin 1989; Altenbockum *et al*. 1993; Truckenbrodt *et al*. 1994); this is compensated by an increased flexion of the first metatarsophalangeal joint (Fig. 24). The ball of the big toe, normally a major weight-bearing area, is spared and atrophies. During walking, body weight is transferred from the heel over the lateral rim to the distal phalanx of the big toe (Altenbockum *et al*. 1993; Truckenbrodt *et al*. 1994) (Fig. 25).

Fig. 23 Pain-relieving position with the feet in neutral or mild dorsal extension in a child with arthritis of both ankles.

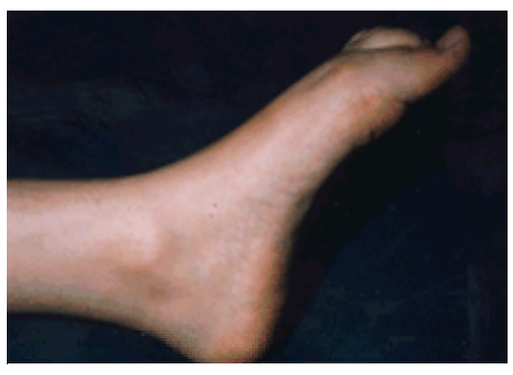


Fig. 24 Rheumatic heel foot: impaired plantar flexion is compensated by flexion of the big toe.

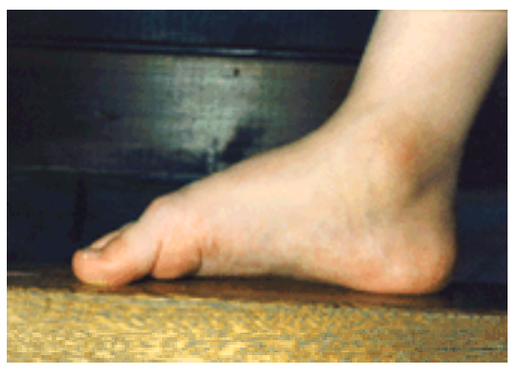


Fig. 25 Loading phase of a heel foot: the weight is shifted from the heel over the lateral rim to the distal phalanx of the big toe; the ball area is spared.

Young children in particular have a physiological tendency towards pes valgoplanus and develop this type of deformity (Swann 1978b; Melvin 1989; Altenbockum *et al*. 1993; Truckenbrodt *et al*. 1994). During walking, impaired mobility is compensated by an outward turning of the leg and rolling over the medial rim (Fig. 26). This incorrect weight distribution increases the valgus deviation and flattening of the longitudinal arch. The forefoot seems to stand in pronation. However, if the heel is corrected into neutral position a supination of the forefoot becomes obvious (Fig. 27). Pronation is always impaired.

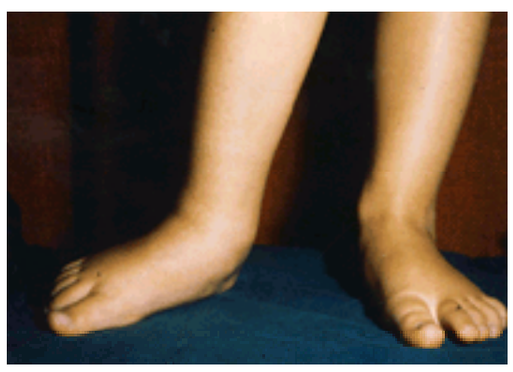


Fig. 26 Pes valgoplanus with outer rotation of the leg and rolling over the medial rim.

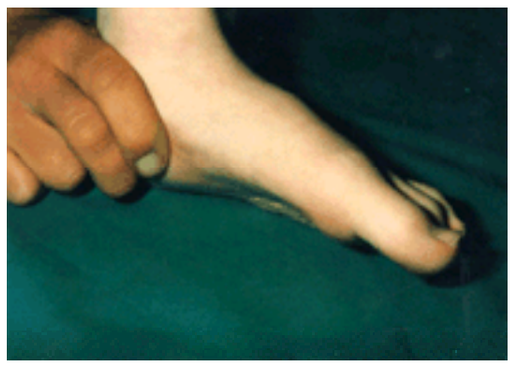


Fig. 27 Pes valgoplanus with erecting of the heel into neutral position. The forefoot stands in supination.

A rheumatic pes cavus develops less often and is seen mainly in older children with involvement of the distal intertarsal joints. These patients react with a reflex tension of the plantar muscles for pain relief (Fig. 28). Heightening of the longitudinal arch occurs together with an increased loading of the ball area. This often induces flattening of the transversal arch and claw toes. When medial intertarsal joints are especially painful, the weight is shifted to the outer rim and the heel glides into a varus position (Swann 1978b; Rana 1982; Gschwend and Ivosevic-Radovanovic 1986; Melvin 1989; Altenbockum *et al*. 1993; Truckenbrodt *et al*. 1994).

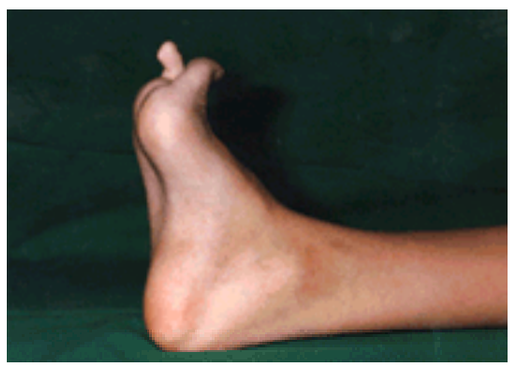


Fig. 28 Rheumatic pes cavus with tension of the plantar muscles.

The three major deformities of the foot can develop alone or in combination. Usually one malposition predominates. In a study of 123 children with foot involvement we have seen pes valgoplanus as the dominant deviation, especially in young children with early-onset oligoarthritis. In the other subgroups heel foot and pes valgoplanus occurred with about the same frequency, while pes cavus was seen less often (Fig. 29).

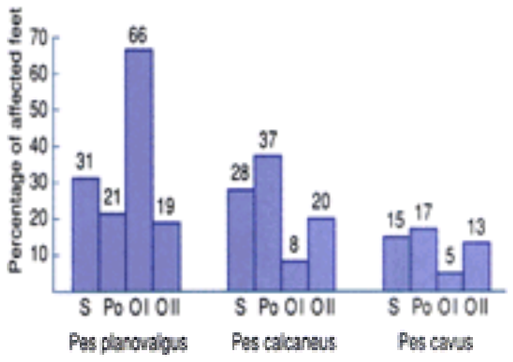


Fig. 29 Frequency of pes valgoplanus, pes calcaneus (heel foot), and pes cavus among the subgroups of juvenile chronic arthritis. (Study from the Children's Hospital for Rheumatic Diseases, Garmisch-Partenkirchen; 123 children, 246 feet). S, systemic juvenile chronic arthritis; Po, seronegative polyarthritis; O I, early onset oligoarthritis; O II, HLA B27-associated oligoarthritis

Other foot deformities concern the forefoot and toes. Arthritis of the first metatarsophalangeal joint often results in a hallux flexus to relieve the painful area of the big toe ball. This deviation resembles the heel foot where a secondary hallux flexus can develop. A compensatory hyperextension of the interphalangeal joint may occur. A hallux valgus is seen mainly in children with polyarthritis and involvement of the metatarsophalangeal joints. These patients also tend to develop claw or hammer toes (Rana 1982; Gschwend and Ivosevic-Radovanovic 1986; Melvin 1989; Altenbockum *et al*. 1993). Forefoot adduction often occurs together with hallux flexus or as a compensatory deviation from arthritis of the knee or ankle joint (Gschwend and Ivosevic-Radovanovic 1986; Melvin 1989; Altenbockum *et al*. 1993) (Fig. 30).

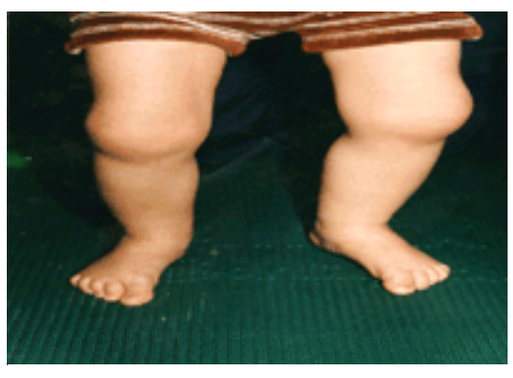


Fig. 30 Forefoot adduction and hallux flexus in a child with arthritis of both knees and ankles.

An equinus position of the foot is very rare in children with chronic arthritis and develops only in patients who have been immobilized in a wheelchair for a lengthy period (Gschwend and Ivosevic-Radovanovic 1986; Melvin 1989).

Therapy

The first step must be to obtain mobility of each of the affected foot and toe joints. To achieve better plantar flexion in the ankle joint the therapist must concentrate in particular on cranial movement of the calcaneus (Fig. 31). Mobilizing forefoot pronation provides a strong fixation of the heel. It is also important to restore extension of the first metatarsophalangeal joint, which is necessary for the physiological loading of the big toe ball. To regain muscular balance the tibialis anterior muscle must be stretched and the triceps surae and peroneus longus muscles be activated.

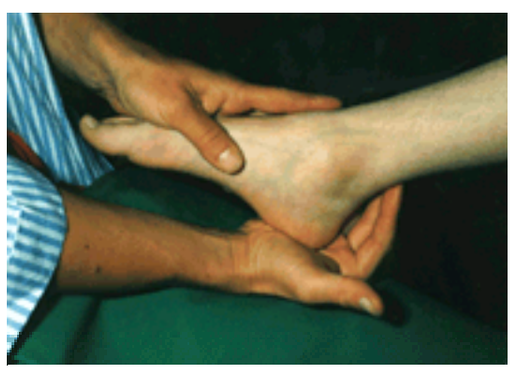


Fig. 31 Mobilization of plantar flexion with cranial moving of the calcaneus.

In a further procedure the combined movement of forefoot pronation with extension of the big toe must be trained (Fig. 32). This can be achieved with another game, painting the big toe ball to use as a stamp for printing patterns on a sheet of paper.

Fig. 32 Training of active toe extension together with forefoot pronation. The child is asked to push against the therapist's thumb.

To remedy the effects of foot involvement, intensive gait training is always necessary (Jarvis 1980; Lechner *et al*. 1987; Melvin 1989; Altenbockum *et al*. 1993; Truckenbrodt *et al*. 1994). A precise heel strike followed by rolling over the big and small toe ball with extended toes is needed. The push-off with plantar flexion of the ankle and extension of the toes is very important to train. The therapist must correct the wrong pattern and assist the re-education of each phase of a normal gait (Fig.

33).

Fig. 33 Gait training with assistance to relearn each single phase of the gait cycle.

Additional aids

Most children with foot involvement require insoles. They should correct the deviation as long as they are adjustable without pain under full weight bearing. Otherwise the insoles must primarily relieve the inflamed joints. It may even be necesssary to support a deviation until the pain-relieving position has improved.

For stabilization of the upper foot, shoes which extend over the ankle are useful. Soft soles lower the transfer of foot-strikes during walking and relieve all joints of the lower extremity. Severe foot deformities may require custom-made orthopaedic footwear. Partial weight bearing is also recommended for children with inflamed foot joints.

The sacroiliac joint

Sacroiliits can develop in children with HLA B27-related oligoarthritis or juvenile spondylitis. In a series of 71 juvenile patients with sacroiliits we detected radiographic changes of the sacroiliac joints between 1 month and 9 years after onset of peripheral arthritis. Only 33 of these patients complained of back pain, which was localized to the sacroiliac joint alone in 4 of them. Fifteen suffered from pain in the sacroiliac joints as well as the lumbar spine and 14 complained only of lumbar spine problems (Häfner 1987).

Since sacroiliitis can develop insidiously without any or only minor pain it is often overlooked. However, careful examination of these patients can detect the typical incorrect positioning even in the early stages. Physiologically the pelvis is positioned in a slight forward tilt, which means that the sacroiliac joints receive oblique pressure forces from the vertically oriented spine. Patients with sacroiliitis avoid such painful forces by raising the pelvis into a vertical position. However, this diminishes the physiological lumbar curve and contributes to lumbar tenderness, which is often a first symptom of sacroiliac involvement.

Therapy

For proper positioning it is important to maintain or restore the pelvic tilt and the lumbar lordosis. If the sacroiliac joints are painful, correct positioning may only be possible in a relief position. When inflammation and pain subside the patient must learn to maintain the correct position during sitting, standing, or walking. Since pain often blocks all movement of the sacroiliac joint, careful manipulation may help restore the physiological function. However, such manipulation may induce progressive pain during acute phases, in which case it should be postponed until the inflammation subsides, when movement can once more be tolerated. Some patients benefit from a tight pelvic belt which stabilizes the sacroiliac area.

Movement and training of the spinal muscles are always indicated to prevent or reduce secondary spine problems caused by sacroiliitis. They are also important prophylactic steps to prevent further spinal involvement, which may otherwise start in a number of these patients during early adulthood.

Juvenile dermatomyositis (see Chapter 5.9.2)

The acute stage of dermatomyositis is characterized by muscle weakness which can be severe enough to render the child immobile. About two-thirds of such patients also complain of muscular pain, with the trunk musculature usually being more severely affected than the peripheral muscles.

The chronic stage of the disease is exacerbated by progressive joint stiffness due to muscular scarring and shortening together with calcification of the skin and subcutaneous tissue.

Joint deformities follow a characteristic pattern. The wrist tends towards dorsal extension with flexed fingers while volar flexion of the wrist is impaired (Häfner and Truckenbrodt 1989) (Fig. 34). This is in contrast to chronic arthritis where extension is usually more limited and wrist flexion is the typical relief position. Elbows, hips, and knees are at high risk of developing flexion contractures. The ankle deviates into an equinus position. Contraction of the lower extremities is promoted by use of a wheelchair. Unfortunately during the acute stage when muscle weakness interferes with walking, children are often placed in wheelchairs and without adequate treatment they remain wheelchair-bound due to progessive muscular wasting and contraction.

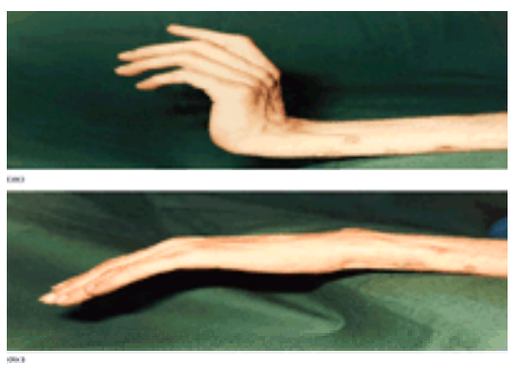


Fig. 34 Range of motion of the wrist in a child with dermatomyositis. Dorsal extension is complete (a) but flexion is severely impaired (b).

Therapy

During the acute phase the main therapy is concentrated on prophylactic treatment of muscle contraction. It includes proper positioning with hips and knees stretched for comfort and the ankle in a neutral position. If pain interferes with optimum positioning, flexed hips and knees must be supported with pillows to relieve muscular tension. Cautious passive movement towards extension becomes very important. Lying prone may help to prevent hip flexion contracture, and sitting with bended hips and knees must be restricted to short periods. Misguided use of the wheelchair can exacerbate the loss of movement.

Careful activated exercises of those muscles which tend to lose tension is very important, otherwise the child will forget what muscular contraction feels like. Muscle

strengthening, however, is not advisable since it can increase muscular damage. Active moving should be encouraged as far as the patient can tolerate without experiencing muscle fatigue.

As soon as joint contractures appear gentle stretching of the muscles is indicated. Use of a plaster cast for a few hours every day helps reduce contractures. Exercises in a sling suspension enable careful active movement with relief from gravity. A further beneficial procedure is hydrotherapy, especially for those children with severe impairment who enjoy the unusual experience of mobility that becomes possible through weight release in warm water.

Childhood scleroderma (see Chapter 5.8.1)

Systemic sclerosis is a very rare disorder in childhood and the paediatric rheumatologist is more often confronted with localized forms of scleroderma. These appear as morphea with round or oval-shaped lesions or in linear form. Fibrotic induration of skin and subcutaneous tissues can impair function depending on the extent of sclerosis. Individual problems arise according to the particular site of damage. Lesions on the face are mainly a cosmetic problem but can have severe psychological impact. Linear forms which cross a joint result in restricted mobility and can be especially disastrous if an entire extremity is involved. In such cases deformities and growth disruption can combine to produce a severe handicap (Fig. 35).

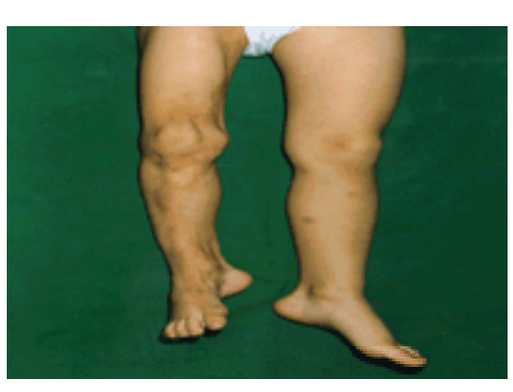


Fig. 35 Widespread localized scleroderma of the right leg with growth disturbance and deformity.

Therapy

The effect of medical intervention in childhood scleroderma is limited so that emphasis should be placed on physiotherapy and physical modalities. The main aspect of treatment aims to prevent or improve contraction. Careful manual stretching of the sclerotic structures is thought to be beneficial but should not exceed the patient's tolerance. Resting splints help to maintain joint position over a prolonged period. Exercises in a sling suspension as well as hydrotherapy in a warm pool represent two possibilities for maintaining or improving mobility without stressing the joints and neighbouring structures.

A beneficial effect can also be expected from manual lymphatic drainage. This process reduces tissue oedema which often dominates in the early phase. Connective tissue massage improves the peripheral blood flow and may contribute to healing of 'rat-bite' necrosis as well as Raynaud's phenomenon in systemic sclerosis (Földi 1991).

Widespread localized forms which involve the whole leg require special shoes with a heel and sole lift to adjust for deformities and growth disturbances.

Parental involvement

For children with chronic progressive disease a regular, often daily, therapy programme is necessary. The best way to guarantee effective treatment at home is to integrate parents into the therapy programme. They must be informed about how joint deformities can develop and learn the principles of physiotherapy. With adequate training most parents are able to undertake the task of daily exercises at home. They must also understand and learn to use the different orthoses and splints as well as the principles of everyday joint protection. Most parents are eager to participate in therapy if they realize that their active role can decisively improve the prognosis for the child's arthritis.

Problems with home therapy usually arise if progressive disease means treatment must be maintained over a prolonged period. In particular where patients suffer frequent relapse or continuous deterioration, general pessimism tends to reduce motivation. Parents and child may feel so discouraged that they abandon therapy. It is important to give them positive encouragement frequently, always acknowledge small achievements, and to assure parents that it is worth maintaining the same state or even accepting a reduction in the rate of progression in severe progressive diseases.

Regular exercises at home are usually necessary for an optimal outcome. It is, however, important that the therapeutic team relieves the family during times of acute stress and accepts temporary interruptions of the treatment routine. The schedule should always take account of holidays and leave enough space for fun and play.

Parental education should also include advice on appropriate leisure activities. Certain sports activities can lead to joint deterioration. On the other hand, sport is an important part of leisure activities in most families. It may be helpful if parents learn to give priority to sports which increase mobility without joint stress such as swimming or cycling, both activities which can have a therapeutic effect for children with rheumatic diseases. Depending on joint involvement and disease activity further sports may be acceptable: table tennis, gymnastics, horse riding, or cross-country skiing can usually be recommended for children with mild disease. Sports requiring a high level of exertion like tennis, skiing, or football should be reserved for patients in remission.

The child with arthritis who cannot participate in sports should be encouraged to gain self-esteem and stimulation in other fields. Parents should be encouraged to help compensate for their child's low level of physical activity with musical education, handicrafts, or playing games at home. Friends can be asked to visit and be integrated into leisure acitvities, which are adapted to the child's capability. This promotes the patient's social development and gives him or her the chance to feel equal in a healthy competition.

Joint protection and adapted devices

Joint protection training is an important task in the therapeutic regimen. The child with arthritis must learn how to reduce stress to the joints in daily activities. Effective joint protection relieves pain and can improve the inflammatory process. Principles of joint protection include:

1. proper positioning of joints to avoid deformities;
2. use of several instead of single joints;
3. transfer of a load from small to large joints or from involved to unaffected areas;
4. avoiding prolonged activity or positioning;
5. planning of rest breaks.

The child will assimilate these principles best if they are demonstrated during work or play. The occupational therapist first watches the patient during an activity and then corrects the incorrect joint position into a proper protective one (Fig. 36). The therapist must also explain the advantages of distributing a load on to stronger or larger areas, for example from the fingers to the palm, or how to use both hands instead of one.

Fig. 36 The occupational therapist corrects a child's hand position during clay crafting.

Joint protection may be supported by use of adapted devices. They help to reduce pain, preserve the proper joint position, minimize joint stress in daily activities, and increase the child's independence. In practice such equipment is prescribed less often for children than adults, since children feel embarrassed to use such appliances and devices with their peers and prefer to use alternative techniques to perform a task or even to ask for help.

Aid appliances, however, may become important for handicapped adolescents who wish to become independent. The most frequently prescribed appliances for children and adolescents include aids for walking and self-care items as well as play and school equipment.

We recommend partial weight release for children with arthritis of the lower extremities. Usually age-appropriate vehicles like tricycles or bicycles are sufficient, but some children may need special aids such as crutches or scooters with a saddle (see Fig. 19).

Self-care equipment should be restricted to severely handicapped patients since most children with arthritis can learn to perform daily tasks without special equipment. Sock cones are sometimes indicated for children with impaired hip and knee flexion, as well as dressing, combing, or hygiene articles with lengthened handles when shoulder elevation and elbow flexion are impaired. Since appliances designed to help adults are often inappropriate for children because they are not adapted to the child's size and reach, the occupational therapist must ensure that any such equipment is adapted for the individual child.

Some equipment for adults may be helpful for children and adolescents as well. Figure 37 and Figure 38 demonstrate the advantage of pot holders and adapted knives for proper wrist positioning. Spring-style scissors and foam material to thicken pencils relieve joint stress during crafting or drawing. Such devices are easy to handle and will be accepted within the peer group.

Fig. 37 Lifting pots with a regular pot holder urges the wrist into ulnar deviation (a). A second pot holder transfers the weight from one to both hands and guarantees proper wrist positioning (b).

Fig. 38 Cutting with a regular knife increases ulnar wrist deviation (a). The malposition is corrected by use of an adapted knife (b).

Integration into school

Most children with rheumatic diseases can and should be integrated into mainstream schools. Classes for the physically handicapped are not the solution since they only serve to increase the child's sense of being abnormal.

Doctors, social workers, and occupational therapists must communicate with educators and school nurses to find appropriate solutions and compromises for a practicable school life.

Problems start with transportation for those children who cannot walk or cycle to school. School buses which pick up the child at home are the best solution but not always available. A car pool with other families may relieve parents of the daily task of driving their child to and from school.

If teachers are informed early enough it is often possible to adapt the timetable to accommodate the child's needs. Less academic subjects could start in the early morning so that children with morning pain and stiffness can arrive later without missing too much. If physical education occurs at the end of the school day those children who are unable to participate can leave early and use the time for home work or physiotherapy. Classes with a handicapped child should preferably be allocated classrooms on the ground floor and moving between classes should be reduced to a minimum.

Keeping an extra set of books at home will save the child from having to carry heavy school bags and for necessary items backpacks should be preferred.

Children with arthritis of the hands may have a major problem with writing. Writing with splints is desirable to avoid ulnar deviation of the wrist but the changed position requires extra training during occupational therapy. Pressure to the finger joints can be avoided if children learn to write holding the pen between index and

middle finger and with the thumb in the opposed position. A thickened pencil also relieves joint stress (Fig. 39).

Fig. 39 Relief to finger joint pressure during writing occurs with an altered writing position (a) or a thickened pen handle (b).

Despite all these aids extended periods of writing may become painful and the child should receive extra time for class work to include rest periods. It is also helpful if the child is allowed to use a typewriter, computer, or dictaphone whenever possible.

Proper positioning during class and whilst doing homework is of the utmost importance, especially for children with back or neck problems. The solution consists of a chair adapted to the child's size, an inclined writing board or bookstand to avoid neck and back flexion, and a slanting cushion to tilt the pelvis and keep the spine erect (Fig. 40).

Fig. 40 Proper positioning during school or homework with an inclined bookstand and a slanting cushion.

All these arrangements can usually be accommodated in a mainstream school if educators are informed and willing to integrate a child with arthritis. The child's classmates receive practical education in social interaction and their support promotes the handicapped child's self-esteem and position in the community.

Guidance for the whole family

Arthritis in a child always has an impact on the whole family. Time-consuming demands, financial loss, and restriction of common activities alter the family's lifestyle; family members become resentful of their personal limitations and added responsibilities, siblings feel neglected and feel that their problems are of less importance or ignored altogether.

This situation requires guidance of the whole family, a task which all the team members may have to help resolve. Diagnosis of a chronic illness is usually devastating and adjustment to the new situation takes time, so it is vital that the health-care team help the families adapt to the altered situation. In the beginning thorough information together with hope for a favourable outcome are important to motivate each family member. Feelings of blame and guilt must be addressed and resolved to guarantee open communication within the family. Practical advice on how to overcome obstacles in daily life and adapt to a new schedule is most helpful.

If disease activity persists over months or years it becomes necessary to guide families towards a lifestyle which is as close to normal as possible.

Overprotection of the ill child and neglect of siblings is a common and understandable parental reaction, but should be avoided by early intervention of educators, psychologists, or other team members. The ill child must learn to cope with frustrations and healthy siblings need support for their problems too.

On the other hand a child with chronic arthritis does need special attention. Pain and physical limitations often exceed what can be tolerated for appropriate social interaction. Regression, aggression, or other behavioural disturbances can result. Parents are often overburdened trying to cope with such problems and should receive early professional advice.

Co-operation within the family facilitates a positive adjustment and parents should be encouraged to share responsibility for all disease-related activities. In many families mothers are left alone to cope with most of the tasks, so fathers must be stimulated to participate in the therapy routine, attend doctors appointments, and help with educational problems. Siblings may also become involved in caring for the ill child. If they receive appropriate recognition for this extra demand they can develop a strong feeling of self-esteem.

Health professionals cannot attend to all the needs of the family and some problems will remain. To deal with these, team members should encourage parents to join in family support groups as exchange of views among families with similar problems can be most helpful. Experienced parents can support families whose child has only recently been diagnosed. The health-care team must, however, be aware of negative influences and uncertainty among the group. This situation can best be avoided if health professionals become regularly involved in support group meetings.

Helping children cope with their disease

Chronic diseases elicit different emotions in different children. Young children usually feel most disturbed by restrictions on their physical activity. In this age group normal motor activity is strongly related to psychosocial development, a process which is endangered in children with rheumatic diseases.

For older children and adolescents the disease has most impact on their social life, as this age group is strongly influenced by peer pressure. The rheumatic disease with its restrictions and special demands does not fit into the life concept of the peer group. Youngsters therefore tend to deny the illness and its consequences.

Different methods of adjustment do exist. Where adjustment is negative children experience their illness as punishment or threat. Emotions of guilt, fear, or helplessness develop, which may result in regression or aggression. A positive adjustment is one in which the child accepts certain limitations but concentrates on personal strengths.

The health-care team together with family, school staff, and friends must encourage such positive adjustment. The children should experience as little limitation on their activities as possible. While treatment discipline is necessary, it should always leave enough time for play and hobbies. Friends and peers should be integrated

into leisure activities. Healthy competition, which is important for the child's development, must be directed towards activities within the child's capabilities. Success at school or competence in music, crafts, games, etc. promote the child's self-esteem.

Loving care is beneficial and especially important during painful relapses and other times of crisis. Overprotectiveness, however, can be devastating and hinder the child's psychosocial development, and children should always be encouraged to become independent. This includes personal care as well as control over medical care and treatment. As they grow up children can learn to determine treatment times and share responsibility for medication, doctor's appointments, etc. Parents are often reluctant to encourage independence and self-care as they fear loss of control over their child's life or tend to underestimate his or her capabilities. Health professionals can mediate between parents and child in this situation and actively support the child's struggle for independence.

Children who experienced pain and physical restrictions from a long-standing rheumatic disease but learned to cope with it often develop a more mature personality compared with others of their age group. Overcoming obstacles which their friends couldn't even dream of can be a source for more self-reliance and positive self-esteem. The adolescent may be earlier and better qualified for an independent life despite some physical limitations. Parents and health-care team members must recognize the child's struggle for freedom and facilitate appropriate self-determination and self-support.

Chapter References

Altenbockum, C. von, Hibler, M., Spamer, M., and Truckenbrodt, H. (1993). *Juvenile chronische Arthritis. Entwicklung von Achsenfehlstellungen an Hand, Knie und Fuss und ihre krankengymnastische Behandlung*. Hans Marseille Verlag, München.

Ansell, B.M. (1978). Rehabilitation. In *Surgical management of juvenile chronic polyarthritis*, (ed. G.P. Arden and B.M. Ansell). Grune & Stratton, New York.

Ansell, B.M. and Kent, P.A. (1977). Radiological changes in juvenile chronic polyarthritis. *Skeletal Radiology*, **1**, 129–44.

Bache, C. (1964). Mandibular growth and dental occlusion in juvenile rheumatoid arthritis. *Acta Rheumatica Scandinavica*, **10**, 142–53.

Bernstein, B., Forrester, D., Singsen, B., Koster King, K., Kornreich, H., and Hanson, V. (1977). Hip joint restoration in juvenile rheumatoid arthritis. *Arthritis and Rheumatism*, **20**, 1099–104.

Blane, C.E., Ragsdale, C.G., and Hensinger, R.N. (1987). Late effects of JRA on the hip. *Journal of Pediatric Orthopedics*, **7**, 677–80.

Chaplin, D., Pulkki, T., Saarimaa, A., and Vaino, K. (1969). Wrist and finger deformities in juvenile rheumatoid arthritis. *Acta Rheumatica Scandinavica*, **15**, 206–23.

Erlandson, D.M. (1989). Juvenile rheumatoid arthritis. In *Pediatric rehabilitation. A team approach for therapists*, (ed. M.K. Logigian and J.D. Ward). Little, Brown and Company, Boston.

Findley, T.W., Halperin, D., and Easton, J.K.M. (1983). Wrist subluxation in juvenile rheumatoid arthritis: pathophysiology and management. *Archives of Physical Medicine and Rehabilitation*, **64**, 69–73.

Földi, M. (1991). Physikalische Therapie bei der Sklerodermie. In *Lehrbuch der Lymphologie*, (ed. M. Földi and S. Kubik). Gustav Fischer Verlag, Stuttgart.

Garcia-Morteo, O., Babini, J.C., Maldonado-Cocco, J.A., Gagliardi, S., and Yabkowski, J. (1981). Remodeling of the hip joint in juvenile rheumatoid arthriris. *Arthritis and Rheumatism*, **24**, 1570–4.

Granberry, W.M. and Mangum, G.L. (1980). The hand in the child with juvenile rheumatoid arthritis. *Journal of Hand Surgery*, **5**, 105–13.

Gschwend, N. and Ivosevic-Radovanovic, D. (1986). Der Kinderfub bei juveniler Polyarthritis. *Orthopäde*, **15**, 212–19.

Häfner, R. (1987). Die juvenile Spondarthritis. *Monatsschrift für Kinderheilkunde*, **135**, 41–6.

Häfner, R. and Truckenbrodt, H. (1989). Die juvenile Dermatomyositis. *Internistische Praxis*, **29**, 307–20.

Hayem, F., Calède, C., Hayem, G., and Kahn, M.F. (1994). Involvement of the hip in systemic-onset forms of juvenile chronic arthritis. A retrospective study of twenty-eight cases. *Revue du Rhumatologie*(English Edition), **61**, 516–22.

Jarvis, R.E. (1980). Treatment of arthritis in children. In *Physiotherapy in rheumatology*, (ed. S.A. Hyde). Blackwell Scientific Publications, Oxford.

Lechner, D.E., Mc Carthy, C.F., and Holden M.K. (1987). Gait deviations in patients with juvenile rheumatoid arthritis. *Journal of the American Physical Therapy Association*, **67**, 1335–41.

Lloyd, S. and Aldrich, S. (1993). Paediatric rheumatology: meeting report. Second workshop of physiotherapy in JCA (Garmisch-Partenkirchen). *British Journal of Rheumatology*, **32**, 425.

Melvin, J.L. (1989). *Rheumatic disease in the adult and child: occupational therapy and rehabilitation*. F.A. Davis Company, Philadelphia.

Rana, N.A. (1982). Juvenile rheumatoid arthritis of the foot. *Foot and Ankle Journal*, **3**, 2–11.

Rombouts, J.J. and Rombouts-Lindemans, C. (1971). Involvement of the hip in juvenile rheumatoid arthritis. A radiological study with special reference to growth disturbances. *Acta Rheumatica Scandinavia*, **17**, 248–67.

Scott, P.J., Ansell, B.M., and Huskinsson, E.C. (1977). Measurement of pain in juvenile chronic polyarthritis. *Annals of the Rheumatic Diseases*, **36**, 186.

Stabrun, A.E., Larheim, T.A., Hoyeraal, H.M., and Rösler, M. (1988). Reduced mandibular dimensions and asymmetry in juvenile rheumatoid arthritis. Arthritis and Rheumatism, **31**, 602–11.

Swann, M. (1978*a*). Management of lower limb deformities. In *Surgical management of juvenile chronic polyarthritis*, (ed G.P. Arden and B.M. Ansell), Grune W Stratton, New York.

Swann, M. (1978*b*). The foot. In *Surgical management of juvenile chronic polyarthritis*, (ed. G.P. Arden and B.M. Ansell), Grune & Stratton, New York.

Truckenbrodt, H. (1993). Pain in juvenile chronic arthritis: consequences for the musculo-skeletal system. *Clinical and Experimental Rheumatology*, **11** (Suppl. 9), 59–63.

Truckenbrodt, H., Häfner, R., and Altenbockum, C. von (1994). Functional joint analysis of the foot in juvenile chronic arthritis. *Clinical and Experimental Rheumatology*, **12** (Suppl. 10), 91–6.

6.5 Corticosteroid injection therapy

Allan I. Binder

Introduction

The injection of local anaesthetic and corticosteroid agents into joints or soft tissue structures constitutes one of the most effective therapeutic options for the treatment of localized painful lesions of articular and soft tissue structures. The value of this form of treatment is widely appreciated (Bamji *et al*. 1990), leading to some concern with regard to the adequacy of the skills and experience of general practitioners and other health-care workers (Phelan *et al*. 1992) who increasingly perform these injections.

This chapter will outline the basic principles of steroid injection therapy, with the technique and special considerations for the more common conditions where it is employed. Sports-related and spinal use is more specialized and will not be considered in detail. The majority of injections can be performed in the clinic, ward, or, if necessary, in the patient's home, using an aseptic no-touch technique.

Effect of steroid on joints and soft tissue structures

Systemic corticosteroid agents have potent anti-inflammatory effects which result in an improvement in pain, stiffness, and systemic illness in patients with active inflammatory arthropathy. They may also significantly reduce the rate of joint damage in early rheumatoid arthritis (Kirwan *et al*. 1995). However, the limiting factor for prolonged systemic steroid therapy is the considerable risk of side-effects such as osteoporosis, reduced resistance to infection, deficient wound healing, accelerated atherosclerosis, and the suppressive effect on the hypothalamic–pituitary–adrenal axis, which can persist for up to a year after corticosteroid therapy has been withdrawn. Atlantoaxial subluxation is also more common in rheumatoid patients treated with systemic steroids.

Local steroid injected into a joint aims to achieve a sustained concentration of the drug in the synovial fluid and to provide the maximum anti-inflammatory effect locally whilst minimizing absorption into the plasma with its attendant risk of systemic side-effects. After injection, the corticosteroid passes from the synovial fluid into the synovial cells, before being gradually released into the circulation to be cleared. There are many steroid agents available and the localized effects depend on the potency, solubility, and dose of the particular agent used.

Hydrocortisone acetate is short acting, has a weak anti-inflammatory effect, is fairly soluble, and is completely absorbed from the joint within hours. The synthetic corticosteroids, such as methylprednisolone acetate, prednisolone acetate, and triamcinolone acetonide, are much more potent and less soluble and hence remain localized in the joint for several weeks (Gray *et al*. 1981). Triamcinolone hexacetonide is the most insoluble steroid, and has the slowest absorption, lowest plasma levels (Derendorf *et al*. 1986), and longest action (Bird *et al*. 1979) of the drugs available.

The benefit of local steroid injection can be observed by the profound reduction in inflammatory change and decrease in the expression of genes that play a role in articular destruction (collagenase, tissue inhibitors of metelloproteinases, complement, and HLA-DR), seen in histological tissue samples following local steroid injection (Firestein *et al*. 1991). Local steroid also leads to an improvement in the relationship between synovial fluid and serum levels of hyaluronan, restoring it towards normal (Pitsillides *et al*. 1994). Corticosteroid crystals which are found in joints immediately after steroid injection, and even a week later, are distinctive, being both intra and extracellular when visualized by polarized light and electron microscopy (Gordon and Schumacher 1979).

The systemic absorption of corticosteroid after local joint injection, may be sufficient to reduce inflammation in non-injected joints, and suppress the hypothalamic–pituitary–adrenal axis for up to 4 weeks after a single dose of a long-acting steroid (Bird *et al*. 1979). However, serious systemic side-effects are rare after a single injection, except in patients with brittle diabetes, infection, or where emergency surgery becomes necessary.

Outside the joints, local steroid injection may also produce beneficial effects in acute or chronic lesions, even where an inflammatory component is difficult to confirm. The injection of local steroid into soft tissue structures produces a characteristic histological picture, with circumscribed deposits of acellular finely granular material that stains faintly with haematoxylineosin. In some cases a mild to moderate reaction of histiocytes, fibroblasts, and lymphocytes may be noted round the edges. Small elongated spaces are usually evident, with occasional weakly birifringent (positive or negative) crystals that have an identical shape and size to the empty spaces (Balogh 1986). There is little data on the absorption of steroid from tendon sheaths, bursas, or other soft tissue structures, although methylprednisolone has been shown to produce measurable plasma levels for a mean of 16 days after soft tissue injection (Mattila 1983). Not all the effects of local steroid on soft tissue structures are beneficial, with impairment in the healing of damaged ligaments (Wiggins *et al*. 1994) and atrophy of up to 40 per cent of skin and subcutaneous tissue thickness, following a single injection of steroid (Gomez *et al*. 1982).

Indications for local steroid injection

Joint puncture permits the aspiration of fluid for microbiological, crystal, and, occasionally, biochemical assessment, and should always precede steroid injection where infection, haemarthrosis, or crystal arthropathy is suspected. The colour, smell, turbidity, and viscosity of the aspirated fluid may help differentiate inflammatory arthropathy from crystal-related or non-inflammatory disease. However, a heavy concentration of polymorphs (pyoarthrosis) is equally likely with acute rheumatoid, infection, or gout. Conventional light and polarized microscopy and, in some cases, microbiological culture of the synovial fluid may be necessary before steroid can be safely injected. Soft tissue lesions, such as olecranon bursitis, are also amenable to aspiration, with fluid from these lesions being subjected to analysis similar to synovial fluid, before steroid is injected.

Local anaesthetic and/or steroid injection also has value as a diagnostic tool where the source of the pain is uncertain. For example the abolition of a 'painful arc' on

abduction, as a result of local anaesthetic injected with the steroid, confirms the diagnosis of rotator cuff tendinitis.

However, the main use of local steroid injection is for treatment of inflammatory and non-inflammatory conditions of joints or soft tissues, whether they are acute or chronic. The main anatomical structures with examples of common conditions amenable to local corticosteroid injection are shown in Table 1. Even acute gout and pseudogout will respond to steroid injection (Gray *et al*. 1981), although this therapy is not usually necessary or always feasible. The treatment of osteoarthritic joints with steroid injection is more controversial (Dieppe 1991), although the presence of an effusion increases the likelihood of short-term reduction in pain (Gaffney *et al*. 1995) in these cases.

Structure	Examples of diseases amenable to injection
Joints	Rheumatoid and other inflammatory arthropathies
Entheses	Tennis and golfers elbow, rotator cuff tendinitis
Tendon sheath	De Quervain's tenosynovitis, trigger finger
Bursae	Trochanteric, olecranon bursitis
Ligaments / muscles	Strains, sprains, sports injuries
Nerve compression	Carpal and tarsal tunnel syndrome
Nodules/ganglia	Rheumatoid arthritis, trauma
Trigger points	Cervical, lumbar
Epidural	Nerve root compression

Table 1 Indications for local steroid injection

Contraindications to local injection

There are some absolute and other relative contraindications to steroid injection therapy as shown in Table 2. Any suspicion of sepsis within the affected joint, on the same limb, or systemically precludes steroid injection, until infection has been excluded or adequately treated. Steroid injection is best avoided in joints previously affected by septic arthritis, as local defence mechanisms within the joint are suspect. Monoarthritis of unknown cause should always be regarded with suspicion and steroid injection delayed, until either a diagnosis is made or at least until infection has been excluded.

Absolute contraindication
Septic arthritis, septicaemia, active sepsis in locality, tuberculosis
Febrile patient, cause unknown
Serious allergy to local anaesthetic, steroid, or previous local injection
Sickle cell disease
Relative contraindication
Monoarthritis of unknown cause
Neutropenia, thrombocytopenia
Anticoagulants or bleeding diathesis

Table 2 Contraindications to local steroid injection

Septic arthritis is particularly easy to overlook in rheumatoid patients who are frail and elderly or being treated with systemic steroid or other immunosuppressive drugs. In these patients infection is often polyarticular, mimicking an exacerbation of the inflammatory disease. Although *Staphylococcus aureus* is the most common infecting organism, patients may not present with fever, weight loss, or other features usually associated with systemic infection (Kraft *et al*. 1985). A mild rise in neutrophil count and a high sedimentation rate may also fail to differentiate infective from inflammatory causes, although an exceptionally high C-reactive protein level is more suggestive of infection. While systemic ill health can be a prominent feature in septic arthritis, it can also occur in patients with rheumatoid arthritis without infection (systemic rheumatoid), and steroids should not be given by any route until infection, including tuberculosis, has been sought and if necessary treated.

Patients need to be asked about previous allergic reactions to local anaesthetic agents (e.g. with dental treatment), steroids, or local injections. An exacerbation of the pain for 24 to 48 h following a previous steroid injection does not suggest allergy but rather a crystal-related phenomenon (Berger and Yount 1990). Neutropenia, thrombocytopenia, and clotting abnormalities constitute relative contraindications to local steroid injection. If patients are on anticoagulants, these should be stopped for at least 24 h before the steroid injection. With other clotting abnormalities or blood dyscrasias, haematological advice should be sought, and appropriate preparations made to reduce the risk of bleeding following the injection. While haemophilic arthropathy has been successfully treated with cautious steroid injection (Shupak *et al*. 1988), this therapy should be avoided in sickle cell disease (Gladman and Bombardier 1987) as a sickle crisis may be precipitated by the procedure.

Potential dangers and side-effects

Local steroid injections are extremely safe, provided that the indication is appropriate, there are no contraindications, and a careful aseptic technique is used. Some side-effects apply to all injections (Table 3), while others are site-specific, and will be discussed below. While the exacerbation of pain for up to 48 h following the injection suggests a crystal-related phenomenon (Berger and Yount 1990), pain persisting beyond this time requires assessment for infection or other complications. Infection following steroid injection constitutes the most important side-effect, although it is very rare, with an estimated sepsis rate of 1 in 14 000 to 50 000 injections (Gray *et al*. 1981). Chronic debilitating illness such as severe diabetes mellitus or rheumatoid (Ostensson and Geborek 1991), or factors like drug abuse and alcoholism (Haslock *et al*. 1995), which suppress general immunity can predispose patients to infection. Postinjection sepsis may result from contamination of the injection equipment (Nakashima *et al*. 1987) or skin, haematogenous spread (Von Essen and Savolainen 1989), or reactivation of a previous infection. Accelerated joint destruction will also follow the injection of steroid into an undiagnosed septic joint.

Exacerbation of pain for 24–48 h (crystal related)
Septic arthritis
Accelerated destruction of an unsuspected septic joint
Subcutaneous tissue atrophy
Depigmentation
Transient facial flushing
Anaphylaxis
Peripheral nerve injury
Tendon rupture
Avascular necrosis, Charcot joint
Loss of diabetic control
Reactivation of tuberculosis
Cartilage damage
Soft tissue calcification

Table 3 Side-effects attributed to local steroid injection

Subcutaneous tissue atrophy (Fig. 1) or depigmentation (Fig. 2) are more likely if excessively large or repeated doses of long-acting steroid are given. While tendon rupture (Fig. 3) or facial flushing is encountered fairly often, other side-effects (Table 3) are rare. The importance of local injection in the development of avascular necrosis or a Charcot joint (Parikh *et al*. 1993) remains uncertain.

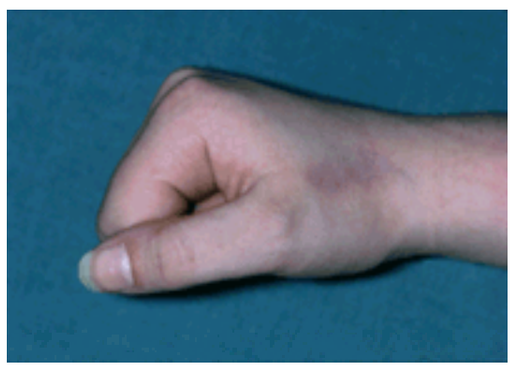


Fig. 1 Subcutaneous tissue atrophy following steroid injection of De Quervain's tenosynovitis lesion.

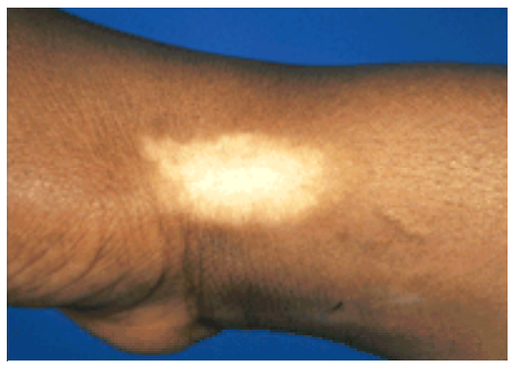


Fig. 2 Depigmentation following steroid injection of De Quervain's tenosynovitis lesion.

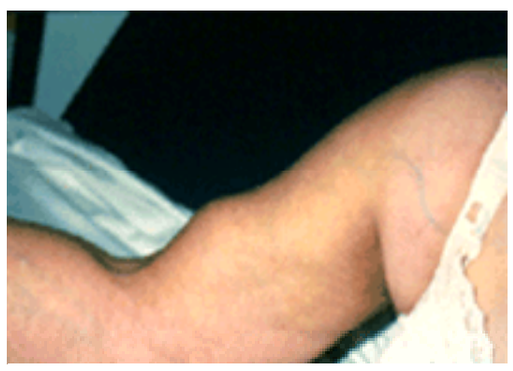


Fig. 3 Rupture of the long head of the biceps tendon following anterior shoulder injection.

Method of steroid injection

The method of injection of local steroid into painful joints and soft tissue structures is similar, although the exact technique may depend on the local anatomy, the size and depth of the lesion concerned, and the need to aspirate fluid before injection. The broad principles of injection therapy will be outlined first with locally determined modification in techniques being considered under the individual conditions.

As the success of local injection therapy depends on the accurate placement of the injected steroid (Jones *et al*. 1993), detailed knowledge of the anatomy at the site of injection is important. Consistency in the choice of equipment (needles and syringes) and adequacy of relaxation of the patient, and especially of the limb to be injected, further improves the feel, and hence reliability, of the injection. While the injection of large fluid-filled joints is relatively easy, other injections may require considerable skill and experience.

There is no single 'correct' method of performing steroid injection and most skilled operators develop their own unique style (Haslock *et al*. 1995), based on training, experience, and the equipment available. Irrespective of the individual variation in method, this should always be based on a consistent, safe, aseptic, 'no-touch' technique. The method presented reflects the author's own practice.

Equipment

Table 4 summarizes the equipment needed for routine steroid injection. Single-dose ampoules of steroid and local anaesthetic agents should always be used to eliminate any possibility of cross infection.

Steroid preparations (individual ampoules)
- short acting — hydrocortisone acetate 25 mg/ml
- long acting — methylprednisolone acetate 40 mg/ml
  - methylprednisolone acetate 40 mg/ml, lignocaine 10 mg/ml
  - triamcinolone hexacetonide 20 mg/ml
  - triamcinolone acetonide 40 mg/ml

Single-dose 1% lignocaine ampoules for injection

Refrigerant spray

Sealed isopropyl alcohol swabs

Other
- sealed single-use needles — 25G, 23G, 21G and 19G
- single-use syringes
- sterile cottonwool
- elasticplast, crepe bandages

Table 4 Equipment required for local steroid injection

Steroid agents

Hydrocortisone acetate is the only frequently used, short-acting agent (Gray *et al*. 1981) and is especially useful for superficial lesions, where there is a risk of subcutaneous atrophy or depigmentation, or in lesions (such as Achilles tendinitis) where tendon rupture may occur. Although there are many long-acting steroid agents available (Table 4), familiarity with one or two agents is adequate, as the differences between agents is not great. Methylprednisolone acetate has the advantage of being available alone or premixed with lignocaine, permitting greater flexibility in the adjustment of both steroid dosage and fluid volume when injected at different sites. Triamcinolone is also widely used as the acetonide or hexacetonide salts. In most circumstances, the long acting steroid agents are superior to hydrocortisone (Blyth *et al*. 1994), leading to a more sustained clinical improvement.

The dosage of steroid used varies according to the lesion injected. For most joint and soft tissue lesions, 25 mg of hydrocortisone, 40 mg of methylprednisolone or triamcinolone acetonide, or 20 mg of triamcinolone hexacetonide is adequate. Large joints, like the knee, may require double and very small joints or lesions half this dose.

Local anaesthesia

Local anaesthesia of the skin can be achieved by the use of a refrigerant spray such as ethyl chloride. Freezing numbs the skin rapidly and is especially useful in the hand and other sensitive parts of the body, or in patients requiring frequent injections. The skin is sprayed only until it turns white to achieve anaesthesia, as beyond this point a painful burn may ensue.

Lignocaine 1 per cent mixes well without flocculation with most steroid agents, and is usually introduced in combination with the steroid. Where the injection is technically difficult or fluid aspiration is to precede injection, the lignocaine is introduced first. The volume of local anaesthetic agent injected will vary according to the size of the joint or lesion being injected. While the use of lignocaine may reduce the immediate postinjection pain in tennis elbow or other soft tissue lesions (Haslock *et al*. 1995), the rationale for routine use in large joint injection is less clear cut (Kirwan *et al*. 1984).

Other equipment

Sealed isopropyl alcohol swabs or other alcohol-based agents provide a safe and cost effective method of achieving skin cleansing. Iodine, chlorhexidine in spirit, and other cleansers increase the cost of the procedure but offer no advantage over the sealed alcohol-based swabs (Cawley and Morris 1992) and can increase the risk of infection (Nakashima *et al*. 1987).

While a 2 or 5 ml sterile syringe is adequate for most steroid injections, 10 ml and 20 ml syringes are required for fluid aspiration. The choice of needle for any given site is also important. Most medium sized joints and soft tissue lesions can be injected with a 23G (blue) needle. A larger 21G (green) needle is used for larger joints, especially where aspiration is required, and for deep soft tissue lesions such as those round the buttocks and thighs. Where the synovial effusion is large or the fluid is thick, fibrinous, or purulent a 19G (white) needle may be needed for successful aspiration. Small hand and foot joints and superficial soft tissue lesions are best injected with a fine 25G (orange) needle.

Preparation

The method of preparation for joint aspiration and steroid injection is shown in Table 5.

Exclude contraindications or allergy
Explain the procedure before starting
Position patient so patient and limb are relaxed
Wash hands and dry carefully
Prepare injection immediately before use
  To aspirate: lignocaine and steroid separate
  To inject: lignocaine and steroid mixture
Mix steroid thoroughly before preparation and before injection
Use single-dose ampoules
Draw up the drugs yourself!
Change needle before injecting
Keep needle covered before use

Table 5 Preparation for steroid injection

The patient must be as relaxed as possible, as tension in surrounding muscles can make the injection difficult or even impossible. Where patient anxiety is a particular concern, most injections can be carried out with the patient lying supine with the head comfortably supported on pillows. This position facilitates the treatment of vasovagal attacks if they occur. The presence of a nurse helps to reassure and position the patient but the nurse must be dissuaded from drawing up the drugs or shaking them out on to a tray or table before the operator is ready. The procedure should be described to the patient in detail before starting the injection, to reduce the risk of air-borne infection. At the same time, the patient can be assessed for sepsis, allergy, or other contraindications to the injection.

Simple hand washing is adequate for the procedure if a no-touch technique is to be employed, and should precede preparation of the injection. Careful hand drying will prevent water running down the needle. While gloves are now recommended for joint aspiration (American College of Rheumatology Council on Rheumatological Care 1992), less than 50 per cent of American (Yood 993) and 10 per cent of British rheumatologists (Haslock *et al*. 1995) currently heed this advice. The injection is prepared with a combination of steroid and local anaesthetic in the same syringe, in most cases. Where the injection is likely to be difficult or where aspiration is to precede injection, the two agents are drawn up separately, the lignocaine being in a larger syringe. As the steroid preparations are crystalline, they need to be thoroughly mixed before being drawn up and again before injection. The needle should ideally be changed following preparation to maintain sharpness, and must remain covered until the injection is given to reduce the risk of infection.

Injection technique

A method for aspiration and/or steroid injection is summarized in Table 6. Talking should be kept to a minimum during the procedure. The exact site for injection is carefully marked with a blunt-pointed object such as a thumbnail or ball-point pen with the point retracted. This mark should only be made once careful positioning of the patient has been achieved, and needs to remain visible even after skin cleansing to permit a no-touch aseptic technique to be used.

Mark exact point of needle insertion with a blunt object
Use no-touch technique
Cleanse injection site with 2–3 alcohol swabs
Use refrigerant spray for skin anaesthesia
Fleswab injection site
Insert needle without touching the metal
For aspiration: keep needle stationary as fluid is tapped, then change syringe and inject steroid
For injection: pull back plunger before injecting
Remove needle carefully and check for bleeding
Use elastoplast if no allergy; crepe bandage after knee aspiration
Dispose of all needles and syringes safely and avoid needlestick injury
Rest or splint joint after injection if possible
Re-emphasize possible side-effects and benefits

Table 6 Technique for steroid injection

Only one or two joints should be injected per session, with the same joint being injected no more than three or four times a year. Patients should be warned of possible side-effects such as a temporary exacerbation of pain or destabilization of diabetic control following the injection. The patients should also be told about the possible beneficial effects of the absorbed steroid on systemic symptoms or other inflamed joints.

Postinjection

The injected part should, where possible, be rested for 24 to 48 h after injection (Neustadt 1992; Chakravarty *et al*. 1994), although admission for absolute bed rest is neither cost-effective nor practical. Splinting the injected joint may also prolong the duration that the steroid remains localized at that site (Dixon and Graber 1983). Where a large effusion has been drained, as from a knee, a crepe bandage should be firmly applied to provide support and take up the slack in the tissues stretched by the effusion. The maximum improvement usually develops in 2 to 4 weeks, and may last many months.

Joint and soft tissue lesions

The most common lesions amenable to steroid injection will be considered on a regional basis (Table 7), describing the method of injection, choice of steroid agent, volume of injection, and particular differences at each site. Several monographs are available outlining joint injection techniques (Dixon and Graber 1983; Doherty *et al*. 1992).

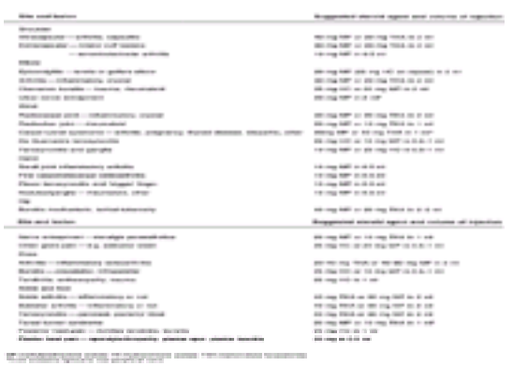


Table 7 Lesions amenable to injection with suggested steroid agent and volume of injection

The painful shoulder

Shoulder pain usually arises from soft tissue lesions, although arthropathies affecting the glenohumeral, acromioclavicular, and sternoclavicular joints also occur, with most lesions being amenable to steroid injection (Table 7). As shoulder structures arise from the C5 dermatome, pain is felt in the upper arm, maximally at the point of insertion of the deltoid muscle. Only the acromioclavicular joint arises from C4, and hence pain radiates to the neck and needs to be differentiated from cervical lesions. Shoulder pain can also result from polymyalgia rheumatica, or can be the site of referred pain from cervical, thoracic, or intra-abdominal pathology.

Injection therapy is much more cost-effective than physiotherapy (Dacre *et al*. 1989) for shoulder lesions but the site of steroid injection needs to be based on anatomical pathology rather than the localization of tender or trigger points (Hollingworth *et al*. 1983). Some studies have failed to confirm the benefit of injection therapy (Adebajo *et al*. 1990) but this may reflect difficulties in placement of the injection (Hollingworth *et al*. 1983).

Glenohumeral joint

This joint communicates with the tendon sheath of the long head of the biceps tendon, but not with the subacromial bursa unless rupture of the rotator cuff has occurred. Aspiration and or injection of the glenohumeral joint can be achieved using the anterior or posterior route, as both provide good access to the joint, although the posterior route is technically easier.

Anterior route (Fig. 4)

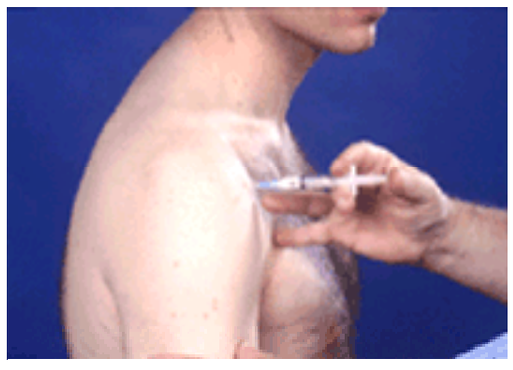


Fig. 4 Injection of the shoulder joint via the anterior route.

This route gives more reliable access in patients with adhesive capsulitis and is better suited for aspiration of large shoulder effusions. With the patient lying supine, the arm is rotated to identify the coracoid process and joint line anteriorly, and acromium posteriorly. Laying the arm across the abdomen to result in partial internal rotation of the shoulder, the injection is made just lateral to the coracoid in the line of the joint margin, with the needle directed towards the acromium. If correctly sited, the injection can be given without resistance.

Posterior route (Fig. 5)

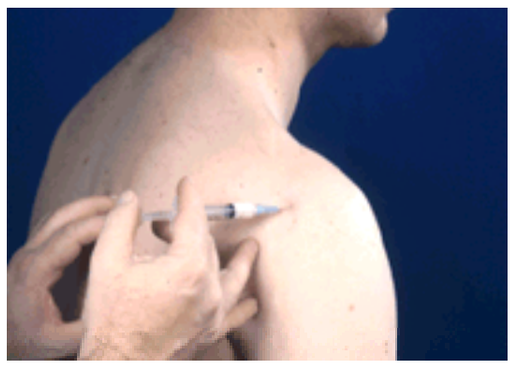


Fig. 5 Injection of the shoulder joint via the posterior route.

This is especially useful where the cause of the pain is uncertain or the operator lacks skill in injection procedure. With the patient seated across the couch and approached from behind, the thumb is used to identify the spine of the scapula which is followed laterally until it bends forward as the acromium. The point just below the acromium is marked. The forefinger is then used to palpate the coracoid process in front, rotating the shoulder if necessary to locate it. The needle is inserted under the acromium and gently advanced without resistance with the needle pointing towards the outer side of the coracoid process. More superficial injection is used to treat cuff lesions.

Subacromial bursa—lateral approach (Fig. 6)

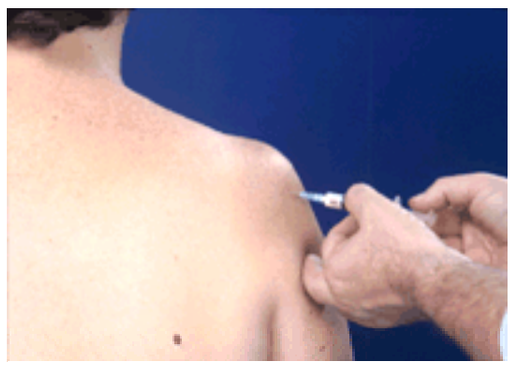


Fig. 6 Injection of the subacromial bursa via the lateral route.

Injection into the bursa is particularly useful for rotator cuff lesions, especially when accompanied by a painful arc on abduction. Although subacromial bursitis is rarely a primary diagnosis, it is frequently secondarily affected by pathological processes arising in the rotator cuff and other nearby structures. The subacromial space is very large, being in continuity with the subdeltoid space and also, with cuff rupture, the glenohumeral joint.

To inject the subacromial space, the patient should be seated and approached from the lateral aspect with the arm hanging vertically down, to use gravity to help enlarge the gap between the acromium above and the humeral head below. This gap is palpated, marked, and the needle advanced into the space, directed medially and slightly posteriorly. A relatively large volume of local anaesthetic (2–10 ml) is mixed with the steroid in view of the large capacity of the space being injected. The local anaesthetic rapidly abolishes the 'painful arc', which both helps confirm the diagnosis of a rotator cuff lesion and the correct siting of the injection. The benefit of this injection is particularly dramatic in patients with acute calcific tendinitis, where the pain is so acute that patients often present as an emergency.

While steroid injection to the glenohumeral joint or subacromial bursa often ameliorates the pain associated with adhesive capsulitis (frozen shoulder), it has less effect on the recovery of range of movement (Rizk *et al*. 1991). Arthrographic capsular distension, which ruptures the capsule, when combined with steroid injection, may lead to earlier recovery in range (Rizk *et al*. 1994) in adhesive capsulitis. However, a home exercise regimen which encourages hourly wall climbing to increase abduction and forward flexion, achieves a satisfactory rate of recovery in most patients with frozen shoulder.

Individual rotator cuff tendons

The individual rotator cuff tendons can also be injected with appropriate expertise (Cyriax 1984). The supraspinatus (superior) and subscapularis (inferior) tendons are anterior, and the infraspinatus tendon posterior, to the joint. Pain on resisted shoulder abduction, internal, and external rotation respectively, helps to identify the rotator cuff tendons involved. These injections are especially useful in patients in whom a 'painful arc' is absent, and hence where subacromial injection may be unsuccessful. Bicipital tendinitis is difficult to distinguish from rotator cuff tendinitis, but often coexists with it. While steroid injection via the anterior route may reduce pain from this source, rupture of the long head of biceps may follow the injection (Fig. 3).

Acromioclavicular joint

The acromioclavicular joint, which is on the anterosuperior aspect of the shoulder, is located by following the clavicle laterally until the joint is reached. Local tenderness confirms joint involvement. The joint is injected from the anterior or superior route, but only accepts 0.5 ml of fluid. Care is necessary to avoid deep injection as a pneumothorax can result from apical lung penetration.

Elbow

Elbow pain is extremely common and usually results from epicondylitis. Care is necessary to exclude pain referred from the cervical spine, brachial plexus, or shoulder, which can closely mimic primary elbow pathology. Most elbow lesions (Table 7) are injected from the lateral side, with the patient sitting in a relaxed manner and with the elbow resting on the examination table at an angle of 90°. Only medial epicondylitis and ulnar entrapment neuropathy require a greater angle and appropriate positioning to permit access to the medial aspect of the elbow.

Epicondylitis

Epicondylitis can result from trauma, lifting heavy weights, or repetitive power rotation movements. It is particularly difficult to cure when it is occupation or sport related.

Lateral epicondylitis (tennis elbow) (Fig. 7)

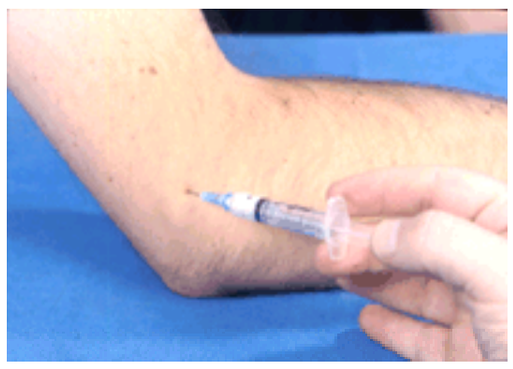


Fig. 7 Injection of a tennis elbow lesion.

In tennis elbow, localized tenderness is found either just distal to the lateral epicondyle or over the radial head. Pain is exacerbated by wrist dorsiflexion against resistance, especially when the elbow is extended and the hand prone. Grip strength is also reduced with the arm held in a similar position. While supination and pronation may be painful, there is no significant loss of elbow range.

Local steroid injection is directed to the site of maximum tenderness, with the injection aimed at 45°, to end near the insertion of the common extensor tendon to bone. A fair amount of pressure is needed to inject at this site, so care should be taken to ensure the needle is firmly attached to the syringe.

Medial epicondylitis (golfer's elbow)

Golfer's elbow causes tenderness very similar to tennis elbow but it is localized to the region of the medial epicondyle. Pain is usually noted on resisted palmar flexion of the wrist against resistance with the hand held supine.

Injection is again directed at the site of maximum tenderness, near the common flexor tendon insertion to bone. Care is necessary to avoid injury to the ulnar nerve which lies in a groove just behind the medial epicondyle.

There is no agreement with regard to the steroid agent of choice to inject tennis and golfer's elbow lesions. While triamcinolone is more effective than hydrocortisone (Price *et al*. 1991), it is more likely to lead to atrophy of the subcutaneous tissue. Both temporary postinjection exacerbation of pain and early relapse of symptoms are common following injection therapy (Dijs *et al*. 1990; Price *et al*. 1991), irrespective of the type of steroid used. The most logical approach is to use a long-acting steroid for the first injection, and hydrocortisone for further injections or for treatment of relapses, but with no more than three injections being given to treat the same lesion. Further injections can lead to long-lasting or irreversible atrophy which may amplify the pain from trivial knocks. Healing of epicondylitis lesions is often slow and diffuse forearm ache may precede recovery.

Elbow joint

The elbow joint is commonly involved in inflammatory arthropathies, causing swelling and painful limitation of elbow movements. The joint is injected using either the posterior or anterolateral route.

Posterior approach (Fig. 8)

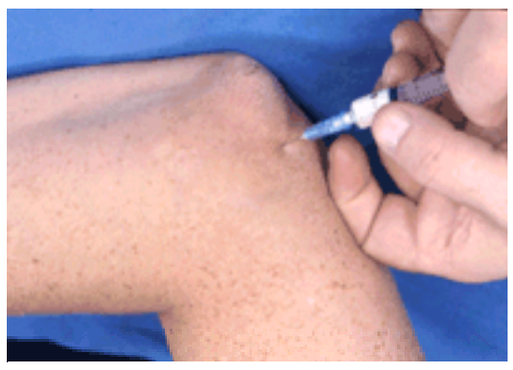


Fig. 8 Injection of the elbow joint via the posterior approach.

This route is easier to use before gross deformity has developed. The posterior joint line can be identified by placement of the thumb on the lateral epicondyle and third finger on the olecranon. The paraolecranon groove between the two fingers identifies the joint line, which runs between the two heads of the triceps tendon. Injection into the joint is just above and slightly lateral to the olecranon, where a bulge can be felt in the presence of a joint effusion.

Anterolateral approach

The thumb is used to palpate the head of the radius at the radiohumeral part of the elbow joint, best identified by rotating the forearm while palpating. An effusion may be evident at this site as a result of synovial extension around the radial head. Once the joint line is identified, the injection is made tangentially just under the capsule, not attempting to reach the centre of the joint.

Olecranon bursitis

Olecranon bursitis can result from infection, trauma, rheumatoid arthritis, or gout. Once infection has been excluded, the bursa can be aspirated and injected with steroid (Smith *et al*. 1989). However, secondary infection (Canoso and Sheckman 1979), chronic local pain, or skin atrophy (Weinstein *et al*. 1984) may follow the injection.

Ulnar nerve entrapment

The ulnar nerve lies in the groove just behind the medial epicondyle, and can be impinged on by local pressure, trauma, or inflammatory synovial tissue. In many cases the cause is unknown. Patients present with pain on the medial side of the elbow, paraesthesia of the ulnar digits, and occasionally sensory or motor abnormalities in the hand. A positive Tinel sign, elicited by local percussion over the nerve behind the medial epicondyle, suggests the diagnosis, which can be confirmed by electromyography. Steroid injection at this site may reduce the symptoms, especially when the lesion has an inflammatory cause, but care is necessary to avoid injury to the nerve during the procedure.

Wrist

Although the wrist and surrounding structures are often involved in the inflammatory arthropathies, there are many other common, non-inflammatory wrist lesions (Table 7), which are amenable to steroid injection therapy. Rheumatoid arthritis has a particular predilection for involvement of the distal radio-ulnar joint near the

ulnar styloid, which is normally separated from the rest of the wrist joint by the triangular fibrocartilage. Radio-ulnar disease, if severe, may lead to attrition and rupture of the fourth and fifth extensor tendons as they pass over the joint. It is also important to differentiate true wrist (radiocarpal) joint involvement from overlying tenosynovitis. Injection around the wrist and hand is best carried out with both the patient and doctor seated, and with the limb resting on the examination couch.

Wrist (radiocarpal) joint (Fig. 9)

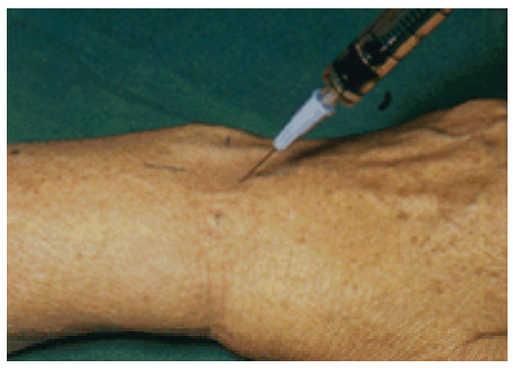


Fig. 9 Injection of the wrist joint.

Injection of the wrist is carried out with the hand held palm down and the joint opened up by palmar flexion over a pillow or similar object. The joint margin is felt as a triangular gap between the lower end of the radius and the lunate and scaphoid bones. The needle is inserted into the joint pointing in a proximal direction at an angle of about 60°. Immobilization of the wrist may prolong the beneficial effects of the injection (Dixon and Graber 1983).

The distal radio-ulnar joint

This joint is identified by gently supinating and pronating the forearm with a finger over the ulnar styloid. Once the joint line is located, it can be injected in a similar manner to the wrist but with the needle inserted almost tangentially into the radio-ulnar joint.

Carpal tunnel syndrome (Fig. 10)

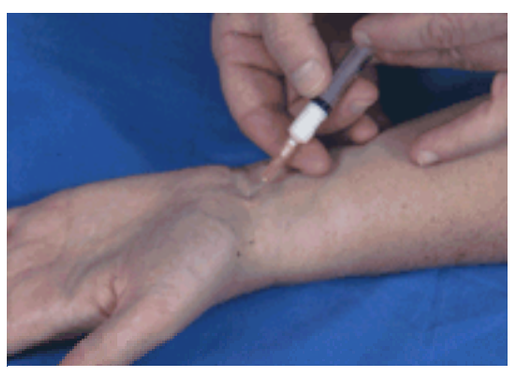


Fig. 10 Injection of carpal tunnel syndrome to the ulnar side of the palmaris longus tendon.

Carpal tunnel syndrome results from median nerve compression in the carpal tunnel of the wrist. Patients present with intermittent or persistent pain and paraesthesia in the thumb, index, and middle fingers, characteristically waking the patient at night. A positive Tinel sign, elicited by local percussion over the median nerve at the wrist, help confirm the diagnosis. Untreated, the lesion can lead to wasting of the muscles of the thenar eminence. Carpal tunnel syndrome can be mimicked by cervical spine or brachial plexus lesions, and electromyography, although not infallible, may assist diagnosis in difficult cases. Even once the diagnosis has been confirmed, a cause, such as inflammatory arthropathy, previous trauma, thyroid disease, or pregnancy should be sought.

The injection is carried out just medial to the midline of the ventral aspect of the wrist, in the first crease, at the junction to the hand. If the palmaris longus tendon is present, it overlies the median nerve and injection should be just medial and parallel to it. A fine 25G needle is inserted, just medial to the midline or palmaris longus tendon, to a depth of 1 cm, directing the needle towards the palm. If the needle is correctly positioned, there should be no resistance, pain, or paraesthesia during the procedure.

Care is necessary to avoid damage to the nerve, which can cause persistent pain and paraesthesia (McConnell and Bush 1990; Frederick *et al*. 1992). Local anaesthetic agents should also be used sparingly to avoid unpleasant, though temporary, paralysis of the median nerve. Long-acting steroid preparations should be used for carpal tunnel injection therapy, as the aim of the injection is to atrophy the soft tissue structures around the nerve. A wrist splint, especially if used at night, may also hasten recovery (Kulick *et al*. 1986). In idiopathic lesions, especially affecting women under the age of 40 years, steroid injection is less likely to be successful (Weiss *et al*. 1994b) and earlier surgery should be considered. The presence of thenar muscle wasting, persistent sensory loss, or a lack of response to steroid injection, is associated with a less favourable outcome even following surgical decompression (Green 1984; Kulick *et al*. 1986).

De Quervain's tenosynovitis (Fig. 11)

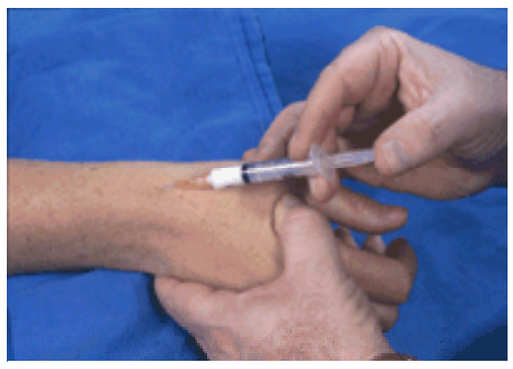


Fig. 11 Injection of De Quervain's tenosynovitis.

Stenosing tenosynovitis of the extensor pollicus brevis and abductor pollicus longus (De Quervain's tenosynovitis) is a common condition which is often occupational and related to repeated minor trauma. It causes pain on gripping especially with use of the thumb. Tenderness is noted in the 'snuffbox' area of the wrist where palpable crepitus may be found. Pain can be increased by ulnar deviation of the wrist against resistance after placing the patients thumb in the palm (Finkelstein's

test), with the arm held in the midprone position.

The injection for this lesion is given at the point of maximum tenderness, using a fine-bore needle which is inserted tangentially along the tendon sheath. When correctly sited, the injection can be given without resistance. If there is doubt with regard to placement of the needle, the syringe can be disconnected and movement of the needle will mirror movement of the thumb, if the needle is correctly sited in the tendon sheath.

Local steroid injections (Weiss *et al*. 1994a) are more likely to be effective in acute lesions, and should be combined with a splint and modification of the activities which caused the lesion. Although long-acting steroids are more effective (Anderson *et al*. 1991), hydrocortisone is the preferred steroid agent as it is less likely to cause complications such as subcutaneous atrophy or depigmentation (Fig. 1 and Fig. 2), especially after repeated injection. Dark skinned patients need to be warned that these cosmetic complications may follow the injection.

Extensor and flexor tenosynovitis

These lesions, if troublesome, can be treated by aspiration and local steroid injection, although recurrence is common and surgical excision is more likely to lead to a permanent cure (Wright *et al*. 1994).

Hand

The hand structures amenable to steroid injection are shown in Table 7.

The small joints in the hands

These joints are frequently involved in inflammatory arthropathies and all are amenable to steroid injection to reduce pain and swelling. However, the injections are painful and it is only possible to inject one or two joints per session, unless general or regional anaesthesia is used. The presence of a joint effusion simplifies the procedure and distortion of the normal anatomy makes it considerably more difficult.

The technique of steroid injection is similar for all the small hand joints and can be performed with the patient seated with the hand palm down across the table or examination bench. The joint margin on the lateral or medial side of the joint is identified by gently flexing and extending the digit. The superior part of the joint line is marked, with the joint flexed to an angle of 45°. By distracting the finger with one hand, it is injected with the other hand, using the superolateral or superomedial approach to avoid injury to the neurovascular structures. The joint will only accept 0.5 to 1 ml of injected fluid, which should contain a combination of long-acting steroid and local anaesthetic. Splintage of the joint may prolong the effect of the injection.

Metacarpophalangeal joint (Fig. 12)

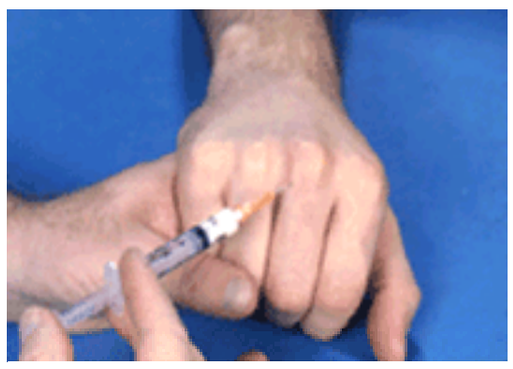


Fig. 12 Injection of the metacarpophalangeal joint.

The joint line, which is located about 1 cm distal to, and not at, the crest of the knuckle, is identified by passive movement of the finger. After marking the joint line, the patients finger is distracted as the joint is injected, tangentially under the extensor expansion.

Proximal (Fig. 13) and distal interphalangeal joint

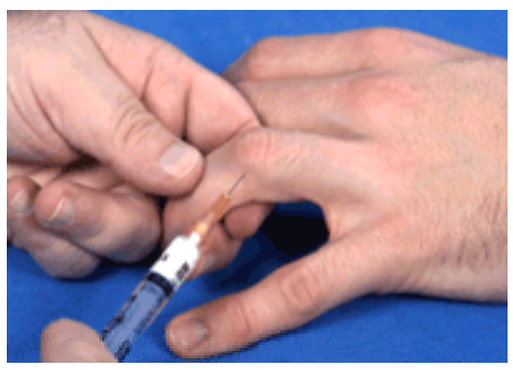


Fig. 13 Injection of the proximal interphalangeal joint.

Injection is also carried out by distracting the joint as the steroid is injected tangentially under the extensor expansion (Evans 1984). The distal joints are technically more difficult to inject unless an effusion is present, as occurs in psoriatic and other seronegative spondylarthropathies.

First carpometacarpal joint (Fig. 14)

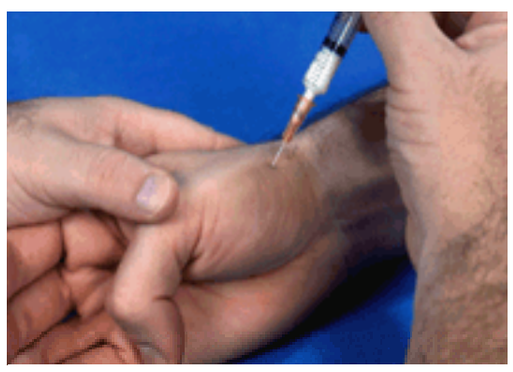


Fig. 14 Injection of the first carpometacarpal joint.

This joint is characteristically affected in primary generalized osteoarthritis and non-inflammatory arthropathies. Patients complain of pain at the base of the thumb on gripping, and tenderness and 'squaring' of the thumb base is noted. For injection, the hand is rested on the couch in the midprone position, with the joint line being identified laterally. By applying pressure on the thumb the joint can be distracted, and the injection facilitated. The patient needs to relax the thumb as the needle is angled slightly distally to enter the joint. Different approaches to the joint may be necessary once gross distortion of the joint has occurred. Despite the degenerative nature of the problem, the response to local steroid injection is often good, although patients need to be warned that a temporary increase in the pain may precede recovery. Patient should also be provided with a thumb pillar splint to protect the joint against further damage.

Flexor tenosynovitis and 'trigger' fingers (Fig. 15)

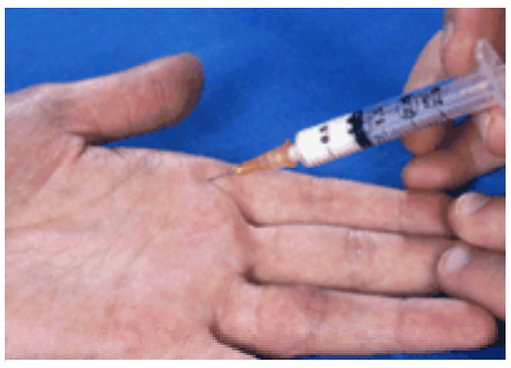


Fig. 15 Injection of second flexor tendon sheath of the hand.

Flexor tenosynovitis is a common cause of poor hand function in inflammatory arthropathies, lupus, or diabetes mellitus but it can also result from trauma or be idiopathic. Nodules may form within the tendon sheaths, making it difficult for the tendon to move freely past the anatomical constrictions, thus leading to 'triggering' of the digit. The nodules can be palpated either attached to the tendon in the pad of tissue adjacent to the proximal phalanx or in the palm opposite the distal palmar crease. They can also occur at the base of the thumb.

Although the tendon nodules can be injected directly, it is more effective to inject the affected tendon sheath. The injection is made with the patient seated with the hand resting palm up on the couch. The needle is advanced tangentially along the tendon sheath in a proximal direction, using a fine-bore needle. Similar injection is possible for thumb tendon nodules.

Steroid injection improves over two-thirds of patients with flexor tenosynovitis (Anderson and Kaye 1991; Lambert *et al*. 1992), and success rates rival surgical intervention (Kraemer *et al*. 1990). Rupture of tendons may rarely follow steroid injection at this site (Tonkin and Stern 1991).

Hip region

Although hip pathology is common, many patients referred with hip problems have pain arising from soft tissue lesions around the hip. The extreme depth of the hip joint and technical difficulty associated with anatomical derangement or osteophyte formation, preclude hip injection as a routine outpatient procedure. Aspiration and, if appropriate, steroid injection, needs to be carried out by an orthopaedic surgeon in theatre or, better, by a radiologist under suitable radiological screening.
Although the hip joint itself is not routinely injected, there are many soft tissue lesions around the hip which are amenable to steroid injection in an outpatient setting (Table 7).

Trochanteric bursa (Fig. 16)

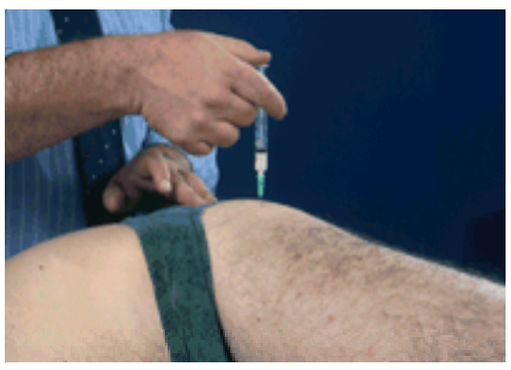


Fig. 16 Injection of trochanteric bursitis.

Patients presenting with hip pain often have pain and local tenderness maximally around the greater trochanter of the femur. It is difficult to differentiate bursitis overlying the greater trochanter from enthesopathy of the muscles inserting at this site. Trochanteric bursitis is usually a non-inflammatory lesion (Ege-Rasmussen and Fano 1985) but is also a common cause of hip pain in rheumatoid patients (Raman and Haslock 1982).

To inject the lesion, the patient is positioned with the painful thigh uppermost and flexed, and the lower leg kept extended. By rotating the affected hip, the prominence of the greater trochanter with its associated tenderness is located, marked, and injected at right angles to the skin. The injection site can be very deep seated in patients with obese thighs, and a longer (5.08 cm) 21G needle may be needed to reach the bursa. While steroid injection therapy has a high success rate in both inflammatory and mechanical lesions, postural advice and review of gait by a physiotherapist is necessary to prevent recurrence of the lesion.

Meralgia paraesthetica

This lesion is an entrapment neuropathy of the lateral cutaneous nerve of the thigh as it traverses the deep fascia, about 10 cm below and medial to the anterior superior iliac spine. Clearly demarcated blunting of pinprick or hyperaesthesia may be found over the anterolateral aspect of the thigh, with tenderness localized to the point where the nerve penetrates the fascia (10 cm below the anterior superior iliac spine). If this point is found, infiltration with steroid and local anaesthetic may abolish the symptoms.

Ischial tuberosity

The ischial tuberosities, located deep in the medial side of the buttocks, have overlying bursas which can become inflamed and cause pain on sitting, especially on a

bicycle seat. These lesions are amenable to infiltration of steroid and local anaesthetic with the patient lying on the lateral side facing away from the examiner. The tender point is identified and injected using a long 21G needle.

Obscure groin pain

Groin pain is common as a result of sport or other injuries. Enthesopathy of the adductor longus tendon, inquinal ligament, and other structures, if identified (Ashby 1994), can be treated with local steroid injection.

Knee

Knee lesions amenable to steroid injection are shown in Table 7.

Knee joint

A knee effusion is a common presenting feature in both inflammatory and non-inflammatory arthropathies, and can seriously impair quadriceps function (Geborek *et al*. 1990). Where the cause of the effusion is uncertain, aspiration of synovial fluid for microbiological and crystal assessment should precede steroid injection.

While there are many approaches to knee aspiration and injection, the most common technique is to use the retropatellar route, via either the medial or lateral approaches. With the patient lying supine, the knee should be extended and the quadriceps muscle relaxed. In the presence of fixed flexion contracture of the knee, support under the knee is necessary to encourage quadriceps relaxation.

Lateral retropatellar approach (Fig. 17)

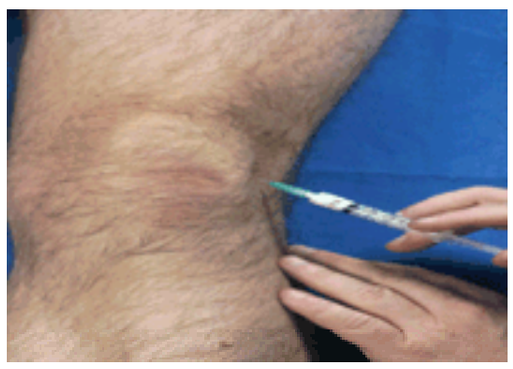


Fig. 17 Injection of the knee joint via the lateral retropatellar approach.

The joint line is marked between the upper and middle third of the patella, with the needle advanced tangentially between the patella and femoral condyle. By pushing on the medial aspect of the patella, the gap between the patella and femur can be increased, facilitating joint penetration. Aspiration as the needle is inserted will reveal fluid as soon as the joint capsule is entered, so reducing the risk of cartilage injury.

Medial retropatellar approach

The site of joint entry is just below the midline of the patella, with the needle advanced tangentially towards the suprapatellar pouch. Aspiration as the needle is introduced will again prevent cartilage injury.

Irrespective of the approach, a 21G or larger-bore needle is used, to remove as much of the inflammatory fluid and debris as possible, before steroid is injected. The introduction of lignocaine before aspiration and 'milking' the fluid towards the needle with controlled pressure from the other hand will facilitate the procedure. Given the large capacity of the joint and synovial surface area, effusions can be massive. The fluid should be viewed and, if purulent, examined in the laboratory to exclude infection before the steroid is injected.

Long-acting steroid agents are more effective than hydrocortisone (Blyth *et al*. 1994), with a relatively large dose of steroid being necessary; 40 to 80 mg of methylprednisolone or triamcinolone acetonide, or 20 to 40 mg of triamcinolone hexacetonide are needed, depending on the size of the effusion. With such a large capacity and synovial surface area, significant absorption of steroid into the circulation can be anticipated and patients need to be informed of both potential benefits to other inflamed joints and possible risks, such as destabilization of diabetic control, following the injection.

A bandage should be firmly applied to support the knee following injection if a large effusion is drained. Bed rest for 24 h following the injection prolongs the beneficial effects (Chakravarty *et al*. 1994), although routine admission to hospital following injection is not cost-effective.

Popliteal (Baker's) cyst

Patients with a knee effusion may develop a popliteal (Baker's) cyst as a result of a one-way valve between the knee joint and semimembranosus or gastrocnemius bursae. The fluid is resorbed from the bursa, leaving a gelatinous material which cannot be aspirated. The cyst can gradually increase in size or can rupture into the calf muscle mimicking a deep vein thrombosis. Ultrasound of the back of the knee and calf or arthrography of the knee joint can demonstrate the cyst, and if present, rupture into the calf muscle.

If a Bakers cyst requires treatment, the knee joint proper should be aspirated and injected as described above. Attempts to inject the cyst directly could result in damage to the neurovascular structures at the back of the knee. For a ruptured Baker's cyst, a below knee support stocking should be used to facilitate sealing of the capsular rupture, following steroid injection into the knee joint proper.

Osteoarthritis of the knee

Some patients with osteoarthritis of the knee have an effusion which might benefit from aspiration and steroid injection, although the improvement is usually short lived (Schnitzer 1993; Gaffney *et al*. 1995). Infiltration of steroid around the patella (Sambrook *et al*. 1989) or other tender trigger spots around the knee, also provides symptomatic relief in some patients, although steroid injection for osteoarthritic joints remains controversial (Dieppe 1991).

Other knee lesions

Non-infected prepatellar bursitis, painful collateral ligaments, and other painful trigger spots around the knee may be amenable to local steroid injection, which is given at the point of maximum tenderness. If an infected prepatellar or other superficial bursa is aspirated, care is necessary to avoid entry into the knee joint as septic arthritis may ensue. Rupture of the patellar ligament may follow steroid injection near the tendon insertion (Alexeeff 1986).

Ankle and hindfoot

The main causes of ankle and heel pain are shown in Table 7. The tendon sheaths around the ankle often communicate with the ankle joint and can be involved in inflammatory or other pathological processes. Most ankle and hindfoot lesions can be injected with the patient lying supine on the couch.

Ankle joint (Fig. 18)

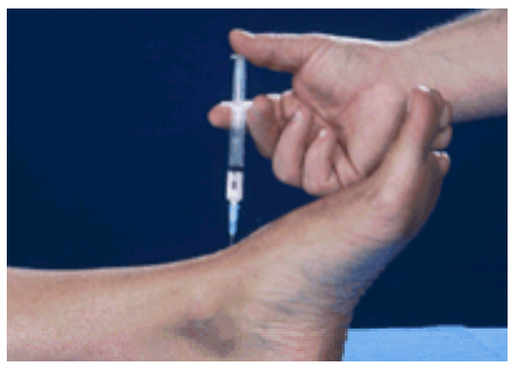


Fig. 18 Injection of the ankle joint.

The ankle joint is located by dorsiflexing the foot to stretch the tibialis anterior ligament and so make it visible. The joint margin which lies between the tibia and talus is then palpated just lateral to the tendon. After marking the injection site, the needle is inserted almost horizontal to the foot, curving over the talus. If an effusion is present, the joint can be aspirated before steroid is injected.

Tendon sheaths

Injection of steroid along swollen and inflamed tendon sheathes is possible behind the medial malleolus (posterior tibial tendon), or lateral malleolus (peroneal tendon), with marked reduction in pain and swelling following successful injection.

Posterior subtalar joint

This joint is commonly involved in inflammatory and degenerative conditions of the ankle and hindfoot, and leads to an increasing valgus deformity below the ankle. The joint often communicates with the ankle joint, but can be injected directly from behind the lateral malleolus, with the patient lying prone on the couch. The needle is angled in the direction of the first metatarsal joint (Beaudet and Dixon 1981).

Tarsal tunnel syndrome

Entrapment of the posterior tibial nerve by the flexor retinaculum can occur behind and below the medial malleolus. Patients present with burning, tingling, and numbness in the distribution of the nerve, most prominently in the toes and distal part of the sole. Tenderness with a positive Tinel sign elicited by percussion over the nerve near the medial malleolus will confirm the diagnosis. Injection under the flexor retinaculum between the calcaneum and medial malleolus may relieve the symptoms.

Posterior heel pain

Achilles tendinitis includes several lesions such as an enthesopathy at the insertion of the tendon, typical of the spondylarthropathies, and also bursitis and peritendinitis which affect the tendon at a distance from its insertion. All these lesions can result from inflammatory disease, or follow sports or other injuries. Rheumatoid nodules, gouty tophi, and xanthomata can also develop along the Achilles tendon. Partial tendon rupture or core necrosis of the Achilles tendon cannot be clinically distinguished from Achilles tendinitis.

The enthesopathy lesion is diffuse and not readily amenable to injection therapy. While steroid injection is feasible with peritendinitis or bursitis, it is not always successful (DaCruz *et al*. 1988), and is only justified once partial tendon rupture has been excluded, using ultrasound or other radiological investigations. Even if partial tendon rupture is not found, patients need to be warned of the risk of tendon rupture following steroid injection (Galloway *et al*. 1992). Short-acting steroid agents lessen the risk of rupture, especially if patients avoid exercise for a few weeks following the injection.

Inferior heel pain (Fig. 19)

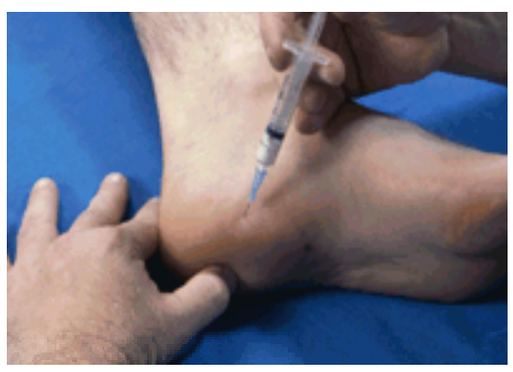


Fig. 19 Injection of plantar fasciitis and plantar heel pain.

A plantar (calcaneal) spur is a common finding on routine radiographs of the heel. These spurs may be asymptomatic or can cause pain under the heel. While most spurs are idiopathic in origin, they are also common in the spondylarthropathies. These lesions can be associated with more diffuse pain radiating up the arch of the foot (plantar fasciitis), especially in patients with reduction in the longitudinal arch of the foot.

Pain under the heel and plantar fasciitis can usually be improved by the use of a cushioned insole, although a local steroid injection under the heel may be indicated for resistant cases. If an injection is needed, the thick skin of the sole should be avoided, with injection being made from the medial side after careful localization of the point of maximum tenderness. The needle is inserted tangentially through the softer skin, so the point of the needle is under the point of maximum tenderness near the bony spur. A cushioned insole should also be provided following injection but recurrence of pain is common. In patients with plantar fasciitis, rupture of the calcaneal origin of the plantar spur may follow steroid injection, with recurrence of pain of a different type after a symptom-free period (Sellman 1994). With very large plantar spurs, surgical excision is necessary.

Forefoot

Metatarsophalangeal joints (Fig. 20)

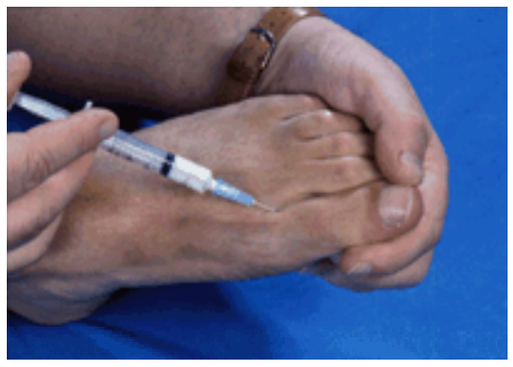


Fig. 20 Injection of the first metatarsophalangeal joint.

Forefoot pain is extremely common in inflammatory and degenerative arthropathies, and steroid injection therapy is often very effective in reducing pain (Helfland 1973) especially in the period before serious deformity has developed. The lateral metatarsophalangeal joints are located by moving the toe between thumb and forefinger. Once the joint line is found and marked, the joint is injected through the dorsum of the toe with the needle entering tangentially from the lateral side, passing under the extensor tendon which overlies the dorsum of the joint. The first metatarsophalangeal joint is sometimes easier to enter from the medial side, using a similar technique to the other joints. The foot joints need to be injected with particular care, as the risk of infection is greater than at other sites.

Other lesions amenable to steroid injection

While many other lesions can be treated with steroid injection, they often require greater expertise. The temporomandibular joint (Ahlqvist and Legrell 1993), sacroiliac joint (Maugars *et al*. 1992), coccyx (Wray *et al*. 1991) and painful trigger-points in the spine (Garvey *et al*. 1989), and many sports injuries are examples of other lesions which may benefit from steroid injection. Subcutaneous rheumatoid nodules (Ching *et al*. 1992) and ganglia can be shrunk by intralesional steroid injection, although this is not widely practised. Steroid injection has particular importance in juvenile chronic arthritis (Evans *et al*. 1991; Honkanen *et al*. 1993) in an effort to limit growth-related deformities, particularly in weight-bearing joints (Eich *et al*. 1994).

Conclusion

Local steroid injection is one of the most useful therapeutic modalities for the amelioration of pain arising from joint and soft tissue structures. While considerable latitude exists with regard to the precise technique employed, the method must be based on a safe and accurate aseptic no-touch technique. A method of joint injection has been described, with consideration of the most frequently used injections described on a regional basis.

Chapter References

Adebajo, A.O., Nash, P., and Hazleman, B.L. (1990). A prospective double blind dummy placebo controlled study comparing triamcinolone hexacetonide injection with oral diclofenac 50 mg TDS in patients with rotator cuff tendinitis. *Journal of Rheumatology*, **17**, 1207–10.

Ahlqvist, J. and Legrell, P.E. (1993). A technique for the accurate administration of corticosteroids in the temporomandibular joint. *Dentomaxillofacial Radiology*, **22**, 211–3.

Alexeeff, M. (1986). Ligamentum patellae rupture following local steroid injection. *Australia New Zealand Journal of Surgery*, **56**, 681–3.

American College of Rheumatology Council of Rheumatological Care (1992). *Guidelines for performing arthrocentesis*.

Anderson, B. and Kaye, S. (1991). Treatment of flexor tenosynovitis of the hand ('trigger finger') with corticosteroids: a prospective study of the response to local injection. *Archives of Internal Medicine*, **151**, 153–6.

Anderson, B.C., Manthey, R., and Brouns, M.C. (1991). Treatment of de Quervain's tenosynovitis with corticosteroids: a prospective study of the response to local injection. *Arthritis and Rheumatism*, **34**, 793–8.

Ashby, E.C. (1994). Chronic obscure groin pain is commonly caused by enthesopathy: 'tennis elbow' of the groin. *British Journal of Surgery*, **81**, 1632–4.

Balogh, K. (1986). The histologic appearance of corticosteroid injection sites. *Archives of Pathology Laboratory Medicine*, **110**, 1168–72.

Bamji, A.M., Dieppe, P.A., Haslock, D.I., and Shipley, M.E. (1990). What do rheumatologists do? A pilot audit study. *British Journal of Rheumatology*, **29**, 295–8.

Beaudet, F. and Dixon, A.S. (1981). Posterior subtalar joint synoviography and corticosteroid injection in rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **40**, 132–5.

Berger, R.G. and Yount, W.J. (1990). Immediate 'steroid flare' from intraarticular triamcinolone hexacetonide injection: case report and review of the literature. *Arthritis and Rheumatism*, **33**, 1284–6.

Bird, H.A., Ring, E.F., and Bacon, P.A (1979). A thermographic and clinical comparison of three intra-articular steroid preparations in rheumatoid arthritis. *Annals of the Rheumatic Diseases*, **38**, 36–9.

Blyth, T., Hunter, J.A., and Stirling, A. (1994). Pain relief in the rheumatoid knee after steroid injection: a single-blind comparison of hydrocortisone succinate, and triamcinalone acetonide or hexacetonide. *British Journal of Rheumatology*, **33**, 461–3.

Canoso, J.J. and Sheckman, P.R. (1979). Septic subcutaneous bursitis. Report of 16 cases. *Journal of Rheumatology*, **6**, 96–102.

Cawley, P.J. and Morris, I.M. (1992). A study to compare the efficacy of two methods of skin preparation prior to joint injection. *British Journal of Rheumatology*, **31**, 847–8.

Chakravarty, K., Pharoah, P.D.P, and Scott, D.G.I. (1994). A randomized controlled study of post-injection rest following intra-articular steroid therapy for knee synovitis. *British Journal of Rheumatology*, **33**, 464–8.

Ching, D.W., Petrie, J.P., Klemp, P., and Jones, J.G. (1992). Injection therapy of superficial rheumatoid nodules. *British Journal of Rheumatology*, **31**, 775–7.

Cyriax, J. (1984). *Textbook of orthopaedic medicine*, *Vol. 2: Treatment by manipulation, massage and injection* (11th edn). Baillière Tindall, London.

Dacre, J.E., Beeney, N., and Scott, D.L. (1989). Injections and physiotherapy for the painful stiff shoulder. *Annals of the Rheumatic Diseases*, **48**, 322–5.

DaCruz, D.J., Geeson, M., Allen, M.J., and Phair, I. (1988). Achilles paratendonitis: an evaluation of steroid injection. *British Journal of Sports Medicine*, **22**, 64–5.

Derendorf, H., Mollmann, H, Gruner, A., Haack, D., and Gyselby, G. (1986). Pharmacokinetics and pharmacodynamics of glucocorticoid suspensions after intra-articular administration. *Clinical Pharmacology and Therapeutics Series*, **39**, 313–7.

Dieppe, P.A. (1991). Are intra-articular steroid injections useful for the treatment of the osteoporotic joint? *British Journal of Rheumatology*, **30**, 199.

Dijs, H., Mortier, G., Driessens, M., De-Ridder, A., Willems, J., and De Vroey, T. (1990). A retrospective study of conservative treatment of tennis elbow. *Acta Belgica Medica Physica*, **13**, 73–7.

Dixon, A. St J. and Graber, J. (1983). *Local injection therapy in rheumatic diseases* (2nd edn). Eular Publishers, Basle, Switzerland.

Doherty, M., Hazleman, B.L., Hutton, C.W., Maddison, P.J., and Perry, J.D. (1992). *Rheumatology examination and injection techniques*. W.B. Saunders, London.

Ege-Rasmussen, K.J. and Fano, N. (1985). Trochanteric bursitis. Treatment by corticosteroid injection. *Scandinavian Journal of Rheumatology*, **14**, 417–20.

Eich, G.F., Halle, F., Hodler, J., Seger, R., and Willi, U.V. (1994). Juvenile chronic arthritis: imaging of the knees and hips before and after intra-articular steroid injection. *Pediatric Radiology*, **24**, 558–63.

Evans, D.M. (1984). The PIP joint. *Clinics of Rheumatic Disease*, **10**, 631–56.

Evans, D.M., Ansell, B.M., and Hall, M.A. (1991). The wrist in juvenile arthritis. *Journal of Hand Surgery*, British Volume, **16**, 293–304.

Firestein, G.S., Paine, M.M., and Littman, B.H. (1991). Gene expression (collagenase, tissue inhibitor of metalloproteinases, complement, and HLA-DR) in rheumatoid arthritis and osteoarthritis synovium: quantitative analysis and effect of intra-articular corticosteroids. *Arthritis and Rheumatism*, **34**, 1094–105.

Frederick, H.A., Carter, P.R., and Littler, J.W. (1992). Injection injuries to the median and ulnar nerves at the wrist. *Journal of Hand Surgery*, American Volume, **17**, 645–7.

Gaffney, K., Ledingham, J., and Perry, J.D. (1995). Intra-articular triamcinolone hexacetonide in knee osteoarthritis: factors influencing the clinical response. *Annals of the Rheumatic Diseases*, **54**, 379–81.

Galloway, M.T., Jokl, P., and Dayton, O.W. (1992). Achilles tendon overuse injuries. *Clinical Sports Medicine*, **11**, 771–82.

Garvey, T.A., Marks, M.R., and Weisel, S.W. (1989). A prospective randomized, double-blind evaluation of trigger-point injection therapy for low-back pain. *Spine*, **14**, 962–4.

Geborek, P., Mansson, B., Wollheim, F.A., and Moritz, U. (1990). Intra-articular corticosteroid injection into rheumatoid arthritis knees improves extensor muscle strength. *Rheumatology International*, **9**, 265–70.

Gladman, D.D. and Bombardier, C. (1987). Sickle cell crisis following intra-articular steroid therapy for rheumatoid arthritis. *Arthritis and Rheumatism*, **30**, 1065–8.

Gomez, E.C., Berman, B., and Miller, D.L. (1982). Ultrasonic assessment of cutaneous atrophy caused by intradermal corticosteroids. *Journal of Dermatology Surgery and Oncology*, **8**, 1071–4.

Gordon, G.V. and Schumacher, H.R. (1979). Elecron microscopic study of depot corticosteroid crystals with clinical studies after intra-articular injection. *Journal of Rheumatology*, **6**, 7–14.

Gray, R.G., Tenenbaum, J., and Gottlieb, N.L. (1981). Local corticosteroid injection treatment in rheumatic disorders. *Seminars in Arthritis and Rheumatism*, **10**, 231–54.

Green, D.P. (1984). Diagnostic and therapeutic value of carpal tunnel injection. *Journal of Hand Surgery*, American Volume, **9**, 850–4.

Haslock, I., MacFarlane, D., and Speed, C. (1995). Intra-articular and soft tissue injections: a survey of current practice. *British Journal of Rheumatology*, **34**, 449–52.

Helfland, A.E. (1973). A clinical study of methylprednisolone acetate (Depo-Medrol) in the treatment of pain and inflammation associated with various foot disorders. *Journal of the American Podiatry Association*, **63**, 287–92.

Hollingworth, G.R., Ellis, R.M., and Hattersley, T.S. (1983). Comparison of injection techniques for shoulder pain: results of a double blind, randomised study. *British Medical Journal Clinical Research Edition*, **287**, 1339–41.

Honkanen, V.E.A., Rautonen, J.K., and Pelkonen, P.M. (1993). Intra-articular glucocorticoids in early juvenile chronic arthritis. *Acta Paediatrica*, **82**, 1072–4.

Jones, A., Regan, M., Ledingham, J., Pattrick, M., Manhire, A., and Doherty, M. (1993). Importance of placement of intra-articular steroid injections. *British Medical Journal*, **307**, 1329–30.

Kirwan, J.R., Haskard, D.O., and Higgens, C.S. (1984). The use of sequential analysis to assess patient preference for local skin anaesthesia during knee aspiration. *British Journal of Rheumatology*, **23**, 210–3.

Kirwan, J.R. and the Arthritis and Rheumatism Council Low Dose Glucocorticosteroid Study Group (1995). The effect of glucocorticoids on joint destruction in rheumatoid arthritis. *New England Journal of Medicine*, **333**, 142–6.

Kraemer B.A., Young, V.L., and Arfken, C. (1990). Stenosing flexor tenosynovitis. *South Medical Journal*, **83**, 806–11.

Kraft, S.M., Panush, R.S., and Longley, S. (1985). Unrecognised staphylococcal pyarthrosis with rheumatoid arthritis. *Seminars in Arthritis and Rheumatism*, **14**, 196–201.

Kulick, M.I., Gordillo, G., Javidi, T., Kilgore, E.S. Jr, and Newmayer, W.L. (1986). Long-term analysis of patients having surgical treatment for carpal tunnel syndrome. *Journal of Hand Surgery*, American Volume, **11**, 59–66.

Lambert, M.A., Morton, R.J., and Sloan, J.P. (1992). Controlled study of the use of local steroid injection in the treatment of trigger finger and thumb. *Journal of Hand Surgery*, British Volume, **17**, 69–70.

Mattila, J. (1983). Prolonged action and sustained serum levels of methylprednisolone after a local injection of methylprednisolone acetate. *Clinical Trials Journal*, **20**, 18–23.

Maugars, Y., Mathis, C., Vilon, P, and Prost, A. (1992). Corticosteroid injection of the sacroiliac joint in patients with seronegative spondylarthropathy. *Arthritis and Rheumatism*, **35**, 564–8.

McConnell, J.R. and Bush, D.C. (1990). Intraneural steroid injection as a complication in the management of carpal tunnel syndrome: a report of three cases. *Clinical Orthopaedics*, **250**, 181–4.

Nakashima, A.K., McCarthy, M.A., Martone, W.J., and Anderson, R.L. (1987). Epidemic septic arthritis caused by *Serratia marcescens* and associated with benzalkonium chloride antiseptic. *Journal of Clinical Microbiology*, **25**, 1014–8.

Neustadt, D.H. (1992). Intra-articular steroid therapy. In R.W Moskowitz, D.S. Howell, H. Makin, and V.M. Goldberg (ed), *Osteoarthritis: diagnosis and surgical management* (2nd edn), pp. 493–510. W.B. Saunders, Philadelphia.

Ostensson, A. and Geborek, P. (1991). Septic arthritis as a non-surgical complication in rheumatoid arthritis: relation to disease severity, and therapy. *British Journal of Rheumatology*, **30**, 35–8.

Parikh, J.R., Houpt, J.B., Jacobs, S., and Fernandes, B.J. (1993). Charcot's arthropathy of the shoulder following intraarticular corticosteroid injection. *Journal of Rheumatology*, **20**, 885–7.

Phelan, M.J., Byrne, J., Campbell, A., and Lynch, M.P. (1992). A profile of the rheumatology nurse specialist in the United Kingdom. *British Journal of Rheumatology*, **31**, 858–9.

Pitsillides, A.A., Will, R.K., Bayliss, M.T., and Edwards, J.C. (1994). Circulating and synovial fluid hyaluronan levels: effects of intra-articular corticosteroid on the concentration and the rate of turnover. *Arthritis and Rheumatism*, **37**, 1030–8.

Price, R., Sinclair, H., Heinrich, I., and Gibson, T. (1991). Local injection treatment of tennis elbow: hydrocortisone, triamcinolone and ignocaine compared. *British Journal of Rheumatology*, **30**, 39–44.

Raman, D. and Haslock, I. (1982). Trochanteric bursitis—a frequent cause of 'hip' pain in rheumatoid arthritis. *Annals of the Rheumatic Disease*, **41**, 602–3.

Rizk, T.E., Pinals, R.S., and Talaiver, A.S. (1991). Corticosteroid injections in adhesive capsulitis: investigation of their value and site. *Archives of Physical Medical Rehabilitation*, **72**, 20–2.

Rizk, T.E., Gavant, M.L., and Pinals, R.S. (1994). Treatment of adhesive capsulitis (frozen shoulder) with arthroscopic capsular distension and rupture. *Archives of Physical Medical Rehabilitation*, **75**, 803–7.

Sambrook, P.N., Champion, G.D., Browne, C.D., Cairns, D., Cohen, M.L., Day, R.O., Graham, S., *et al*. (1989). Corticosteroid injection for osteoarthritis of the knee: peripatellar compared to intra-articular route. *Clinical Experimental Rheumatology*, **7**, 609–13.

Schnitzer, T.J. (1993). Osteoarthritis treatment update. Minimising pain while limiting patient risk. *Postgraduate Medicine*, **93**, 89–95.

Sellman, J.R. (1994). Plantar fascia rupture associated with corticosteroid injection. *Foot and Ankle International*, **15**, 376–81.

Shupak, R., Teitel, J., Garvey, M.B., and Freedman, J. (1988). Intraarticular methylprednisolone therapy in haemophilic arthropathy. *American Journal of Hematology*, **27**, 26–9.

Smith, D.L., McAfee, J.H., Lucas, L.M., Kumar, K.L., and Romney, D.M. (1989). Treatment of nonseptic olecranon bursitis: a controlled, blinded prospective trial. *Archives of Internal Medicine*, **149**, 2527–30.

Tonkin, M.A. and Stern, H.S. (1991). Spontaneous rupture of the flexor carpi radialis tendon. *Journal of Hand Surgery*, British Volume, **16**, 72–4.

Von Essen, R. and Savolainen, H.A. (1989). Bacterial infection following intra-articular injection: a brief review. *Scandinavian Journal of Rheumatology*, **18**, 7–12.

Weinstein, P.S., Canoso, J.J., and Wohlgethan, J.R. (1984). Long-term follow-up of corticosteroid injection for traumatic olecranon bursitis. *Annals of the Rheumatic Disease*, **43**, 44–6.

Weiss, A.P., Akelman, E., and Tabatabai, M. (1994a). Treatment of de Quervain's disease, *Journal of Hand Surgery*, American Volume, **19**, 595–8.

Weiss, A.P., Sachar, K., and Gendreau, M. (1994b). Conservative management of carpal tunnel syndrome: a re-examination of steroid injection and splinting. *Journal of Hand Surgery*, America Volume, **19**, 410–5.

Wiggins, M.E., Fadale, P.D., Barrach, H., Ehrlich, M.G., and Walsh, W.R. (1994). Healing characteristics of a type I collagenous structure treated with corticosteroids. *American Journal of Sports Medicine*, **22**, 279–88.

Wray, C.C., Easom, S., and Hoskinson, J. (1991). Coccydynia. Aetiology and treatment. *Journal of Bone and Joint Surgery*, British Volume, **73**, 335–8.

Wright, T.W., Cooney, W.P., and Ilstrup, D.M. (1994). Anterior wrist ganglion. *Journal of Hand Surgery*, American Volume, **19**, 954–8.

Yood, R.A. (1993). Use of gloves for rheumatology procedures. *Arthritis and Rheumatism*, **36**, 575.

6.6 Sports medicine

Mark Harries

Introduction

This chapter provides a basis for understanding the essential metabolic processes involved in muscle movement and how these translate to human performance. The adaptive responses of the cardiorespiratory and musculoskeletal systems to exercise are also explained. The chapter concludes with a description of the consequences of overreaching these adaptive processes.

Fundamentals of metabolism

The energy for muscular activity derives from a cycle of chemical reactions in which ingested carbon-based fuels are burned with oxygen liberating carbon dioxide. This forms carbonic acid in the tissues which dissociates to bicarbonate and hydrogen ions; a reaction that is catalysed by carbonic anhydrase. Carbonic acid is in dynamic equilibrium with its dissociation products and with CO_2. The high solubility of CO_2 compared with O_2 (more than 20 times as great), coupled with the wide distribution of carbonic anhydrase in the tissues, means that the rate at which CO_2 is mopped up and voided from the lungs can always keep pace with its rate of production. In other words, CO_2 can never accumulate, even during extremes of exertion.

The tendency is always toward a metabolic acidosis, but this is completely overwhelmed by the capacity of the tissues to absorb CO_2 and the capacity of the lungs to dispose of it. The equivalent loss of hydrogen ions by the lung compared with the kidney is in the order of 250:1 and so even during moderate exercise, pH does not change. However, as exercise levels increase, lactate begins to accumulate and the subsequent metabolic acidosis stimulates ventilation with the result that arterial CO_2 may actually fall (Fig. 1).

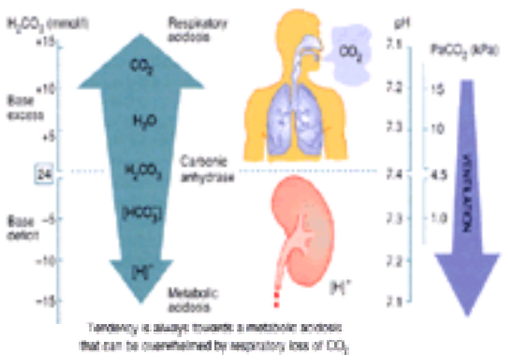


Fig. 1 Metabolism: an acid generator. Oxidation of carbon-based fuels causes a metabolic acidosis due to the conversion in the tissues of CO_2 to carbonic acid by carbonic anhydrase. This tendency is overwhelmed by the capacity of the tissue to absorb CO_2 and the capacity of the lungs to dispose of it. The ratio of the loss of hydrogen ions via the lungs and kidney is of the order of 250:1. Thus, CO_2 can never accumulate, and during moderate levels of exercise, arterial pH does not change. Tendency is always to towards a metabolic acidosis that can be overwhelmed by respiratory losses of CO_2.

How muscle uses its fuel

Energy for movement comes from the hydrolysis of adenosine triphosphate (**ATP**) forming adenosine diphosphate (**ADP**) and liberating 7.6 kcal/mol of ATP. Only around 40 per cent of the energy yield goes into creating tension between molecules of actin and myosin. All is eventually lost as heat. Though the store of ATP is small (only around 10 mmol/kg wet muscle weight), turnover is very rapid and is virtually inexhaustible during sustained moderate exercise. The two principal fuels for ATP synthesis are glucose and long-chain fatty acids: these are available directly fom the diet, or can be liberated from stores—glucose from glycogen in muscle or liver and fatty acids from triglycerides stored in fat cells.

Glycolysis and b-oxidation

The metabolism of both glucose and fatty acids to acetyl coA is an aerobic process. Glycolysis of 1 mol of glucose to form acetyl coA yields 2 mol of ATP (3 mol if the starting point is glycogen). Long-chain fatty acids are metabolized to acetyl coA by b-oxidation, each pair of carbon atoms cleaved from the chain yielding 5 mol of ATP and 1 mol of acetyl coA. Thus, b-oxidation of 1 mol of palmitic acid ($CH_3.(CH_2)_{14}.CO_2H$), would generate 35 mol (5 × 7) of ATP and 8 mol of acetyl coA (Fig. 2).

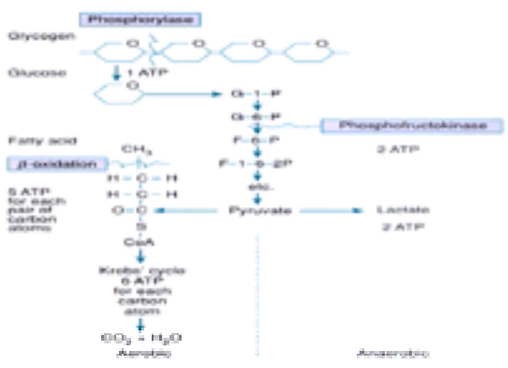


Fig. 2 Glycolysis and b-oxidation. 'Fat burns in a carbohydrate flame'; trained athletes have a greater dependence on fat as a fuel source than the untrained, thereby conserving carbohydrate stores (glycogen). The greatest energy yield comes from mitochondrial respiration. Higher energy demands are met with a switch to anaerobic metabolism and the production of lactate. The rate at which lactate is metabolized is exceeded by its rate of production at around 4 mmol/l. Higher levels can be tolerated, but only for short periods of time. Inherited deficiency of muscle phosphorylase or phosphofructokinse limits performance by restricting glycolysis.

Glycogen accumulates in muscle and liver.

The rest of the oxidative process takes place within the mitochondria in the citric acid (Krebs') cycle and is the source of the greatest metabolic energy yield. Acetyl coA donates two carbon atoms to each turn of the cycle. An *aide mémoire* is that 6 mol of ATP is generated for each carbon atom of the fuel source, such that 1 mol of glucose ($C_6H_{12}O_6$) yields 36 mol of ATP (giving a total of 38 mol when considering the 2 mol generated from glycolysis, or 39 mol if the starting point is glycogen). One mole of palmitic acid ($C_{16}H_{32}O_2$) gives 96 mol of ATP which, when the 33 mol generated from b-oxidation are added, provides a total yield of 129 mol of ATP.

During aerobic exercise, glucose is the preferred fuel, with a gradual switch to fatty acids as glycogen stores are depleted. This rather complex picture is encompassed by the adage 'fat always burns in a carbohydrate flame'. A feature of training is a greater dependence on fatty acids than in the untrained state resulting in a more efficient use of glycogen, which is in limited supply in comparison with the vast stores of fat.

Anaerobic metabolism

For higher energy demands, glucose becomes the exclusive fuel and is metabolized anaerobically with the production of lactate. The energy cost of doing this is enormous, with only 2 mol of ATP generated for each mole of glucose consumed (about 38 mol are produced by aerobic glycolysis). Anaerobic metabolism cannot be sustained because lactate production soon exceeds its rate of metabolism, a point reached when plasma lactate rises beyond a threshold level of 4 mmol/l. When this anaerobic threshold is attained, trained athletes are able to exercise closer to their maximal aerobic capacity (max $\dot{V}O_2$) than sedentary people. In the short term they are also able to tolerate much higher plasma lactates, which may reach 15 to 20 mmol/l in power athletes such as sprinters.

Dietary considerations for energy sources

It follows from the above that glucose is the principal fuel for heavy exercise and fat for low-intensity endurance work. This is reflected in the dietary advice given to the majority of athletes engaged in aerobic sports. Carbohydrate should form 60 per cent of the diet with protein comprising 15 per cent and the rest being fat. The greater the glycogen store, the longer glucose can be burned before fat stores are utilized. This is of great importance in events such as the marathon, triathlon, and cross-country skiing. The concentration of muscle glycogen is in the range 60 to 150 mmol glucosyl units/kg wet weight. Trained athletes can increase this by gorging carbohydrate foods in the few days before competition (Coyle 1991). Simple sugars eaten in such quantities are unpalatable and starch-containing foods such as pasta or potatoes are preferred. It is now accepted that the abundance of phosphocreatine stores is also important in sustaining force during muscle contraction. Ingestion of around 20 g of creatine per day will add to whole body creatine.

Inherited disorders of glycolysis

Deficiency either of muscle phosphorylase (McArdle's syndrome) or phosphofructokinase (Tarui's disease) causes a failure of glycolysis, with the result that glycogen cannot be accessed and instead accumulates in muscle and liver. Of the two, Tarui's disease is much the rarer (Haller and Lewis 1991). Both defects share an autosomally recessive mode of inheritance and so there is often evidence of consanguinity. Both are more common in boys but neither is sex linked. The clinical presentation is that of excessive fatigue on exertion associated with muscle pain, raised serum levels of creatine kinase and blilirubin, and sometimes myoglobinuria. Symptoms may not appear until the second or even third decade of life. Glycogen storage disease type II, deficiency of a $_1$-glucosidase (known variously as acid maltase deficiency or Pompe's disease), presents with respiratory failure as late as the mid-forties. The diagnosis is made on periodic acid–Schiff staining of the muscle biopsy, which shows massive accumulations of glycogen (Fig. 3). Since glycolysis is blocked, pyruvate is never synthesized and there is failure to generate lactate, or in other words, there is a relative incapacity for anaerobic power. The principal fuel then becomes fatty acids, with a slower energy yield, hence fatigue on anything but light work rates.

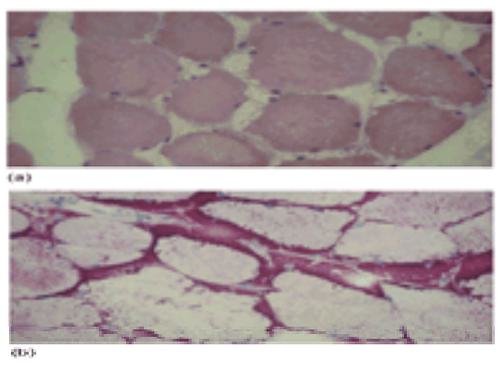


Fig. 3 Inherited disorders of glycolysis. A 17-year-old Royal Naval recruit complained of fatigue and muscle stiffness on exercise although he had been a good sportsman while at school. An exercise test was terminated early due to exhaustion. Samples taken throughout showed that his blood lactate failed to rise. Glycogen in muscle stains pink with periodic acid–Schiff reagent and normally shows as just a few granules around each muscle fibre (a). By contrast, the biopsy sample of the recruit (b) is full of glycogen, which is unavailable for glycolysis due to an inherited deficiency of myophosphorylase (McArdle's syndrome).

Limitations to aerobic work rates

During vigorous exercise, muscle burns all the oxygen available to it. Further demands are met by a switch to anaerobic metabolism with its concomitant limitations imposed by lactic acidosis. The principal factor containing aerobic capacity involves the rate at which oxygen can be delivered, which is itself dependent on three factors: haemoglobin concentration, minute ventilation (pulmonary output), and cardiac output.

The importance of haemoglobin concentration

In the laboratory, the oxygen combining power of haemoglobin is quantified at 1.39 ml of oxygen for each gram of haemoglobin; but in vivo, haemoglobin passing though the lungs picks up only around 1.306 ml/g of haemoglobin. A shift in haemoglobin concentration within the normal range, say from 11 to 18 g/dl, would increase oxygen delivery by a factor of 1.6. It has been well documented that raising haemoglobin beyond the normal range translates directly to an increase in maximal oxygen consumption and an improved performance (Brien and Simon 1987).

Altitude training gives rise only to modest increases in haemoglobin—1 to 2 g/dl at most. For a bigger advantage illicit means have been sought. At the 1984 Olympic Games in Los Angeles, the American cycling team came from nowhere, and contrary to expectations and form, swept the board. Some months later it was revealed that seven of the successful team members had received a transfusion of blood the night prior to competition (Klein 1985). Erythropoietin is now being used with the same ends in mind and has the added advantage that it is virtually undetectable. Paradoxically, highly trained athletes may have a haemoglobin at or below the lower limit of normal: this reflects an increase in plasma volume due to heavy training and is not a true anaemia. If serum ferritin levels are normal, no treatment is indicated in these cases.

Minute ventilation and aerobic performance

While exercising at the maximum sustainable ventilatory capacity (**MSVC**), ventilation can be increased still further for short periods to reach maximum voluntary ventilation (**MVV**). The MSVC is found to be 55 to 80 per cent of MVV (Freedman 1970) or, put another way, assuming that the lungs are normal, the capacity to increase ventilation is so great that the key factor that most limits aerobic performance is likely to be the cardiac output rather than minute ventilation.

Pulmonary conditions limiting ventilatory capacity

The relationship between maximum aerobic capacity (max VO_2) and MSVC is roughly linear: the higher the max VO_2 the higher the ventilatory capacity. This holds true whether VO_2 and MSVC are measured in absolute terms (l/min), or independently of body size (ml/kg per min) (Fig. 4). Minute ventilation is the product of the respiratory rate and breath volume. While exercising at maximum aerobic capacity, respiratory rates are remarkably consistent. Young men take around 60 breaths and women 55 breaths a minute. Given a respiratory rate of 60/min, breath volume clearly cannot exceed the forced expiratory volume (**FEV_1**). Elite oarsmen access only around 50 per cent of FEV_1 each breath compared with middle-distance runners who reach around 60 per cent of FEV_1. This accords well with the formula used to estimate ventilatory capacity from FEV_1 (MSVC = [FEV_1 × 30] + 23). Hence it follows that any condition of the lung in which FEV_1 is significantly reduced will have an impact on aerobic capacity by reducing minute ventilation.

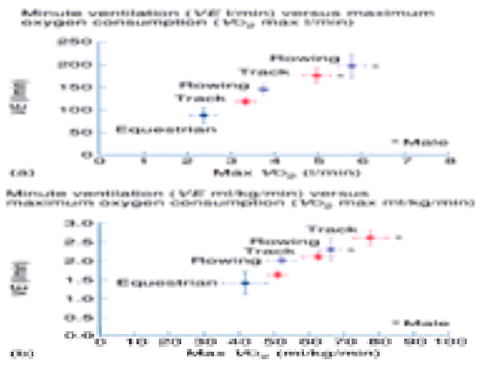


Fig. 4 The importance of high ventilatory capacity. Minute ventilation (VE) at maximum oxygen consumption (max VO_2) is measured in 120 Olympic class athletes grouped according to their sport. The relationship between ventilation and oxygen consumption is linear, whether measured in absolute terms or corrected for body weight. The higher the oxygen consumption, the higher must be the minute ventilation. The highest values are recorded in men, with track athletes at the top according to body weight. Equestrian competitors engage in a more sedentary sport and represent values closer to those which would be expected in normal fit individuals in their mid-twenties.

Asthma

Although chronic bronchitis and emphysema are diseases that reduce ventilatory capacity, they are diseases of older people. Asthma, on the other hand, occurs at all ages and is also much commoner, accounting for around 20 per cent of complaints concerning underperformance received from athletes attending the British Olympic Medical Centre.

The most important feature of asthma for the sportsman is that it worsens during exercise and therefore impacts both on performance and on training. Treatment is highly effective and, therefore, a reproducible diagnostic test is important. However, there is no agreed protocol for exercise testing, although certain ingredients appear to be important. Running in the open air is a more potent stimulus to bronchial constriction than exercising on a cycle or treadmill ergometer. The reasons for this are complex and related in part to climatic conditions. Cold dry air causes more bronchial constriction than warm moist air. The exercise must be rigorous, enough to raise the heart rate to around 80 per cent of the maximum that can be achieved (approximately 220 minus age in years). The duration of the test is also important. It should last at least 3 min but not much longer than 5 min.

Normal subjects may show a short-lived bronchial dilatation on stopping, but a fall in FEV_1 or peak flow of more than 15 per cent occurring 5 to 10 min after exercise is a positive test. Treatment is with an inhaled steroid such as budesonide turbohaler 200 µg twice daily (Fig. 5) (Harries 1994).

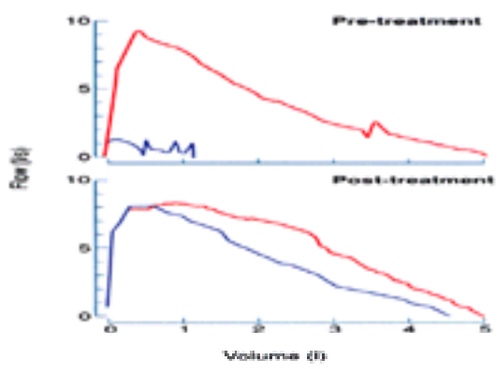


Fig. 5 Exercise-induced asthma. This example of severe bronchial constriction induced by exercise was recorded in an athlete with a clear history of sleep disturbance due to wheeze and tightness in the chest. Flow is plotted against the volume in a forced expiratory manoeuvre from full inspiration to full expiration. Two expiratory efforts are shown; one measured before exercise (red line) and the second around 5 min after a 3-min run (blue line). Peak flow measured 9 l/s before exercise and less than 1 l/s afterwards, representing a fall of more than 80 per cent. Oscillations in flow during expiration are caused by coughing. The athlete was give prednisolone at 30 mg daily for 2 weeks, then continued on beclomethasone dry powder at 400 µg twice daily. One month later the exercise test was repeated under identical conditions (post-treatment). There was no change in peak flow, but more important, pretest FEV_1 had risen 20 per cent from 4.03 to 5.3 l. The exercise test was still marginally abnormal with a small fall in FEV_1 post-exercise and a larger fall in mid-expiratory flow. However, the athlete was sufficiently improved to win a medal at a major international competition just 2 months later.

Anatomical and physiological adaptations to exercise

Elite athletes differ markedly in their physical make-up to sedentary individuals. Whether or not these differences are inherited, or are the result of training, has been the subject of much debate. However, recent evidence suggests that genetic factors are the more important in determining athletic prowess (Bouchard *et al*. 1992).

Adaptations of muscle

The relationship between muscle size and strength is poor. Increase in size reflects fibre hypertrophy and fibre numbers remain unchanged. The greatest hypertrophy results from high-intensity isometric exercise in which muscle is required to contract against a resistance. Strength, on the other hand, may be due in part to fibre recruitment; in other words to an increase in the number of fibres that are available for contractile activity.

Training improves fibre recruitment and chronic endurance training favours the transformation of fast-twitch glycolytic (type IIb) to slow-twitch (type 1) fibres (Dudley *et al*. 1982). Muscle biopsy samples show that capillary density and enzyme activity both increase through training. The mitochondria also appear to increase in number, although this finding may reflect mitochondrial hypertrophy and folding rather than any absolute increase (i.e. it is an artificial observation, reflecting only the cross-section of the biopsy specimen). The net result is an improvement in oxidative potential of muscle with an increase in maximal oxygen consumption in the order of 15 to 25 per cent over time.

Cardiac adaptations

Echocardiography shows that both ventricular dilatation and hypertrophy occur with training. The greatest dilatation is seen in rowers and cyclists, who also show the greatest degree of cardiac hypertrophy; a finding that runs counter to the expectation that strength athletes should develop the most hypertrophy.

Dilatation results in the greatest increase in heart size, such that the cardiac silhouette can be found to occupy more than half the transverse diameter of the chest when observed on a chest radiograph. The diameter of the left ventricle at the end of diastole may reach 7 cm (compared with the normal size of around 5.7 cm). Although this represents only a 10 to 15 per cent increase in diameter, it translates to a 30 per cent increase in stroke volume. A large stroke volume means that adequate cardiac output can be maintained with a lower heart rate. A slow resting pulse is therefore a constant feature of training.

Hypertrophy of the septum and left ventricular wall also occurs but contributes very little to overall change in heart size. Normal wall and septal thickness is from 7 to 11 mm, ranging up to 13 mm in cyclists and rowers. A thickness beyond 16 mm is considered abnormal, particularly if chamber size is not also increased (Pelliccia *et al*. 1991). Disproportionate hypertrophy of the septum with respect to the left ventricular wall may give rise to obstruction of the outflow tract during heavy exercise (see obstructive cardiomyopathy below).

The electrocardiogram is often grossly abnormal with widespread T-wave inversion across the lateral chest leads as far as V4. Multiple nodal (supraventricular) ectopic beats are also common. Both abnormalities tend to correct on exertion. Unfortunately this may also be true of myocardial ischaemia. The resting electrocardiogram, especially recordings made at night, may show bradyarrhythmias with sinoatrial block (Ector *et al*. 1984).

Skeletal adaptations to exercise

Bone density increases in response to exercise in all subjects except amenorrhoeal women. Rise in mineral content is greatest in bone subject to the most stress, such as the femur of runners, the lumbar spine of rowers, and the radius of gymnasts and racket players (Wolman *et al*. 1990). Women who train intensively develop amenorrhoea. The result is a loss in bone density exactly analogous to that which occurs following the menopause. The earliest changes are seen in bone with the highest rate of turnover, and appear first in the trabecular bone of the lumbar spine. Bone loss is detectable after as few as 6 months of amenorrhoea. Losses great enough to result in pathological fracture are usually only seen in women with an associated eating disorder such as anorexia nervosa. Treatment is with oestrogen in the form of hormone replacement therapy (Drinkwater *et al*. 1984; Wolman *et al*. 1991).

Clinical syndromes associated with overexertion

The health benefits of taking regular exercise are beyond dispute. The risk of osteoporotic fracture later in life is reduced (Heinonen *et al*. 1996). Furthermore, there is now direct evidence of a reduced risk of acute myocardial infarction in men and exercise is a major factor in controlling obesity and diabetes (Helmrich *et al*. 1991; Lakke *et al*. 1994), although evidence that exercise benefits hypertension is less strong (Blumenthal *et al*. 1991). Nevertheless, in the United States of America it is recommended that 'every adult should accumulate 30 min or more of moderate-intensity physical activity on most, preferably all, days of the week' (Anonymous 1995).

Sudden death

Although rare, sudden death is almost always due to a cardiac cause. It is assumed that the terminal event is ventricular fibrillation, and limited forensic evidence often reveals an underlying abnormality, such as coronary atheroma or a congenital abnormality of the coronary tree. Even moderate levels of exercise have been implicated; for instance, 12 deaths occurred over 6 years among the joggers of Rhode Island, all but one of these were found to be due to coronary artery disease, with the mortality estimated to be seven times the expected mortality in a sedentary population (Thompson *et al*. 1992). With the rising popularity of veteran events, it is clear that coronary artery disease in sport is going to pose an ever-increasing problem.

Hypertrophic obstructive cardiomyopathy

Obstructive cardiomyopathy, also known as muscular subaortic stenosis, is an absolute contraindication to strenuous dynamic or static activity, because it is a condition that may lead to sudden unexpected death from ventricular fibrillation at a young age. The outflow tract of the left ventricle becomes obstructed in systole by the interventricular septum, which is disproportionately hypertrophied with respect to the ventricular wall. The ratio of the thickness of the septum to the ventricular wall, which is normally 1.3:1, exceeds 1.5:1 (Fagard *et al*. 1984). Death occurs, often in the second decade of life, during the course of vigorous physical activity. Faintness or syncope occurring immediately following exercise is one of the few symptoms, providing an early indication for urgent investigation. The electrocardiogram is unhelpful because T-wave changes suggesting left ventricular hypertrophy are often seen in normal highly trained individuals. The diagnosis is made by echocardiography (Stewart Hillis *et al*. 1994).

The condition is inherited as an autosomal dominant, so all first-degree relatives should be screened. There was once a vogue for treatment with surgery which involved shaving the septum, but medical treatment aimed at suppressing the ventricular arrhythmias is now more established. The drug of choice has been amiodorone, though inhibitors of angiotensin-converting enzyme are currently being assessed.

Both aortic and pulmonary stenosis also cause outflow tract obstruction which may have the same effects as hypertrophic obstructive cardiomyopathy during vigorous exertion. On the other hand, tricuspid and pulmonary regurgitation are very common in the highly trained athlete, occurring in over 90 per cent. The murmurs produced may be difficult to distinguish clinically and expert advice should be sought.

Hyperpyrexia (heat stroke)

During extremes of exercise, rectal temperature may reach 41°C. Heat loss by radiation through peripheral vasodilatation is limited and further losses can only be achieved by evaporating sweat. Once these homeostatic mechanisms are rendered ineffective, such as in a very humid environment or with sweat failure, core temperature rises unchecked. Heat stroke is rare amongst experienced athletes except where there has been stimulant abuse. The usual clinical setting is one of an undertrained individual, competing for the first time and with inadequate fluid intake (Clowes and O'Donnell 1974). A similar picture is also seen in people who have taken amphetamine (speed) or MDMA (Ecstasy) tablets and who have been dancing to exhaustion.

Confusion or coma is the rule with a clinical presentation very similar to septicaemic shock. The skin may be clammy and cold, contrasting with the high rectal temperature. Grand mal seizures occur with decerebrate posturing. Rhabdomyolysis develops early with the creatine kinase often reaching over 100 000 U/l. Paradoxically in the face of dehydration, serum sodium may be low, sometimes below 110 mmol/l. Disseminated intravascular coagulation, renal failure, and liver failure may occur within 24 h (Sutton 1994). Rapid intravenous infusion can be life-saving with an initial infusion of 4 l of saline given in the first hour. Broad-spectrum antibiotics effective against Gram-negative organisms should also be given.

Immune deficiency

There is some evidence that those suffering fatigue syndrome due to overtraining seem to develop frequent infections of the upper respiratory tract. Furthermore, the infecting agent occasionally proves to be unusual (such as *Toxoplasma gondii*). These and other observations have led to the suggestion that overtraining leads to an immunosuppressed state (Khansari *et al*. 1990). Some support for this concept is obtained from studies of plasma glutamine levels. Glutamine is an essential fuel for lymphocytes: it is synthesized by muscle and is found to be low in chronic fatigue states (Newsholme *et al*. 1991).

Drugs to avoid

Amphetamine-like substances and drugs with a-adrenergic (stimulant) actions are all banned. These include isoprenaline, adrenaline, noradrenaline, and phenylpropanolamine. For example, adrenaline may not be given by local injection mixed with an analgesic such as lignocaine. Over-the-counter cold cures often contain one, or more, of these agents.

Selection of an analgesic or anti-inflammatory agent that will not fall foul of the International Olympic Committee's (**IOC**) list of banned substances is made more

difficult by the fact that not only do the rules change from time to time, they also lack consistency. For example, opiate analgesics and pentazocine are banned, but codeine and dihydrocodeine have recently been reinstated: yet dextropropoxyphane remains a banned substance! Corticosteroids may not be given by parenteral injection, which effectively outlaws all depot preparations, but hydrocortisone may be given by intra-articular injection or into soft tissue injuries such as a tennis elbow. At the time of writing, the position of systemically acting corticosteroids is under review by the IOC's Medical Commission.

Such is the state of confusion that some sporting bodies, particularly those not affiliated to the IOC, are beginning to go their own way. That said, there is broad agreement that drugs such as anabolic steroids and amphetamine should be avoided, if only because there is no clinical indication for taking them. However, a blanket ban on so many other medicaments, particularly those that are therapeutically useful, is crude and unworkable.

Chapter References

Anonymous (1995). Physical activity and public health: a recommendation from the Centres for Disease Control and Prevention and the American College of Sports Medicine. *Journal of the American Medical Association*, **273**, 402–7.

Blumenthal, J., Siegel, W., and Applebaum, M. (1991). Failure of exercise to reduce blood pressure in patients with mild hypertension: results of a randomized controlled trial. *Journal of the American Medical Association*, **266**, 2098–104.

Bouchard, C., Dionne, F., Simoneau, J.-A., and Boulay, M. (1992). Genetics of aerobic and anaerobic performances. *Exercise and Sport Science Reviews*, **20**, 27–58.

Brien, A. and Simon, T. (1987). The effects of red blood cell infusion on 10-km race time. *Journal of the American Medical Association*, **257**, 2761–5.

Clowes, G. and O'Donnell, T. (1974). Heat stroke. *New England Journal of Medicine*, **291**, 564–7.

Coyle, E. (1991). Timing and method of increased carbohydrate intake to cope with heavy training, competition and recovery. *Journal of Sports Science*, **9** (Suppl.), 29–52.

Drinkwater, B., Nilson, K., Chestnut, C., Bremner, W., Shainholtz, S., and Southworth, M. (1984). Bone mineral content of amenorrhoeic and eumenorrhoeic athletes. *New England Journal of Medicine*, **311**, 277–81.

Dudley, G., Araham, W., and Terjung, R. (1982). Influence of exercise intensity and duration on biochemical adaptations in skeletal muscle. *Journal of Applied Physiology*, **53**, 844–50.

Ector, H. *et al*. (1984). Bradycardia, ventricular pauses, syncope, and sports. *Lancet*, **ii**, 591–4.

Fagard, R., Aubert, A., Staessen, J., Vanden Eynde, E., Vanchees, L., and Amery, A. (1984). Cardiac structure and function in cyclists and runners. Comparative echocardiographic study. *British Heart Journal*, **52**, 124–9.

Freedman, S. (1970). Sustained maximal voluntary ventilation. *Respiration Physiology*, **8**, 230–44.

Haller, R. and Lewis, S. (1991). Glucose-induced exertional fatigue in muscle phosphofructokinase deficiency. *New England Journal of Medicine*, **324**, 364–69.

Harries, M. (1994). Pulmonary limitations to performance in sport. *British Medical Journal*, **309**, 113–15.

Heinonen, A., *et al*. (1996). Randomized controlled trial of effect of high-impact exercise on selected risk factors for osteoporotic fractures. *Lancet*, **348**, 1343–7.

Helmrich, S., Ragland, P., Leung, R., and Paffenbarger, R. (1991). Physical activity and reduced occurrence of non-insulin-dependent diabetes mellitus. *New England Journal of Medicine*, **325**, 147–52.

Khansari, D., Murgo, A., and Faith, R. (1990). Effects of stress on the immune system. *Immunology Today*, **11**, 170–4.

Klein, H. (1985). Blood transfusion in athletes. Games people play. *New England Journal of Medicine*, **312**, 854–6.

Lakke, T., Venalainen, J., Rauramaa, R., Salonen, R., Tuomilehto, J., and Salonen, J. (1994). Relation of leisure-time physical activity and cardiorespiratory fitness to the risk of acute myocardial infarction in men. *New England Journal of Medicine*, **330**, 1549–54.

Newsholme, E., Parry-Billings, M., McAndrew, N., and Budgett, R. (1991). A biochemical mechanism to explain some characteristics of overtraining. In *Advances in nutrition and top sport* (ed. F. Brouns), Vol. 32, 79–93. Medical Sports Science, Basel.

Pelliccia, A., Maron, B., Spataro, A., Proschan, M., and Spirito, P. (1991). The upper limit of physiological cardiac hypertrophy in highly trained elite athletes. *New England Journal of Medicine*, **324**, 292–302.

Stewart Hillis, W., MacIntyre, P., Maclean, J., Goodwin, J., and McKenna, W. (1994). Sudden death in sport. *British Medical Journal*, **309**, 657–60.

Sutton, J. (1994). Physiological and clinical consequences of exercise in heat and humidity. In *Oxford textbook of sports medicine* (ed. M. Harries *et al*.), pp. 231–8. Oxford University Press.

Thompson, P., Funk, E., Carleton, R., and Stumer, W. (1992). Incidence of death during jogging in Rhode Island from 1975 through 1980. *Journal of the American Medical Association*, **247**, 1535–8.

Wolman, R., Clark, P., McNally, E., Harries, M., and Reeve, J. (1990). Menstrual state and exercise as determinants of spinal trabecular bone density in female athletes. *British Medical Journal*, **301**, 516–18.

Wolman, R., Faulmann, L., Clark, P., Hesp, R., and Harries, M. (1991). Different training patterns and bone mineral density of the femoral shaft in elite female athletes. *Annals of the Rheumatic Diseases*, **50**, 48–79.

6.7 Sports injuries

J. R. Jenner and M. Shirley Emerson

Introduction

Regular aerobic exercise has been shown to have a beneficial effect on several aspects of health including:

1. reducing the risk of cardiovascular disease, particularly coronary artery disease and hypertension;
2. reducing the incidence of osteoporosis;
3. helping with weight loss by increasing the metabolic rate;
4. inducing a feeling of general well-being and reducing depression (Report of the Royal College of Physicians 1991).

As a result of these findings, the government has actively encouraged the general population to participate in sporting activities. Unfortunately, one side-effect of sport is injury. If these injuries are not treated promptly and appropriately, the injury may become recurrent or chronic.

Different sports have different rates of injury and different injury profiles (Fig. 1) but the fundamental principles underlying the treatment of acute injuries, as well as their rehabilitation, apply across sports. The major cause of injury in sport is acute trauma to soft tissues such as muscle, tendon, or ligaments. Chronic sporting injuries result from either adverse body mechanics, resulting in excess strain, or overuse, leading to fatigue. Children and teenagers with immature or rapidly growing musculoskeletal systems are prone to injuries not encountered in the adult and these injuries deserve special mention.

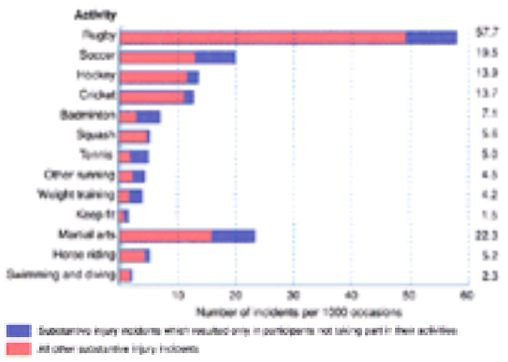


Fig. 1 New substantive injury incident rates by activity in a random sample of 17 564 people aged between 16 and 45 in England and Wales (Nicholl *et al*. 1991)

Injury prevention

Prevention of injury involves several strategies.

Modification of the environment

Many accidents are due to poor surfaces and inadequate equipment. Uneven pitches, collapsing gymnastic equipment, and worn landing mats have all been shown to contribute to injury. These injuries are definitely preventable. Approximately 10 per cent of sports injuries in the United Kingdom are due to collisions with the 'furniture' of the playing field or pitch (Nicholl *et al*. 1991; Coleman *et al*. 1996; Wyatt *et al*. 1996).

Education and behaviour modification

1. Changes in the rules of some games may be necessary to reduce injury levels. Banning the dangerous practices of 'spearing' in American football has resulted in a reduction in the number of injuries producing permanent cervical quadriplegia (Torg *et al*. 1985).
2. Good referees are needed to prevent deliberate fouls, which often lead to injury.
3. Continuing player education is also important to persuade players to obey the rules and not to indulge in 'sport rage'.
4. Trained coaches should have a good knowledge of exercise physiology and be able to advise on good techniques, diet, and to help players avoid overuse injuries through improved training techniques.
5. Matching of young players in team games should be done by size and not by age.
6. Relevant protective equipment should be used. Mouth guards have been shown to reduce the number of dental and oral injuries but many players do not like using them (Jennings 1990). Protective goggles in squash and shin pads in football are examples of the equipment that can reduce injury.

Personal strategies

1. Training to build up muscle strength and endurance. Many injuries occur during the second half of a game, when fatigued muscles do not respond quickly to new situations. This is seen particularly in weekend athletes who do no other training.
2. Warm up and cool down. There are differing opinions on the benefit of warm ups and much of the evidence is anecdotal but the majority of athletes and coaches

are in favour of some mobility exercises and gentle exercise to raise the heart rate and body temperature. After an intense effort, continuing light exercise does seem to allow an enhanced rate of removal of lactic acid and also prevents blood 'pooling' in the lower limbs. Improving flexibility both before and after exercises is reported to reduce injury levels but there is little hard evidence to support this.

3. Diet. There is overwhelming evidence as to the importance of adequate carbohydrate intake while competing and training to maintain and replace muscle glycogen. Carbohydrate should provide 50 to 60 per cent of the calorie intake and the athlete may need to take some of this as a carbohydrate drink (Costill and Hargreaves 1992). Cyclists on the *Tour de France* need to consume about 7000 calories/day and this would be impossible to eat as solid food.
4. Hydration. Most athletes are now aware of the importance of adequate rehydration but children need to be encouraged to drink as part of the game. Immediate postexercise replacement of fluid and carbohydrate, within the first hour, has been shown to restore rapidly muscle glycogen. Reduced glycogen levels lead to muscle fatigue and injury and reduced muscle glycogen is associated with a reduction in muscle and plasma glutamine levels.
5. Avoiding overtraining. Many athletes and sports players are convinced that 'more is better' and are not prepared to incorporate rest and recovery into their schedule. The overtraining syndrome produces symptoms of fatigue, sleeplessness, and loss of appetite (Budgett 1995). Intense and prolonged exercise has also been shown to produce immune system depression for 6 to 20 h after exercise (Newsholme 1994). This may explain why an athlete is more susceptible to injury and illness, including opportunistic infections such as toxoplasmosis and why injuries may take longer to heal.

Acute injuries

Acute management

Trauma results in bleeding. Blood in the extra vascular space is an extreme irritant causing an acute inflammatory response with swelling, an increase in pressure, and, if this pressure becomes great enough, tissue necrosis occurs (Fig. 2). Acute treatment of soft tissue trauma is aimed at limiting the bleeding and its subsequent deleterious affect (Fig. 3). This is achieved by:

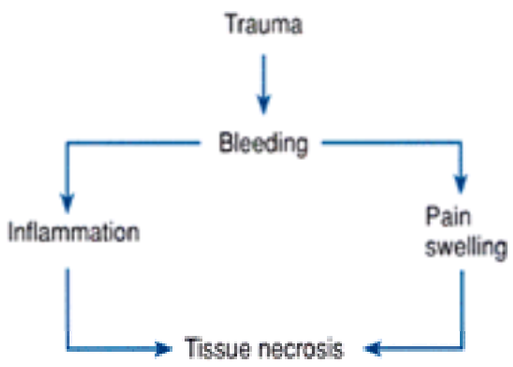


Fig. 2 Flow chart of sequence of events after soft tissue trauma.

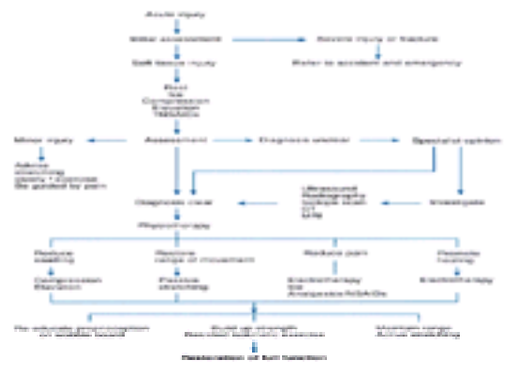


Fig. 3 Flow chart for the treatment of soft tissue injuries after acute injury.

1. rest
2. ice
3. compression
4. elevation
5. non-steroidal anti-inflammatory agent.

This treatment is often given the mnemonic of RICE or NICER.

Rest

Rest is the fundamental treatment for an acute injury. It is essential to limit the trauma and to minimize the bleeding. Once an injury has been sustained, the injured player must immediately leave the field of play. This allows an immediate assessment of the injury to be made and for prompt first aid measures to be instituted. This early assessment can be vital as injuries such as ruptured cruciate ligaments can quickly be masked by bleeding and effusions in the knee and delay an accurate diagnosis.

Ice

Ice has become the traditional agent to help stop bleeding and reduce pain. The mechanism by which ice influences the underlying vasculature is complex. While superficial skin blood flow is reduced, there is an initial short period of increased blood flow in the underlying tissues followed by alternating vasospasm and dilatation. It has been suggested that this 'pumping' action may be beneficial in removing oedema from the muscles.

Compression

Compression is probably the most effective means of limiting bleeding. The application of a firm bandage such as 'Tubigrip' over the whole of a limb may be helpful where direct pressure is difficult or inappropriate. It is important that pressure is not applied so tightly that it interferes with arterial or venous blood flow and therefore cause further damage.

Elevation

Tissue injury causes an acute inflammatory response followed by the development of oedema. The increased tissue pressure can reduce blood flow, increase pain, and delay healing. Swelling can be reduced by elevating the injured limb or, in the case of an injured leg, just by lying down.

Non-steroidal anti-inflammatory drugs

The treatment of soft tissue injury with non-steroidal anti-inflammatory drugs is controversial. Oral non-steroidal anti-inflammatory drugs will reduce pain and

inflammation but probably only have a marginal beneficial effect on the healing process. They also have significant potential side-effects. If prescribed for acute soft tissue injuries, non-steroidal anti-inflammatory drugs are best prescribed for no more than 2 to 3 days after the injury. Topical non-steroidal anti-inflammatory drugs have a lower side-effect profile and exert a small beneficial effect in soft-tissue injuries, helping with an early return to sport.

Haematomas

Muscle haematoma may be due to either extrinsic or intrinsic injury (Table 1). After injury, there is disruption of the muscle fibres and capillaries with associated bleeding. The torn ends retract from the injury leaving it filled with blood.

	Extrinsic	Intrinsic
Mechanism	Direct trauma or compression	Overstretching or overload
Pathology	Muscle compressed against bone	Sudden contraction of a muscle against resistance
Cause	Contact sports	Stretching or sprinting

Table 1 Muscle haematoma

An intact muscle sheath results in an intramuscular haematoma—a swelling within the muscle—causing pain and a considerable amount of disability which resolves slowly over many weeks. An intermuscular haematoma results when there is a tear in the fascial sheath or there is a tearing of vessels between the muscle fascicles. In this situation the blood disperses by gravity, to some distance away from the injury and, although the appearance is quite spectacular with considerable bruising, these injuries tend to recovery within 2 to 3 weeks (Fig. 4).

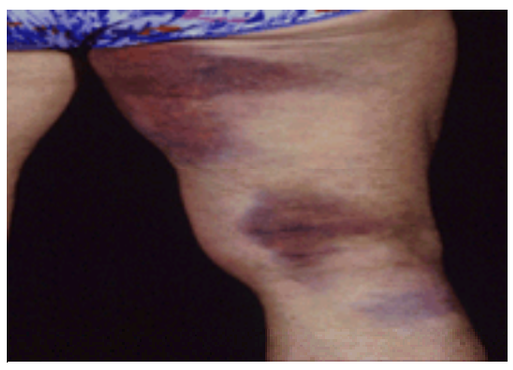


Fig. 4 Dramatic bruising in posterior thigh of rugby player caused by sprinting.

Symptoms depend to some degree on the type of injury. In a compression injury such as a rugby tackle, when the quadriceps muscle is compressed against the femur, the bleeding may be slow allowing the player to continue and only after the game does the increased swelling and accompanying pain and disability become apparent. An overstretching injury is sudden and acutely painful immediately. The commonest presentation is seen in sprinters who overstretch suddenly and tear hamstring muscles. Here the sudden acute pain feels like a blow on the back of the leg and may be severe enough for the athlete to fall to the ground. The diagnosis is usually fairly straightforward. Patients with a severe intramuscular haematoma of the hamstrings often present with a flexion deformity. There is some evidence that the greater the loss of flexion, the longer the time required for healing and return to sport (Renstrom 1988).

The clinical appearances and history are usually diagnostic. The presence of an intramuscular haematoma can be confirmed and treatment monitored by ultrasound.

Rehabilitation

Except in the most severe haematomas, early mobilization is indicated as it:

1. speeds the return of muscle strength;
2. improves the orientation of regenerating muscle fibres;
3. encourages recapillarization;
4. prevents disuse atrophy.

After 48 h of rest and ice, a rehabilitation programme should begin. This should initially consist of gentle stretching with the aim of restoring a full range of joint movement. The physiotherapist may also use various modalities of electrotherapy, such as pulsed magnetic energy, to assist healing. Stretching and strengthening programmes for all the muscle groups of the involved limb are essential, as some injuries may be due to a pre-existing lack of elasticity either in the injured muscle or its antagonist. The importance of balance between the muscle groups in the prevention of injury has been recognized and may be measured using an isokinetic machine, both for assessment and muscle retraining.

Cardiovascular fitness can be maintained by cycling (real bike or exercise bike), swimming, and running in water using a specially designed 'wet vest' to keep the athlete upright. Apart from the physical benefits of this exercise, being active in this way helps to maintain the sanity of the injured athlete. Many committed athletes become anxious and even depressed if exercise is not permitted.

Return to sport

A muscle haematoma can be considered to be completely healed when there is full and pain free muscle contraction. An intermuscular haematoma will usually resolve in 2 to 4 weeks. An intramuscular haematoma may take 8 weeks or more to completely heal, but there is a great variation depending on the site and severity of the injury. Before returning to competition, a further programme of sport-specific training must be carried out. This includes sprints, rapid deceleration, twists, and sharp turns. Co-ordination and proprioception are impaired by injury and the athlete must relearn specific techniques with the help of a coach.

Myositis ossificans

A severe crush injury may be followed by the development of heterotopic ossification or myositis ossificans. This is more likely to occur if the injury is aggravated by continuing activity and further bruising. Applications of heat and vigorous massage all appear to irritate an already irritable muscle, although the exact cause of the calcification is not clear. Suspicions are raised if the patient is unable to achieve full contraction of the affected muscle. The calcification may be seen on radiographs and ultrasound.

Various treatments apart from rest have been used, including high doses of indomethacin and diphosphonates (DeLee and Drez 1994), but with doubtful efficacy. The use of a calcium channel blocker—Diltiazem—may offer a safer alternative (Palmieri *et al*. 1995). Most cases resolve symptomatically over a period of 3 months, although the calcification may persist radiologically. Intervention is rarely required unless mobility or persistent flexion deformity persists for over 6 months.

Muscle ruptures

Muscles may rupture and, although the defects look dramatic (Fig. 5), they rarely need surgical repair unless the rupture is complete.

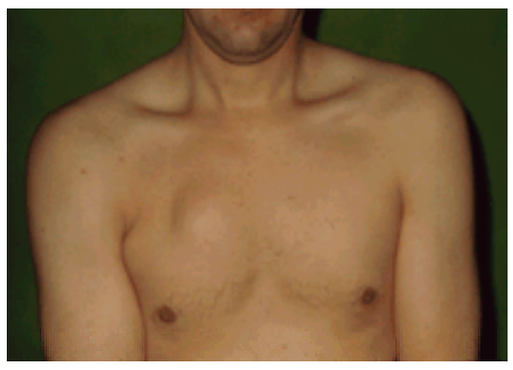


Fig. 5 Partial rupture of pectoralis major resulting from pulling against resistance in a rugby maul.

Chronic injuries

The vast majority of chronic sports injuries are to the lower limb, predominantly affecting the knee. A description of the majority of these injuries and their treatment can be found in other sections of this book, particularly the section on soft tissue rheumatism (Chapter 5.14). In this section a few of the commonest problems encountered in a sports injury clinic are described—further information can be found in specialist textbooks (Lachmann and Jenner 1994; Harries *et al*. 1994). A careful history is important for diagnosis, particularly to ascertain if there has been overuse or a sudden unaccustomed use of the lower limbs. Type and age of footwear as well as training surface must be enquired after, as well as taking a full history of the pain problem.

Examination must always start with the spine and include the whole upper or lower limb. Leg length inequality of more than 1 cm should be noted. Chronic lower limb sporting injuries are frequently associated with the 'malicious malalignment syndrome' with a broad pelvis, patellas that look towards each other, internal tibial torsion, and flat or hyperpronated feet. Examination should include the joints, soft tissues, and neurological system as well as skin and circulation. The examination is completed by inspecting the running shoes for signs of wear.

Correcting leg length inequality with a simple insole may resolve back problems. Simple shock absorbing insoles may also help many chronic overuse problems. The use of more elaborate insoles to correct hyperpronation and other foot problems are also popular but their success is unpredictable.

Stress fractures

In 1855, a German military physician, called Breihaupt, was the first to describe fractures of the foot bones in new recruits, not caused by direct trauma but by the stress of marching. Stress fractures can occur in any loaded bone and are seen most commonly in lower limbs, particularly in the tibia and metatarsals but they also occur in the upper limb, such as in the ribs of rowers and the arms of gymnasts (Fig. 6) (Matheson *et al*. 1987).

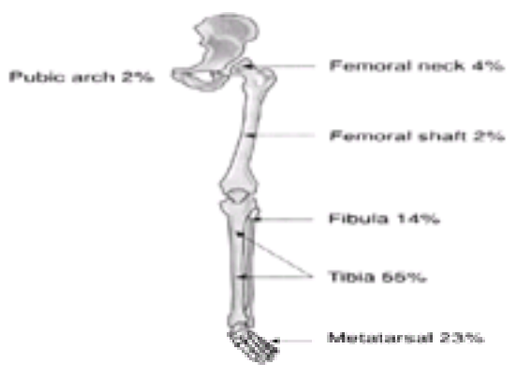


Fig. 6 Percentage of stress fractures occurring at different sites in the lower limb.

Stress fractures occur when bone fails to adapt to new or unusual loading. Normally, microdamage stimulates new bone where needed. When microdamage outpaces the development of new bone, a stress fracture results. Foot biomechanics may influence the development of stress fractures, such as rigid, poorly adapting feet or conversely hypermobile feet. Running on hard surfaces using 'collapsed' running shoes or racing flats certainly has an influence. Women appear to be at greater risk of developing a stress fracture. In a survey of 218 consecutive patients presenting with a stress fracture to our sports injury clinic almost half were women, although women only accounted for a quarter of all the patients seen in the clinic (Fig. 7). Some of the fractures may be due to an inadequate calorie and calcium intake in sports where leanness is favoured, especially if low food intake results in amenorrhoea and secondary osteoporosis. These factors do not entirely account for the high incidences in women.

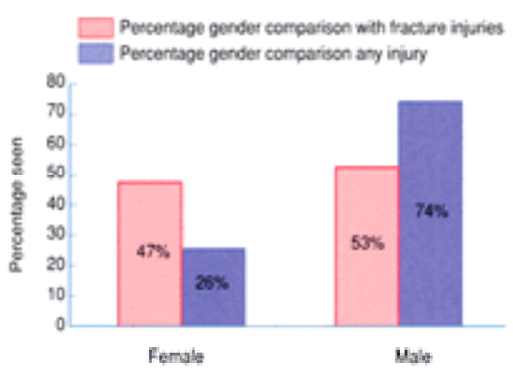


Fig. 7 Comparison of the incidence of stress fractures and all injuries between females and males attending a sports injury clinic.

Clinical features

In vigorous sports, such as high jumping and gymnastics, the onset of pain is sudden and acute, completely preventing the continuation of the activity. In other less explosive sports, such as long distance running, the onset maybe insidious over several days or even weeks, with pain present at first only during exercise and weight bearing but eventually aching at rest and during the night. The site of the pain is very localized and local pressure produces exquisite pain over the fracture, causing the patient rapidly to pull the injured limb away. A stress fracture may be diagnosed on these clinical grounds, but in order to persuade an exercise-addicted athlete to rest there is frequently a need for some corroboration.

Radiographs are not very helpful as they do not show the injury until it is healing and maybe not even then. Bone scans are extremely sensitive and can differentiate between osteitis of the tibia and a stress fracture. Although not absolutely necessary for diagnosis, a positive scan does help to convince the athlete of the nature of his problem and need for rest (Fig. 8).

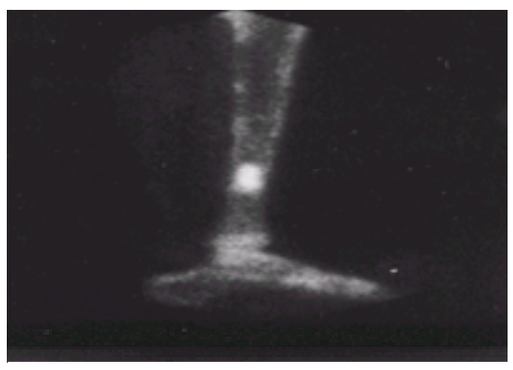


Fig. 8 Tc-99m bone scan showing high focal uptake typical of a stress fracture of the tibia.

Although the majority of stress fractures respond to conservative measures, stress fractures of the femoral neck have the potential for serious complications, including avascular necrosis and persistent non-union, and may need a surgical assessment (Fullerton and Snowdy 1988). These present with vague groin pain, pain in the anterior thigh and the knee, and are most commonly seen in long distance runners.

Management

Reduction of weight bearing is essential, usually with the help of elbow crutches. Modification of activity should continue until symptoms are absent. This can take 6 to 8 weeks and during this time the athlete can do some form of non or partial weight-bearing exercise, such as swimming, water walking, and cycling. Sport can be resumed when there is no pain on weight bearing and must be gradually increased. A stretching and strengthening programme should be carried out.

Using this regime, the majority of athletes will be able to return to sport in 6 to 8 weeks. Problems arise when the athlete is not prepared to rest. Under these circumstances the fracture can continue to cause problems for several months and sometime the only way to prevent continuing disability is to apply a plaster of Paris case and insist on no weight bearing. If not properly dealt with, a stress fracture can persist for 6 months or more.

Anterior knee pain

A patient presenting with poorly localized anterior knee pain, without any history of trauma, is a very common occurrence in a sports injury clinic. The patient is often teenage and a sports enthusiast, playing or training on a daily basis. However, older patients responding to advice to increase their exercise also present with this problem.

Aetiology

Patellofemoral joint stability depends on many factors (Table 2). Any of these factors can cause an imbalance of the patellofemoral joint. It has long been assumed that patellofemoral pain was caused by abnormal patella tracking but all of these anatomical variations are common in the general population and only cause problems when the patellofemoral joint is subjected to chronic overload by repetitive overuse and obvious eccentric contraction strains as in deep squats (Thomee *et al*. 1995).

Torsional deformities of tibia/femur (winking or frog eye patella)
High or lateral patella
Weak or absent vastus medialis obliquus
Tight muscles on the lateral aspect — vastus lateralis and iliotibial band
Increased quadriceps angle
Excessive foot pronation
Weak ankle dorsiflexors
Reduced range of motion at the ankle
Tight hamstrings

Table 2 Factors predisposing to patellofemoral pain syndrome

Clinical features

The patient complains of generalized anterior knee pain and may exhibit the 'grab' sign—holding the whole of the front of the knee. There is no history of trauma but more of a gradual increase in pain over several weeks, often affecting both knees. The pain is made worse by activity, either during sport or going up and down stairs. Sitting with the knee in flexion causes acute discomfort, such as sitting in a cinema, giving the so-called 'movie sign' (Insall 1982). The knee may feel unstable or 'give way'; this is due to reflex quadriceps inhibition rather than a true instability. Clicking and crepitus is common. There is pain and sometimes crepitus if the patella is compressed and moved proximally and distally. There is usually some tenderness along the very anterior joint line and at the attachment of the medial and lateral retinacula. Patella inhibition is elicited by holding the patella at the proximal pole and flexing the knee to about 20°. Straightening the leg causes pain and the patient stops the movement.

Patella subluxation, which can present with similar symptoms, is excluded by the patella apprehension test. The patella is usually very mobile and pushing it sideways produces great anxiety. Treatment is aimed at stabilizing the patella with appropriate quadriceps exercises.

Management

The patient, and often parents, need to be reassured that this complaint is common and does not mean indefinite incapacity. Radiographs may be required to exclude

any more serious pathology, especially if the problem affects one knee only. In mild cases, some reduction of the level of activity is all that is necessary—the patient decides what is tolerable. Activities not causing pain, such as swimming, can be substituted to maintain fitness. Local applications of ice after exercise and a course of non-steroidal anti-inflammatory drugs can hasten recovery.

Rehabilitation

When the acute symptoms have been alleviated, a strengthening and stretching programme involving hamstring and quadriceps is begun under the care of a sports-orientated physiotherapist. The treatment involves selectively strengthening the vastus medialis obliquus, while at the same time reducing the pull of the vastus lateralis. Strengthening of the quadriceps as a whole has been replaced by work only on vastus medialis obliquus (Shelton 1991). There is also some evidence that patellofemoral dysfunction may be related to abnormal timing in the firing of the various quadriceps components (McConnell 1986; Voigt 1991). Muscle re-education using biofeedback techniques during various exercises can be used where the patient is taught to fire the vastus medialis obliquus earlier and more efficiently. McConnell also developed external support for the patella by means of taping. Using these methods, while working on strengthening techniques, often provides immediate and fairly long-lasting relief.

Although the vast majority of patients will improve using these measures, it is important that they understand this is not a cure. Adolescent sufferers may well improve with the passage of time, as some of the imbalance may be associated with growth spurts. If there are persistent anatomical features, the patient will have to be prepared to continue the muscle strengthening exercises indefinitely in order to control their pain. Ten per cent of patients do not improve with the above measures; some may be prepared to modify their lifestyles but others may request surgery. Various surgical options are available, the most common being lateral release (Table 3).

Stage 1	*2–3 weeks*
Reduce symptoms by cutting down activity. No exercise involving deep squats. Relieve the pain with ice and non-steroidal anti-inflammatory drugs.	
Stage 2	*3–4 weeks*
With help of physiotherapist, strengthening exercises for vastus medialis obliquus. Flexibility of hamstrings, quadriceps, iliotibial band, gastrocnemius–soleus complex.	
Stage 3	
Gradual return to sport and then full training	
Stage 4	
Maintenance exercises for vastus medialis obliquus, at least 3 times weekly.	

Table 3 Rehabilitation programme for patellofemoral pain syndrome

Anterior shin pain induced by exercise ('shin splints')

Chronic pain in the shins induced by exercise (commonly known as 'shin splints') is a frequent occurrence in sportsmen and women but is rarely seen in any other group of patients. This syndrome embraces a variety of conditions:

1. stress fracture of the tibia or fibula;
2. fasciitis of tibialis posterior;
3. compartment syndrome;
4. popliteal artery stenosis;
5. referred pain from the spine (cord claudication);
6. peripheral vascular disease (intermittent claudication).

Fasciitis of tibialis posterior

Hyperpronation of the feet is often seen in association with other common malalignment problems. This problem puts extra stress on the tibialis posterior muscle which is responsible for inversion of the foot and inserts onto the posterior aspect of the tibia and fibula. The tendon of the muscle runs behind the medial malleolus and the flat hyperpronated foot is easily overstretched by repetitive activity resulting in a traction injury at the fascial insertion along the posterior border of the tibia.

Patients complain of pain in the shins which starts after starting to run. Often they can run through the pain only for it to return with great severity on ceasing activity. The pain often takes days to wear off. Examination reveals diffuse tenderness along the medial border of the tibia. The diagnosis may be confirmed with a bone scan.

Treatment is difficult. Rest is vital and is followed by a gradual return to sport with appropriate footwear. Orthotic supports to correct the hyperpronation and to provide shock absorption can be tried.

Compartment syndrome

Acute swelling of muscles within a muscle compartment leading to muscle ischaemia and even necrosis is a well recognized complication of acute trauma. The chronic form of this condition is less well known and occurs almost exclusively in endurance sports such as long distance running. As the muscles are exercised they swell within a tight fascial compartment raising the pressure sufficiently to reduce tissue perfusion. There are four separate compartments in the lower limb (Table 4). Compartment syndrome commonly follows a sudden increase in training. Symptoms are absent at rest but pain in the shin commences at a variable time after the onset of exercise and invariably increases until exercise is stopped or moderated. If the limb is rested the pain usually wears off within a few hours. Examination is often unremarkable although highly-developed calf muscles that feel tense at rest may be noted. A Tc 99 bone scan will exclude a stress fracture or fasciitis. Pressure studies can be performed by inserting catheters into each compartment and measuring the pressure rises. This is an invasive investigation and only available at specialist centres. A promising alternative is isotope scanning employing a technetium isotope Tc-99m-methylisobutylisonitrile (MIBI) which is used routinely to detect cardiac ischaemia and can be modified to look at peripheral muscle ischaemia. If the symptoms do not settle with rest, the appropriate compartments can be released by a surgical fasciotomy (Miles *et al*. 1992) (Fig. 9).

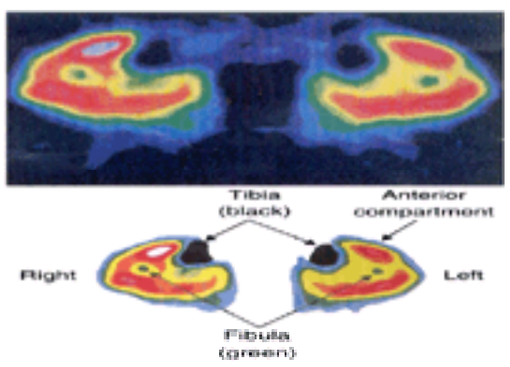


Fig. 9 Methylisobutylisonitrile emission tomogram of mid calf after exercise showing reduction in uptake in left anterior compartment compared with an annotated drawing below.

Compartment	Muscle
Anterior	Tibialis Anterior
Posterior	Gastrocnemius/soleus
Deep posterior	Tibialis posterior
Lateral	Peronei

Table 4 Compartments of the lower limb

Popliteal artery entrapment

Rarely, similar symptoms can be caused by entrapment of the popliteal artery by a fibrous band or hypertrophied head of gastrocnemius. Arteriography may be required to confirm the diagnosis.

Children and sport

Children are prone to similar injuries to adults but should not be regarded as 'little adults' as their physiology and anatomy are very different, resulting in injuries not encountered in the adult:

1. They take shorter, more shallow breaths and need more oxygen for their activity than an adult.
2. They also have lower glycogen levels and essential muscle enzymes and waste more energy.
3. Children perceive exercises as less fatiguing than adults and can almost exercise to destruction, but they also have the capacity to recover very quickly.
4. Muscle, tendon, and ligaments are stronger than bone until bony maturity is reached at 18 to 21 years.
5. Bones are still growing with active apophyses and epiphyses which can be the site of pathology.
6. Growth occurs in spurts, resulting in tight muscles and loss of flexibility making adolescents particularly prone to injury.

Avulsion fractures

Tendons, ligaments, and muscles, especially trained muscles, are stronger than bone and more so at the epiphyseal junctions. Sudden, intense loading of a muscle does not produce a muscle tear as in an adult but is more likely to cause an avulsion fracture where the bony attachment of the muscle or ligament is torn away. Common sites for avulsion fractures are the growth zones around the pelvis, including the attachments at the ischial tuberosity and the attachments of rectus femoris and sartorius. The less severe injuries can sometimes present only as a pain in the groin, with loss of function of the appropriate muscle. These injuries do not show up on radiographs and a bone scan is necessary. The usual treatment is rest followed by physiotherapy treatment, with an emphasis on restoring strength to the injured muscle. Healing may take up to 6 months.

Overuse injuries of the apophysis

In children and adolescents the muscle tendon attachment to bone or the apophysis presents as a 'high-risk' area for overuse injuries (Renstrom 1988). Overuse caused a traction apophysitis and occurs at various sites (Table 5).

Eponym	Site of injury
Osgood–Schlatters	Tibial tubercle
Sinding–Larsen, Johannason	Lower pole of patella
Severs disease	Achilles tendon attachment to calcaneum

Table 5 Sites of apophysitis

Osgood–Schlatter disease follows overloading of the patellar tendon at its attachment to the tibial tubercle. Very active boys, between the ages of 14 to 16 years and girls at a slightly younger age, present with swelling, pain, and acute tenderness over the tibial tubercle. These children are usually playing not only for the school, but taking part in other competitive sports on most days of the week.

The treatment of all the apophyseal overuse injuries is by modifying activity, playing only to a bearable level of discomfort, and using ice before and after exercise. Most respond to this regime but this may take several months and some persist until growing has ceased.

Osteochondritis

This is a collection of conditions affecting various sites of uncertain aetiology and is thought to be due to avascular necrosis of bone and subsequent flattening of the affected bones (Fig. 10; Table 6).

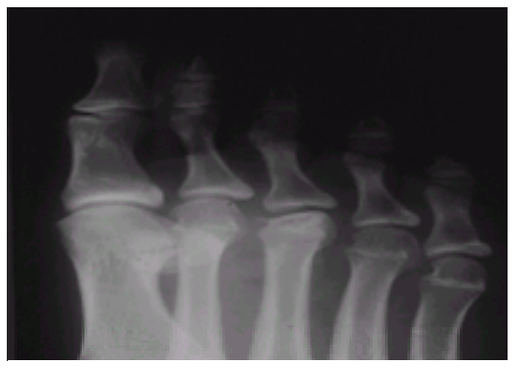


Fig. 10 Freiberg's disease showing flattening of the head of the 3rd metatarsal.

Name	Site
Perthes' disease	Hip
Scheuermann's disease	Vertebral ring epiphysis
Kohler's disease	Navicular bone
Freiburg's disease	2nd or 3rd metatarsal
Panner's disease	Elbow capitellum

Table 6 Sites of osteochondritis

Perthes' disease and the potentially disastrous unrelated hip problem of slipped femoral epiphysis may present with a limp and a painful knee. This can cause diagnostic problems as the child may present following an injury, drawing attention away from the real source of the problem. When faced with this history a radiograph is mandatory, as both conditions require urgent orthopaedic referral.

Scheuermann's disease mainly affects the thoracic vertebrae and because the anterior border of the vertebrae are subjected to deforming forces a kyphosis occurs. Often this is asymptomatic, although there may be some discomfort after hard exercise. In mild cases, flexibility exercises for the hip flexors and hamstrings and abdominal strengthening are sufficient until the disease ends spontaneously. Increasing kyphosis demands more serious intervention.

Osteochondritis dissecans

Osteochondritis dissecans is a disease of unknown aetiology, occurring in young people between the ages of 12 and 16 years. There is destruction and subsequent disintegration of cartilage and bone. Popular theories are trauma or continuous repetitive damage (Fairbanks 1993), ischaemia, and genetic factors. In the United States, the condition is seen as one of the entities in 'little league elbow', seen in young baseball pitchers.

The knee is a common presenting site and the history is of a diffuse knee ache, made worse by activity and accompanied by occasional effusion. When a fragment of bone becomes loose in the joint the patient may present with an episode of locking and effusion. The diagnosis can be confirmed by radiography. Surgical replacement of the loose fragment can give excellent results (Fig. 11).

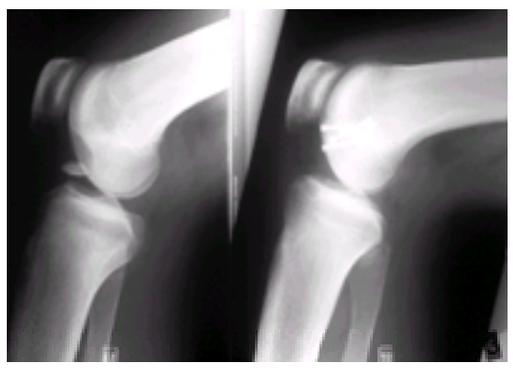


Fig. 11 Osteochondritis of femoral condyle before and after surgical intervention.

Chapter References

Budgett, R. (1995). *The overtraining syndrome*. *ABC sports medicine*. British Medical Journal Publications, London.

Coleman, P, Munro, J., Nicholl, J., Harper, R., Kent, G., and Wild, D. (1996). *The effectiveness of interventions to prevent accidental injury to young persons aged 15 to 24*. Medical Care Research Unit, University of Sheffield.

Costill, D.L. and Hargreaves, M. (1992). Carbohydrate nutrition and fatigue. *Sports Medicine*, **12**, 86–92.

DeLee, J.C. and Drez, D. (1994). *Orthopaedic sports medicine*. Vol II. WB Saunders, Philadelphia.

Fairbanks, H.A.T. (1993). Osteochondritis dissecans. *British Journal of Surgery*, **21**, 67.

Fullerton, L.R. and Snowdy, H.A. (1988). Femoral neck stress fractures. *American Journal of Sports Medicine*, **16**, 365–77.

Harries, M., Williams, C., and Stanish, W., eds. (1994). *Oxford textbook of sports medicine*. Oxford University Press, Oxford.

Insall, J. (1982). Chondromalacia patellae. Current concepts and review patella pain. *Journal of Bone and Joint Surgery*, **64(A)**, 147.

Jennings, D.C. (1990). Injuries sustained by users and non users of gum shields in local rugby union players. *British Journal of Sports Medicine*, **24**, 159–65.

Lachmann, S.M.L. and Jenner, J. (1994). *Soft tissue injuries in sport*. Blackwell Science, Oxford.

Matheson, G.O., Clement, D.B., McKenzie, D.C., Tavnton, J.E., Lloyd Smith, D.R., and MacIntyre, J.G. (1987). Stress fractures in athletes. *American Journal of Sports Medicine*, **15**, 46–58.

McConnell, J. (1986). The management of chondromalacia patellae: a long term solution. *Australian Journal of Physiotherapy*, **2**, 215–23.

Miles, K. *et al*. (1992). Leg muscle scintigraphy with 99Tc MIBI in the assessment of peripheral vascular arterial disease. *Nuclear Medicine Communications*, **1**, 593–603.

Newsholme, E.A. (1994). Biochemical mechanisms to explain immunosuppression in well-trained and overtrained athletes. *International Journal of Sports Medicine*, **15** (Suppl. 3), 5142–7.

Nicholl, J.P., Coleman, P., and Williams, B.T. (1991). *Injuries in sport and exercise*. London Sports Council.

Palmieri, G.M., Secen, J.I., Aelion, J.A., Moinuddin, M., Ray, M.W., Wood, G.C., and Leventhal, M.R. (1995). Treatment of calcinosis with diltiazem. *Arthritis and Rheumatism*, **38**, 1646–54.

Renstrom, P. (1988). *Olympic text book of sports medicine*. Vol I. Blackwell Scientific Publications.

Report of the Royal College of Physicians (1991). *Medical aspects of exercise. Benefits and risks*. Royal College of Physicians of London, London.

Shelton, T. (1991). Rehabilitation of patellofemoral dysfunction; A review of the literature. *Journal of Orthopedic Sports Physiology and Therapy*, **144**, 243–9.

Thomee, R., Renstrom, P., Karlsson, J., and Grimby, G. (1995). Patellofemoral pain syndrome in young women. *Scandinavian Journal of Medicine and Science in Sports*, **5**, 237–44.

Torg, J.S., *et al*. (1985). The national football head and neck injury registry. *Journal of the American Medical Association*, **254**, 3439–43.

Voigt, W. (1991). Comparative reflex response times of vastus medialis obliqus and vastus lateralis in normal subjects and subjects with extensor mechanism dysfunction. *American Journal of Sports Medicine*, **19**, 131–7.

Wyatt, P. *et al*. (1996). A prospective study of rock climbing injuries. *British Journal of Sports Medicine*, **30**, 148–50.